Early Numerical Experiences

Amy Bennett

A thesis submitted for the degree of

Doctor of Philosophy

Loughborough University

September 2017

Abstract

There are large individual differences in children's mathematical abilities when starting formal schooling and these differences can have lasting consequences. One factor that could lead to differences in children's mathematics skills is the home numeracy environment. This thesis examines the home numeracy environment, firstly as a whole concept and then more in-depth of one area of the home numeracy environment, number books.

The home numeracy environment section starts by presenting a systematic review of the home numeracy environment literature and draws conclusions about the inconsistency of the results. The studies presented in this section investigate both methodological and theoretical questions surrounding the home numeracy environment. A novel text message method to measure the home numeracy environment is presented and the relationship between three different measures of the home numeracy environment (questionnaire, observation and text messages) is investigated, as well as their relationships to mathematics skills. This section has two key findings: firstly the self-report measures of the home numeracy environment are not related to the observation measure and secondly all three measures (apart from child number talk in the observation) were not related to mathematics skills.

The second section of this thesis focuses on number books. Number books are often used in the home to teach young children number symbols. They primarily use multiple concrete pictures, but the benefits (or costs) to using these types of images are not known. The next three studies investigate the use of abstract and concrete images to teach children number symbols using an artificial symbol learning paradigm. It is concluded that there is a cost to using multiple representations when teaching children number symbols, and therefore number books should use a single picture throughout for children to benefit the most from the book.

Overall the findings from this thesis show that the home numeracy environment is very broad and future research should change the way the home numeracy environment is measured and conduct more in-depth analysis of areas of the home numeracy environment.

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost I would like to thank my supervisors Camilla Gilmore and Matthew Inglis for their support, guidance and encouragement throughout my whole PhD. I am thankful for the huge amount of time and effort they have put in to my development as a researcher and I could not have asked for better supervisors.

Thank you to the whole mathematics education department for creating such a friendly and positive place to work and in particular the Mathematics Cognition Group (MCG) for all your support and feedback.

A special thank you to all the 500+ parents and children that took part in my studies. Without their support and enthusiasm this research would not be possible. Also many thanks to Nottingham University for giving me the opportunity to run studies at their 'Summer Scientist Week' event.

Finally my thanks go out to my family and friends, in particular, my mum, Debby, and dad, Dave and my sisters, Katie and Holly for always believing in me. A special thanks goes to my husband, Matt, for his endless supply of support, encouragement and patience throughout my whole PhD experience.

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 for this research. I will start by presenting why mathematics skills are important and particularly why we should investigate mathematics skills in young children. Next, I will present the factors that are commonly used to explain individual differences in young children's mathematics skills. Finally I will present an overview of the aims and structure of this thesis.

1.1 The importance of mathematics skills

The society we live in requires everybody to understand number. Everyday, when working, shopping or just communicating, we make decisions based on quantitative information. Number skills can have an impact on quality of life and job prospects, with higher risks of depression and unemployment linked to low numeracy skills (Parsons & Bynner, 2005). However, given the importance of numeracy for gains in our society, as well as personal health and well being, numeracy skills are still an area for concern in the UK today.

Researchers have shown that children begin formal schooling with different abilities, particularly in numeracy, (Ginsburg, Lee, & Boyd, 2008) and these differences predict later mathematical achievement (Duncan et al., 2007; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Taggart, Sylva, Melhuish, Sammons, & Siraj, 2015; Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Melhuish, Sylva, et al., 2008). Children who start school with poor numeracy knowledge are unlikely to catch up with their peers (Jordan, Kaplan, Locuniak, & Ramineni, 2007). Therefore, in recent years there has been an increased focus on children's mathematical skills in their early years.

Heckman's 2013 book, 'Giving Kids a Fair Chance' emphasised that

> "To foster individual success, greater equality of opportunity, a more dynamic economy and a healthier society, we need a major shift in social policy toward early intervention "(p. 6)

This statement helps to underline the importance of research into developing early years interventions.

Over the years there has been an increase in government intervention in young children's development with the production of the Early Years Foundation Stage (EYFS) framework in 2008, with further updates in 2012 and 2014 (Department for Education, 2014). The EYFS framework sets standards for the learning, development and care of children from birth to 5 years old. This includes a section on numeracy detailing the skills that children should have developed by the age of 5:

> Numbers: children count reliably with numbers from 1 to 20, place them in order and say which number is one more or one less than a given number. Using quantities and objects, they add and subtract two single-digit numbers and count on or back to find the answer. They solve problems, including doubling, halving and sharing.

As well as the government intervention, private charities are focusing on numeracy for young children. For example, National Numeracy are running a project focusing on parental engagement in numeracy in a bid to improve the number skills of children throughout the UK (National Numeracy, 2015).

Even though the importance of early numeracy skills is starting to be recognised, there is still a lot to be done in terms of research. It is not clear what interventions would be most beneficial to improve early years numeracy skills. The individual differences in performance when starting school requires further investigation.

1.2 Individual differences in mathematics skills

Because individual differences in young children's mathematics skills appear to have lasting consequences, it is important we try to understand the root of the differences. Current research has highlighted four factors:

1. Domain-specific cognitive abilities
2. Domain-general cognitive abilities
3. Dispositional factors
4. Environmental factors

1.2.1 Domain-specific cognitive abilities

Domain-specific factors are cognitive skills which specifically predict mathematics achievement. Over the past 20 years, the main focus of the research in this area has shown a link between symbolic and non-symbolic numerical magnitude processing skills and mathematics performance (De Smedt, Nol, Gilmore, & Ansari, 2013), in particular focusing on the Approximate Number System (ANS). Tasks involving the ANS are solved using approximate representations of quantity, instead of exact calculations (see Fazio, Bailey, Thompson, & Siegler, 2014 for a review). The ANS is usually measured by a dot comparison task. However, it has been suggested that the dot comparison task involves domain general skills (Clayton & Gilmore, 2014) and therefore this could be driving the relationship. This area of research has received a lot of attention in the last few years.

1.2.2 Domain-general cognitive abilities

Domain general factors are those cognitive skills which predict achievement in many subject areas. It is well-established that formal mathematics abilities are influenced by domain-general skills such as IQ and executive functioning (Bull, Espy, & Wiebe, 2008; Sowinski et al., 2015). Executive functioning is a term used for a range of cognitive processes such as working memory, inhibitory control and planning.

1.2.3 Dispositional factors

Dispositional factors are internal attributes such as attitudes, beliefs and motivations. One dispositional factor, specifically related to numeracy, is children's Spontaneous Focus On Numerosity (SFON). SFON is a developing concept, but research has shown links between SFON and mathematics achievement (Hannula & Lehtinen, 2005; Hannula, Lepola, & Lehtinen, 2010). Another dispositional factor that can be influential on mathematics performance is maths anxiety. This is a well-established area of research which has shown links to mathematics achievement. However, research is continuing into the development of maths anxiety, specifically in early childhood (Ramirez, Gunderson, Levine, & Beilock, 2013).

1.2.4 Environmental factors

Environmental factors are aspects of a child's environment that can influence mathematics skills. The role of environmental factors has been shown to be important for children developing cognitive skills (Wachs, 1996; Petrill, Pike, Price, & Plomin, 2004; Cunha & Heckman, 2007).

Much of the work in this area has focused on language and literacy development. It has also been well-established that early experiences with language at home are associated with language growth. Sénéchal and LeFevre (2002) developed a home literacy environment model, showing two distinct pathways linking children's experiences to their literacy skills. One pathway showed informal home experiences, such as shared reading with parents, relates to vocabulary knowledge and indirectly to reading ability. The other pathway showed formal home experiences, such as teaching children specific skills (e.g. letter recognition), predicted word reading. This research has developed over the years (Evans & Shaw, 2008; Rodriguez & Tamis-LeMonda, 2011; Niklas & Schneider, 2013b; Sénéchal & LeFevre, 2014) clearly showing the importance of the home literacy environment to young children's literacy skills. This has led to intervention schemes such as 'The Reading Agency' (Reading Agency, 2013) and the 'National Literacy Trust' (Trust, 2017).

However, there are far fewer studies looking at the home numeracy environment and the evidence for a link between the home numeracy environment and mathematics achievement is mixed (LeFevre et al., 2009; Kleemans, Peeters, Segers, & Verhoeven, 2012; Skwarchuk, Sowinski, & LeFevre, 2014; Segers, Kleemans, & Verhoeven, 2015; Ramani, Rowe, Eason, & Leech, 2015).

In addition to looking into the home environment, researchers have also investigated the impact of pre-school on children's number performance (Sylva, Melhuish, Sammons, Siraj-Blatchford, Taggart, Smees, et al., 2004; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006; Anders et al., 2012). It has been shown in a longitudinal European study, entitled Effective Provison of Pre-School Education (EPPE), that pre-school experience enhances children's development. Further to this, intellectual development, independence, concentration and sociability are all linked to the age children start pre-school. In particular, disadvantaged children can benefit significantly from high quality pre-school experiences. The quality of pre-school centres is directly related to better intellectual, cognitive, social and behavioural development in children (Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2004). With the rise of government funding for children to attend pre-school and the introduction of the Early Years Foundation Stage (EYFS), research into pre-school education is becoming more and more important.

The EPPE project investigated the impact of the home learning environment and demographic variables, as well as pre-school education. They found that pre-school attendance had a significant effect on mathematics performance up to the age of 16. However, the home learning environment, measured at age 4, had a bigger effect than pre-school attendance at age 7, 11, 14 and 16 (having an effect of .45 at age 16 compared to the .21 effect of pre-school attendance) (Taggart et al., 2015).

This shows the importance of the home learning environment over and above the effects of pre-school, emphasing the need for further research specifically into the home numeracy environment. It is important that this area of research is developed further to confirm the relationship between the home numeracy environment and numeracy skills, and to develop this research area to mirror the effectiveness of the home literacy environment research. This will be the focus for this thesis.

1.3 Overview of current thesis

This thesis investigates the relationship between the home numeracy environment and mathematics achievement in early childhood years. The thesis is split into two parts with the first part investigating the relationship between the general home numeracy environment and mathematics achievement and the second part providing a more in-depth analysis of one part of the home numeracy environment and it's effectiveness in teaching mathematics skills.

Part II of this thesis focuses on the home numeracy environment and its relationship with mathematics performance. Firstly, Chapter 2 presents a systematic review of the home numeracy environment literature and discusses the findings from this literature. Next, Chapter 3 analyses the relationship between home numeracy environment and mathematics achievement using data from the Millennium Cohort Study. Chapters 4 and 5 present empirical studies that investigate a new measure of the home numeracy environment and compare this to existing measures, while also evaluating the relationship between these measures and mathematics performance. Finally, Chapter 6 presents a confirmatory factor analysis of the questionnaire measure and re-analyses the studies presented in Chapter 4 and 5 using the formal and informal factors of the questionnaire. Following these empirical chapters, Part II closes with a general discussion of the home numeracy environment in Chapter 7.

Part III of this thesis focuses on a more in-depth analysis of the home numeracy environment by investigating children's number books. Chapter 8 evaluates the literature in the area of number books, abstract and concrete representations and children's early number learning. Chapter 9, 10 and 11 present empirical findings on the use of abstract and concrete representations for children learning number symbols. Chapter 12 then reviews and discuss the findings from Part III.

In Part IV, Chapter 13 concludes this thesis by bringing together the findings from Part II and III and outlines future directions for the home numeracy environment research.

Part II

Home Numeracy Environment

Chapter 2

Home numeracy environment literature review

The aim of this chapter is to present an overview of the literature on the home numeracy environment, in order to provide a background for studies presented later in this part of the thesis. I will begin by describing the current literature relating to the home numeracy environment and how it has been measured. I will move on to discuss variables that have been shown to impact the home numeracy environment. I will then present a systematic review of the relationship between the home numeracy environment and mathematics achievement, including a p-curve analysis. I will then discuss intervention studies that have looked at improving the home numeracy environment. Finally, I will describe the aims and research questions addressed in the first part of this thesis.

2.1 Defining home numeracy environment

To begin, it is important to define the term "home numeracy environment" because this term can be used to cover a wide variety of activities. This ranges from help parents give children with homework to number activities parents do with their children at home. I will use the term "home numeracy environment" to refer to number activities that parents do with their young children before they start formal schooling. This does not include any formal teaching children may receive from nurseries or pre-schools.

2.2 Measuring the home numeracy environment

The majority of current research into the home numeracy environment focuses on the frequency of activities in the home. Some researchers have looked at the combined frequency of literacy and numeracy activities as one home learning environment measure. It has been suggested that the home literacy environment is just as important as the home numeracy environment for developing mathematics skills, because children need to develop language skills in order to develop numeracy skills. Lefevre, Clarke, and Stringer (2002) combined literacy and numeracy into a parent teaching variable and found parent teaching positively correlated with counting and a number recognition task. Furthermore, Melhuish, Phan, et al. (2008) used a combined home learning environment measure and reported an effect size of .65 of the home learning environment in a model predicting numeracy skills for 5 years-old and an effect size of .50 for the model predicting numeracy skills at 7 years-old. These results suggest a rich general home learning environment is beneficial to mathematics performance but it is not certain which aspects of a general home learning environment are beneficial for mathematics. However, Segers et al.'s (2015) study looked into this specifically and found that the home literacy environment was not a predictor of early numeracy, when child factors, such as phonological awareness and working memory were taken into account. Therefore, it is important that the frequency of the activities in the home numeracy environment are considered separately to the frequency of activities in the home literacy environment. The following studies specifically investigate number activities.

The home numeracy environment is usually measured by self-report because this is a quick and convenient method. However, there are problems related to this method (see Chapters 3 & 4).

One of the first self report methods to quantitatively measure the frequency of number activities in the home numeracy environment was used by Blevins-Knabe and Musun-Miller (1996). Parents were interviewed over the phone with regards to how often they had taken part in certain number activities with their child in the last week. They found a significant positive correlation between the frequency of four home numeracy activities and mathematics performance on a standardised test. However, there was also a significant negative correlation between four number activities and mathematics performance. This led to mixed opinions about the benefits of a rich home numeracy environment. Blevins-Knabe, Austin, Musun, Eddy, and Jones (2000) conducted a similar experiment using the telephone interview method. This study failed to show a relationship between reported frequency of number activities and mathematics performance on a standardised test.

LeFevre et al. (2009) developed a home numeracy environment questionnaire designed to measure the frequency of number activities, as well as parents' attitudes towards mathematics and their expectations for their child. A factor analysis revealed a distinction between formal activities (directly teaching their child specific number skills) and informal activities (indirectly involving their child in numerical content). This study showed that the frequency of informal activities, such as games, were significantly positively related to children's mathematics knowledge and fluency. However, the frequency of formal activities, such as teaching number skills, were not significantly correlated to mathematics performance. In contrast, the frequency of number book activities was significantly negatively correlated with mathematics fluency.

Adaptations of the questionnaire have been used in further studies of the home numeracy environment. Both Kleemans et al. (2012) and Vandermaas-Peeler and Pittard (2014) found a significant positive correlation between frequency of all number activities (measured by an adapted version of LeFevre et al.'s (2009) questionnaire) and mathematics performance (r = .47, r = .57, respectively). However, Missall, Hojnoski, Caskie, and Repasky (2014) also used an adapted measure of the questionnaire, but failed to find a significant correlation between the frequency of activities and mathematics performance on both mathematics measures they used.

Some studies have also split the activities into formal and informal activities, as in the original questionnaire. For example, Skwarchuk et al. (2014) found that informal numeracy activities (game exposure) significantly correlated with non-symbolic arithmetic (e.g. asking children 'how many animals have been put into the barn altogether?' without the child seeing inside the barn), and advanced formal numeracy activities correlated significantly with symbolic number knowledge (e.g. counting, ordering numbers, number identification). However, basic formal numeracy did not significantly correlate with non symbolic arithmetic or symbolic number knowledge. Ramani et al. (2015) also spilt the activities in the questionnaire into categories (direct teaching, applications of number, and games). They found a significant correlation between direct teaching activities with both foundational (r = .55) and advanced number knowledge (r = .46). There was also a significant correlation between games and foundational number knowledge (r = .35). However, all other number activities did not correlate significantly with number knowledge.

There have also been longitudinal studies into the effects of frequency of number activities on future mathematics performance. In Greece, Manolitsis, Georgiou, and Tziraki (2013) measured the frequency of number activities using a questionnaire at the beginning of kindergarten. They found that the frequency of number activities predicted children's counting performance at the beginning of kindergarten. However, frequency of number activities did not predict counting at the end of kindergarten. This suggests that there were no longitudinal effects. On the other hand, in China, Ciping, Silinskas, Wei, and Georgiou (2015) also used a questionnaire to measure the home numeracy environment and showed that there was a negative relationship between formal home numeracy environment activities and mathematics skills at Grade 1 ($r = -.18$) and that children's mathematics skills in Grade 1 negatively predicted frequency of formal home numeracy activities in Grade 2. These studies show the uncertainty about the long-term benefits of the home numeracy environment and that the benefits may vary in different countries.

Many studies have investigated variables that could affect the relationship between the frequency of number activities and mathematics performance. Below, I will discuss the evidence for relationships between demographic variables, parents' expectations and attitudes and the home numeracy environment as well as cultural differences in the home numeracy environment.

2.3 Demographic variables, parents' expectations and attitudes, cultural differences and the home numeracy environment

Demographic variables, parents' expectations for their children and parents' attitudes towards number have all been shown to influence the home numeracy environment. Many papers have looked at demographic variables, expectations and/or attitudes to investigate any mediating effects the home numeracy environment may have between these variables and mathematics skills.

2.3.1 Demographic variables

Demographic variables, in particular Socio-Economic Status (SES), have been linked to mathematics performance (Ribner, 2014). However, the home learning environment could mediate the relationship between demographic variables and mathematics performance.

Galindo and Sheldon (2012) conducted a hierarchical regression investigating the specific demographic variables which predicted the general home learning environment. They found that parents' educational level, language, age at kindergarten entry, gender and number of siblings all predicted the frequency of the home learning activities. Dilworth-Bart (2012) also found that maternal education, income and SES were related to the general home learning environment. These studies show that demographic variables can influence the general learning environment, but it is important to see if the same demographic variables are related to the more specific home numeracy environment.

Many studies investigating the home numeracy environment have included an SES measure. DeFlorio and Beliakoff (2014) showed that children from a high SES background participated in more number activities at home than children from a low SES background. Other studies have shown a positive relationship between SES and frequency of home numeracy activities (i.e. children from a high SES background do more number activities at home) (Levine, Suriyakham, Rowe, Huttenlocher, & Gunderson, 2010; Kluczniok, Lehrl, Kuger, & Rossbach, 2013). Furthermore, several studies have found a specific link between informal home numeracy activities and SES (Skwarchuk et al., 2014; Ciping et al., 2015) reporting that children from a high SES background do more informal number activities than children from a low SES background. In contrast, there are also several studies that failed to find a relationship between SES and the home numeracy environment (LeFevre, Polyzoi, Skwarchuk, Fast, & Sowinski, 2010; Lukie, Skwarchuk, LeFevre, & Sowinski, 2013; Manolitsis et al., 2013; Ramani et al., 2015) and one, to my knowledge, that found a negative relationship (Niklas & Schneider, 2013a).

The conflicting findings could be due to variation in measurement of SES. All of the studies mentioned have used different combinations of SES measures. Liberatos, Link, and Kelsey (1987) stated there isn't one best measure of SES and they found the most commonly used indicators are occupation, education and income. Furthermore, it has been noted that SES is usually measured on individual characteristics, however it is important to also consider the neighbourhoods of the individuals as another measure of SES. The neighbourhood can influence the social resources available to the individual and characterise aspects of living conditions that are not captured by individual measures (Bornstein & Bradley, 2014). Smith and Graham's (1995) study found that in family research, one SES variable may have considerably more power than another for a specific outcome. Therefore, it is important that future studies consider carefully which measure of SES to use in relation to the home numeracy environment.

Child age has also been considered when investigating frequency of activities. Some papers have shown that older children tend to do more complex number activities (Skwarchuk, 2009; Skwarchuk et al., 2014). However, most papers fail to control for age when correlating the home numeracy environment and mathematics achievement (Esplin et al., 2016; Ramani et al., 2015; Skwarchuk et al., 2014; Niklas, Cohrssen, & Tayler, 2016; LeFevre et al., 2009).

Other demographic variables that have shown to be significantly related to the frequency of number activities are gender (LeFevre et al., 2010) and age at entry to non-parental care (Kluczniok et al., 2013). Age at entry into non-parental care is negatively correlated with frequency of activities, suggesting the earlier children go to child care, the less number activities parents do with their child at home. This could either be because parents have less time to do activities with their child , or because they believe the child is receiving sufficient educational input elsewhere.

Links between demographic variables and the home numeracy environment varies between studies, with some papers finding significant relations that other papers fail to find. Therefore, it is important to consider all possible demographic variables and ways in which they are measured in future studies.

2.3.2 Expectations

Parents' expectations for their children could influence the home numeracy environment. Parents' expectations are often measured by asking parents their beliefs about what children should be able to do before starting school. For example, questionnaires may ask parents how important they believe it is for children to be able to count to 10, know simple sums or know multiplying before starting school. If parents have high expectations for their child they are more likely to do more number activities to help them improve their number skills. Several studies have found links between parents' general curriculum expectations for their child and the frequency of number activities in the home (Lefevre et al., 2002; Kluczniok et al., 2013; Skwarchuk et al., 2014). Other papers have measured specific numeracy expectations and have also found a link to frequency of activities (Segers et al., 2015). Galindo and Sheldon (2012) showed that several demographic variables also predicted parents' general expectations for their child, however no other papers measuring both expectations and demographic variables have found any significant correlation, suggesting expectations are not related to demographic variables and should be considered individually when measuring the home numeracy environment.

2.3.3 Attitudes

Parents' attitudes towards mathematics are also an important aspect of the home numeracy environment. In some studies parents are asked about their own feelings towards mathematics and how important they feel mathematics to be. It has been found that the more positive parents' attitudes towards mathematics, the more number activities they do at home (Musun-Miller & Blevins-Knabe, 1998; Sonnenschein et al., 2012; LeFevre et al., 2010). Skwarchuk et al. (2014) found that parents' attitudes were only related to basic formal activities, suggesting that parents' attitudes are only important to formal teaching of numbers and do not influence informal activities.

2.3.4 Cultural differences

Cross-cultural studies into the frequency of number activities have shown differences between countries and languages. Pan, Gauvain, Liu, and Cheng (2006) conducted a study comparing American and Chinese parents reported frequency of activities, showing that Chinese parents spend more time doing mathematics activities with their children than American parents. LeFevre et al. (2010) compared Greek and Canadian parents reported frequency of activities, showing that Canadian parents reported greater frequency of sorting activities and using computer software, whereas Greek parents reported greater frequency in playing board or card games. Furthermore, Lefevre et al. (2002) compared French- and English-speaking Canadian parents and found that English speaking parents reported higher frequencies of number activities. This shows that there are key differences in the amount of number activities parents do with their children in these countries, and highlights the importance and need for specific research into the home numeracy environment in the UK.

2.3.5 Summary

There are many studies investigating the impact of different variables on the home numeracy environment and most studies have found some relationship between the frequency of activities and mathematics performance. However, the strength of this relationship varies between studies, with some researchers even reporting negative correlations. More work needs to be done to determine the strength of the overall relationship between the home numeracy environment and mathematics achievement. Therefore in the next section, I will present a systematic review of the home numeracy environment literature that investigates the relationship between the home numeracy environment and mathematics achievement.

2.4 Systematic review of the relationship between the home numeracy environment and mathematics achievement

The aim of this systematic review is firstly to identify and combine all the current available research on the home numeracy environment and mathematics performance, and evaluate the evidential value of this research using a p-curve analysis. Secondly, the review should provide evidence for the relationship between the home numeracy environment and mathematics achievement, and how the relationship differs depending on the measures used, age of the participants and country the study took place.

There is currently only one published review of the home numeracy environment literature (Dunst, Hamby, Wilkie, & Dunst, 2017) which is included in the book "Engaging Families as Children's First Mathematics Educators". This review includes a meta-analysis of the relationship between home and family experiences and numeracy learning. Dunst et al.'s (2017) review used 11 studies and found an overall effect size of .46 for the relationship between the home numeracy environment and mathematics achievement.

However, this is a very small sample of studies to conduct a meta-analysis and there were key studies that were not included in the review but met the authors' inclusion criteria, such as Missall et al. (2014) and LeFevre et al. (2009). Both of these studies find non-significant results which could have impacted the overall effect size. Furthermore, the review included combined home numeracy and literacy measures whereas, for my systematic review, I am only interested in home numeracy measures.

Finally, the majority of research papers in this area have more than one effect size and there are several ways these could be combined. In Dunst et al.'s (2017) meta-analysis it is not clear which effect sizes they have used from each study and how the overall effect sizes have been calculated (Lakens, 2017). Therefore, it is possible that a meta-analysis of the same studies could produce a different overall effect size.

Aside from the flaws of this meta-analysis, it has been debated if a meta-analysis in general is a valid method for calculating a 'true' effect size. The goal of a meta-analysis is "to estimate the overall, or combined effect" (Borenstein, Hedges, Higgins, & Rothstein, 2011) and meta-analyses have been used to combine effect sizes across the mathematics education literature in high-profile research areas (Higgins et al., 2013; Hattie, 2009). Meta-analyses have the advantage of giving an overall effect size, which is often called the true effect size, for a particular relationship or effect. However, while this is a useful piece of information to have, researchers are starting to debate how the technique is used in the education literature, where studies in the same area of research are often combined even when they are using different methods to answer slightly different research questions. Simpson (2017) has argued that effect sizes are open to researcher manipulations and, cruicially, researchers employing more sensitive designs may produce bigger effect sizes. Therefore, combining these effect sizes would not give a true effect size, it would just show how sensitive the designs used to address the specific questions are. Furthermore, when conducting a meta-analysis the effect sizes from the studies are given a weight and deciding how the weights should be assigned is problematic. Two meta-analyses using the same studies could produce different results depending on the weights given to the studies (Simonsohn, 2015). Another major criticism of meta-analysis research is that they can never fully control for publication bias. Many studies that find non-significant results are not published and therefore are not included in the meta-analysis resulting in a skewed view of the research in that area. Because of these flaws with the meta-analysis technique in the education sector, I have decided not to conduct a meta-analysis for the home numeracy environment studies, instead I conducted a p-curve analysis to show the evidential value of the studies in the home numeracy environment literature.

A p-curve analysis aims to investigate if there is evidential value in a set of studies (Simonsohn, Nelson, & Simmons, 2014; Simonsohn, Simmons, & Nelson, 2015). In other words, it tells us if we were to repeat the studies in exactly the same way the authors did if we would detect an effect. This is achieved by graphing the p-values of the studies. If the null hypothesis is true, we expect the p-values to have a uniform distribution. However, if the null hypothesis is false the p-values would be right skewed (more lower p-values that higher p-values). This is true for all p-values between 0 and 1 but also true for the p-values between 0 and .05. Simonsohn et al. (2015) suggests that by investigating the shape of the distribution of p-values, between 0 and .05, from a set of studies we can determine if those studies have evidential value.

The benefit of p-curve analysis over meta-analysis is that we do not need to worry about publication bias. For meta-analysis studies to be completed correctly the authors should collate all research in the area, however a lot of studies that have found non-significant results are not published and therefore not accessible to the author conducting the meta-analysis, giving an incomplete picture of the research area. The p-curve analysis avoids this problem by only looking at the p-values between 0 and .05 (i.e. the studies that have found significant results).

The main research question for this review will be "What is the relationship between the home numeracy environment and mathematics achievement?". I looked for the evidential value of the results of these studies and also investigate how this relationship varies when different measures are used.

2.4.1 Method

2.4.1.1 Inclusion / Exclusion Criteria

Before beginning the search, I decided upon the inclusion and exclusion criteria for the studies. These criteria were decided upon so that I had the most relevant research in order to answer the research question. For a study to be included in the review the study must meet the following criteria:

1. The study must be published after 1995.
2. The study must include a quantitative measure of the home numeracy environment.
3. The study must include at least one mathematics measure.
4. The study must statistically report the relationship between the home numeracy measure and the mathematics measure.
5. The participants of the study must be under 8 years old.
6. The study must be written in English.
7. The study must be published and accessible.

It was decided to only include studies published after 1995 as I wanted to include the most recent studies on the home numeracy environment. The numeracy activities available to children are constantly changing with changing technology, therefore I decided a cut off that only included the last 20 years in order to capture the most recent activities that parents do with their children, while also including key studies in the area, such as Blevins-Knabe and Musun-Miller (1996). Furthermore, in order to investigate the relationship between the home numeracy environment and mathematics achievement the studies needed to include a quantitative measure of the home numeracy environment, this could be anything from a questionnaire to an observation, as long as the measure was quantified. The studies also had to include at least one type of mathematics measure. In order to investigate the relationship between the home numeracy environment and mathematics achievement the studies must also report a test statistic that related to this relationship. With regards to the age criteria, ideally the sample would include all studies where the children have yet to start formal schooling. However, with formal schooling starting at different ages in different studies, this would be too difficult to evaluate and therefore the inclusion criteria stated that children should be under the age of 8 years old. For practical reasons, and as is the case in many systematic reviews, the studies needed to be published, accessible and written in English.

2.4.1.2 Data Sources and Search Strategies

The studies included in the systematic review were located through a comprehensive search of publicly available literature, mostly through manual electronic searches of two databases: PscyhInfo and ERIC. I searched the databases for all possible combinations of the following terms:-

> math* OR number OR numeracy OR numerical
> AND
> environment OR home
> AND
> play OR activities
> AND
> child OR toddler OR preschool OR early years OR early childhood

The search was conducted on 2nd March 2017. The search produced 3863 results using the search terms, including 260 duplicates, which were deleted. The remaining studies were screened based on the title and abstract using the inclusion/exclusion criteria detailed above. 86 studies met the inclusion criteria based on the title and abstract and were then screened on the full paper. 24 studies met the full inclusion criteria. The references of the 24 studies were then reviewed for any studies that were not included in the search but met the inclusion criteria. One further study was added.

2.4.2 Results

25 studies were identified that met the inclusion criteria. These studies varied in the sample size, age of children, home numeracy measure, mathematics measure and most importantly in the range of correlations produced

between the home numeracy environment and mathematics achievement. Table 2.1 summarises these differences. In this results section, I will start by evaluating the evidential value of these studies using a p-curve analysis. I will then move onto to a narrative review of how the strength of the correlations vary based on the measures used and age and country of origin of the children. Finally, I will discuss the direction and causality of the relationship between the home numeracy environment and mathematics performance.

Author and date	Sample size	Age range of children	Type of HNE measure	Categories of HNE measure	Type of mathematics measure	Correlations between mathematics and HNE
Anders et al. (2012)	532	2 - 6 year olds	Interview and Observation (10 measures)	Combined	Experimenter made measure	Latent growth curve analysis Intercept $B = .14$*,Slope $B = .09$
Blevins-Knabe & Musun-Miller (1996)	49	4.5 - 6 year olds	Telephone Questionnaire (33 activities)	Combined and individual activities	Standardised measure (TEMA-2)	Individual significant correlations $r =$ -.36* to $r = .42$*, combined correlation $r = .09$ (provided via email)
Blevins_Knabe, Austin, Musun, Eddy, & Jones (2000)	64	3 - 6 year olds	Telephone Questionnaire (22 activities)	Combined	Standardised measure (TEMA-2)	Correlation not significant (not reported)
Ciping, Silinskas, Wei, & Georgiou (2015)	177	6 - 7 year olds	Questionnaire (4 activities)	2 categories: formal and informal (by factor analysis)	Experimenter made measure	$r =$ -.18* to .12

Dearing et al. (2012)	127 (all girls)	6 - 7 year olds	Telephone Questionnaire (16 activities)	Combined	Experimenter made measure	r = .29*
DeFlorio (2011)	26	3 - 4 year olds	Observation	4 categories: activity type, minutes including maths, total number of math occurrences, number of math occurrences per minute	Experimenter made measure	No significant correlations (correlations not reported)
DeFlorio & Beliakoff (2014)	178	3 - 4 year olds	Questionnaire (12 activities)	Combined	Experimenter made measure	r = .17*
Esplin et al. (2016)	89	3.5 - 5 year olds	Questionnaire (37 activities)	Four categories: playing with numbers, working with numbers, measuring and comparing and counting and technology (from factor analysis)	Standardised measure (TEMA-3)	r = .03 to .39**

Huntsinger, Jose, & Luo (2016)	200 (T1)	4 - 5 year olds (T1)	Questionnaire (23 activities)	3 categories: Informal, formal and fine motor activities (from factor analysis)	Standardised measure (TEMA-2)	*r* = .00 to .40**
	97(T2)	5 - 6 year olds (T2)		3 categories: Informal, formal and games, blocks and toys (from factor analysis)		*r* = -.14 to .27**
Kleemans, Peeters, Segers, & Verhoeven (2012)	89	5 - 7 year olds	Questionnaire (4 activities)	Combined	Experimenter made measure	*r* = .47**
Kleemans, Segers, & Verhoeven (2013)	150	5 - 7 year olds	Questionnaire (6 activities)	Combined	Experimenter made measure (2 variables)	*r* = .64**, .48**
LeFevre et al. (2009)	146	5 - 8 year olds	Questionnaire (20 activities)	Four categories: number skills, number books, games and applications (from factor analysis)	Experimenter made measure (2 variables)	*r* = -.19* to .27*

LeFevre, Polyzoi, Skwarchuk, Fast, & Sowinski (2010)	204	5 year olds	Questionnaire (12 activities)	Three categories: direct, indirect and speeded activities (from factor analysis)	Experimenter made measure	r = .02 to .38** (provided via email)
Levine, Suriyakham, Rowe, Huttenlocher, & Gunderson (2010)	44	1 - 3 year olds	Observation	Two categories: parent cumulative number talk and child cumulative number talk	Experimenter made measure	r = .47**, .34*
Manolitsis, Georgiou, & Tziraki (2013)	82	4 - 5 year olds	Questionnaire (5 formal activities)	Combined (formal activities only)	Experimenter made measure (3 variables)	r = -.02 to .28*
Ramani, Rowe, Eason, & Leech (2015)	33	3 - 5 year olds	Questionnaire (20 activities)	3 categories: direct teaching of number skills, applications of number, games	Experimenter made measure (2 variables)	r = -.22 to .71**
			Observation	4 categories: foundational and advanced math talk for parent and child		r = -.31 to .46**
Missall, Hojnoski, Caskie, & Repasky (2014)	72	3 - 5 year olds	Questionnaire (36 activities)	Combined	Experimenter made measure (5 variables)	r = -.13 to .21

2

Niklas & Schneider (2013)	609	5 - 8 year olds	Questionnaire (3 activities)	Combined	Experimenter made measure (3 variables)	r = .01 to .15*
Niklas, Cohrssen, & Tayler (2016)	113	4 year olds	Questionnaire (10 activities)	Combined	Experimenter made measure and standardised test (Woodcock Johnson Applied Problems)	r = .26*, .23*
Segers, Kleemans, & Verhoeven (2015)	60	5 - 7 year olds	Questionnaire (5 activities)	Combined	Experimenter made measure	r = .41**
Skwarchuk (2009)	25	4 - 5 year olds	Questionnaire (48 activities)	Two categories: Basic (4 activities) and Complex (6 activities)	Standardised test (Woodcock Johnson Quantitative subtest)	r = .52** for complex, n.s. for basic (not reported) β = -.598* (basic) and .937 (complex)
Skwarchuk, Sowinski, & LeFevre (2014)	121	5 - 6 year olds	Questionnaire (13 activities plus number game exposure)	Two categories: basic and advanced activities (by factor analysis)	Experimenter made measure (2 variables)	r = -.08 to .3*
Swick (2007)	179	5 - 6 year olds	Questionnaire (2 activities)	Combined	Standardised test (Woodcock Johnson Applied Problems)	r = .03
Vandermaas-Peeler & Pittard (2014)	18	4 - 5 year olds	Questionnaire (10 activities)	Combined	Standardised test (TEMA 3)	r = .57*

Zippert & Ramani (2016)	43	3 - 5 year olds	Questionnaire (activities not specified)	2 categories: conventional and advanced activities	Experimenter made measure (2 variables)	r = -.10 to .41**

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

Table 2.1: Overview of all 25 studies included in systematic review

2.4.2.1 The relationship between the home numeracy environment measures and mathematics measure

From Table 2.1, we can clearly see that there is a wide range of effect sizes for the relationship between the home numeracy environment and mathematics achievement, both between and within studies. Some studies found medium strength positive correlations (Vandermaas-Peeler & Pittard, 2014; Kleemans, Segers, & Verhoeven, 2013), others found non-significant correlations (Blevins-Knabe et al., 2000; Missall et al., 2014) and even some found significant negative correlations (Ciping et al., 2015; LeFevre et al., 2009). Many of the studies report multiple correlations between different categories of the home numeracy measures and different mathematics measures, and these correlations also vary within studies with some studies finding both positive and negative significant results (Blevins-Knabe & Musun-Miller, 1996, LeFevre et al., 2009). In the next section I will discuss the evidential value of these studies using a p-curve analysis.

2.4.2.2 P-Curve analysis

To conduct a p-curve analysis, one test statistic from each paper associated with the hypothesis of interest (in this case the relationship between the home numeracy environment and mathematics achievement) needs to be selected. I followed Simonsohn et al.'s (2015) method to conduct the p-curve analysis. As seen in Table 2.1, the studies either report on the relationship between a combined measure of the home numeracy environment and mathematics achievement or the relationship between different categories of the home numeracy environment and mathematics achievement. It was decided that separate p-curve analyses should be carried out for combined home numeracy measures and separate home numeracy activities. The separated home numeracy activities p-curve analysis was then separated further into correlations relating to direct/formal activities and correlations relating to indirect/informal activities. The definition for formal activities given in the papers is "activities that engage children in explicit teaching of numbers and counting skills" (Ciping et al., 2015, p. 3) which mirrors the definition given for direct activities which is direct activities "are focused on numbers and typically are used by parents for the explicit purpose of developing quantitative skills" (LeFevre et al., 2009, p. 56). Informal activities are defined as activities that involve "incidental exposure to numeracy through real-world tasks" (Ciping et al., 2015, p. 3) which again is mirrored by the definition of indirect activities which is that indirect activities are "real-world tasks (e.g., playing card or board games that involve numbers, cooking, or carpentry) for which the acquisition of numeracy is likely to be incidental" (LeFevre et al., 2009, p. 56). Therefore, even though the categories are given different names between studies they appear to have the same definition and therefore it was decided to combine the p-values for formal and direct home numeracy categories and to combine the p-values for informal and indirect categories. If more than one correlation was presented for different categories of the home numeracy environment, I picked the correlation that was most closely related to the formal and informal categories in other studies. Furthermore, some studies contained multiple mathematics measures, if this was the case the most general mathematics measure was selected. No studies reported both a combined and a formal/informal test statistics, therefore studies were either included in the combined p-curve analysis or the formal and informal p-curve analyses. Appendix A shows the disclosure table for the p-curve analysis.

I analysed the test statistics using the p-curve app v4.052 (http://www.p-curve.com/app4/). I will start by presenting the p-curve analysis for the combined home numeracy measures then the formal home numeracy environment measures and finally the informal home numeracy environment measures.

Combined Home Numeracy Measures.

The distribution of p-values for studies investigating the relationship between combined home numeracy environment measures and mathematics achievement is shown in Figure 2.1. Of the 12 p-values entered, 9 were statistically significant with 7 being below .025. As 7 out of 9 p-values (78%) were below .025, this was higher than the 4.5 p-values expected to be below .025 if there was no effect, (one-tailed binomial test, p = .089). Stouffer's method (Simonsohn et al., 2015) computes *pp*-values for each test statistic. The *pp*-value is the probability of at least as extreme a p-value conditional on $p < .05$. These are then converted to z-scores and added together for all the studies (in this case, 9) and then divided by the square root of the number of studies. This z-score and it's corresponding p-value can then be used to indicate the evidential value of the studies. This z-score is calculated for the full curve (0 to 0.05) and the half curve (0 to 0.025). For the studies to have evidential value, either both the half and full curve z-score should be right skewed with a p-value less than .1 or the half curve should be right skewed with a p-value less than .05. For the combined home numeracy measures the full curve (z = 4.83, $p < .001$) and the half curve (z = 4.46, $p < .001$) both have p-values less than .1 so this shows there is evidential value for these studies.

The p-curve analysis also calculates an estimate of the power of the studies. It does this by calculating the p-curve for studies with different levels of power from 5% to 99%. It then compares these p-curves with the observed p-curve (using the Stouffer method) to find the p-curve that most closely resembles the observed p-curve. Figure 2.2 shows the fit for the observed p-curve with different levels of power. The estimated power for these studies is 88% with a confidence interval of 58% to 97%. In conclusion, the studies with combined home numeracy measures have evidential value of a relationship between the home numeracy environment and mathematics achievement and the studies are well powered.

Formal Home Numeracy Measures.

The same analysis as above was repeated for the studies which included a formal home numeracy environment measure. The distribution of p-values for studies investigating the relationship between formal home numeracy environment measures and mathematics achievement is shown in Figure 2.3. Of the 10 p-values entered, 8 were statistically significant with all of them being below .025. As 9 out of 9 p-values (100%) were below .025 this was higher than the 4 p-values expected to be below .025 if there was no effect (one-tailed binomial test, $p = .004$). The Stouffer method (detailed above) showed that these studies had evidential value (full curve z = -5.68, $p < .001$; half curve z = -4.5,$p < .001$). The estimated power of these studies was calculated (as detailed above) to be 91% with a confidence interval of 71% to 98%. In conclusion, the studies with formal home numeracy measures have evidential value of a relationship between the home numeracy environment and mathematics achievement and the studies are well powered.

Informal Home Numeracy Measures.

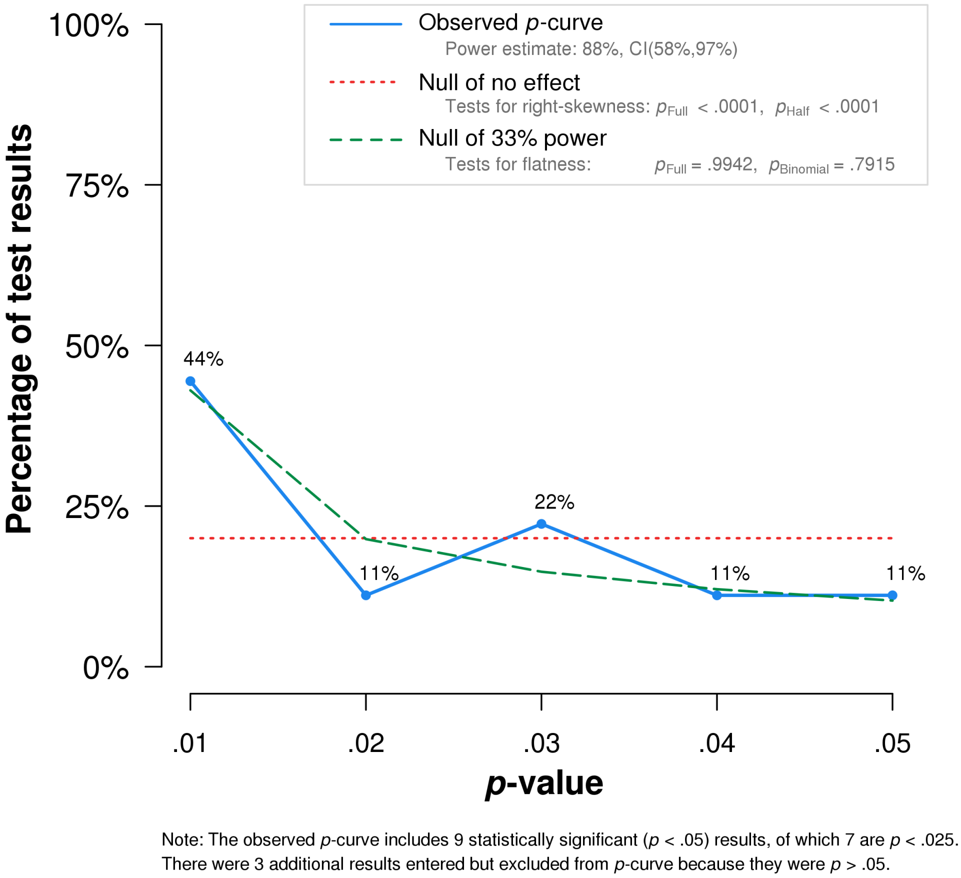

Figure 2.1: The distribution of p-values for studies finding a relationship between combined home numeracy environment measure and mathematics performance.

Finally, the same analysis was repeated for the studies which included an informal home numeracy environment measure. The distribution of p-values for studies investigating the relationship between informal home numeracy environment measures and mathematics achievement is shown in Figure 2.4. Of the 8 p-values entered, only 2 were statistically significant with only 1 of them being below .025. As only two studies produced significant results it shows that the relationship between the informal home numeracy environment and mathematics achievement has often be found to be non-significant. I will continue to report the p-curve analysis but keep in mind

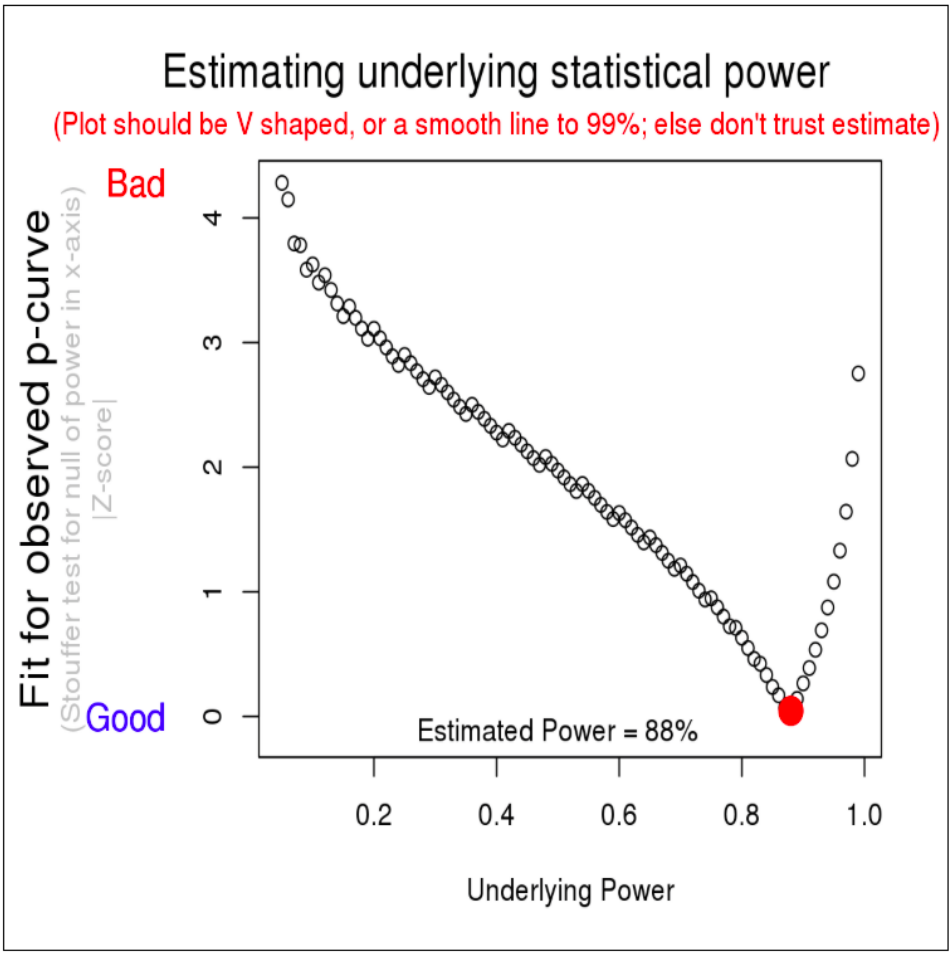

Figure 2.2: The fit of the observed p-curve plotted against the p-curves with differing levels of power for studies finding a relationship between combined home numeracy environment measure and mathematics performance.

that it is based on just two significant results. As 1 out of 2 p-values (50%) were below .025 this was the same as the 1 p-value expected to be below .025 if there was no effect, (one-tailed binomial test, p = .75). The Stouffer method (detailed above) showed that these studies had evidential value (full curve z = -1.27, p = 1.01; half curve z = -1.76, p = .039). The estimated power of these studies was calculated (as detailed above) to be 49% with a confidence interval of 5% to 95%. In conclusion, the studies with informal home numeracy measures have evidential value of a relationship between the home numeracy environment and mathematics achievement but the studies are not well powered and the analysis is only based on two studies that

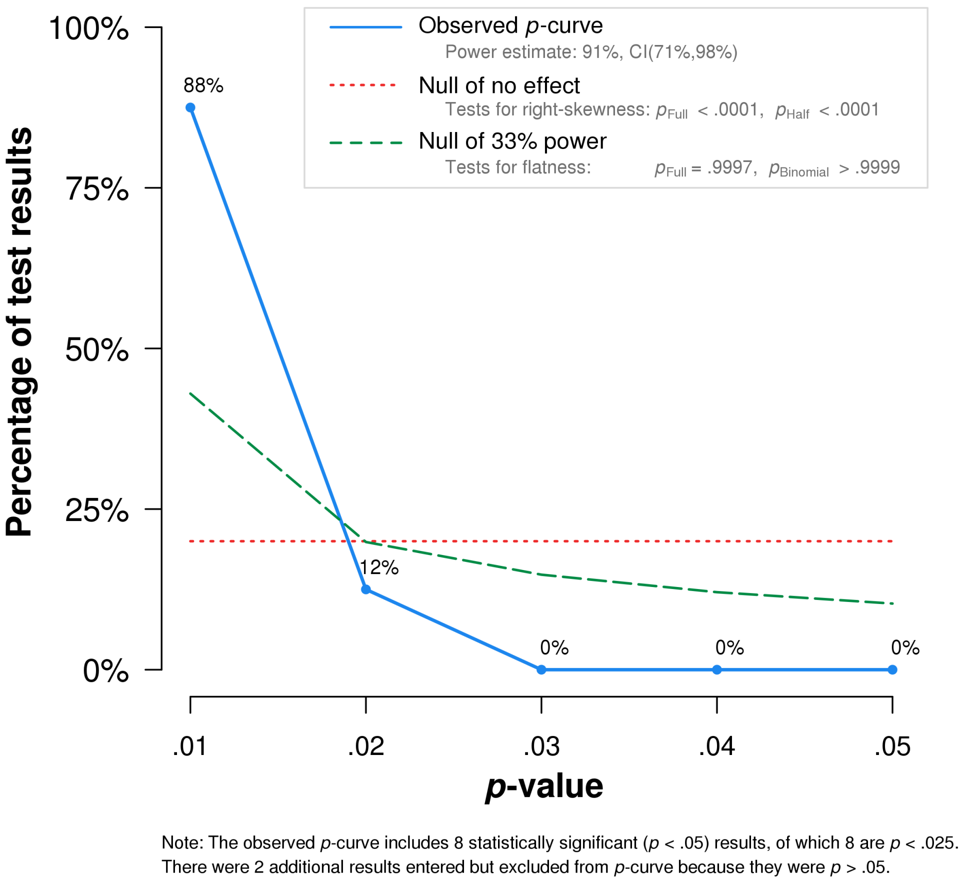

Figure 2.3: The distribution of p-values for studies finding a relationship between formal home numeracy environment measure and mathematics performance.

found a significant relationship.

In summary the p-curve analyses have shown that there is evidential value for the relationship between the home numeracy environment and mathematics achievement for all types of home numeracy measures. However, only the combined measure and formal measure studies were well-powered and the informal measures of the home numeracy environment often found non-significant relationships and the ones that did find a significant relationship are estimated to be underpowered. The lack of the relationship between the informal home numeracy environment and mathematics

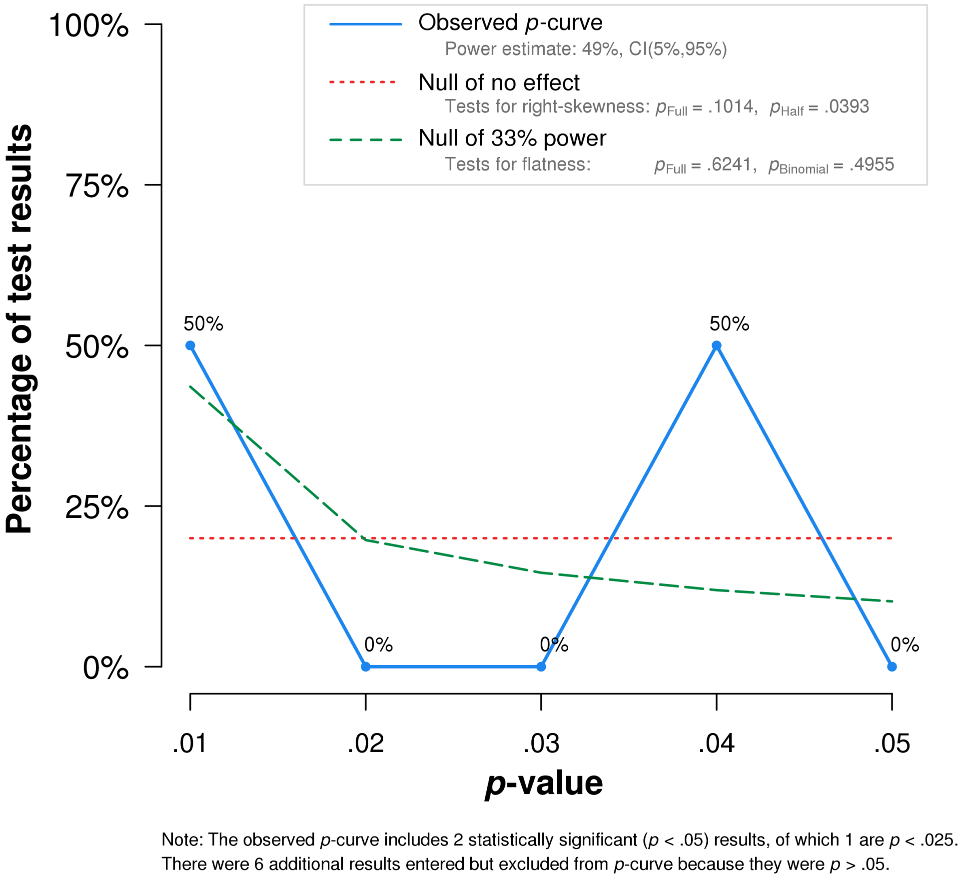

Figure 2.4: The distribution of p-values for studies finding a relationship between informal home numeracy environment measure and mathematics performance.

achievement is discussed in more detail below.

2.4.2.3 Narrative review of the relationship between the home numeracy environment and mathematics performance

In this section I will discuss how the strength of the correlations between the home numeracy environment and mathematics performance vary based on the home numeracy environment measures used, the mathematics measures used and the age and country of origin of the children.

2.4.2.3.1 Home numeracy environment measure In the papers identified in the review, there were only two different methods used to measure the home numeracy environment. These were questionnaires and observations. Even though only two types of measures were used, the questionnaire and observation measures varied between studies.

Questionnaires were used in 22 out of 25 studies and observations were used in 4 studies. Only one paper (Ramani et al., 2015) used both a questionnaire and observation method. In their study, some of the correlations between the observation measure and the mathematics measures significantly differed to some of the correlations between the questionnaire and the mathematics measures. However they did not look at the relationship between the two measures. This correlation would have helped to determine if both measures are actually measuring the same concept. Missall, Hojnoski, and Moreano (2016) (not included in this review as did not include a mathematics measure) compared parent's responses to a home numeracy questionnaire with an observation of parent and child talk, both of which were designed to measure the home numeracy environment. They found that the measures were not significantly correlated, suggesting the measures were tapping into different aspects of the home numeracy environment. Unfortunately, Missall et al. (2016) did not include a mathematics measure in their study, so we don't know how these two measures related to mathematics achievement. Future research should investigate how different measures of the home numeracy environment relate to each other and to mathematics performance. This is investigated in Chapter 5 of this thesis.

2.4.2.3.1.1 Questionnaires Typically a questionnaire measure of the home numeracy environment includes a list of mathematics activities that parents may do with their child at home, and parents are asked to select how often they do each of the activities. However there is variability between the questionnaires with regards to the question asked at the start of the questionnaire, the list of activities given and how these activities are combined into a home numeracy measure. All these differences will be discussed below while also referring to the strength of the correlations found.

The question asked at the beginning of the list of activities can vary between studies. Firstly, the time frame that researchers ask the parents to think about when answering the question varies. Some studies ask how often parents do the activities in general (Kleemans et al., 2013), some ask how often they did them in the last month (LeFevre et al., 2009) or some ask how often they did the activities in the last week (Blevins-Knabe & Musun-Miller, 1996). This could cause a difference in the home numeracy measures as there may be a difference between activities parents report doing with child in general compared to the activities they report doing in the last week. Furthermore, parents may be able to recall activities they have done in the last week more easily than those they have done in the last month. However if the last week was an abnormal week for number activity in the home, this could provide unreliable results. Asking parents to report generally how many number activities they do may not provide consistent results between parents in the same study, due to their interpretation of the word 'general' in terms of the time frame the researchers are referring to.

Secondly, the question may be phrased so that parents just report the frequency of the activities that they have personally done with their child (Kleemans et al., 2013), or it may ask how often the child was involved in the activities, suggesting that the activities could have been done with anyone in the home (including siblings, grandparents etc.) (Dearing et al., 2012; Skwarchuk, 2009). In many of the studies the question asked is not specific regarding whether the parent should report on just their activities or any activities in the home. This could provide inconsistent results within studies which could affect the overall results. This highlights the importance that researchers should be clear on how the question is phrased and who they want parents to report the frequency of activities for.

Of the 22 studies that used a questionnaire method, three of these were telephone questionnaires. These were very similar to the paper questionnaires used in the other studies, but instead of answering how many activities they do on paper they are asked over the phone. The telephone interview studies appear to produce similar mixed results to the paper questionnaires and will therefore all be included together for further discussion.

The number of activities listed in the questionnaires varied between studies. One study asked parents to report on only two mathematics activities, (Swick, 2007) whereas another study asked parents to report on 36 mathematics activities (Missall et al., 2014). Most of the studies that combined less than 30 activities into one home numeracy environment measure had a positive significant correlation between the questionnaire and a mathematics measure. However once the list was longer than 30 activities more non-significant results are present. This indicates that giving parents a long list of activities to consider may result in smaller effect sizes, perhaps due to boredom effects of filling in the questionnaire or the activities may become too specific that parents find they don't do most of the activities in the list.

However, it is not just the number of activities in the list that could have an impact on the results, it is just as important to consider how these activities have been combined into a measure of the home numeracy environment. Many of the studies with a shorter list of activities have used a combined measure of the home numeracy environment. For a combined measure researchers put all activities into one category. Many of the studies with a longer list of activities have split the activities into different home numeracy categories.

Twelve studies used a combined home numeracy environment measure. Figure 2.5 shows the distribution of the correlations for these studies. The majority of these studies (7 out of 12) used less than 10 activities and all found at least one significant positive correlation (except Swick, 2007). However when more than 20 activities have been combined into one home numeracy measure, the correlations are all non-significant. By including more activities in the list, researchers may think they are getting a more comprehensive view of the home numeracy environment but they are also adding more noise to the measure by including more activities and this could be the reason for the lack of correlations when many activities are combined. Furthermore, all the activities may not fit together into one home numeracy environment measure. All the papers that used less than 10 activities chose to ask about similar activities, which were counting activities, counting games, practising numerical conceptual knowledge (ordering, more/less etc.), counting rhymes and talking about money, weight, temperature etc.. This suggests that these activities could be classed as general mathematics activities that link to mathematics skills and should be included in future questionnaires. However, the activities included in Swick's (2007) combined measure were more specific ('how often does your child measure something' and 'how often does parent show child to add and subtract'). These activities could be used to explain Swick's (2007) non-significant correlation in this study as these activities are measuring a small area of the home numeracy environment and these two activities alone may not have an enough of an impact on children's general mathematics skills to be detected in this study. The list of activities mentioned above appear to lead to significant

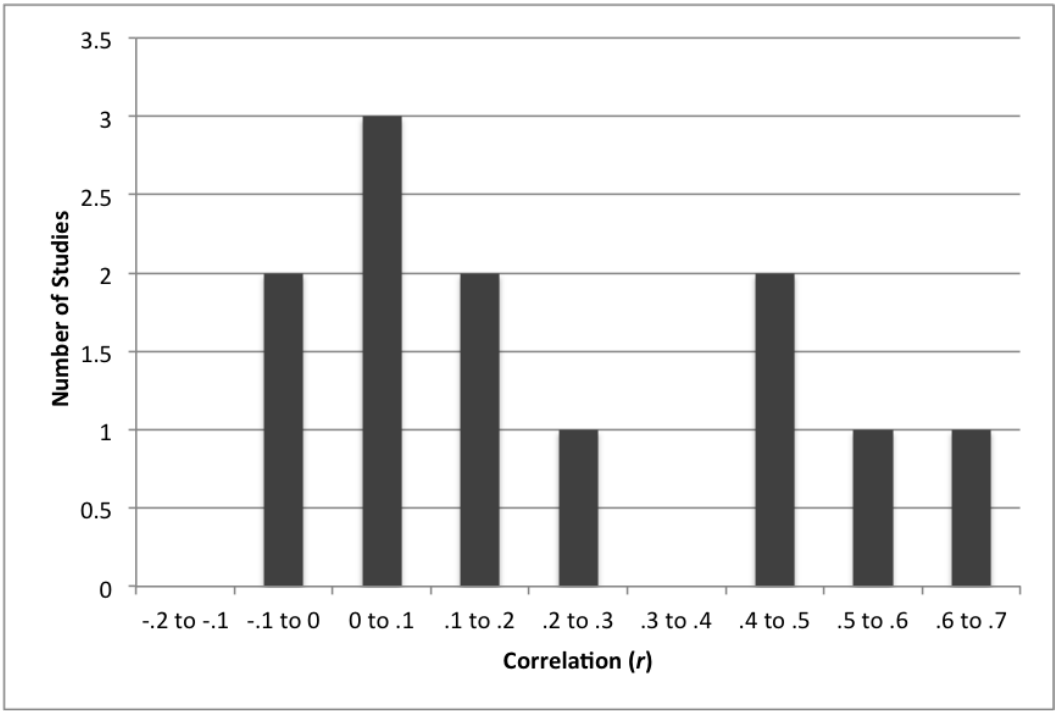

Figure 2.5: The distribution of correlations for studies investigating the relationship between combined home numeracy environment measure and mathematics performance.

results when combined, however the activities seem general, open to interpretation and are not tapping into all types of number activity. Therefore, some researchers have included more mathematics activities and combined them into different categories. Most of the researchers that have separated the activities into separate categories have used a factor analysis, but some chose the categories based on previous research (Ramani & Eason, 2015; Skwarchuk, 2009; Zippert & Ramani, 2016).

The most common categories to divide the activities into is formal and informal (Ciping et al., 2015; Huntsinger, Jose, & Luo, 2016; Manolitsis et al., 2013) or direct and indirect (LeFevre et al., 2010; LeFevre et al., 2009). As discussed in Section 2.4.2.2 the definitions for formal and direct activities and informal and indirect activities are very similar and therefore from now on I will use the term formal and informal to refer to these categories. Figure 2.6 shows the distribution of correlations for studies using formal home numeracy environment measures with mathematics measures and Figure 2.7 shows the distribution of correlations for studies using informal home numeracy environment measures with mathematics measures. Looking at the correlations between mathematics performance and the formal and informal categories, significant positive results tend to be found between formal activities and mathematics skills and non-significant results tend to be found between informal activities and mathematics skills (Huntsinger et al., 2016; Manolitsis et al., 2013; LeFevre et al., 2010). These results are also shown in the p-curve analyses. However, this is not always the case, Ciping et al. (2015) found a significant negative relationship between the direct activities and mathematics performance and LeFevre et al. (2009) also found a significant negative correlation between one of their formal activities categories (number books) and mathematics performance, and a non-significant relationship between the other formal activities category (number skills) and mathematics performance.

Furthermore, some studies have split the activities into basic and advanced activities (Skwarchuk, 2009; Skwarchuk et al., 2014; Zippert & Ramani, 2016). In all three studies the basic activities category was not significantly correlated to the mathematics measure and the advanced category of activities was significantly correlated with at least one of the mathematics measures used.

Other studies have split the categories further to separate out game activities. When games have had their own category (Huntsinger et al., 2016; LeFevre et al., 2010; Ramani et al., 2015) there is a significant positive correlation between the games category and the mathematics measure in all of these studies.

Even though there seems to be patterns in results between the categories, the activities listed in each category differ between papers, so it is hard to pinpoint which activities parents should be promoting. Furthermore, a certain activity, say recognising printed numbers, will not always be taught in a formal way and whether this activity is advanced depends on the age of the child, therefore the activities in the formal/informal and basic/advanced categories could change between studies. Huntsinger et al.'s (2016) paper includes two separate factor analyses for the same list of activities, but for different aged children, just one year apart, and showed that different categories were created for different age groups with the activities in the

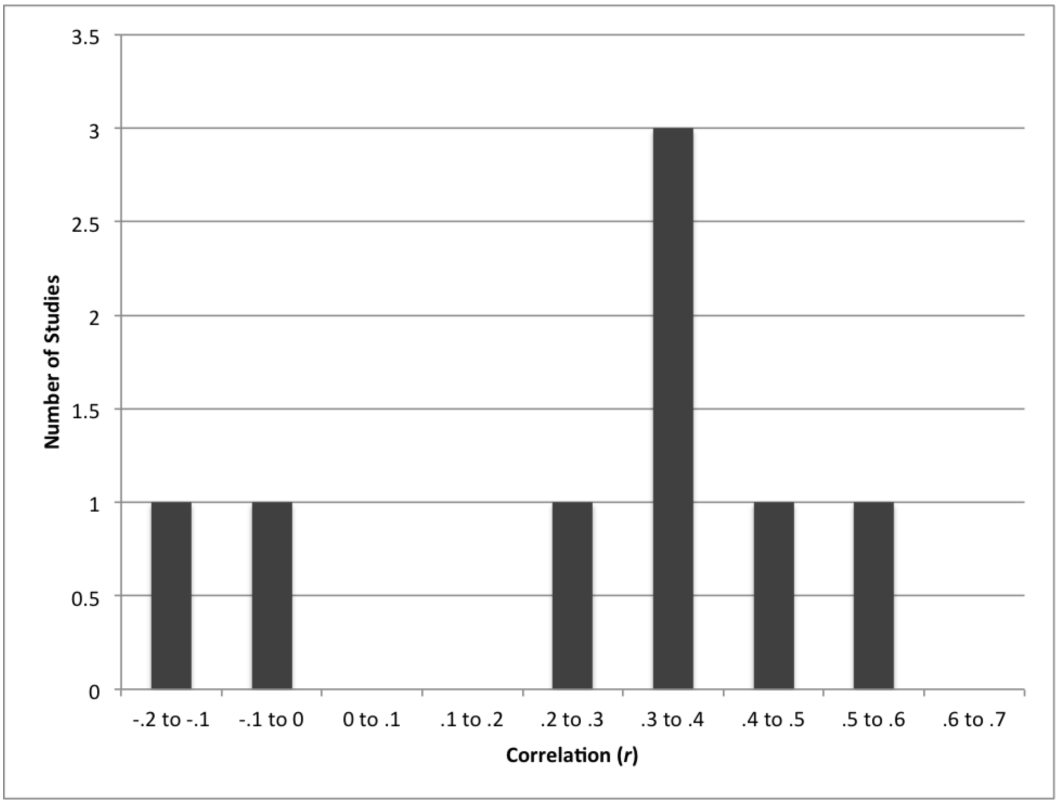

Figure 2.6: The distribution of correlations for studies investigating the relationship between formal home numeracy environment measure and mathematics performance.

formal and informal categories also changing. This highlights how these categories could change depending on the age of the participants (see below for further discussion on age).

2.4.2.3.1.2 Observations Only 4 studies used observation techniques to measure the home numeracy environment (Anders et al., 2012; DeFlorio, 2013; Levine et al., 2010; Ramani et al., 2015). All four studies looked at mathematics talk during a parent and child interaction, however the type and place of interaction varied between all measures.

Levine et al. (2010) used the most naturalistic observation of the home environment by recording 5 blocks of 90 minutes of day-to-day life between parent and child and measured the amount of mathematics talk during mealtimes. Anders et al. (2012) also observed parents at home and asked parents to read a book with their child. They also recorded the amount of toys in

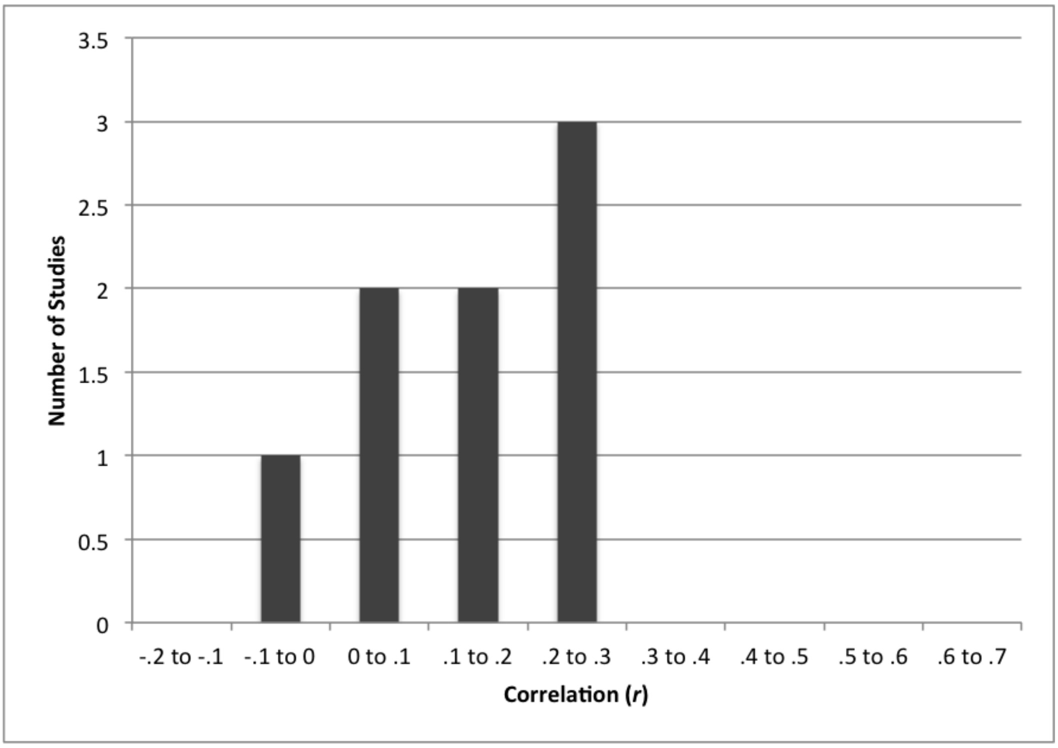

Figure 2.7: The distribution of correlations for studies investigating the relationship between informal home numeracy environment measure and mathematics performance.

the home that were related to mathematics and combined this with the amount of stimulation of mathematics parents provided when reading the book. DeFlorio (2013) also used a home environment to gather data, but it was less natural, as they asked parents to demonstrate two mathematics activities of their choosing. In contrast, Ramani and Eason (2015) tried to mirror the home environment in the lab and asked parents to play with 3 different activities with their child as they would at home. Even though Ramani and Eason (2015) did not use the home environment for their study, they still found positive significant results showing that a relationship is still present between mathematics achievement and the home numeracy environment when measured at home.

In Anders et al. (2012), Levine et al. (2010), Ramani et al.'s (2015) studies the parents were not aware that the researchers were measuring the home numeracy environment. All three of these studies found positive significant results between home numeracy environment and mathematics measures.

In contrast to this, in DeFlorio's (2013) study the parents were asked to demonstrate two mathematics activities with their child so parents were aware that the mathematics in their activity was being measured. Interestingly DeFlorio (2013) found no significant correlation between the home numeracy environment measure and mathematics skills. This could be because parents were talking about mathematics as they had been told to do, even though for some of the parents this may not have been natural and may not occur at home on a regular basis.

Furthermore, researchers divided the home numeracy measure into different categories. Anders et al. (2012) used just one combined home numeracy environment measure. This was based on both parent and child mathematics talk when reading a book and amount of mathematics toys in the home. They found a significant positive result. DeFlorio (2013) used 4 different categories from the observation of parent and child mathematics activities. These categories were activity type, minutes including mathematics, total number of mathematics occurrence and number of mathematics occurrence per minute. All these categories included a combined measure of child number talk and parent number talk. They found no significant correlations. Ramani and Eason (2015) and Levine et al. (2010) separated the amount of number talk for the parent and child. Levine and colleagues found a significant positive correlation for both parent and child number talk. Ramani and Eason (2015) separated the categories even further to include the level (foundational and advanced) of parent and child number talk. Interestingly not all types of parent number talk correlated with the child's mathematics skills. Parents' advanced number talk significantly correlated with one of the mathematics measures, but parent's foundational number talk did not significantly correlate to mathematics performance. However, both child's foundational and advanced mathematics talk were significantly correlated with the mathematics measures.

Even though all four studies used an observation technique, it is difficult to draw conclusions from combining these studies as they are all measuring the home numeracy environment in different ways. We can state that three out of four of these studies found positive significant results which shows more consistency than the questionnaire method. However, it is possible that the positive relationship between the home numeracy environment and mathematics achievement is being driven by other factors such parents' IQ or SES.

2.4.2.3.2 Mathematics measure Table 2.1 also shows that there are variances between the studies with regards to the type and number of mathematics measures used. For this analysis I have spilt the mathematics measures into standardised mathematics measures (such as the Test of Early Mathematical Ability (TEMA)), subtests from standardised mathematics measures (such as the applied problems subtest from the Woodcock Johnson standardised test) and experimenter made measures and will compare the relationship between the home numeracy environment and mathematics ability when using these different types of measures.

Six out of the 25 studies used a standardised test to measure mathematics ability. All six of these studies chose to use the Test of Early Mathematical Ability (TEMA). The TEMA is quick and easy to administer to the participants. However, even though all the studies used the same measure, there is a mixture of results. Two studies found no significant correlation between the mathematics measure and home numeracy environment (Blevins-Knabe & Musun-Miller, 1996; Blevins-Knabe et al., 2000), one study found a positive correlation between a combined measure of the home numeracy environment and mathematics ability (Vandermaas-Peeler & Pittard, 2014) and three studies found positive significant correlations with some categories of the home numeracy environment and mathematics ability, but non-significant correlations between other categories of the home numeracy environment and mathematics ability. This shows that even if all studies investigating the home numeracy environment used the same measure of mathematics ability, there would still be a mixture of results.

Six studies used a subtest of standardised measures. All six studies used different subtests of different standardised tests and produced a mixture of results. All of the subtests tap into different types of mathematics ability, so for further analysis I combined these with the experimenter made measures and investigate the relationship between home numeracy environment and different types of mathematics ability.

One key category that many of the studies use as a measure of mathematics skill is addition and subtraction problems (Anders et al., 2012; Niklas et al., 2016; Dearing et al., 2012; Kleemans et al., 2013; LeFevre et al., 2009). Interestingly when addition and subtraction skills are correlated with the home numeracy measures, they all produce at least one significant correlation. However, Skwarchuk et al. (2014) used a measure of non-symbolic arithmetic and found non-significant results.

In contrast, those studies that tap into a wide range of mathematics skills and combine all mathematics skills into one measure find a mixture of results (DeFlorio, 2013; DeFlorio & Beliakoff, 2014; Kleemans et al., 2012; Ciping et al., 2015; Segers et al., 2015; LeFevre et al., 2010). These studies measure a variety of mathematics skills but use different measures, so it is difficult to tell if the difference in correlations is due to how mathematics being measured or other variables.

Some studies decided to use multiple mathematics variables and look at the difference in correlations depending on type of mathematics skills. Both Zippert and Ramani (2016) and Skwarchuk et al. (2014) used a non-symbolic mathematics measure, in which children were shown a number of objects, the objects were then hidden behind a card, the researcher would then show the child them adding more objects and the child was asked how many objects were behind the card. Both studies found no significant correlations between any home numeracy measure category and non-symbolic mathematics. Zippert and Ramani (2016) and Ramani and Eason (2015) both used foundational (counting and number recognition) and advanced (numerical magnitude comparison, number line estimation and symbolic arithmetic) mathematics skills categories and both studies found that the advanced skills were correlated significantly with direct/advanced home numeracy activities. Ramani and Eason (2015) also found that basic mathematics significantly correlated with direct and indirect home numeracy activities. LeFevre et al. (2009) used two different categories of mathematics skills, mathematics knowledge and mathematics fluency, however there were no significant differences between the correlations for the two measures. Similarly Missall et al. (2014) used 5 different measures of mathematics skills and found all five had a non-significant relationship with the home numeracy environment measure.

Some studies used different types of mathematics measures for different age groups of children (Manolitsis et al., 2013; Niklas & Schneider, 2013a). It is important to consider the age of the children that are being tested, as the age range of children in these studies differs. The same mathematics measure could not be used in all studies. However, it is difficult to determine if the difference in correlations is due to the different mathematics measures or another variable such as age or country.

2.4.2.3.3 Age and country The children in these studies ranged in age from 1 - 8 years. Ideally the sample would include all studies where the children have yet to start formal schooling. However with formal schooling starting at different ages in different countries, this would be complex to evaluate and result in different aged children being included from different countries and therefore the inclusion criteria stated that children should be under the age of 8 years old.

Huntsinger et al. (2016) showed that the categories of activities from the home numeracy environment questionnaire differed between children age 4 years old and 5 years old, which suggests that parents do different activities with their children depending on their age. Some studies have combined children with a four year age gap (LeFevre et al., 2009) which may make it difficult to draw firm conclusions about the home numeracy environment as the activities a 5 year old may do at home will differ significantly to the activities a 8 year old may do.

Some studies have used longitudinal studies to investigate the effect of the home numeracy environment over time (Anders et al., 2012; Ciping et al., 2015; Huntsinger et al., 2016; Manolitsis et al., 2013; Niklas & Schneider, 2013a). Both Manolitsis et al. (2013) and Niklas and Schneider (2013a) took one measure of the home numeracy environment and measured mathematics abilities for several years afterwards. Manolitsis et al. (2013) took a home numeracy environment measure at time 1 (beginning of kindergarten) and a mathematics measure at time 1 (beginining of kindergarten), time 2 (end of kindergarten) and time 3 (end of Grade 1). They found that there was only one significant correlation between the home numeracy measure and their maths measure at time 1. Niklas and Schneider (2013a) used a home numeracy measure at time 2 (end of kindergarten) and a mathematics measure at time 1 (beginning of kindergarten), time 2 (end of kindergarten), time 3 (begining of Grade 1) and time 4 (end of Grade 1). They found that the home numeracy measure was significantly correlated to all time points, apart from time 1, suggesting the opposite to Manolitsis et al. (2013) that home numeracy activities in kindergarten are related to mathematics skills at the end of Grade 1. However, even though Niklas and Schneider's (2013a) correlations are significant (r = .09 to .15), they are very small effect sizes due to the large sample size (N = 609) (J. S. B. T. Evans, 2005, p. 63-65) and should be treated with caution.

Anders et al. (2012), Ciping et al. (2015) and Huntsinger et al. (2016) all measured the home numeracy environment and mathematics ability at the same time points. Anders et al. (2012) measured the home numeracy environment and mathematics ability in children's first, second and third year in pre-school. The mean scores from the three home numeracy environment measures was used to predict the development of mathematics skills. The latent growth curve showed that the home numeracy environment explained substantial variance in numeracy skills at time 1 and this was maintained over the next two years. Ciping et al. (2015) measured the home numeracy environment and mathematics ability at Grade 1 and at Grade 2. They found that the only significant relationships were between formal mathematics activities and mathematics skills at both Grade 1 and Grade 2. However in contrast to other studies, this relationship was negative. Huntsinger et al. (2016) measured both the home numeracy environment and mathematics skills at time 1 and time 2 (1 year later). They found that formal activities at time 1 and mathematics achievement at time 1 and formal activities and games at time 2 and mathematics achievement at time 2 were positively significantly related. However they also found that home numeracy activities at time 1 were still related to mathematics skills at time 2. These studies show the uncertainty about the long-term benefits of the home numeracy environment.

The huge age difference between participants in the studies also affects the mathematics measures used. Older children are more likely to score higher on the mathematics measure than younger children in the same way that they are more likely to do more advanced home numeracy activities. In children this young, there can be large differences in both these variables between a child that is three years-old and a child that is five years-old, and therefore, age should be considered when looking at the relationship between home numeracy activities and mathematics abilities to identify whether the relationship is solely driven by age differences. However, none of the studies in this systematic review controlled for age when looking at the correlation between home numeracy environment and mathematics skills. Five studies reported the relationship between age and the home numeracy environment and mathematics measures and all five papers (Esplin et al., 2016; Ramani et al., 2015; Skwarchuk et al., 2014; Niklas et al., 2016; LeFevre et al., 2009) found a significant relationship between age and the home numeracy environment or age and the mathematics measure or both. This highlights that future studies should control for age when investigating the relationship between the home numeracy environment and mathematics achievement.

Studies included in this systematic review include children from a variety of countries. The studies were conducted in the USA (13 studies), Netherlands (3 studies), Canada (3 studies), Germany (2 studies), China (1 study), Greece (1 study) and Australia (1 study). Even though there may not be consistency in results within countries due to the different measures used, there may also be cultural differences in the activities that parents do with their children at home and at what age these activities start. Therefore, comparisons between countries may be difficult. Only one study included in this review has compared the relationship between home numeracy environment and mathematics measures between countries. LeFevre et al. (2010) compared Canadian and Greek parents' home numeracy activities and even though they found that there were differences between the type of activities reported, the correlations between the home numeracy measures and the mathematics skills were very similar for each country. However, this is only looking at two countries in which the education systems are similar and it is important to note that the home numeracy environment may not have the same relationship to mathematics performance in all countries.

2.4.2.3.4 Direction and causality of relationship Even though all these papers report the correlation between the home numeracy environment and mathematics skills, many do not consider the direction and causality of the relationship. Most papers researching the home numeracy environment are correlational in nature and, therefore, no causation can be identified. However, most papers that find a positive correlation assume that doing more number activities causes the child to develop better number skills. Nonetheless, it is possible that because the child enjoys and is good at mathematics, then this causes the parents to do more number activities at home. A few papers have mentioned the problems associated with doing correlational studies and not being able to determine the causality of the relationship in their discussion section (Blevins-Knabe & Musun-Miller, 1996; Ramani & Eason, 2015), however, no papers, to my knowledge, have directly looked at the direction of the relationship.

There are a few papers that have found a negative relationship between the home numeracy environment and mathematics performance (Ciping et al., 2015; LeFevre et al., 2009). One way to explain the negative relationship would be that the child's number skills predict the amount of activities they engage in. In other words, if the child struggles with number skills, then parents do more number activities at home in order to help the child. Ciping et al. (2015) examined cross-lagged relationships between home numeracy environment and mathematics achievement. They found a negative correlation between the formal home numeracy environment at Grade 1 and mathematics performance at Grade 1 ($r = -.18$). They also found a negative relationship between mathematics in Grade 1 and home numeracy environment in Grade 2 ($r = -.15$). They argue this relationship shows that the amount of activities parents do at home depends on their child's mathematics achievement. However, this study was conducted with Chinese children in the first years of formal schooling where parents are getting feedback on their child's ability and, therefore, a similar effect may not be found before children start formal schooling. The alternative way to explain this relationship would be, if parents do more number activities with their child, the child gets bored of the activities and then develops negative attitudes towards mathematics, resulting in poor mathematics achievement. However, this alternative explanation has not been considered in any papers finding a negative relationship.

2.4.3 Discussion

This systematic review has shown that evidence for the relationship between the home numeracy environment and mathematics skills is not clear. A full range of correlations from significant negative to significant positive relations have been found across studies, and sometimes within the same study. The p-curve analysis showed that there was evidential value for a relationship between the home numeracy environment and mathematics performance when considering the literature as a whole. However, the further analysis has shown that some categories of the home numeracy environment may be more strongly linked to mathematics performance such as formal activities and game activities, and observation studies appear to find more significant results. It has also shown that addition and subtraction problems appear to be linked to the home numeracy environment more than general mathematics measures. Furthermore, it has also highlighted that future studies should consider how far age is able to explain observed relationships.

The methods used to investigate this topic vary making it difficult to evaluate if it is the methods used to measure the relationship that cause varying results, or if the home numeracy environment is complex and other variables effect the strength of the relationship with mathematics.

2.5 Intervention studies

Even though questions remain about which aspects of the home numeracy environment are important, some researchers have begun doing intervention studies. The interventions are mainly games which parents play with their children in order to improve their number skills. There are a range of studies by Ramani and colleagues (Ramani & Siegler, 2008; Siegler & Ramani, 2008) in which they have developed a linear board game that has been shown to improve children's mapping between symbolic and non-symbolic representations of number. They have highlighted the importance of the game being linear – and not circular – (Siegler & Ramani, 2009) and also showed that children from low income backgrounds benefit from playing the game more than children from middle-income backgrounds (Ramani & Siegler, 2011).

Even though it has been shown that children benefit from playing this board game, evidence for long-term effects is not clear. It is also important to consider the frequency of playing the game, the quality of the interaction and the variety of activities needed to improve general number skills.

Finally, Niklas et al. (2016) ran an intervention study in which parents attended a meeting that encouraged them to do more number activities with their children at home. They claim that this lead parents to do more number activities, however the measures were all self-report.

2.6 Summary and research questions

Research to date has shown that the home numeracy environment is a complex area of research, that does not appear to be as clean cut as the research on home literacy environment. At present it is not clear if the mixed results are due to differing measures used to measure the home numeracy environment, or differing demographic variables that affect the relationship. However research in the area is growing substantially with studies trying to answer these questions. It is important to note that Lakens and Etz's (2017) recent paper argues that an area of research that produces significant and insignificant results is likely to be showing evidence for the alternative hypothesis. Therefore, this suggests that the mixture of results in the home numeracy environment is to be expected and as long as the studies are well powered, the mixture of results could provide evidence for the relationship between the home numeracy environment and mathematics performance.

A notable gap in the research is that the home numeracy environment within the UK has not been investigated. The UK's formal education system starts earlier than many other countries and it would be interesting to see the impact that the home numeracy environment has for UK children.

With regards to measures of the home numeracy environment, most studies have used a questionnaire method to measure the home numeracy environment as it a quick and convenient method. However the use of the questionnaire method could have contributed to the mixed findings. Questionnaires involve problems with social desirability, remembering the activities done and doing activities that are not on the list (see section 4.2 for a more detailed discussion). Therefore, there appears to be a gap in the research for the development of another measure of the home numeracy environment that could avoid some of these problems. Future research should also investigate the relationship between different measures of the home numeracy environment and how they relate to mathematics achievement.

Therefore within this section of the thesis, I aim to research the methods used to measure the home numeracy environment and develop a novel measure of the home numeracy environment. I will also investigate the correlation between the home numeracy environment and maths performance, as well as the relationship with demographic variables, parents' expectations and attitudes.

In the following chapters I present 2 studies to investigate these areas, before discussing findings in relation to the existing home numeracy environment literature.

Chapter 3

Home numeracy environment evidence from the Millenium Cohort Study

Before starting my own studies investigating the home numeracy environment, I explored evidence for the relationship between the home numeracy environment and mathematics achievement in the UK using data from a single question from the Millennium Cohort Study.

3.1 Introduction

The home numeracy environment is often measured through a questionnaire measure and other studies have used data from nation/state wide studies to evaluate the relationship between the home numeracy environment and mathematics achievement (Niklas & Schneider, 2013a; Swick, 2007). Niklas and Schneider (2013a) found a positive but small relationship between 3 items asking parents about number games and children's mathematics at the end of grade 1 ($r = .15$). However, Swick (2007) found a non-significant relationship between how often parents reported doing measuring, adding and subtracting activities with their child and children's mathematics performance ($r = .03$). This shows that this approach has resulted in mixed findings.

The UK-wide Millennium Cohort Study (MCS) follows the lives of 19,000 children born in 2000 - 2001. The study is a multi-disciplinary research project which plans to follow these children to adulthood. The study collects information on a wide range of topics from parents, siblings and child's health, wealth and housing to child's behaviour, schooling and development. There have been six surveys carried out so far, these were completed when the children reached the ages of 9 months, three, five, seven, eleven and fourteen years old.

In the survey, when the children were three years old, the primary carer was asked about activities their child did at home. This included questions such as "Does anyone at home ever teach 'your child' numbers or counting?" and "How often does someone at home try to teach your child numbers or counting?". These questions tap into the home numeracy environment. Furthermore the MCS also includes measures of children's mathematics performance when starting school at age 5. Therefore, the aim of this analysis was to test if the home numeracy environment question correlated with mathematics performance when starting school and to investigate the reliability of measuring the home numeracy environment in this way.

3.2 Method

The data from the MCS study is available online and was downloaded from http://www.cls.ioe.ac.uk. The files for the three year old surveys and five year old surveys were merged by participant number and any participant with missing data for the home numeracy questionnaire or the mathematics measure was removed. I also removed any twins or triplets from the data, because parents' time would be divided between the children and not reflect a typical sole child's environment. Furthermore, independent parent data is not available for these children. This resulted in a final sample size of 7497 children.

The question that asked parents to rate "How often does someone at home try to teach your child numbers or counting? was used as a measure of the home numeracy environment. Parents were asked to select their answer on a scale from 1 - Occasionally/less than once a week to 6 - constantly/7 days a week. The mathematics measure was taken from the children's foundation stage profiles, which was reported by their teachers at age 5. For mathematics performance, I selected the mathematical development variable from the foundation stage profile, which is a composite of the 3 variables: ‘numbers as labels and for counting’, ‘calculating’ and ‘shape, space and measure’. Teachers were asked to report if the child was competent in the given skill. A list of 9 skills for each of the three mathematics topics were assessed, resulting in a score between 0 and 18 for each child.

3.3 Results

Figure 3.1 shows the reported frequency for the question “How often does someone at home try to teach your child numbers or counting?”. 50% of parents reported teaching their child counting constantly/ 7 days a week. This data does not follow a normal distribution and therefore a Spearman rank correlation was used.

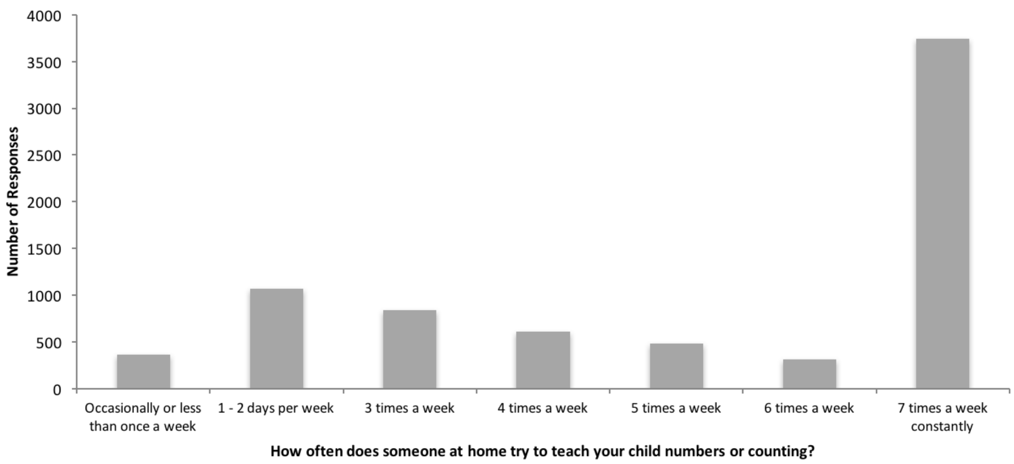

Figure 3.1: The distribution of responses to the question “How often does someone at home try to teach your child numbers or counting?”.

The correlation between the mathematics development measure and the home numeracy question was very small (r_s=.072, $p < .001$). Even after removing the ceiling responses to the home numeracy question (i.e. those who answered 7 times a week), the correlation was still small and significant (r_s=.064, $p < .001$).

3.4 Discussion

The correlation was significant but this was due to the size of the sample (J. S. B. T. Evans, 2005 p.63-65). A study that has a very big sample size has the power to detect very small effects. In these cases, the size of the effect should be considered above the significance. Past research in the home numeracy environment has found a much higher correlation between mathematics performance and home numeracy activities. However, as no previous research has investigated the link between the home numeracy environment and mathematics achievement in the UK, we cannot be sure if this is the true relationship for the home numeracy environment for the UK or just a poor measure of the home numeracy environment. More specifically, the very small correlation between the reported home numeracy activities and mathematics performance could be due to the question that was asked, because it only focuses on parents specifically teaching their child counting and not a range of number activities.

Other questions in the Millennium Cohort Study asked about the literacy activities parents did with their child at home. As detailed in Chapter 1, research into the home literacy environment has found a relationship with reading and literacy performance. Included in the MCS, there were two questions specific to the home literacy environment, one question was “How often do you help your child learn the alphabet?” and another question asking “How often do you read to your child?”. The first question is very specific about learning the alphabet whereas the second question is more general and similar to the questions found in home learning environment questionnaires. If the lack of correlation between the counting question and mathematics performance is due to the wording of the question, we would expect a low correlation between the alphabet question and literacy performance and a stronger correlation between the reading question and literacy performance. I selected the communication language and literacy variable from the foundation stage profile as a measure of literacy performance. This was a composite of the ‘reading’, ‘linking sounds and letters’ and ‘writing’ variables. The correlation between the alphabet question and composite literacy score was very low (r_s = .069, $p < .001$). The correlation between the question about reading and the composite literacy score was a lot higher (r_s = -.240, $p < .001$) (this is negative as parents were asked to rate 1 - Everyday through to 6 - not at all). This suggests the questions about teaching counting and learning the alphabet respectively are not sufficient to measure the home numeracy and literacy environment alone.

This short analysis demonstrates that there is need for a more robust measure of the home numeracy environment to be used to evaluate the relationship between the home numeracy environment and mathematics performance in the UK. A more detailed questionnaire could be used to measure the home numeracy environment but, as highlighted in the next chapter, there are problems with the questionnaire measure.

Chapter 4

Developing a novel home numeracy environment measure (Study 1)

The aim of this chapter is to present and evaluate the reliability and validity of a novel measure of the home numeracy environment whilst also investigating the relationship between the home numeracy environment and mathematics performance.

4.1 Introduction

Measuring the frequency of number activities using a questionnaire has been one of the main ways to measure the home numeracy environment. However, as mentioned in Chapter 2 and shown in my analysis of the Millennium Cohort Study data, this has led to mixed results. Therefore, in this study I will present a novel measure of the home numeracy environment. I will start by discussing the problems with these questionnaires, I will then detail a new method designed to measure the home numeracy environment. Following that I will discuss how to measure validity and reliability. I will then present the research questions for this study.

4.1.1 Problems with using a questionnaire to measure the home numeracy environment

Figure 4.1 shows an example of a typical questionnaire used to measure the home numeracy environment. There are many problems which can be associated with using a questionnaire to measure the home numeracy environment and I will discuss these below:

This is a list of activities that might take place in your home. Please indicate how often you do these activities with your child. I understand some these activities are not all relevant for the age of your child so please bear this in mind when answering these questions.

	Rarely /Never	Monthly	Weekly	Several days per week	Most days per week
I help my child learn simple sums (e.g. 2+2).	1	2	3	4	5
I encourage my child to do maths in his/her head.	1	2	3	4	5
We talk about time with clocks and calendars.	1	2	3	4	5
I help my child weigh, measure and compare quantities.	1	2	3	4	5
We play games that involve counting, adding or subtracting.	1	2	3	4	5
I teach my child to recognise printed numbers.	1	2	3	4	5
We sort and classify by colour, shape and size.	1	2	3	4	5
I ask about quantities (e.g. How many spoons?).	1	2	3	4	5
We play board games or cards.	1	2	3	4	5
I encourage collecting (e.g. cards, stamps, rocks).	1	2	3	4	5
I help my child recite numbers in order.	1	2	3	4	5
We sing counting songs (e.g. five little monkeys).	1	2	3	4	5
I encourage the use of fingers to indicate how many.	1	2	3	4	5

Figure 4.1: Example of a questionnaire used to measure frequency of activities in the home numeracy environment taken for Skwarchuk et al. (2014)

1. Social desirability bias. This is the effect that parents may be tempted to overestimate the activities they do when responding to the questionnaire, as they believe they should report the amount of activities that appears to be socially acceptable.

2. It is difficult for parents to select an average amount of time they spend on an activity during the course of a month. For example, the above questionnaire asks how often do you talk about time with clocks and calendars. During the weeks near Christmas, birthdays and holidays, parents may use calendars everyday to countdown to a big event, however in other weeks of the year, parents may not use a calendar at all. This would make it difficult for the parent to judge whether they do this activity most days a week, several day a week, weekly or monthly when answering the questionnaire.

3. It can be difficult for parents to remember how often they do each of these activities in a month and can only provide a retrospective account.

4. Parents may do activities with their child that are not listed on the questionnaire and therefore it may appear to the researcher that they don't participate in number activities at home. This may result in an underestimate of the frequency of number activities.

These are just a few problems associated with the questionnaire method. Although these problems have been recognised by researchers, there have been few attempts to address these problems. Plewis, Mooney, and Creeser (1990) highlighted some of these problems with the questionnaire method and attempted to use a time budget approach to measure the frequency of home activities, including mathematics activities. They telephoned parents on 3 days and asked them to report everything their child had done the previous day and for how long. However, they found that this method was unreliable, particularly for mathematics activities as many parents did not report any specific mathematics activities. They suggested that in order to reach a good level of reliability the phone calls would need to happen everyday for 10 days, which they concluded would not be practical.

Therefore, in order to address some of these problems, there is a need for a new valid and reliable measure of the home numeracy environment. I used a type of ecological momentary assessment to measure the home numeracy environment.

4.1.2 Ecological Momentary Assessment (EMA)

Ecological Momentary Assessment (EMA) is the "repeated sampling of subjects' current behaviours and experiences in real time, in subjects' natural environments" (Shiffman, Stone, & Hufford, 2008, p. 1). EMA is not just a single research method; there are many methods to gather this data from electronic diaries to physiological sensors. The benefits of using EMA is that it minimises recall bias and maximizes ecological validity. EMA is popular in clinical psychology to gather data in many areas from drug use to pain (Shiffman et al., 2008). For the purpose of measuring of the home numeracy environment I used a text message method of EMA. The text message method involved sending participants a daily text message asking about the number activities they have done at home. The text messages asked the question"Have you done any number activities with your child today? Reply yes or no."

Using text messages as a EMA measure is a relatively new method that has only been used in the last 10 years (Alfvn, 2010; Axn, Bodin, Kongsted, et al., 2012; Christie, Dagfinrud, Dale, Schulz, & Hagen, 2014). This is because of the rise in the number of people having a mobile phone. The text message method involves the researcher sending participants a text message which they have to reply to in order for the researcher to gather data. The text messages can be sent frequently to gather real time data. As the text message method is a relatively new method it is often compared to existing questionnaires, usually correlating significantly with the questionnaire data (Kuntsche & Robert, 2009; Johansen & Wedderkopp, 2010). Many studies have also highlighted advantages of using text messages. This includes a higher response rate when using text messages compared to questionnaires (Alfvn, 2010; Christie et al., 2014) and that it is more cost effective than sending individual questionnaires (Johansen & Wedderkopp, 2010). It has also been shown that responses to text messages were not affected by baseline characteristics, as had been the case with questionnaires (Axn, Bodin, Bergstrm, et al., 2012). Therefore, it has been shown in the clinical psychology literature that text messages can be used successfully to gather data.

To my knowledge, the text message method has not previously been used to gather data on mathematics education. Therefore, this is a novel measure of the home numeracy environment. The text messages will capture day-today activities reducing parents' recall and judgement of how often they do activities in a typical month. It will also not limit parents' responses to a set list of activities. Therefore this new method should address the majority of the problems with the existing questionnaire. However, when using a new method it is important to consider the reliability and validity of the method.

4.1.3 Reliability and validity

Reliablity and validity are two related concepts that are important for evaluating measures used in research. In general terms, reliability is concerned with the measure being consistent (i.e. if the construct was to be measured again would the same results be achieved, provided the construct is not changing). Validity is a term used to evaluate if what is being measured is an accurate representation of the construct it should be measuring (Bryman, 2012). There are different types of reliability and validity. Below, I will briefly mention the different types of reliability and validity with reference to the measurement of the home numeracy environment.

4.1.3.1 Reliability

There are three main types of reliability: internal reliability, inter-rater reliability and test-retest reliability.

Internal reliability. One way to measure reliability is by estimating how well the items on a multiple item measure reflect the same construct. When a measure uses multiple items in which participants' responses to each item are then aggregated to give an overall score, it is important that each of the items relate to each other. For example, consider a questionnaire designed to measure the home numeracy environment. All the questions should be measuring the number activities that parents do with their child at home. If this is the case then the items will be highly correlated with each other and show high internal reliability. However if some of the questions are measuring a more general home learning environment then these items will be less highly correlated and the questionnaire will have low internal reliability.

One way internal reliability can be measured is by using split-half reliability. Split half reliability is measured by randomly dividing the items on a measure into two halves and the correlation between the participants' scores on these two halves is calculated. If the measure has good internal reliability then the correlation should be high because participants who score high on one of the sets of items, should also score highly on the other set of items. A more sophisticated measure of internal reliability has been developed known as *Cronbach's alpha*. Cronbach's alpha calculates the average of all possible combinations of split-half reliability coefficients. A measure is considered to have good internal reliability if the Cronbach's alpha is 0.7 or above.

With regards to measuring the home numeracy environment, internal reliability is an important consideration because both the commonly used questionnaire and the novel text message methods aggregate several responses from participants to give an overall home numeracy environment score. The questionnaire methods usually report high Cronbach's alphas and therefore it is important that the text messages also achieve a high level of internal reliability.

Test-retest Reliability. One way to measure reliability is to evaluate the stability of the measure over time. Measuring test-retest reliability involves administering the same measure to the same sample on another occasion.

We should expect that if the measure is stable then there should be a high correlation between the measure at time 1 and time 2. A critical issue with regards to the test-retest measure of reliability is the length of time between administering the measures. Generally it is better to have a shorter gap between the measures than a longer gap because if the gap is too long then the correlation could be lower due to a change in the construct over time and not due to the reliability of the measure. However if you leave too short a gap between the measures the results may be affected by practise or boredom effects.

With regards to the home numeracy environment, it is important to consider if the home numeracy environment is expected to be stable over time. I believe that test-retest reliability would be difficult to measure with the home numeracy environment measures, used in this study, because the home numeracy environment is variable over time. The amount of number activities can vary between weeks and months and many other variables can change which in turn could effect the number activities done at home. Therefore, it is difficult to determine if the correlation would be due to the reliability of the home numeracy environment measure or the change in the number activities parents are doing with their child.

Inter-rater Reliability. This is another way to measure the reliability of a measure when a great deal of subjective judgement is involved. This can happen in multiple situations, for example, when open questions need to be categorised or during observation when participant behaviours need to be classified. The data is often reviewed by two or more independent raters and the extent to which they agree gives a measure of inter-rater reliability. The data used to measure the home numeracy environment in this study is objective, therefore there is often no need to measure inter-rater reliability. However, when using an observation method to measure the home numeracy environment, inter-rater reliability becomes relevant (see Chapter 5, Section 5.3.2.1).

4.1.3.2 Validity

There are many types of validity ranging from internal validity (how sure can we be that the results are due to the variable that the researcher intended to study?), external validity (can the result be generalised beyond the research context?), ecological validity (are the findings applicable to everyday life?) and measurement validity (does the measure actually measure the intended concept?). While all types of validity are important, I will focus on measurement validity in this discussion as I am interested in whether the novel text message method is actually measuring the home numeracy environment. Measurement validity is also commonly referred to as construct validity.

I will follow Trochim's (2006) framework as he splits measurement validity into two areas: translation validity and criterion-related validity.

Translation validity. Translation validity investigates if the operalisation reflects the construct. There are two types of translation validity.

- **Face Validity.** Face validity is a judgement as to whether, based on face value, the operationalisation seems like a good translation of the construct. There is no strict measure of face validity and it is an intuitive process when developing a new measure. With regards to measuring the home numeracy environment, we can judge that both the questionnaire and text messages appear to measure the home numeracy environment but the text messages solve some of the problems associated with the questionnaire measure and therefore is judged to have better face validity.

- **Content Validity.** Content validity involves checking the operationalisation against the relevant content domain for the construct. Some areas of research may have a content domain, which is a list of items a measure must include to be classed as measuring a certain construct. However for other areas of research it is harder to decide on a content domain. It would be difficult to construct a content domain for the home numeracy environment as it is still a newly developing area of research and we are not sure what exactly should be included in all measures of the home numeracy environment.

Criterion-related validity. Criterion-related validity is concerned with predicting how the operationalisation will perform based on previous theory of the construct. Trochim (2006) splits this further into 4 types of criterion-related validity: predictive, concurrent, convergent and discriminant.

- **Predictive Validity.** One way to measure criterion-related validity of a measure is if the measure can predict something that in theory it should be able to predict. With regards to the home numeracy environment, it is thought and has been shown in some previous research (but not consistently) that children who do more number activities at home have better numerical skills than those who do less number activities at home. However, Chapter 2 showed how much the relationship between the home numeracy environment and mathematics performance can vary and therefore we should not use the relationship between the new measure and mathematics performance as a measure of validity.

- **Concurrent Validity.** Another way to measure the validity of a measure is to determine if two measures of the same construct produce the same pattern of results. This is achieved by administering two measures concurrently: an already existing valid measure and the new measure which you want to assess its validity. If the new measure produces the same results as the already valid measure then it would have good concurrent validity. With regards to the home numeracy environment, the questionnaire measure can be used to compare the results to the text message method. Both measures of the home numeracy environment should have the same relationship with other variables. If the questionnaire measure of the home numeracy environment relates to a variable then the text messages measure of the home numeracy environment should also be related to this variable.

- **Convergent Validity.** Convergent validity evaluates if different measures all intended to measure the same construct converge. For example, the questionnaire measuring the home numeracy environment should be related to the text message responses as both methods are aiming to measure (or converge on) the same construct, the home numeracy environment. If this is the case then parents' responses to the different measures should be positively correlated.

- **Discriminant Validity.** This type of validity, is the opposite to convergent validity and evaluates if the measure can discriminate between constructs that are theoretically different. In terms of the home numeracy environment it is important that the measures diverge from constructs such as the home literacy environment. Therefore, the home numeracy measures should not correlate highly with measures of the home literacy environment. However, previously it has been shown that the home literacy questionnaire correlates highly with home numeracy questionnaire, showing either that the home numeracy questionnaire does not have good discriminant validity or that parents who do lots of number activities also do lots of literacy activities.

4.1.4 Research questions

My main research questions for this study are:

1. Is using text messages to measure the frequency of activities a valid and reliable measure of the home numeracy environment?
2. Does the frequency of home numeracy activities relate to demographic variables, parents' expectations or attitudes?
3. Does the frequency of number activities relate to children's counting performance?

4.2 Method

4.2.1 Participants

82 parents of children aged 2 to 5 years old were recruited from 13 different playgroups (a mixture of church playgroups and SureStart centres) in the Nottinghamshire, Derbyshire and Leicestershire area. Only the primary caregiver was eligible to take part.

Participants were excluded from the study if they did not reply to a total of three text messages during the 3-week period. 8 participants were removed for this reason. Three participants were removed because their children had already started formal schooling.

Characteristics of Final Sample

The final sample was 71 parents. Parents completed the study between January and April 2015. The children ranged in age from 2:3 (years: months) to 4:6 (M = 3:5, $SD = 0.64$; 60% female). Responding parents ranged in age from 23 to 46 years ($M = 34$ years, $SD = 5.06$). One parent failed to provide their age. The final sample was entirely made up of mothers. One father was involved in the study but did not complete the text messages and was therefore excluded. All families spoke English at home, with 6% speaking another language (Cantonese, Thai, Chinese). 41% of parents reported having a GCSE qualification, 20% having A-Level qualification, 30% having bachelors degree, 8% having masters degree and 1% having PhD as their highest qualification. SES was calculated by postcode using the 'Office for National Statistics' website (see *http://www.neighbourhood.statistics.gov.uk/dissemination/*). All 32,482 neighbourhoods in England have been ranked on a range of deprivation topics. The rank is then divided by 32,482 to give a percentile of deprivation for the area, with 0 representing very low SES and 1 representing very high SES. The SES rank for this study ranged from .005 to .990 ($M = .448, SD = .270$). Parents also reported the amount of time their child spent in childcare (nursery, pre-school or child-minder). This ranged from 0 to 27 hours ($M = 12, SD = 7.76$), with 21% reporting no childcare hours (i.e. they were cared for entirely at home). One parent failed to report childcare hours.

There were no significant differences between the final sample and the group that were removed with regards to child age, child gender, parent age, other languages spoke at home and SES. There was however a significant difference between childcare hours,$t(79) = -3.511, p = .001$, with parents who completed the study reporting more childcare hours ($M = 12, SD = 7.76$) than parents who did not complete the study ($M = 4, SD = 6.17$).

4.2.2 Procedure

The study consisted of two questionnaires for the parent to complete and text messages to reply to every day for a three-week period. Participants were approached and asked to take part in the study at local playgroups. Parents completed a demographic questionnaire at the playgroup and then were given details and a leaflet about the text messages. The text messages commenced on the day of sign up and lasted for 21 days. The text messages were sent for 21 days based on the average length time other ecological momentary assessment studies had recorded data for (Alfvn, 2010; Axn, Bodin, Kongsted, et al., 2012; Christie et al., 2014). Each text message during the study was sent at 8pm. Once the text messages were complete, the participant received the parent questionnaire exactly one week later. Participants were given the choice when they signed up to the study as to whether they received the parent questionnaire by post or email. Upon completion, participants were posted a £10 voucher as an inconvenience fee.

4.2.3 Measures

Demographic Questionnaire

This questionnaire was given to participants at the playgroup and included questions on the child's age, number and age of siblings, childcare arrangements and hours spent at preschool, child-minders and nursery. It also included questions about the parent with regards to age, highest qualification (GCSE, A level, Bachelors degree, Masters degree, PhD) and highest mathematics qualification (GCSE, A level, Bachelors degree, Masters degree, PhD). The questionnaire asked for postcodes to calculate SES and languages spoken at home. Participants also provided their mobile phone number in this questionnaire in order to send the text messages and indicated whether they wished to receive the last questionnaire by post or email.

Text Messages

The text messages were set up and sent using an online program called Text Tank (Texttank, 2017). Participants' mobile numbers were entered and a text message was sent to the participant everyday for 21 days, including weekends. The text message asked the participants "Have you done any number activities with your child today? Reply yes or no." The text messages were sent every night at 8pm. In the leaflet given to parents at the beginning of the study, they were given a definition of number activities. The definition was "Any activities that you may do with your child, in which your child uses numbers. This could include counting, singing a counting song, measuring ingredients for a cake or telling the time. These are just a few examples, there are lots more". The definition tried to be as a general as possible so participants could decide for themselves what to class as a number activity. The activities that parents are more likely to be aware of and report on are direct formal activities where mathematics is explicit. In the systematic review (Section 2.4.2.3) I showed that direct/formal activities are more likely to be linked to mathematics performance and therefore it is important that the definition is not too prescriptive for parents to include all types of formal direct number activities. The leaflet also emphasised the importance of the participant only reporting the activities that they had personally done with their child, and not to include any activities that their child may have done with any other family members or activities their child may have done in preschool or nursery. This was also emphasised in the questionnaire.

Parent Questionnaire
Parents received the questionnaire in paper or electronic format exactly one week after completing the text messages. If the questionnaire was received by post, participants were supplied with a stamped envelope to post the questionnaire back. If the participant received the questionnaire by email, they were given a link to the survey to complete online. Most of the questions were taken from the questionnaire used by Skwarchuk et al. (2014). This questionnaire was chosen as this is a well cited paper investigating the relationship between the home numeracy environment and mathematics performance (Ciping et al., 2015; Niklas et al., 2016; Huntsinger et al., 2016).

- Home Numeracy and Literacy Activities.
  This included a list of 28 home learning activities, 13 of which focus on numeracy, 11 of which focus on literacy and 4 of which do not have specific literacy or numeracy content. Parents indicated how often they personally did each of these activities with their child on a rating from 1, indicating rarely/never to 5, indicating most days per week. These items were taken from Skwarchuk et al.'s (2014) questionnaire. It was emphasised that the parents should only report activities that they had personally done with their child. Furthermore, parents were only asked to report how often they had done these activities in the last month.

- Academic Expectations.
  Parents were asked to indicate how important they thought it was for their child to reach certain benchmarks before starting school. They were given a list of 13 benchmarks (6 applying to numeracy and 7 applying to literacy). Parents were asked to rate these from 1 (unimportant) to 5 (extremely important). Some items were extremely advanced for this age group (such as read a chapter of book or count

to 1000) to minimise response biases. These items were taken from Skwarchuk et al.'s (2014) questionnaire.

- Literacy and Numeracy Attitudes.
  Parents were asked to rate their agreement with 4 statements ("I find maths/writing enjoyable", "I avoid situations involving maths/writing") on a scale from 1 "strongly disagree" to 5 "strongly agree".

- Counting.
  Parents were also asked to report how high their child could count. They were also asked if they had asked their child to count in order to answer the question. This question has been used in a previous study (Blevins-Knabe & Musun-Miller, 1996) and it was found that the number the parents gave was accurate when compared to children's actual counting ability.

- Text Messages.
  Parents were asked if answering the text messages had made them do more number activities with their child than they would normally do? Parents responded 'yes' or 'no' to this question.

4.3 Results

In this section I will start by detailing how all the data was reduced for the analysis. I will then present some preliminary analysis of the text messages to show that the text message method works as a measure of the home numeracy environment. Moving on from this I will investigate the relationship between the two measures of the home numeracy environment, followed by an investigation of the relationship between the home numeracy environment measures and demographic variables, parents' mathematics expectations and attitudes. Finally, I will present the relationship between the home numeracy environment measures and children's counting skills. The key hypotheses are :-

1. The text messages will be significantly positively correlated to the home numeracy environment questionnaire.

2. The frequency of home numeracy activities (measured by both the questionnaire and text messages) will be significantly correlated to

demographic variables, parents' expectations and attitudes.

3. The frequency of home numeracy activities (measured by both the questionnaire and text messages) will be significantly correlated to children's counting performance.

4.3.1 Data reduction

4.3.1.1 Questionnaire

For the home numeracy and literacy activities from the parent questionnaire, two composite scores were created by summing the scores of the different activities relating to numeracy and literacy. Both the home numeracy environment composite and the home literacy environment composite had good internal reliability (Cronbach's alpha = .82, .89 respectively). Two composite scores were also created for the parents' mathematics and literacy expectations. Two further composites were created for parents' mathematics and literacy attitudes. Table 4.1 shows the means, composite means and Cronbach's alphas for the individual activities, expectations and attitudes.

4.3.1.2 Text messages

There was some missing data when participants had failed to reply to messages (maximum 3 days of missing data per person),therefore a mean was calculated from the messages that participants had replied to. This meant that no assumptions were made about what may have happened on the days when parents failed to reply. The observed internal reliability (Cronbach's alpha) was .72 for the 21 days the text messages were sent. This shows that the text messages have good internal reliability.

4.3.2 Preliminary analysis of the text messages

Before analysing the relationship between the text message measure of the home numeracy environment with the questionnaire measure, demographic variables, parents' expectations and attitudes and children's mathematics skills it is important to evaluate if the text messages have worked as a

	Mean	Standard Deviation	Cronbach's Alpha
Home Numeracy Environment (HNE)	**41.22**	**8.843**	**.818**
I help my child learn simple sums (e.g. 2+2).	1.96	1.247	
I encourage my child to do maths in his/her head.	1.51	1.073	
We talk about time with clocks and calendars.	2.68	1.452	
I help my child weigh, measure and compare quantities.	2.11	1.222	
We play games that involve counting, adding or subtracting.	3.68	1.262	
I teach my child to recognise printed numbers.	3.94	1.027	
We sort and classify by colour, shape and size.	3.46	1.263	
I ask about quantities (e.g. How many spoons?).	4.01	1.089	
We play board games or cards.	3.09	1.359	
I encourage collecting (e.g. cards, stamps, rocks).	2.13	1.287	
I help my child recite numbers in order.	4.37	.975	
We sing counting songs (e.g. five little monkeys).	4.23	1.058	
I encourage the use of fingers to indicate how many.	3.99	1.270	
Home Literacy Environment (HLE)	**34.15**	**10.035**	**.886**
I help my child read words.	2.73	1.558	
I ask my child to point to words/letters when we read.	2.94	1.548	
I teach my child to recognise printed letters.	3.58	1.284	
I help my child to print words.	2.46	1.350	
We identify words on signs (e.g. stop, exit)	2.49	1.351	
I teach my child the sound of letters.	3.75	1.262	
I introduce new words and their definitions to my child.	3.38	1.324	
I help my child to sing/recite the alphabet.	3.54	1.329	
We make up rhymes in songs (e.g. down by the bay).	3.17	1.444	
I ask questions when we read together.	3.76	1.325	
We visit the library for children?s books.	2.35	1.001	
Math Expectations	**14.23**	**4.115**	**.835**
Count to 10	4.34	.970	
Count to 100	2.14	975	
Read printed numbers up to 100	2.17	956	
Know simple sums (e.g. 2+2)	2.63	1.045	
Count to 1000	1.39	.686	
Know multiplying (e.g. 2 x 6)	1.51	.826	
Literacy Expectations	**20.15**	**6.112**	**.905**
Know some alphabet letters	3.89	1.115	
Write his/her name	3.65	1.160	
Know all 26 alphabet letters	3.06	1.252	
Write all 26 alphabet letters	2.37	1.059	
Read a few words	2.72	1.031	
Read simple picture books	2.99	1.201	
Read chapter books	1.49	.772	
Mathematics Attitudes	**5.31**	**.935**	
I find math enjoyable.	3.31	1.090	
I avoid situations involving math.	2.00	1.134	
Literacy Attitudes	**5.55**	**.875**	
I find writing enjoyable.	3.85	1.70	
I avoid situations involving writing.	1.70	.852	

Table 4.1: Means, standard deviations and Cronbach's alphas of activities, expectations and attitudes taken from the questionnaire

measure of the home numeracy environment. Therefore, I will start by investigating if there is a weekend effect when sending the text messages, then I will investigate if answering the text messages made parents report doing more number activities and finally I will investigate how long the text messages should be sent for in order to reach a good level of reliability.

Is there a weekend effect?

One concern was that there may have been a difference in responses for weekdays and weekends (i.e. parents may do more number activities in the week than at weekends or visa-versa) and this may affect the reliability of the measure. The mean was calculated for each participant for weekends and weekdays. However, there was no significant difference in the mean number of activities reported at the weekend ($M = .660, SD = .247$) compared to the activities reported in the week ($M = .673, SD = .208$), $t(70) = .496, p = .621$. Therefore, the measure is not be affected by the day of the week.

Did answering the text messages make participants report doing more number activities than normal?

If answering the text messages made parents do more number activities than they would normally do, then the text messages may not provide a valid measure of the frequency of activities. 60% of participants reported that the text messages led them to do more number activities with their child than normal. A two way ANOVA, with group (more activities, not more activities) as between groups factor and time (first 10 days, last 10 days) as within groups factor, was conducted to examine this effect on the first 10 days compared with the last 10 days. The ANOVA revealed a significant main effect of time, $F(1,\ 69) = 6.51$, $p = .013$, $\eta_p = .086$, with higher number of reported activities in the first 10 days than the last 10 days ($M =$ .70, $SD = .197$; $M = .64$ $SD = .224$, respectively). There was no significant main effect of group, $F(1,\ 69) = .440$, $p = .510$, $\eta_p = .006$, and the group by time interaction effect also did not reach significance, $F(1,\ 69) = 2.37$, $p = .128$, $\eta_p = .033$. It was important to investigate the effect over time and the interaction between group and time because for the first 10 days the text messages are still new to the parents and if the text messages did lead a group of parents to do more number activities you would expect this to have a greater effect in the first 10 days compared to the last 10 days, when the text messages have become more routine. However, this was not the case as both the main effect of group and interaction between group and time were non-significant. This shows that even though some parents felt that responding to the text messages led them to do more number activities, they were actually not reporting any more activities than the other participants.

How long do the text messages need to be sent for in order to reach a good level of reliability?
The measure had shown to be reliable when sending the text messages for 21 days, so the next question was how long do the text messages have to be sent to reach a good level of reliability? Cronbach's alpha was calculated for the first 5 days of the text messages being sent and then a day was added and the Cronbach's alpha was recalculated until 21 days was reached. Figure 4.2 shows there is a gradual increase in the observed reliability. The longer the text messages are sent for, the more reliable the method is and there is no obvious point before 21 days where the reliability stabilises. This suggests the longer the text messages are sent for the more reliable the measure.

Furthermore, a test-retest analysis was conducted to measure reliability between the 3 weeks the text messages were sent for. The text message data was separated into 3 weeks and the correlation between the three weeks was measured. Table 4.2shows the correlations between the three weeks. This test-retest analysis shows that the responses in all three weeks were significantly related to each other showing good reliability, however they are not not strongly related because of the nature of the variable being measured, the number of activities a parent does will differ from week to week.

4.3.3 Is there a relationship between the measures of the home numeracy environment?

Both the questionnaire and the text messages try to measure the home numeracy environment, therefore it is expected that they should be highly correlated. This is also known as convergent validity. The correlation between the questionnaire and text messages was $r = .37, p = .002, (r = .48, p < .001$, controlling for age). Both correlations (with and without controlling for age)

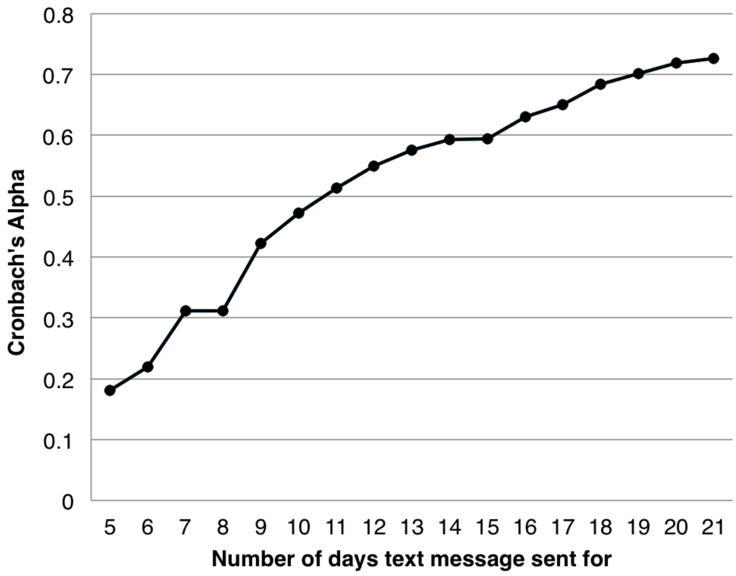

Figure 4.2: Reliability of text messages for a varying number of days

	First Week	Second Week	Third Week
First Week	—	.65**	.59**
Second Week		—	.46**
Third Week			—

Table 4.2: Correlations between text messages responses for Week 1, 2 and 3

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

were significant which suggests the two methods were related in some way, but they were not measuring exactly the same construct because if they were measuring the same concept we would expect higher correlations.

On the questionnaire, 58 parents (82%) reported that they did at least one numeracy activity "most days" with their child. If this was accurate then we would expect these participants to report that they were doing number activities most days when replying to the text message. Assuming most days might be considered to be 5 days out of 7 (70% of the time), 47% of parents who reported on the questionnaire that they did some activities "most days" reported doing activities less frequently than "most days" when replying to the text messages, with some parents reporting as low as 2 days a week (35%) (see Figure 4.3). This shows an inconsistency between the measures, which suggests either that they are not measuring the exact same construct or that they do not both do so accurately. When answering the text messages parents report daily the number activities they have done compared to recalling activities in a typical month on the questionnaire. Therefore, based on face validity, the text messages appear to be a more accurate and honest measure of the home numeracy environment, given the problems previously mentioned with the questionnaire.

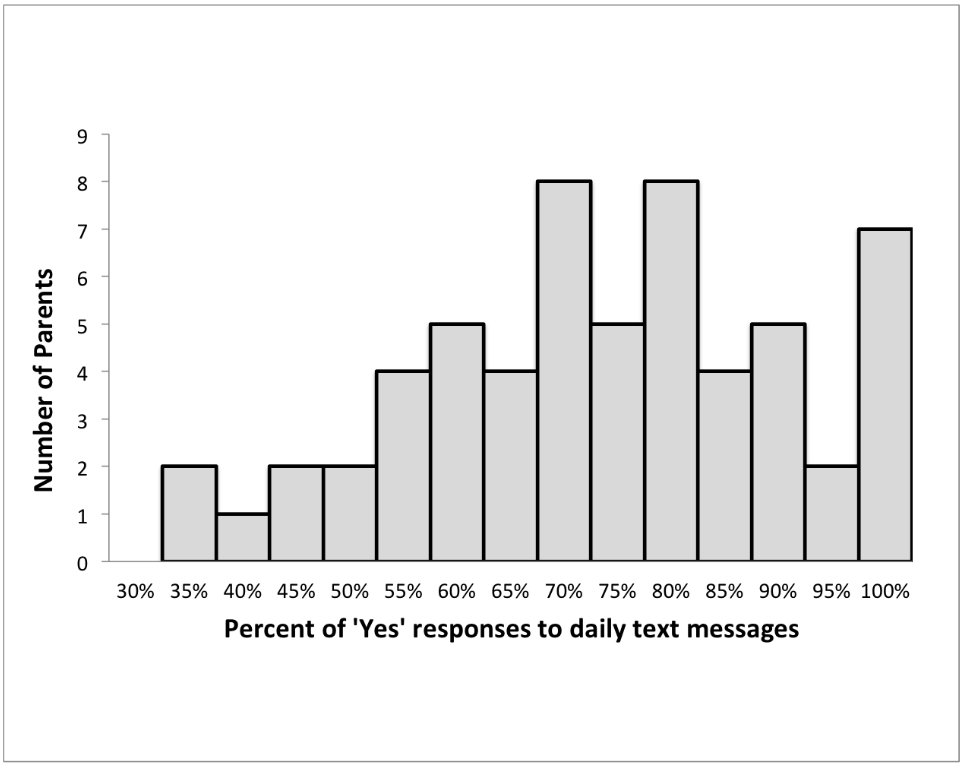

Figure 4.3: Response to text messages by parents who indicated that they did number activities most days on the questionnaire.

4.3.4 Does the frequency of activities in the home numeracy environment relate to demographic variables, parents' expectations and attitudes?

Table 4.3 shows the correlations between the questionnaire, text messages, demographic variables (age, childcare hours, SES, parents' highest qualification, parents' age), parents' mathematics expectations and parents' mathematics attitudes. The first two columns of table 4.3 show the results for the key hypothesis, the remaining columns show exploratory data. Many of the variables were not significantly correlated with either measure of the home numeracy environment (SES, parents' highest qualification, parents' age and parents' mathematics attitudes). Where variables were significantly correlated with only one of the home numeracy measures, I used Williams-Steiger tests to explore whether the strength of the correlations significantly differed.

There was a significant correlation between the questionnaire and parents' mathematics expectations ($r = .32, p = .002$) but the correlation was not significant between the text messages and parents' mathematics expectations ($r = .21, p = .06$). However the correlations were not significantly different ($p = .23$).

There was no significant correlation between hours in childcare and the questionnaire ($r = .20, p = .089$), but there was a significant correlation between hours in childcare and the text messages ($r = -.24, p = .043$). These correlations were significantly different, $t(67) = 3.56, p < .001$.

There was a significant correlation between child's age and the questionnaire ($r = .30, p = .01$), but there was no significant correlation between child's age and the text messages ($r = -.22, p = .07$). These correlations were significantly different, $t(68) = 4.24, p < .001$. However, there was also a significant correlation between child's age and hours in childcare because older children spend more time in childcare ($r = .57, p < .001$). This highlights the importance of controlling for age as many of the other variables are also significantly correlated with age (see Table 4.3). The children in this study ranged from 2 to 4 and a half years old and there are large differences between the activities a 2 year old may do compared to a 4 year old. Therefore in Table 4.3, the correlations controlling for age are also reported above the diagonal.

Once controlling for age, the only significant correlations with measures of the home numeracy environment were parents' mathematics expectations and SES. There was a significant correlation between the questionnaire and parents' mathematics expectations ($r = .37, p = .002$) but the correlation was not significant between the text messages and parents' mathematics expectations ($r = .23, p = .06$). However the correlations were not significantly different ($p = .23$). There was also a significant correlation between the text message method and SES ($r = .24, p = .048$) but there was not a significant relationship between the questionnaire and SES ($r = .03, p = .795$). However, again these correlations were not significantly different from each other($p = .08$).

Therefore, once controlling for age, there were no significant difference between the correlations for the text message method and the correlations for the questionnaire method. This shows that once age has been taken into account, the text message method has the same pattern of correlations as the home numeracy questionnaire. In other words, the text message method has good concurrent validity, this will be discussed in more detail in the discussion.

	1	2	3	4	5	6	7	8	9	10
1 Text Messages	—	.48**	.45**	−.12	.24*	−.09	.02	.23	.01	.26*
2 Questionnaire (Number Activities)	.37**	—	.76**	.07	.03	−.22	−.01	.37*	−.18	.32**
3 Questionnaire (Literacy Activities)	.33**	.78**	—	.14	.11	−.16	−.11	.35**	−.14	.36**
4 Childcare Hours	−.24*	.20	.27*	—	−.08	.19	.14	.07	.13	.04
5 SES	.12	.09	.17	.12	—	.40**	.27**	.08	.05	.19
6 Parents' Highest Qualification	−.11	−.07	−.04	.23	.37**	—	.34**	−.06	.07	.26*
7 Parents' Age	.04	.01	−.11	.14	.24*	.33**	—	−.11	.08	−.01
8 Parents' Mathematics Expectations	.21	.32**	.30*	.05	.08	−.09	−.12	—	−.01	−.01
9 Parents' Attitudes	.12	.16	.15	.12	.09	.19	.12	.11	—	−.05
10 Counting	−.04	.41**	.43**	.45**	.31**	.35*	.08	.10	.17	—
11 Child's Age	−.22	.30*	.31**	.57**	.27*	.23	.06	−.03	.05	.64**

Table 4.3: Correlations between the questionnaire, text messages, demographic variables, parents' expectations and attitudes. Below the diagonal are zero-order correlations with partial correlations controlling the age above the diagonal.

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

4.3.5 Does the frequency of number activities relate to children's counting performance?

In this study I was unable to measure children's mathematics skills directly and therefore a proxy measure for counting was used in this study by asking parents to report how high their child could count. Before controlling for age, there was a significant correlation between how high the parent reported the child could count and the questionnaire ($r = .41, p<.001$). However, there was no significant correlation between how high the parent reported their child could count and the text messages ($r = -.043, p = .724$). These correlations were significantly different, $t(68) = 3.444, p<.001$. However, because counting ability can vary from 2 to 4.5 year olds it is important we control for age. When controlling for age the correlations between the text messages and counting ($r = .26, p = .035$) and the correlation between the home numeracy questionnaire and counting ($r = .32, p = .009$) were both significant and not significantly different from each other ($p = .61$). Therefore both measures of the home numeracy environment were related to children's counting ability.

4.4 Discussion

The study presented in this chapter had three main research aims: firstly to investigate the reliability and validity of the text message method; secondly to explore the relationship between the home numeracy environment, demographic variables, parents' expectations and attitudes and thirdly to extend previous findings in investigating the relationship between the home numeracy environment and mathematics performance. I will now discuss the findings of this study in order to answer the research questions posed.

With regards to the first aim, I measured several different types of reliability and validity for the text message method. As discussed in Section 4.1.3.1, the only type of reliability that was relevant to the study was internal reliability. The text message method had a high Cronbach's alpha and it was shown that the longer the text messages were sent for, the more reliable the measure was. There were several different measures of validity that I used to assess the text message method. Firstly based on the face validity of the measures, the text messages appear to be the more accurate measure because parents were not relying on retrospective accounts of activities.

Secondly, the measures showed convergent validity, because the questionnaire and text message method were significantly correlated but the strength of the correlation was quite small $r = .48$. If the two measures were accurately measuring the same construct, the home numeracy environment, then it might be expected that they would be highly correlated. Nevertheless, because there are problems with the questionnaire method (see Section 4.1.1) these could lead to the questionnaire not accurately measuring the frequency of activities and therefore the questionnaire may not be strongly correlated with a new, more accurate measure. Alternatively, the two measures could be measuring different types of number activities. The questionnaire is limited to a list of activities, which contains both formal and informal activities, whereas the type of activities measured by the text messages depends on parents' interpretations of the text message question. When asked about number activities parents are more likely to report on formal activities, where the inclusion of numbers is explicit. Therefore the low correlation could be because the text messages are measuring more formal number activities whereas the questionnaire is measuring both formal and informal activities. In conclusion, the two measures are either not measuring exactly the same construct or one measure is just more sensitive than the other. The histogram (Figure 4.3) adds support to the fact that they are not measuring exactly the same construct as there is variance in the amount of number activities reported using the two measures.

The next measure of validity that was assessed was concurrent validity. This was evaluated by considering the relationship both measures of the home numeracy environment had with other variables. If the methods were measuring the same construct, it would be expected that they would correlate with the same variables. As highlighted in Chapter 2 Section 2.4.2.7, there can be huge differences between home numeracy activities a two year old may do and the home numeracy activities a four year old may do, and because the questionnaire focuses on a list of activities that may not be relevant for all age ranges, it is important we take age into account. Age was significantly correlated with the questionnaire measure, showing that older children do more activities on the list. Furthermore, other demographic variables were significantly correlated with age such as childcare hours and SES. Therefore, I only discuss the correlations that have been controlled for age. Once controlling for age there was no significant differences between the correlations for the text message method and the questionnaire method for any of the other variables.

The last measure of validity that can be evaluated is discriminant validity. The home literacy environment is a separate construct to the home numeracy environment and therefore measures of the home numeracy environment should not be strongly correlated with measures of the home literacy environment. However, the text messages were significantly correlated with the home literacy environment measure $r = .45$. showing that they did not discriminate completely between the home numeracy and home literacy environment. However, these results could also occur if parents who do lots of number activities also do lots of literacy activities and therefore may not be the best measure of discriminant validity.

Overall, the text message method appeared to be a reliable measure with good internal reliability. The validity of the measure also appeared be good with high concurrent validity, good convergent validity but low discriminant validity.

With regards to the second aim of the study, correlations were calculated between the home numeracy environment measures and the demographic variables, parents' mathematics expectations and parents' attitudes. All but one of the demographic variables were non-significantly correlated with the home numeracy environment. This is not overly surprising as the relationship between the home numeracy environment and demographic variables has not been consistent across other studies (See Chapter 2, Section 2.3.1). SES was significantly correlated with the text message measure of the home numeracy environment ($r = .24$). SES was measured using a neighbourhood variable (calculated by postcode) and an individual variable (parents' highest qualification). The variables were significantly correlated, but not strongly, showing the importance of using multiple measures of SES. SES (measured by postcode) was significantly correlated with the text message method but not with the questionnaire method. Many previous studies have shown that SES is linked to the home numeracy environment (DeFlorio & Beliakoff, 2014; Levine et al., 2010; Kluczniok et al., 2013; Skwarchuk et al., 2014; Ciping et al., 2015) and if we are to believe the text message method is more valid and reliable measure of the home numeracy environment than the questionnaire, then we can also add evidence for the link between SES and home numeracy environment.

Parents' mathematics expectations correlated significantly with the questionnaire, but not the text messages. There was no significant difference in the correlations. As the home numeracy questionnaire is administered with the questions about parents' mathematics expectations for their child, it could be that parents' are influenced by their reported number activities to have higher expectations for their child. Furthermore, some parents may have been influenced by the questions that were included to minimise response bias. Items such as read a chapter book and count to 1000 were included in the expectations question on the questionnaire, these were intended to minimise parents just answering 'important' to all skills and these are not realistic expectations for a 5 year old. However, some parents may have answered the question without thinking realistically about their child's abilities when starting school and therefore selected important for these items. Furthermore, three parents reported that they do not like to have expectations for their child and therefore reported that none of the activities were important. Both these types of parents may have skewed the data.

Both measures also failed to find a significant correlation with parents' mathematics attitudes. Several papers have also failed to find a relationship between the home numeracy environment and parents' attitudes (Skwarchuk, 2009; Vandermaas-Peeler & Pittard, 2014). This could be due to only asking two basic questions about parents' feelings towards mathematics. This should be extended on in future research.

The final aim of the study was to investigate the relationship between the home numeracy environment and mathematics performance. Both measures of the home numeracy environment were significantly correlated with children's counting ability after controlling for age. The correlations were weak but still significant and add to the mixed findings on the relationship between the home numeracy environment and mathematics performance.

It is important to remember that this study only used a proxy measure for counting, asking the parents' how high their child could count. Even though this has been shown to be a reliable measure of how high the child could actually count (Blevins-Knabe & Musun-Miller, 1996), we are still only measuring one area of mathematics ability. As discussed in Chapter 2 Section 2.4.2.4, the relationship between the home numeracy environment and mathematics achievement can vary depending on the mathematics measure used. Therefore, investigating the relationship between a text message measure of the home numeracy environment and a standard measure of mathematics achievement would be worthwhile.

4.4.1 Implications for future research

The text message method has shown to be a valid and reliable method to measure the frequency of activities in the home numeracy environment. Furthermore, the text message method has improved ecological validity compared to the questionnaire. However a limitation to the text message and questionnaire method is that they are both self-report measures of the home numeracy environment and could be subject to social desirability bias. One other measure, that has been used previously to measure the home numeracy environment, is an observation measure (Anders et al., 2012; DeFlorio, 2013; Levine et al., 2010; Ramani et al., 2015). The observation measure reduces social desirability bias but it is time consuming to conduct. Questionnaire and observation measures are often used interchangeably to measure the home numeracy environment but it has been suggested they may not be measuring the same concept (Missall et al., 2016) and in order to evaluate the text message method further it is important to investigate its relationship to an observation measure.

It is also important that the findings reported in this study can be replicated. In the high profile Open Science Collaboration paper (OpenScienceCollaboration, 2015), it was highlighted that as many as two thirds of psychology studies failed to replicate their original findings. Therefore, before we can be sure that the text message method is a valid and reliable measure of the home numeracy environment and can be used in future research to inform us further about the home numeracy environment, we should try to replicate the findings with a new sample.

Finally, as mentioned previously a more general mathematics measure should be used as counting is only a small area of mathematics ability.

This leads to the design of my next study with the aim to replicate these findings, to compare the text message method with an observation measure and further investigate the link between all three home numeracy measures and general mathematics performance.

Chapter 5

Measuring the home numeracy environment (Study 2)

In the previous chapter I presented a novel text message method to measure the home numeracy environment. The key finding was that the text message method was shown to be a valid and reliable method, however there were a few limitations to the study that should be addressed. Therefore this chapter will present my next study designed to replicate and expand the findings from Study 1.

5.1 Introduction

The home numeracy environment has predominately been measured by two methods: questionnaires and observations. However, we are not clear about the relationship between these two measures of the home numeracy environment. As shown in Chapter 2 Section 2.4.2.3, there are different versions of these measures and they have shown varying relationships with mathematics performance. I have suggested that the differences in this relationship could be due to the type of measure used, however we do not know how the measures are related to each other. The different measures could be measuring different aspects of the home numeracy environment, which in turn may relate to different areas of mathematical ability.

Ramani et al. (2015) used both a questionnaire and an observation to measure the home numeracy environment. However they did not compare the two methods. They did observe how the two measures related to mathematics performance but they divided both the questionnaire and observation measures into different categories. Therefore from this study we cannot be sure how the two measures taken as a whole related to mathematics achievement.

Only one study, to my knowledge, has directly compared the home numeracy questionnaire and a home numeracy observation measure. Missall et al. (2016) used a home numeracy questionnaire that they had developed themselves in a previous study (Missall et al., 2014) and compared the results to an observation measure. The observation measure involved inviting parents and children into the lab and asking them to play with toys as they normally would at home for 20 minutes. The measure of the home numeracy environment was the amount of number talk from the parent. However, they found that the correlation between the questionnaire and observation measure was not significant ($r = .20$). This suggests that the observation measure is assessing a different aspect of the home numeracy environment than the questionnaire. However there are limitations to this study because they used a questionnaire that they had developed themselves and in a previous study had shown that the questionnaire did not relate to children's mathematics performance, therefore there may have been problems with the validity of the questionnaire and that could result in a lack of correlation between the two measures. Furthermore, Missall et al. (2016) did not measure children's mathematical ability. Therefore, even though the study shows that the measures are not related, we do not know which of the measures, if either, relate to mathematics performance.

There are many different observation measures used to measure the home numeracy environment, from observing parents and children in everyday life in their home to observing parents and children playing with toys in a lab. I used a similar observation measure to Ramani et al. (2015) because this method showed a significant relationship with mathematics performance. The method used by Ramani and colleagues (2015) was to invite parents and children into the lab and ask them to play with three different activities. The three activities were a book, a puzzle and a game. All the activities gave the parent the opportunity to discuss numbers. Parent and child number talk was recorded as a measure of the home numeracy environment.

Therefore in this study I will compare three measures of the home numeracy environment: a commonly used questionnaire, an observation measure and the novel text message measure. I will also evaluate how each of the measures relate to children's mathematics performance.

5.1.1 Research questions

My research questions for this study were:-

1. Are all three measures of the home numeracy environment valid and reliable and do the three measures of the home numeracy environment relate to each other?
2. How is the home numeracy environment related to demographic variables, parents' expectations and attitudes?
3. How is the home numeracy environment related to mathematics performance?

5.2 Method

5.2.1 Participants

81 parents and their 3 to 4-year old children participated in the study (*M* = 4:0, *SD* = .561, 50% male). 36 families were recruited from Summer Scientist Week at Nottingham University. Summer Scientist Week is an annual event run during the school holidays. Parents and children are invited to the University for half a day (3 hours) to take part in various research activities and games. 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. Another 45 families were recruited from local nurseries, playgroups and soft play centres to take part in the study in the toddler lab at Loughborough University. This was approved by Loughborough University's ethics committee. The procedure and set up was the same in the two settings. Participants were excluded from the study if they failed to complete one or more of the measures. With regards to the text message method if they failed to reply to 3 or more messages during the 3-week period this was classed as not completing the text message section.

6 parents failed to complete the text message section of the study. 4 children failed to complete all three activities in the observation measure and 3 children failed to complete the mathematics measure.

Characteristics of Final Sample

The final sample was 68 parent and child dyads. 26 families were from the Summer Scientist Week Sample and 42 families were from the lab sample. Demographic information was collected through a questionnaire. Parent reports indicated that all of the families reported speaking some English at home (87% only speaking English and 13% speaking one other language at home). The final sample included 65 mothers and 3 fathers. The majority of parents reported having a Batchelors degree or equivalent. Specifically, 1% reported having no qualifications, 12% had GCSE qualifications, 11% had A-Level qualifications or equivalent, 21% had a Batchelors degree, 13% had a masters degree and 10% had a PhD as their highest qualification. SES was calculated by postcode using the 'Office for National Statistics' website (see http://www.neighbourhood.statistics.gov.uk/dissemination). All 32,482 neighbourhoods in England have been ranked on a range of deprivation topics. The rank was divided by 32,482 to give a percentile of deprivation for the area, with 0 representing very low SES and 1 representing very high SES. The SES rank for this study ranged from .056 to .997 (M = .689, SD = .270). Parents reported the amount of time their child spent in childcare (nursery, pre-school or child-minder). This ranged from 0 to 45 hours (M = 19.5; SD = 10.08), with 3% reporting no childcare hours (i.e. they were cared for at home entirely). There was no significant difference between the Summer Scientist Week sample and the lab sample in relation to demographic variables apart from age, because the children that took part at Summer Scientist Week were all 4 years old.

5.2.2 Procedure

The study consisted of one parent-child interaction, a mathematics task, one questionnaire and 21 text messages to reply to every day for a three-week period. The study was advertised as a study focusing on play so that parents did not know that I was looking at number talk when they were taking part in the observation. Parents and children were invited to take part in the study. They completed the parent-child interaction in a room set up to feel like a home environment. After the interaction the child had a short break while the experimenter explained the text messages to the parent and asked them to complete the questionnaire. The child then took part in the mathematics task with the researcher. This was all done in one visit to the lab or one session at the Summer Scientist Week event. The parent then received a text message every day for 21 days. Each text message was sent at 8pm. Upon completion, parents were posted a £10 voucher as an inconvenience fee and children received a small teddy-bear when leaving the lab session.

5.2.3 Measures

5.2.3.1 Parent-child interaction

Parents and children were videotaped during a semi-structured play interaction. Dyads were seated on a blanket surrounded by cuddly toys. Dyads were told that there were three different activities for them to do together and they were to play with each activity for at least 3 minutes. The order of the three activities was counter balanced and each activity was presented to the dyad after they completed the previous one. After the instructions the experimenter sat quietly out of sight of the family. Each of the activities were chosen to elicit talk about numbers, although number talk was not needed to complete the all of the activities.

The three activities were a book, LEGO and a board game. The book was 'The Surprise' by Sylvia Van Ommen (Ommen, 2007). This book is a picture book that gives parents several opportunities to include number, but this is not essential. Parents and children were told "This book is called 'The Surprise'. This is a Spanish version of the book but that's ok because the book has no words. Together, use the pictures to tell the story". The LEGO activity was a bag of LEGO Duplo and four instruction cards. They were told "In the bag you will find some LEGO and instruction cards. Choose an instruction card and make the object on the card". The third activity was a board game named 'The Great Race' (Ramani Siegler, 2008; Siegler Ramani, 2008). The board included 10 identically sized spaces, alternating in a pattern of red and blue and had the numbers 1-10 listed consecutively. They were also given a spinner with the number 1 on one half of the spinner and the number 2 on the other half of the spinner. They also had 10 animal figures to choose from. On the game it had the following instructions "Place your characters on start to begin. Take turns spinning the spinner and move the number of spaces spun for each turn. The first character to reach the number ten space wins the game". Parents and children were told "For this game, each choose an animal for the race and then follow the instructions on the game". These three activities were chosen as they would hopefully be familiar to the children and similar to activities parents and children may do at home.

5.2.3.2 Parent questionnaire

This questionnaire included questions on demographics, the home learning environment and parents expectations and attitudes.

- Demographic Information.
  The first part of the questionnaire was on demographics. This included questions on the child's age, number and age of siblings, childcare arrangements and hours spent at preschool, child-minders and nursery. It also included questions about the parent such as age, highest qualification (GCSE, A level, Bachelors degree, Masters degree, PhD) and highest mathematics qualification (GCSE, A level, Bachelors degree, Masters degree, PhD). The questionnaire asked for postcodes to calculate SES and languages spoken at home. Participants also provided their mobile phone number in this questionnaire in order to send the text messages.

- Home Numeracy and Literacy Activities.
  This included a list of 28 home learning activities, 13 of which focus on numeracy, 11 of which focus on literacy and 4 of which do not have specific literacy or numeracy content. Parents indicated how often they personally did each of these activities with their child on a rating from 1, indicating rarely/never to 5, indicating most days per week. These items were taken from Skwarchuk et al.'s (2014) questionnaire. It was emphasised that the parents should only report activities that they had personally done with their child. Furthermore parents were only asked to report how often they had done these activities in the last month.

- Academic Expectations.
  Parents were asked to indicate how important they thought it was for their child to reach certain benchmarks before starting school. They were given a list of 13 benchmarks (6 applying to numeracy and 7 applying to literacy). Parents were asked to rate these from 1 (unimportant) to 5 (extremely important). Some items were extremely advanced for this age group (such as read a chapter of a book or count to 1000) to minimize response biases. These items were taken from Skwarchuk et al.'s (2014) questionnaire.

- Literacy and Numeracy Attitudes.
  Parents were asked to rate their agreement with 10 statements on a scale from 1 (Strongly disagree) to 5 (Strongly agree). The statements asked about their feelings towards mathematics and reading. In Study 1 mathematics attitudes were assessed using only two statements. Here I included a more comprehensive measure of parents' attitudes. These items were taken from LeFevre et al's (2009) questionnaire.

5.2.3.3 Mathematics measure

The arithmetic subtest of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) (Wechsler, 1967) was administered in accordance to standard procedure. There was a total of 20 questions of increasing difficulty and questions were asked until 4 questions were answered incorrectly or the end of the test was reached. The first 4 questions required the child to make non-symbolic comparisons on size or quantity (e.g. Here are some sticks. Point to the longest one). The next four questions required the child to count blocks (e.g. How many blocks are there? Count them with your finger). The last 12 questions required the child to solve word problems mentally (e.g. Johnny had 3 marbles. He lost one. How many did he have left?). Children received a raw score out of 20.

5.2.3.4 Text messages

The text messages were set up and sent using an online program called TextTank (TextTank, 2015). Participants' mobile numbers were entered and a text message was sent to the participant every day for 21 days, including weekends. The text message asked the participants "Have you done any number activities with your child today? Reply yes or no". The text messages were sent every night at 8pm. Parents were given a leaflet which included a definition of number activities. The definition was "Any activities that you may do with your child, in which your child uses numbers. This could include counting, singing a counting song, measuring ingredients for a cake or telling the time. These are just a few examples, there are lots more". The definition tried to be as a general as possible so participants could decide for themselves what to class as a number activity. The activities that parents are more likely to be aware of and report on are direct formal activities were the mathematics is explicit. In the systematic review (Section 2.4.2.3) I showed that direct/formal activities are more likely to be linked to mathematics performance and therefore it is important that the definition is not too prescriptive for parents to include all types of formal direct number activities. The leaflet also emphasized the importance of the participant only reporting the activities that they had personally done with the child, and not to include any activities that the child may have done with any other family members, or activities the child may have done in preschool or nursery. This was also emphasized when answering the questionnaire.

5.3 Results

In this section I will start by detailing how all the data was reduced for the analysis. Firstly, I will investigate the relationship between the three measures of the home numeracy environment. Secondly, I will report the relationship between the home numeracy environment measures and demographic variables, parents' mathematics expectations and attitudes. Finally I will present the relationship between the home numeracy environment measures and children's mathematics skills.

5.3.1 Data reduction

5.3.1.1 Observation

All speech from the parent and the child during the first three minutes of the interaction with each activity was transcribed. The number of occurrences of number words was recorded separately for the three activities for parents and children. Number words were classed as any symbolic numbers mentioned, apart from the number one, because one can have multiple meanings. The word one could be used in a numerical sense “you have one piece of LEGO” or in a non-numerical “we need that one” and therefore to avoid ambiguous uses of one, it was decided to not include the number one in the measure of number talk.[1] A total parent number talk variable and a total child number talk variable was created, which summed the amount of number talk across all three activities. Both parent and child number talk were used as a measure of the home numeracy environment.

The mean number of occurrences of number words from the parents was 16.75 ($SD = 7.26$) and the mean number of occurrences of number words of occurrences of number words from the child was 7.93 ($SD = 7.45$). This shows that there was good variation in the measure. Furthermore, Figure 5.1 and Figure 5.2 show the frequency distribution of volume of number talk for the parent and child.

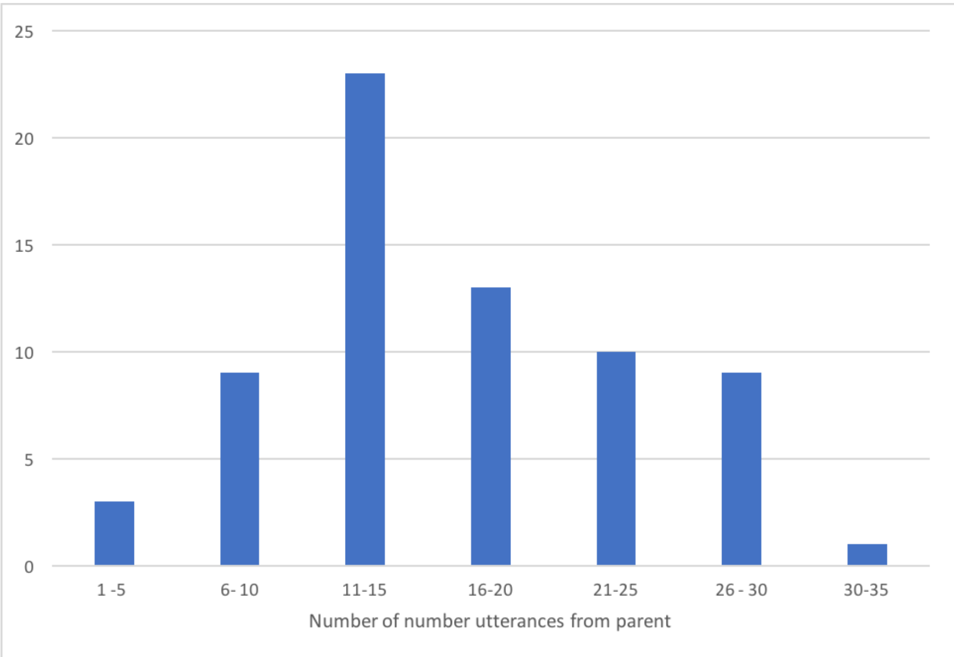

Figure 5.1: Frequency distribution of number words from parents

To measure the inter-rater reliability of the measure, an independent

[1]The data was also analysed with any occurrences of the number one, that were deemed to be numerical included, and these results did not significantly differ to the results reported.

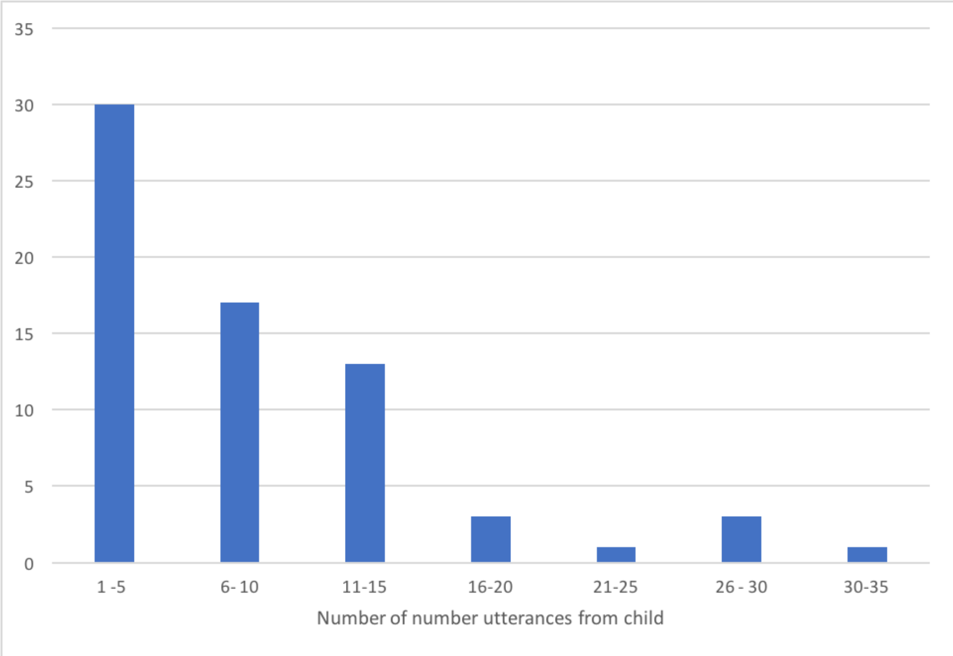

Figure 5.2: Frequency distribution of number words from children

observer coded a random subset consisting of 10% (7) of the observations. There were high levels of agreement between the coders with a Pearson r correlation of 1 for the parent number talk and .98 for the child number talk. This showed the observation measure had good inter-rater reliability.

5.3.1.2 Questionnaire

For the home numeracy and literacy activities from the parent questionnaire, two composite scores were created by summing the scores of the different activities relating to numeracy and literacy. Two composite scores were also created for the parents' mathematics and literacy expectations. Two further composite scores were created for parents' mathematics and literacy attitudes. The observed internal reliability for the questionnaire measure was .81 which shows good internal reliability.

5.3.1.3 Text messages

There was some missing data when participants had failed to reply to messages (maximum 3 days of missing data per person), therefore, a mean was calculated from the messages that participants had replied to. This meant

	1	2	3	4
1 Text Messages	—			
2 Questionnaire	.46**	—		
3 Parent Number Talk	.10	.07	—	
4 Child Number Talk	.18	.22	.00	—

Table 5.1: Partial correlations between the questionnaire, text messages and observation measures controlling for age

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

that no assumptions were made about what may have happened on the days when parents failed to reply. The observed internal reliability (Cronbach's alpha) for the text messages was .65. This was slightly lower than the Cronbach's alpha for Study 1 (.72), but still shows acceptable internal reliability.

5.3.2 Is there a relationship between the measures of the home numeracy environment?

In Study 1, the relationship between the questionnaire and the text message method was $r = .48$, controlling for age. As highlighted in Chapter 2 it is important to control for age because the questionnaire measure is highly correlated with age. Therefore when calculating the relationship between three measures I will report the correlations controlling for age. Table 5.1 shows the correlations between the three measures of the home numeracy environment.

Table 6.4 shows that the correlation between the text messages and the questionnaire ($r = .46, p$<.001) replicated from Study 1 ($r = .48, p$<.001) and both measures are converging on the same concept. However both variables (parent number talk and child number talk) from the observation measure failed to correlate with either the questionnaire ($r = .07, .22, p = .554, .071$) or the text message measure ($r = .10, .18, p = .442, .136$) suggesting that the self report measures do not measure the same concept as the observation measure. Furthermore, the two observation measure variables (parent number talk and child number talk) weren't correlated with each other ($r = .00, p = .972$). This suggests that the two observation measure variables also do not measure the same aspect of the home numeracy environment.

5.3.3 How is the home numeracy environment related to demographic variables, parents' expectations and attitudes?

Table 5.2 shows the correlations between the home numeracy environment, demographic variables, parents' mathematics expectations and attitudes. As in Study 1 all but one demographic variable failed to correlate with any measure of the home numeracy environment. This study found a significant relationship between the home numeracy environment and childcare hours but it was only significant between the questionnaire measure and childcare hours ($r = -.33, p = .01$). Furthermore there were no significant differences between the correlations for the home numeracy measures with the demographic variables.

Parents' mathematics expectations were significantly related to the questionnaire measure ($r = .36, p = .004$), as in Study 1, and child number talk ($r = .26, p = .045$). Furthermore, there was a significant difference between the correlation for the questionnaire and mathematics expectations and the text messages and mathematics expectations ($t(65) = 3.24, p = .002$) and parent number talk and mathematics expectations ($t(65) = 2.27, p = .027$).

Even though the correlations were not significant between the home numeracy environment measures and parents' mathematics attitudes, there was also a significant difference in the correlations with parents' mathematics attitudes for the two self report measures ($t(65) = 2.62, p = .011$).

These results will be discussed in more detail below.

5.3.4 How is the home numeracy environment related to mathematics performance?

Table 5.2 shows that mathematics achievement was only significantly correlated with child number talk ($r = .27, p = .03$). This fails to replicate the finding that the text messages and the questionnaire method were significantly related to mathematics achievement. Study 1 used parent-reported counting as a mathematics measure, however this study used a standard

	Text Message	Questionnaire	Parent Number Talk	Child Number Talk
Home Literacy Environment	.27*	.61**	.02	.21
Childcare Hours	−.23	−.33**	−.08	−.14
SES	−.05	−.20	−.02	−.13
Parents' Highest Qualification	−.06	.10	−.02	−.11
Parents' Age	.04	.02	−.08	.11
Mathematics Expectations	−.02	.36**	.00	.26*
Mathematics Attitudes	−.08	.24	.03	.10
WPSSI	−.12	.13	−.14	.27*

Table 5.2: Partial correlations for the questionnaire, text messages and observation measures with demographic variables, mathematics expectations and attitudes and mathematics performance, all controlling for age. WPPSI = Wechsler Preschool and Primary Scale of Intelligence

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

mathematics measure and could be the reason for the difference. Furthermore, there were significant differences between the correlations with the mathematics measure and the home numeracy measures. There was a significant difference between the correlation for the parent number talk measure and mathematics performance and the correlation for the child number talk measure and the mathematics measure ($t(65) = 2.45, p = .017$). There was also a significant difference between the correlation for the text message method and mathematics performance and the correlation for the child number talk variable and the mathematics performance ($t(65) = 2.59, p = .012$).

5.4 Discussion

The study presented in this chapter had three main aims: firstly to investigate the reliability and validity of the three measures of the home numeracy environment and investigate if they are all measuring the same construct. Secondly, to explore the relationship between the home numeracy environment, demographic variables, parents' expectations and attitudes. Thirdly, to extend previous findings in investigating the relationship between the home numeracy environment and mathematics performance. Below I will review the findings in relation to these aims.

5.4.1 Are all three measures of the home numeracy environment valid and reliable and how do they relate to each other?

With regard to the first aim, all the measures had good reliability showing that the measures were consistent. However, the measures did not show good convergent validity suggesting that they may be measuring unrelated aspects of the home numeracy environment.

The self-report measures (questionnaire, text messages) were significantly correlated, showing that they were converging on the same construct, but they were not strongly correlated, suggesting one measure could be more sensitive than the other. The observation measures (parent number talk, child number talk) did not correlate with either of the self report measures, showing that the frequency of number activities reported in the self-report measures is not related to how much parents and children talk about number when completing activities. This replicates Missall et al.'s (2016) finding that observation and self report measures are not related. But why are three measures all designed to measure the home numeracy environment not related?

One possibility is simply that different aspects of the home numeracy environment are not related to each other. The questionnaire and text messages measure the frequency of activities in the home numeracy environment, whereas the observation measure measures how much parents and children talk about number in a snapshot view of the home numeracy environment. Therefore it is possible that parents may report doing lots of number activities when answering the questionnaire and the text messages, but when play activities are observed, the activities that parents do, do not involve a lot of number talk. If this is true, it suggests that in future research we not only need to measure the amount of activities in the home numeracy environment, but also the quality of number talk within these activities.

Another possiblity for the lack of correlation is that the measure of parent and child number talk could also be seen as tapping into Spontaneous Focus On Numerosity (SFON). SFON is a recently developed construct that refers to an individual's tendency to focus on numerosity in the environment around them and it has been shown to link to children's mathematics ability (Hannula & Lehtinen, 2005; Hannula et al., 2010; Batchelor, 2014). Because parents and children are not aware that the focus of the study is on number talk, any number talk from the parent or child is spontaneous. Batchelor (2014) found that SFON was not related to the home numeracy environment but concluded that issues with the home numeracy questionnaire may have been the reason for the lack of correlation. However I have shown that the new text message measure of the home numeracy environment also is not related to SFON. This adds weight to the conclusion that SFON and the home numeracy environment are unrelated.

Another interesting finding from the observation measure is that the amount of number talk from the parent is not related to the amount of number talk from the child showing that both these variables are not converging on measuring the home numeracy environment. There are two possible explanations for this. Firstly one possibility is that there may be two groups of parents: those parents who may talk more about numbers when their child struggles with number, which would result in a negative relationship between parent and child number talk and those parents who talk a lot about numbers because their child is interested in numbers, which would result in a positive relationship between parent and child number talk. These two groups of parent and child dyads would balance out the relationship and result in a zero correlation. To test this I separated the results by those children who had a number talk score above the median (12) and those children who had a number talk score below the median and re-ran the correlation between parent number talk and child number talk. I found that both correlations were still small and non-significant (r = -.05, .21, p = .715, .458) showing that this was not the case. Another reason why parent and child number talk are not related may be that these measures are tapping into SFON and it has been shown that parental SFON is not related to child SFON (Batchelor, 2014).

The questionnaire measure was shown to have low discriminant validity when attempting to discriminate between the home numeracy and home literacy environment.

The next measure of validity that was assessed was concurrent validity. This was evaluated by considering the relationship both measures of the home numeracy environment had with other variables. If the methods were measuring the same construct, it would be expected that they would correlate with the same variables. The measures showed good concurrent validity in regards to demographic variables as the correlations were not significantly different between the measures. However, the measures had low concurrent validity when comparing the correlations for parents' expectations, attitudes and children's mathematics performance as there were significant differences between these variables and the three home numeracy environment measures. However, because the three measures are not converging on the same construct, it is possible that the different areas of the home numeracy environment relate to different variables.

The last measure of validity that can be evaluated is divergent validity. The home literacy environment is a separate construct to the home numeracy environment and therefore measures of the home numeracy environment should not be strongly correlated with measures of the home literacy environment. As seen in Table 5.2, the questionnaire measure correlated highly with the home literacy environment measure ($r = .61, p<.001$) and the text message method was also significantly correlated to the home literacy environment ($r = .27, p = .03$), showing low discriminant validity. However, the parent and child number talk observation variables were not significantly correlated with the home literacy environment ($r = .02, p = .852, textitr = .21, p = .098, respectively$), showing good discriminant validity.

In answer to the first research question, it appears that while all three measures appear reliable, it is harder to judge their validity. However, we can conclude that the self report measures are not measuring the same aspect of the home numeracy environment as the observation measures.

5.4.2 How is the home numeracy environment related to demographic variables, parents' expectations and attitudes?

The second aim of the study was to investigate the home numeracy environment and its links to demographic variables, parents' attitudes and expectations. The only demographic variable that was significantly linked to the home numeracy environment was childcare hours, suggesting the more hours a child spends in childcare the less number activities they do at home. The correlation was not significant in Study 1 but the variance in childcare hours was higher in this study. However, it is important to remember that this correlation was not significantly different to the non-significant correlations between the other home numeracy environment measures and childcare hours.

Parents' mathematics expectations were significantly correlated to two measures (questionnaire and child number talk). Firstly, this replicates Study 1 in which parents' expectations were significantly related to the questionnaire measure (r = .32). Parents' mathematics expectations for their child was not related to all home numeracy measures, but it is possible that parents' expectations for their child can increase certain measures of the home numeracy environment but not others. For example, parents' mathematics expectations may be linked to how good their child is at mathematics, therefore, if a child talks about numbers a lot, the parent may have higher expectations for them. Furthermore, parents who do a lot of number activities at home are probably reporting more number activities because they have high expectations for their child. The text message method only gives a yes or no measure to show if the child has done number activity that day, whereas some parents that have high expectations for their child may have done multiple number activities with their child and, therefore, the text message measure may not be sensitive enough to detect the relationship between mathematics expectations and home numeracy environment.

Even with an extended measure of parents' mathematics attitudes it was still not related to the home numeracy environment suggesting that how parents' feel about mathematics is not impacting the home numeracy environment for this age group.

5.4.3 How is the home numeracy environment related to mathematics performance?

Mathematics performance was only related to one of the variables: child number talk (r = .27). Study 1 showed a significant relationship between the two self-report measures and counting, however this study used a more general mathematics measure and showed that the self-report measures of the home numeracy environment were not related to a general mathematics measure. This could reflect the true correlation between the home numeracy environment and mathematics achievement in the UK because no previous studies have investigated this relationship within the UK. Alternatively it could be due to the mathematics measure used. The home numeracy environment may not be related to all areas of mathematics and previous studies using a general mathematics measure have also failed to find a positive correlation (DeFlorio, 2013; Ciping et al., 2015). Furthermore, Mutaf, Sasanguie, De Smedt, and Reynvoet (2016) showed that the association between formal home numeracy environment activities and children's calculation skills was mediated by basic processing skills (e.g. mapping). This suggests that the home numeracy environment may be related to basic processing skills, which in turn improve certain areas of mathematics performance. Therefore, future research should consider the relationship between home numeracy environment and basic processing skills.

There was no significant correlation between how much parents talked about number and mathematics achievement but the correlation was significant for how much the child talked about number. This suggests that what is important for mathematics achievement is not what the parent talks about or how many number activities they do with their child, but their child's focus on numbers. This relates back to the observation measure tapping into child's SFON and therefore it may be this aspect of the measure that is driving the relationship with mathematics.

5.4.4 Limitations

One limitation of this study is that the observation measure is a lab-based snapshot of activities. Furthermore, the activities given to the parents for the observation measure may not reflect how parents typically teach their child about number. Therefore, if the observation measured naturally occurring activities in the natural home environment there may be a stronger relationship with the self-report measures.

Another limitation is that this study looked at the three measures assessing a general home numeracy environment but, as discussed in Chapter 2 Section 2.4.2.3, some researchers argue that there are different types of number activities in the home (formal and informal) and have divided the questionnaire measure to reflect this. They have also shown that formal activities have higher correlations with mathematics achievement than informal activities. Therefore, the lack of correlation between the self-report measures and the home numeracy environment could be due to not distinguishing between formal and informal activities. This will be investigated in Chapter 6.

5.4.5 Summary of findings

This study used a multiple measures approach to assess the home numeracy environment and found that self-report measures of the home numeracy environment are not measuring the same concept as observation measures. Therefore, future studies should consider using several measures to investigate the home numeracy environment. It was also predicted that all measures of the home numeracy environment would relate to mathematics performance, however the only variable related to mathematics performance was child's number talk. It is possible that this measure also measured children's SFON which could drive the relationship with mathematics performance. Future research should investigate the relationship between the home numeracy environment and mathematics performance within the UK, to determine if the lack of correlation is due to the mathematics measure used or if this is the true relationship between the home numeracy environment and mathematics performance in the UK.

Chapter 6

Formal and informal home numeracy environment

This chapter investigates the distinction between formal and informal home numeracy environment. I conducted a confirmatory factor analysis using the questionnaire data from Study 1 and Study 2 and used the result of this to determine if the questionnaire should be divided into formal and informal categories for the analysis of Study 1 and Study 2.

6.1 Introduction

Questionnaire measures of the home numeracy environment are often split into formal and informal categories. These categories are sometimes given different names such as direct/indirect and basic/advanced, but typically include similar types of activities. As discussed in Chapter 2 Section 2.4.2.3, researchers have typically found stronger correlations between formal home numeracy activities and mathematics performance.

The questionnaire measure of the home numeracy environment used in Study 1 and Study 2 was taken from (Skwarchuk et al., 2014). In their paper they conducted a factor analysis which revealed two factors within their questionnaire which were named as basic and advanced home numeracy activities. However, because the activities are similar to activities defined in other papers to be formal and informal, I will use these terms in the rest of the chapter to fit with the typical terms used in this literature. Skwarchuk et al. (2014) found that mathematics performance was only related to formal numeracy activities. Therefore, the lack of correlation between the home numeracy questionnaire and mathematics achievement in Study 2, could be due to using a combined measure of home numeracy or the text message method could just measure informal number activities.

Before re-analysing the results of Study 1 and Study 2 with the questionnaire divided into formal and informal categories, it was important to investigate if this was an appropriate fit for the data. This was investigated using a confirmatory factor analysis.

6.2 Confirmatory Factor Analysis (CFA)

To investigate if Skwarchuk's two factor model is a better fit for the home numeracy questionnaire than a one factor model, a confirmatory factor analysis was conducted. A confirmatory factor analysis is a structural equation modelling technique, which is based on theory. Typically a researcher will have a theory about how different items load onto latent variables and a confirmatory factor analysis measures how well the proposed model accounts for the correlations between variables in the dataset. There are two necessary conditions that must be met before running a CFA. Firstly, the scale of the latent variable must be defined and secondly there must be at least as many observations as free parameters (i.e. positive degrees of freedom).

Therefore, the aim of this section is to conduct a confirmatory factor analysis on the questionnaire data from Study 1 and 2 combined to indicate if dividing the questionnaire measure into two categories would be a better fit for the data.

6.2.1 Method

For the confirmatory factor analysis I took all the questionnaire responses from Study 1 and Study 2 and combined them into one data file. This was to ensure that the sample size was sufficient in order to conduct a confirmatory factor analysis. In total there were 155 questionnaire responses. The same home numeracy questionnaire was used in Study 1 and 2 and the parents responding to the questionnaire all had children of similar ages and therefore I was able to combine the data from both studies.

I constructed two confirmatory factor analyses: one for the one factor model including all the home numeracy activities (CFA 1) and a two factor model using the categories given by Skwarchuk (CFA 2). Skwarchuk had two factors (formal and informal), but did not include all the activities listed in the questionnaire in these two factors. Both models are non-nested.

CFA analyses were conducted using the MPlus analytical software (Version 8). Model fit was assessed according to commonly accepted cut-off criteria (Hu & Bentler, 1999; Schermelleh-Engel, Moosbrugger, & Müller, 2003). These include Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) being close to .95 or greater, the Root Mean Square Error of Approximation (RMSEA) and the Standardized Root Mean Residual (SRMR) being close to .05 or smaller. There are problems with χ^2 fit statistic when using small sample sizes and, therefore, the incremental fit was assessed with the χ^2/df fit value (Schermelleh-Engel et al., 2003), in which values closer to 2 or smaller indicate good model fit. These goodness of fit measures are used to measure the fit of the two models. The Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Sample-size Adjusted BIC are used to compare the models, if the models are good fit for the data. The lower the value of these statistics the better the model.

To define the scale of each latent variable, the latent variables for each model were standardised and there were positive degrees of freedom for both models (df = 65, 11). Therefore, the two necessary conditions to run a CFA were met.

6.2.2 Results

Table 6.1 shows the fit statistics for both models. Both CFA 1 (with a single factor) and CFA 2 (with two factors) demonstrated a poor fit. The factor loadings are shown in Figure 6.1 and 6.2. The R-squared statistics shows how much of the latent variables' variance is explained by each of the observed variables. The R-squared statistic for both models is shown in Table 6.2.

As both models were a bad fit for the data, it was inappropriate to compare these models or draw any implications from the statistics given.

6.2.3 Discussion

From this analysis, we can conclude that neither of the factor structures are a good fit for the questionnaire. I will discuss possible reasons for the bad

Model		Fit Indices					
Measurement	χ^2	*df*	χ^2/df	*CFI*	*TLI*	*RMSEA*	*SRMR*
CFA 1	241.022	65	3.708	0.636	0.564	0.132	0.102
CFA 2	76.476	19	4.025	0.745	0.624	0.140	0.093
Acceptable fit			$\leq$5.0	$\geq$.90	$\geq$.90	<.08	$\leq$ .10
Good fit			$0\leq \chi^2/df \leq 2$	$\geq$.95	$\geq$.95	<.05	.00$\leq$SRMR$\leq$.05

Table 6.1: Fit Statistics for the Confirmatory Factor Analysis and the Corresponding Fit Criteria

Observed Variable	R-squared CFA 1	R-squared CFA 2
Simple Sums[1]	.218**	.698***
Maths in head[1]	.099	.698***
Time with clocks & calendars[1]	.217**	.153*
Weigh, measure & compare quantities[1]	.178**	.151*
Games involving add & subtract	.402***	N/A
Recognise printed numbers	.423***	N/A
classify by colour, shape & size	.407***	N/A
Ask about quantities	.369***	N/A
Board games & cards[2]	.189**	.419**
Encourage collecting[2]	.095	.345**
Recite numbers in order[2]	.239**	.076
Sing counting songs[2]	.178**	.132
Encourage use of fingers to count	.175**	N/A

Table 6.2: R-Square statistics for both CFA 1 and CFA 2. [1] Formal activities factor, [2] Informal activities factor.

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

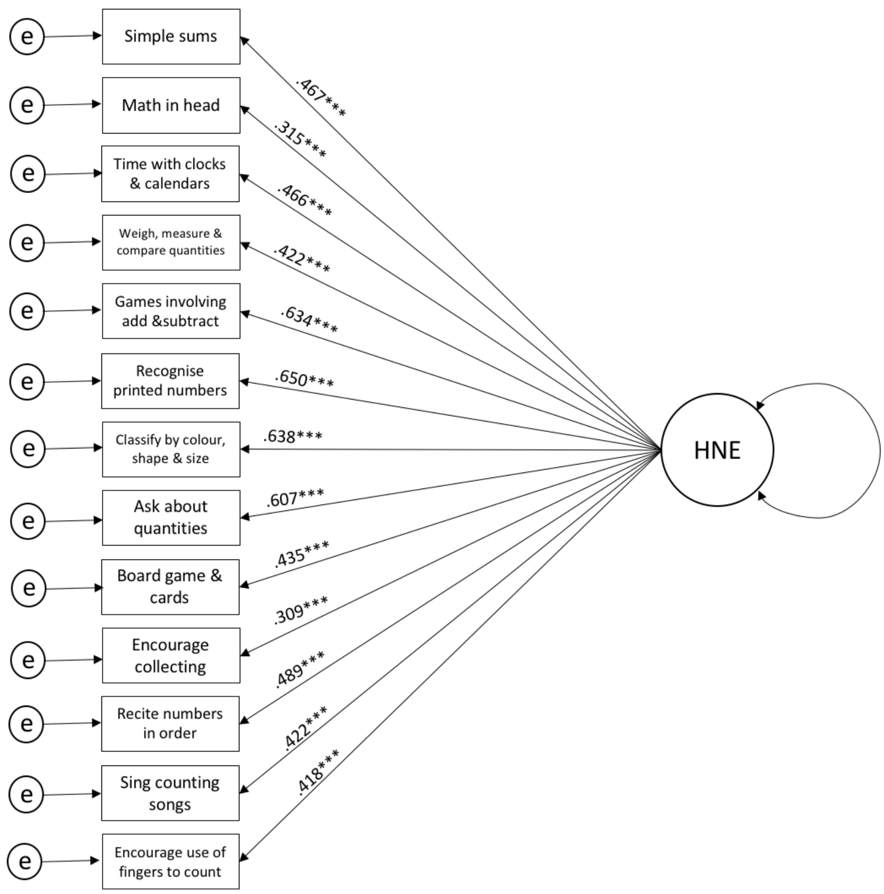

Figure 6.1: CFA 1 measurement model of the home numeracy environment. *** p <.001

model fit below and evaluate if we should use formal and informal categories.

One reason for poor fitting models may be due to sample size. A typical rule in the structural equation modelling literature is the N:q rule which is the ratio of cases (N) to the number of model parameters (q). Jackson (2003) suggests that the minimum ratio should be 10:1. For both CFA 1 and CFA 2 this criteria was met (155:14, 155:10 respectively). Furthermore, it has been shown the structural equation modelling techniques can be done with samples as small as 22 (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). Therefore, sample size should not be an issue in this case.

If the model fit has not been influenced by sample size, then it shows that the one factor and two factor model are not good models for the data. This adds support to my argument that the questionnaire is not a good

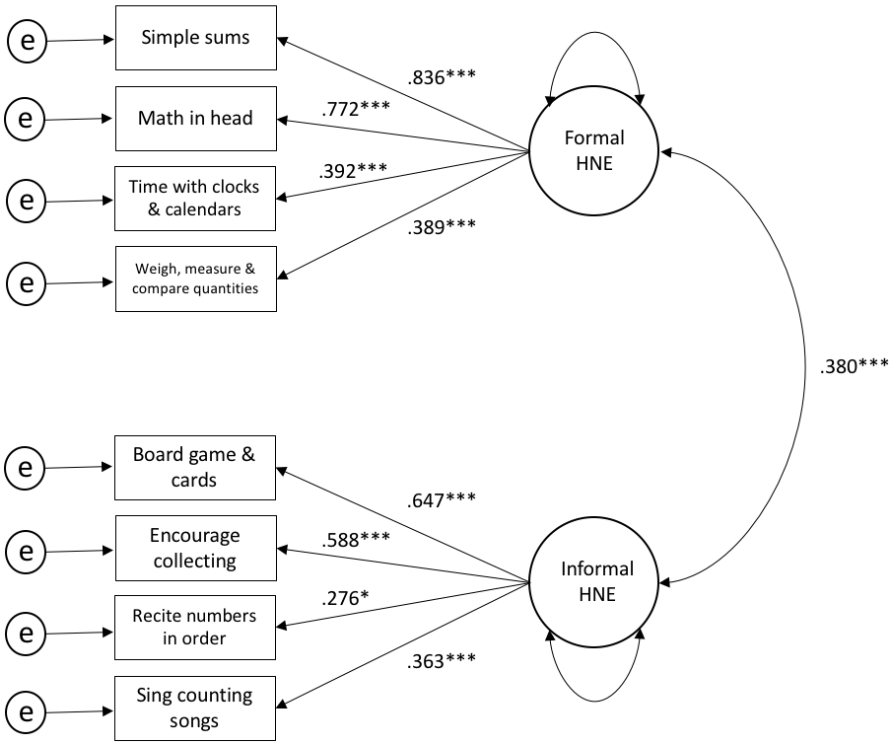

Figure 6.2: CFA 2 measurement model of the formal and informal home numeracy environment. *** p <.001

measure of the home numeracy environment (Chapter 4, Section 4.1.1). Looking at the activities listed in the questionnaire, not all activities will necessarily include numbers and therefore parents' answers to this question may not reflect number activities. Furthermore, many of the activities could be done in a range of environments were parents could engage their child in explicit teaching of numbers (the definition of formal) or the child could have incidental exposure to the activities (the definition of informal). For example encouraging collecting (an activity on the questionnaire) could be done without mention of number and mathematics (i.e let's collect pencils), however, this could also be done with an informal focus on numbers (i.e let's collect 2 more) or with a formal focus on numbers (i.e if we have collected 5 pencils, how many more do we need to collect so we have 10 pencils?). Therefore, it is not clear if this activity should be classed as formal, informal or should even be classed as a home numeracy activity. This illustrates one reason why the model may be a bad fit as parents may interpret the questions in different ways.

It would be possible to investigate the factor structure of the questionnaire further to try and find a good fitting model by changing the number of factors, their relations to the observed variables and patterns of measurement error correlations. However, it is important that any of these changes are based on theoretical reasoning (Kline, 2015). In the home numeracy environment theory the only typical divide is formal and informal activities therefore, I have no theoretical reason to change the model.

Even though the CFAs showed that the formal and informal factor structure was not a good model fit, it also showed that the one factor structure, used in Study 1 and Study 2, was also a bad fitting model. Furthermore many other researchers have used these categories when using a questionnaire method and the formal and informal structure could help explain the non-significant correlations in Study 2. Therefore, I decided to re-run the analysis of Study 1 and Study 2 with the questionnaire split into the formal and informal categories used by Skwarchuk et al.(2014).

6.3 Re-analysis for Study 1

Study 1 investigated the use of the novel text message measure of the home numeracy environment. To evaluate the reliability and validity of the text messages it was compared to a questionnaire measure. However, in Study 1 the questionnaire measure was used as combined measure of the home numeracy environment. In this re-analysis I aimed to investigate if the text message method is still a valid and reliable measure when compared to both the formal and informal categories of the questionnaire, or if the text messages only tap into one type of number activities (formal or informal). Further to this, I investigated how the relationship between the home numeracy environment, demographic variables, parents' mathematics expectations and attitudes and children's counting ability differs when the questionnaire measure is divided into formal and informal activities. Previous research suggests that the correlation between the home numeracy environment and mathematics achievement should be higher when correlating formal number activities and mathematics achievement compared to informal number activities and mathematics achievement.

6.3.1 Results

Table 6.3 shows all the correlations for Study 1 both with and without controlling for age. Throughout the results I will only discuss the correlations that control for age, because the formal number activities category in the questionnaire measure is significantly correlated to age, along with other variables.

The reliability of the text message method remains the same as Study 1 but the internal reliability of the questionnaire is lower for the formal and informal categories (Cronbach's alpha= $.69, .55$ respectively) which is not surprising because there are fewer items in these categories compared to the combined measure in Study 1.

6.3.1.1 Is there a relationship between the measures of the home numeracy environment?

Table 6.3 shows the relationship between the text messages with the formal and informal questionnaire measure. Both the formal and informal measures were significantly correlated to the text messages ($r = .32, .41, p = .009, .001$, respectively) indicating that the text messages are picking up on both formal and informal activities.

6.3.1.2 Does the frequency of activities in the home numeracy environment relate to demographic variables, parents' expectations and attitudes?

The second question in Study 1 evaluated how the two measures of the home numeracy environment related to different variables. Study 1 found only two significant correlations: one between the text messages and SES ($r = .24, p = .048$) and one between the questionnaire and parents' mathematics expectations ($r = .37, p = .002$). Table 6.3 shows that both the formal and informal categories of the questionnaire failed to correlate significantly with any of the demographic variables and parents' attitudes, as expected from Study 1. However, parents' mathematics expectations was significantly correlated with informal activities ($r = .35, p = .005$) but not significantly correlated with formal activities ($r = .21, p = .094$). This suggests that the higher expectations parents have for their child the more informal number activities they do with their child. Most researchers find the opposite to this, that the higher expectations a parent has for their child the more formal activities they will do. However, these correlations do not significantly differ ($p = .284$). This will be discussed further in the discussion.

In the re-analysis there were no significant differences between the correlations for the text messages and the correlations for the formal questionnaire category. There were also no significant differences between the correlations for the text messages and the correlations for the informal questionnaire category. Furthermore, there were no significant differences between the correlations for the formal questionnaire category and the correlations for the informal questionnaire category. This suggests that even though it is thought that the formal and informal categories measure different areas of the home numeracy environment, in this study the different categories did not have different relations with demographic variables, parents' mathematics expectations and attitudes and children's counting ability.

	1	2	3	4	5	6	7	8	9	10	11
1 Text Messages	—	.32**	.41**	.45**	−.11	.24*	−.10	.01	.24	.01	.26*
2 Formal Questionnaire	.16	—	.36**	.51**	−.05	−.03	−.15	−.06	.21	−.14	.42**
3 Informal Questionnaire	.34**	.37**	—	.63**	.03	.06	−.13	.06	.35**	−.14	.22
4 Questionnaire (Literacy Activities)	.34**	.56**	.60**	—	.14	.11	−.16	−.11	.35**	−.14	.36**
5 Childcare Hours	−.24*	.19	.08	.27*	—	−.04	.20	.15	.06	.15	.03
6 SES	.12	.04	.07	.17	.12	—	.40**	.28*	.09	.05	.20
7 Parents' Highest Qualification	−.11	.03	−.09	−.04	.23	.37**	—	.33**	−.07	.09	.27*
8 Parents' Age	.04	−.03	.05	−.11	.14	.24*	.33**	—	−.12	.09	−.01
9 Parents' Maths Expectations	.21	.17	.34**	.30*	.05	.08	−.09	−.12	—	−.01	−.01
10 Parents' Attitudes	.02	−.13	−.14	−.12	.14	.07	.05	.10	−.01	—	−.06
11 Counting	.08	.53**	.23	.46**	.33**	.31**	.32**	.01	−.03	−.02	—
12 Child's Age	−.22	.41**	.07	.31**	.57**	.27*	.23	.06	−.03	.05	.54**

Table 6.3: Correlations between the questionnaire, text messages, demographic variables, parents' expectations and attitudes. Below the diagonal are zero-order correlations with partial correlations controlling for age above the diagonal.

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

6.3.1.3 Does the frequency of number activities relate to children's counting performance?

The final question from Study 1 evaluated the relationship between the home numeracy environment measures and counting. Counting was a proxy measure used for mathematics achievement in this study. Study 1 found that both the text message and questionnaire measure significantly correlated with counting ($r = .26, .32, p = .035, .009$ respectively). In the re-analysis, it was predicted that the formal activities would be more strongly correlated to mathematics performance based on the literature. The correlation between formal activities and counting was significant ($r = .42, p = .001$) while the correlation between informal activities and counting was non significant ($r = .22, p = .086$). However these correlations do not significantly differ ($p = .114$).

6.3.2 Discussion

The re-analysis of Study 1 shows that the text messages are measuring both formal and informal activities. There was only one correlation that was significant for informal activities and not significant for formal activities. This was the correlation with parents' mathematics expectations. As mentioned in the results this correlation is normally stronger for formal activities than informal activities (Segers et al., 2015). However, it is important to note that the correlations were not significantly different from each other and therefore before drawing conclusions from this it is important to see if this correlation is replicated in the re-analysis of Study 2. Finally, the formal activities were more strongly correlated to counting ability as predicted. This mirrors other findings that formal activities are related to mathematics achievement more than informal activities (Skwarchuk, 2009; Skwarchuk et al., 2014; Zippert & Ramani, 2016).

This re-analysis showed that by dividing the questionnaire into formal and informal categories the results were broadly similar. However, Study 2 investigated different measures of the home numeracy environment and found that they were not related to the questionnaire and text messages. It is possible that the observation measure was only measuring the formal or informal home numeracy environment and therefore it is important I re-analyse Study 2 using the formal and informal categories.

6.4 Re-analysis of Study 2

Study 2 used a multiple measures approach to measure the home numeracy environment using three different measures: text messages, questionnaire and an observation. The main finding from Study 2 was that the two self report measures were not related to the observation measures and furthermore only child's number talk was related to mathematics performance. However, the lack of correlation between the measures could be because the observation measure was measuring number talk in activities that would typically be classed as informal number activities. If this is the case then in the re-analysis I expected parent number talk and child number talk to be significantly related to informal number activities. Furthermore, as with the re-analysis of Study 1, it was predicted that mathematics performance would be more highly correlated with formal activities than informal activities.

6.4.1 Results

I will present the results in the same structure as Study 2 by firstly investigating how the formal and informal categories are related to other measures of the home numeracy environment. I will then investigate how the formal and informal categories relate to demographic variables, parents' expectations and attitudes. Finally, I will review the relationship between the formal and informal categories of the questionnaire with mathematics performance.

The internal reliability of the formal and informal categories of the questionnaire (alpha = .72, .55 respectively) was lower than the Cronbach's alpha for the combined activities in Study 2 (alpha = .81) This shows that the internal reliability is reduced for these two categories but again this is not surprising because there are fewer items in these categories compared to the combined measure in Study 2.

6.4.1.1 Is there a relationship between the measures of the home numeracy environment?

Study 2 showed that the self report measures were not significantly correlated to the observation measures. Table 6.4 shows the correlations between the three measures with the questionnaire measure divided into formal and informal activities. This re-analysis replicates the re-analysis of Study 1 showing that the text message method was significantly related to both

	1	2	3	4	5
1 Text Messages	—				
2 Formal Activities (Questionnaire)	.39**	—			
3 Informal Activities (Questionnaire)	.25*	.34**	—		
4 Parent Number Talk	.01	.02	−.01	—	
5 Child Number Talk	.18	.30*	.02	.00	—

Table 6.4: Partial correlations between the questionnaire, text messages and observation measures controlling for age

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

formal and informal number activities. It was predicted that the informal activities may be related to parent and child number talk, because the activities in the observation could be classed as informal activities. This was not the case as the correlations were non-significant. However, there was a significant correlation between formal number activities and child number talk ($r = .30, p = .014$). This suggests that the more formal number activities parents do with their children at home, the more the child speaks about number, however we can not be sure on the causality of this relationship. This will be discussed in more detail in the discussion.

6.4.1.2 How is the home numeracy environment related to demographic variables, parents' expectations and attitudes?

In Study 2, childcare hours and parents' mathematics expectations were significantly correlated with the questionnaire. When separated into formal and informal activities, childcare hours was no longer significantly correlated with either formal or informal activities. However, SES was significantly correlated to formal activities ($r = -.28, p = .025$) suggesting that if children are from a higher SES area they do less formal number activities. However as in Study 2, there were no significant differences between the correlations for the formal and informal categories with the demographic variable measures. None of the other demographic variables were significantly correlated to the formal and informal categories of the home numeracy environment questionnaire.

Parents' mathematics expectations were significantly related to formal home numeracy activities ($r = .38, p = .003$), but not informal activities ($r = .24, p = .065$), which is the opposite finding to the re-analysis of Study 1 and will be discussed in more detail in the discussion. Study 2 found that there were significant difference between the correlation for the questionnaire and mathematics expectations and the text messages and mathematics expectations ($t(65) = 3.24, p = .002$) and parent number talk and mathematics expectations ($t(65) = 2.27, p = .027$). The re-analysis shows that the significant difference in correlations was only for the formal activities in the questionnaire; there was a significant difference between the correlation for formal activities and mathematics expectations and the correlation for text messages and mathematics expectations ($t(65) = 3.21, p = .002$) and the correlation for parent number talk and mathematics expectations ($t(65) = 2.27, p = .027$).

Furthermore, Study 2 found that there was a significant difference in correlations for parents' mathematics attitudes with the two self-report measures and again in the re-analysis this difference was only significant for the formal activities and the text messages ($t(65) = 2.04, p = .045$).

6.4.1.3 How is the home numeracy environment related to mathematics performance?

It was predicted that formal number activities would be more highly correlated to mathematics achievement than informal number activities. Table 6.5 shows that this was not the case. Both formal and informal activities were not significantly correlated to mathematics achievement.

	Text Message	Formal Questionnaire	Informal Questionnaire	Parent Number Talk	Child Number Talk
Home Literacy Environment	.27*	.41**	.52**	.02	.21
Childcare Hours	−.23	−.24	−.14	−.08	−.14
SES	−.05	−.28*	−.08	−.02	−.13
Parents' Highest Qualification	−.06	.05	.17	−.02	−.11
Parents' Age	.04	−.02	.20	−.08	.11
Mathematics Expectations	−.02	.38**	.24	.00	.26*
Mathematics Attitudes	−.08	.19	.18	.03	.10
WPSSI	−.12	.04	.18	−.14	.27*

Table 6.5: Partial correlations for the questionnaire, text messages and observation measures with demographic variables, mathematics expectations and attitudes and mathematics performance, all controlling for age

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

6.4.2 Discussion

The re-analysis of Study 2 aimed to explore if the observation measures of the home numeracy environment measured more formal or informal activities and to investigate how the formal and informal activities related to other variables, particularly mathematics achievement.

The formal and informal categories were both significantly related to the text messages showing again that the text messages were measuring both formal and informal activities. Interestingly, there was a significant correlation between formal number activities and child number talk. However, because this is only a correlation we can not be sure if more formal number activities causes the child to talk more about number, or if the child talking more about number causes the parent to do more formal number activities. However, this important finding shows that one area of the self report measures is related to one area of the observation measure. The only variable to relate to mathematics achievement is child number talk. This could suggest that the home numeracy environment is related to mathematics achievement, but only through how much a child talks about number. However, as we cannot be sure on the causality of the relationship between child number talk and formal number activities, we can not make this claim.

The formal activities category was negatively related to SES and positively related to parents' mathematics expectations, whereas the informal category was not significantly related to any variables. Furthermore, the re-analysis of Study 1 found a significant correlation between mathematics expectations and informal activities suggesting that this relationship is not consistent across studies. The reason for this could be the measure of parent's mathematics expectations because this used a questionnaire method and may not be accurately measuring parents' expectations for their child (see Chapter 5, Section 5.4).

The lack of correlation between the formal and informal number activities and mathematics achievement is not surprising, because in the original analysis the questionnaire was not significantly related to mathematics achievement. The lack of correlation can be explained in two ways. Firstly, it could be that the mathematics measure is not a good measure and a measure that previously found a relationship with the home numeracy environment should be used in future research in order to test this claim. Secondly, it could be the case that the home numeracy environment (measured by questionnaire, text messages and observation) is not related to mathematics performance in the UK in this age group of children as no previous research has investigated the home numeracy environment in the UK.

6.5 Conclusion

This chapter aimed to investigate how the formal and informal categories of the home numeracy environment would fit the UK data, and investigate if by dividing the questionnaire into formal and informal categories the non-significant relationships in Study 2 could be explained.

The first main finding from this chapter is that both a combined one factor model, and the formal and informal two factor model of the home numeracy environment were bad fitting models. The two factor model was used to re-analyses the data for Study 1 and Study 2 as this model has been used in previous research. The re-analysis of Study 1 showed that the text message method measured both formal and informal number activities and that formal activities were more strongly correlated to counting performance than informal activities. The re-analysis of Study 2 showed that formal number activities were significantly related to children's number talk in the observation measure, showing that one area of the self-report measures are related to one area of the observation measures. Finally, the mathematics measure was not related to formal or informal activities, suggesting that either the mathematics measure is not a good measure of mathematics or that the home numeracy environment is not related to general mathematics achievement. In conclusion, if we are to support the formal and informal distinction of the home numeracy environment, this re-analysis can expand my findings from Study 1 and Study 2, but the CFA suggests that we should not be confident with these two categories as demonstrsated bad model fit and therefore these results should be approached with caution.

Chapter 7

Discussion on home numeracy environment

The aim of this chapter is to bring together all of the findings from the home numeracy environment studies presented in Chapters 2 - 6. I will firstly present an overview of the findings. Next I will review the methodological and theoretical conclusions that can be drawn from these findings. Finally I will consider the direction of future research in the home numeracy environment, before concluding this part of the thesis.

7.1 Overall findings

The research presented in Chapters 2 - 6 in this thesis aimed to further our understanding of the home numeracy environment. The home numeracy environment is thought to be related to children's early mathematical abilities. Through a systematic review, analysis of a large national data set and two empirical studies, I aimed to examine methodological and theoretical questions within this body of research. The main findings are summarised below.

Firstly, I conducted a systematic review of the relationship between the home numeracy environment and mathematics performance. A p-curve analysis showed that there was evidential value for all studies. Further analysis of the studies included in the systematic review showed that there were different strengths and directions of the relationships between the home numeracy environment and mathematics achievement for different measures, but there wasn't a consistent pattern. This makes it difficult to evaluate the true relationship between the home numeracy environment and mathematics performance.

In view of this, the aim of my studies was to investigate the relationship between the home numeracy environment and mathematics performance. Firstly, I investigated if a question that could be seen as measuring the home numeracy environment within the Millennium Cohort Study was related to mathematics achievement. The relationship was very small and it was concluded that this was due to the measurement type.

Chapter 4 went on to present a novel measure of the home numeracy environment to combat the problems with the questionnaire measure. The novel measure was using text messages to ask parents if they had done number activities with their child that day. The text message method was compared to an existing questionnaire to evaluate its reliability and validity. The novel measure showed good reliability and validity and based on face value appeared to be a better measure of the home numeracy environment. Furthermore, both measures were related to children's counting ability.

One limitation to Study 1 was that both measures were self report measures and I was interested to see if the two typical measures used to measure the home numeracy environment (questionnaire and observation) were related to each other as well as the new text message method. Furthermore, Study 1 had only used a proxy measure for mathematics achievement, so it is important that the relationship was tested with a general mathematics measure. Study 2 addressed these concerns by measuring the home numeracy environment using a questionnaire, text messages and observation, as well as measuring mathematics achievement with a general mathematics measure. I found that the self report measures were not related to the observation measures and only child number talk during play was related to mathematics performance. Together these findings indicated that different measures of the home numeracy environment are not tapping into the same concept and furthermore the home numeracy environment did not appear to be related to mathematics performance.

Chapter 6 attempted to answer why there was no relationship between the measures. Many researchers believe that there are two types of number activities: formal and informal. It was hypothesised that the observation measure only measured informal home numeracy activities and therefore did not correlate with the combined questionnaire measure. I conducted a confirmatory factor analysis to determine if a two factor (formal and informal) model for the questionnaire data was a good fit. The model was a bad fit but a better fit than one factor (combined activities) model and therefore Study 1 and Study 2 were re-analysed with the questionnaire divided into formal and informal categories. The re-analysis of Study 1 showed that the text messages measured both formal and informal activities, and that formal activities were more highly correlated with children's counting than informal activities. The re-analysis of Study 2 showed the opposite to what was predicted: child number talk from the observation measure was related to formal number activities. Furthermore, child number talk was still the only variable related to mathematics achievement, indicating that the home numeracy environment may not be directly related to children's mathematics performance.

7.2 Methodological conclusions

One of the main aims of these studies was to investigate how the home numeracy environment is measured and this has given rise to some novel methodological conclusions. It is evident from Chapter 3 that using one simple question to measure the home numeracy environment is not a comprehensive measure. It has also been highlighted by researchers that using a questionnaire to measure the home numeracy environment is problematic because it relies on parents' long-term re-call of activities, judgement of how often the activities typically occur and many home numeracy activities may not be on the list of activities on the questionnaire. Furthermore, Chapter 6 of this thesis showed that both a one factor (combined) and two factor (formal and informal) model of the activities in the questionnaire were both bad-fitting models, giving more evidence that the questionnaire is a poor measure. The novel text message method used to measure the home numeracy environment presented in Chapter 4 combats many of these problems.

There are two key conclusions from the text message method. First, it showed good reliability and validity, second, it measured both formal and informal number activities. It offers the advantage over the questionnaire that parents are not required to recall how often they do number activities over a long time period, they are not limited to a list of activities and they do not need to make a judgement on how often activities take place. Furthermore, by only requiring a quick one word reply to the text message it makes it easy for parents to reply.

Even though many researchers have used questionnaire methods, a few have also used observation methods to measure the home numeracy environment. Frequently researchers combine findings from questionnaire studies with findings from observations studies in their evaluation of the home numeracy environment. Study 2 showed that observation and self-report measures (text messages and questionnaires) are not related. All measures were shown to be reliable but they were not all measuring the same general home numeracy environment. While the self -report measures assessed the frequency of number activities in the home, the observation measures assessed what typically happens during these activities and these two areas are not related. This suggests that a parent could report doing lots of number activities at home, but when a typical number activity is observed they do not frequently mention number. Therefore suggesting there is a difference between the quantity of number activities parents do and the quality of the activities and it could be that the quality of the number activities is more important than the quantity. The text message method provides a more accurate measure of the quantity of activities as it is more inclusive of different types of activities and also parents will only report number activities in which they have the intention to mention number, whereas with the questionnaire the parent may do some of the activities without having the intention to use/ teach number skills. This suggests that the text message method is a more accurate measure of the activities that parents do at home that have an intention to include number therefore tapping into the quality of the activities as well as the quantity. Furthermore, as discussed in Chapter 5 Section 5.4, observations measuring the home numeracy environment could also be tapping into parents' and childrens' Spontaneous Focus On Numerosity (SFON). SFON is a measure of how much people tend to focus on number in the environment around them. If a parent/child are very aware of numbers around them (high SFON) then they will tend to talk more about numbers in number activities. These findings have important implications for future research which will be discussed below.

7.3 Theoretical conclusions

The findings add to many of the theoretical issues within the home numeracy environment literature. Firstly the main question surrounding the home numeracy environment is how does it relate to mathematics performance. There is mixed evidence with regards to this relationship and Study 1 and 2 add to these. Counting was related to the home numeracy environment showing that more number activities relates to children's counting skills. However, most home numeracy environment measures used in Study 2 were not related to a general measure of mathematics performance. Some researchers have found a relationship between the home numeracy environment and a general mathematics measures (DeFlorio & Beliakoff, 2014; Kleemans et al., 2012; Segers et al., 2015; LeFevre et al., 2010) whereas others have not (DeFlorio, 2013; Ciping et al., 2015). This indicates that the home numeracy environment might have unique relationships with different areas of mathematics. We already know that basic number processing skills, such as non-symbolic number processing, symbolic number processing and mapping, have different relationships with children's mathematics achievement (De Smedt et al., 2013, Mundy & Gilmore, 2009).Mutaf et al. (2016) showed that the association between formal home numeracy environment activities and children's calculation skills was mediated by basic processing skills. This suggests that how the home numeracy environment relates to basic processing skills should also be investigated.

General mathematics achievement was significantly linked to children's number talk. As children's number talk could be seen as also measuring SFON, this could drive the relationship with mathematics performance. Furthermore children's number talk was linked to formal number activities. This suggests that the relationship between the amount of formal number activities in the home and children's mathematics performance may be mediated by child's number talk. However as all this research is correlational, no conclusions can be made about the causality of these relationships. There are still questions over the direction and causality of the relationship between the home numeracy environment and mathematics achievement that can not be answered by my research. But it is important to highlight that we should not assume that a positive relationship means doing more number activities causes the child to develop better number skills. Nonetheless, it is possible that because the child enjoys and is good at mathematics, then this causes the parents to do more number activities at home. In order to answer these questions about direction and causality, different research methods are needed and will be discussed in more detail in the next section.

The findings from these studies add support to the theory that the amount of number activities (formal or informal) in the home numeracy environment are not directly related to general mathematics performance. This could be just within the UK because no previous research has investigated this relationship within the UK. The relationship does vary between different countries, and it has also been shown to vary in different areas within individual countries and therefore we can not be sure if cultural differences cause the null relationship found here.

Another theoretical issue is how the home numeracy environment is affected by other variables, such as SES, childcare hours, parents' expectations and attitudes. None of these relationships were consistent between measures and studies and therefore this research cannot add to this theoretical issue and can only conclude that these relationships are not consistent.

7.4 Implications for future research

These findings have important implications for future research into the home numeracy environment. These are discussed below.

7.4.1 Measuring the home numeracy environment

These studies highlight that the questionnaire measure of the home numeracy environment has many problems and the activities in the questionnaire are not well modelled by the one or two factor structure typically used. Therefore future research should conduct a CFA to check the factors used are a good fit for the data if using a questionnaire method, but should, if possible, avoid using a questionnaire. The novel text message method presented in this thesis can be a good alternative to the questionnaire.

The text message method can be used in future research to measure the amount of number activities in home numeracy environment and can be adapted to suit many different research questions. In the current studies the question asked was kept simple to test the method, however it could be used in future to ask about specific number activities (formal or informal), how many number activities parents are doing or could even be sent to multiple family members to get an overview of the whole home numeracy environment.

However, even though the text messages can give a valid and reliable measure of the amount of number activities in the home, it does not show what is happening during those activities and it is important that if researchers want a full representation of the home numeracy environment, they should also conduct an observation measure of the activities. As shown in Study 2, the activities parents report doing at home are not related to how much they talk about number during typical home activities. An observation measure could be conducted using technology by asking parents to send pictures or videos of typical home numeracy activities or to video their child's routine for an hour each day. This way the observation would be more natural as it occurs in the home but would be harder to analyse as it would be less structured.

7.4.2 Investigating the relationship between the home numeracy environment and mathematics performance

Most researchers are interested in how the home numeracy environment effects children's mathematics performance and, as this thesis has shown, this relationship is not consistent. Much of the research has used correlational studies to investigate this relationship and it has been suggested that the difference in results could be due to the measures of the home numeracy environment, the measures of mathematics performance, parents' mathematics expectations for their child, the type of number activity (formal and informal), child's SES, child's age, child's country of origin and many other variables. As there are so many variables that could effect this relationship and it would be very difficult to determine which combination of these variables results in a positive relationship with mathematics, therefore I believe there are two ways forward in this area of research.

The first way would be to run intervention studies that impact what parents do with their children at home. Intervention studies can provide clear evidence whether certain activities impact mathematics performance. There are already some intervention studies using board games (Ramani & Siegler, 2008; Siegler & Ramani, 2008), however this could be expanded to interventions that encourage parents to read number books to their child, interventions that encourage parents to pick out numbers in the environment around them or even interventions that give parents specific number activities to do with their child. This method would indicate more clearly if the interventions result in increased number skills, leading to increased confidence in which activities parents do at home impact mathematics performance, without worrying about all the other potential variables that could be influencing mathematics performance. However, in order to run interventions decisions on how long interventions should run for in order to see an effect, how often parents should do the activities and which type of activities are expected to have an effect on mathematics performance need to be made. As intervention studies are time consuming to run, these decisions should be made with confidence before running any intervention studies.

A second way to progress in the home numeracy environment research would be to focus on a small area of the home numeracy environment and investigate how the features of that area helps children learn mathematics. With the questionnaire and observation measures of the home numeracy environment it is assumed that all activities in the home numeracy environment are effective at improving mathematics performance. More in-depth analysis of how these typically occurring activities and their effectiveness in teaching mathematics is important. Researchers could analyse specifically how parents ask about quantities, which counting songs parents sing or how number books are used in the home environment and see how different features of each of these activities relate to children's mathematics skills. Furthermore, if it is found that a certain type of activity is effective then an intervention study encouraging parents' to do that specific activity could be conducted. In the remaining part of this thesis I build on this idea by conducting an in-depth analysis of number books, specifically investigating how children learn number symbols from number books.

7.5 Summary

To summarise, the work presented in Part II of this thesis focused on the home numeracy environment, specifically how to measure the home numeracy environment and how it relates to mathematics performance. The findings from two empirical studies were reported alongside a systematic review of the literature and analysis of cohort data. These findings add to the current literature on the home numeracy environment both methodologically and theoretically. The findings have advanced the tools used for measuring the home numeracy environment in a valid and reliable manner and highlighted the importance of using both self-report and observation measures. They also show that there is no relationship between the home numeracy environment and general mathematics performance, and the reason for the lack of relationship could be due to a variety of factors. Therefore, in order to overcome this issue, new approaches should be used in the home numeracy environment research field.

Part III

In-depth Analysis of Learning Number Symbols from Number Books

Chapter 8

Literature review

The first part of this thesis showed that the relationship between the home numeracy environment and mathematics achievement is inconsistent both between and within studies. This could be due to numerous factors and it was decided that in order to develop the field further there was need for in-depth analysis of certain areas of the home numeracy environment. Therefore in this part of the thesis I report an in-depth investigation of one area of the home numeracy environment: number books.

One of the most common teaching activities in the home is reading (Shanahan & Lonigan, 2010; Niklas & Schneider, 2013b) and there are many schemes that have been developed to encourage parents to read to their child (BookTrust, 2017; Trust, 2017; Reading Agency, 2013; Wade & Moore, 2000). For example, BookTrust are a UK charity that gives story books to children at age 1 and age 3 in order to encourage parents to read books with their child.

With regards to the home numeracy environment it has been shown that parents often read number books to their child (LeFevre et al., 2009). Number books are books that have the primary purpose of teaching children counting. Typically they present number symbols, in order, alongside a pictorial representation of the number. Through reading number books parents aim to help their child attach meaning to number words and symbols. However how young children attach meaning to number words and symbols is a question that has drawn attention from researchers in developmental psychology and mathematics education alike, and despite considerable research interest we still do not know what gives number words and symbols their meaning (known as the 'symbol grounding problem', Leibovich & Ansari, 2016) or how young children come to acquire number words. One consequence of this gap in our knowledge is that we don't know how to develop the most effective resources to support preschool children's early number learning. Therefore it is important that number books that are used to teach children number symbols are investigated in detail and are designed in the most effective way for children to attach meaning to number words and symbols.

In this chapter, I will start by discussing current research into children's number books and highlighting the key features of number books. Secondly, I will then discuss the extensive literature which has explored the use of abstract or real world examples in mathematics learning more generally. Moving on from this, outline the theory behind how young children learn numbers. Finally I will draw upon this literature to outline how I will test the benefits and costs of using concrete or abstract representations to support children's learning of novel number symbols.

8.1 Number books

The notion that books support mathematical learning is not new (Doig, 1989; Gailey, 1993; Hong, 1996; Jenner, 2002; Lewis, Long, & Mackay, 1993; Roth McDuffie & Young, 2003; Mudkiff & Cramer, 1993; Thatcher, 2001). It has been shown that more exposure to mathematics related books leads to children's increased interest in mathematics activities (Jennings, Jennings, Richey, & Dixon-Krauss, 1992), a deeper mathematical knowledge (Jennings et al., 1992; Young-Loveridge, 2004) and more frequent use of mathematics vocabulary (Hassinger-Das, Jordan, & Dyson, 2015). Furthermore, Young-Loveridge (2004) showed that the increase in mathematical knowledge was still evident 15 months after an intervention that encouraged parents to read number books.

More in-depth research has been conducted into the interactions that take place when parents read number books with their child. Anderson, Anderson, and Shapiro (2004) asked 4 parent and child dyads to read the book "One Snowy Night" by Butterworth (1989). This book had opportunities for mathematics discussion but it was not explicit. They found that each parent and child dyad shared the book in a unique way with varying amounts of mathematics talk. Hojnoski, Columba, and Polignano (2014) also found that parents engage in informal mathematics talk when reading books, but the ways in which number was discussed was not consistent between dyads. Vandermaas-Peeler, Nelson, Bumpass, and Sassine (2009) found that the interactions during number book reading vary according to the socio-economic status (SES) of the family, with less initiating and scaffolding mathematics talk with lower SES families than higher SES families. These studies show that how number books are read varies greatly but, as discussed above, something about children's interactions with the number books apparently improves mathematics skills in children. Furthermore, if number books are influencing children's mathematical thinking it is important to consider the influence that different features of the books can have on their mathematical thinking.

Very few studies have focused on the content of pre-school number books. A selection of papers have aimed to designed a coding tool for these books (Halsey, 2005; Hellwig, Monroe, & Jacobs, 2000; Hunsader, 2004; Schiro, 1997) but it has been shown to be unreliable (Nesmith & Cooper, 2010). Only two research papers to my knowledge have investigated the content and features of number books objectively. Powell and Nurnberger-Haag (2015) conducted a review of 160 trade books that related to number. Their main findings were that the majority of books (68%) only used the numbers 1-10 and that the books did not include multiple representations of the number (e.g. word, symbol and pictorial representation). Furthermore, the nature of the items to be counted typically varied throughout the book (e.g. one horse, two sheep, three pigs etc.); this was true for 72% of the books reviewed. Ward, Mazzocco, Bock, and Prokes (2017) expanded Powell and Nurnberger-Haag's (2015) study by investigating a wider variety of structure and content features of number books. They evaluated the structure and content of 120 children's number books. Their main conclusion was that "the features coded in this study reveal the vast degree of variability in the nature of counting books, but not their worthiness as tools to promote numeracy. The latter issue must be systematically addressed by research targeting select features of books" (p. 60). This quote highlights the need for further research into the specific features of number books. One interesting finding from their study is that 96% of books included at least one real-world set of items to be counted (e.g. apples), and 87% included only real-world pictures. Moreover, the nature of the items to be counted typically varied throughout the book (e.g. one horse, two sheep, three pigs etc.); in only 33% of the books reviewed did the identity of the items remain consistent (one horse, two horses, three horses, etc.). This is a select feature of the counting books that appears to be a common occurrence but as detailed above we don't know the worthiness of this feature. There is reason to question whether such real-world representations are best to support children's ability to learn numbers.

There is some evidence that using real-world pictures may introduce difficulties when children learn number words. Huang, Spelke, and Snedeker (2010) suggested that when children first learn the meaning of number words, they struggle to generalise from the real-world context in which it was taught. Huang et al. (2010) taught 16 children about the number three. These children knew the meaning of the number one and two but had not yet mastered the meaning of the number three. They were trained using pictures of dogs and told, "This card has (does not have) three dogs". Children were then asked to select the card with three objects. They were successful when the test cards were pictures of dogs (in a different configuration or breed of dog to the training cards) but they could not successfully identify three when the pictures were of sheep. This study suggests that children may attach the meaning of number words to the context in which they were taught, at least for smaller numbers, and consequently there may be a cost to using pictures of real-world items to teach children the meaning of number words.

8.2 Learning from abstract and concrete representations

Decades of research in mathematics education has debated whether real-world or abstract representations are more effective when teaching abstract mathematical concepts. This debate is often characterised as a choice between 'abstract' or 'concrete' representations. However, researchers do not agree upon a definition of abstract and concrete (Sarama & Clements, 2009; Wilensky, 1991) and it is often argued that there is in fact a continuum between abstract at one extreme and concrete representations at the other. Here I adopt Fyfe, McNeil, Son, and Goldstone's (2014) approach: concrete materials are defined as those that "connect with learners' prior knowledge, are grounded in perceptual and/or motor experiences, and have identifiable correspondences between their form and referents" (p. 1) while abstract materials are defined to "eliminate extraneous perceptual properties, represent structure efficiently, and are more arbitrarily linked to their referents" (p. 1-2). For example, a concrete representation of the fraction $\frac{1}{2}$ might be (a picture of) half a pizza whereas an abstract representation might be a square with half of the area shaded.

For many years it was generally accepted that concrete representations were beneficial when introducing mathematical concepts to learners, particularly young children. Piaget (1971) and Montessori (1917) suggested that young children's cognitive abilities are not mature enough to fully engage in abstract thinking and, therefore, that concrete materials are necessary to aid their learning. Bruner (1966) went further and argued that all learners, not only young children, benefit from being presented with new information in a concrete form before being introduced to the abstract form. These theories resulted in a general acceptance that children should learn about mathematical concepts through concrete representations.

Two key advantages to the use of concrete materials have been proposed. First, concrete materials allow learners to activate real world knowledge to help them solve problems or understand mathematical ideas (Kotovsky, Hayes, & Simon, 1985; Schliemann & Carraher, 2002). Second, concrete examples improve memory and understanding by giving the learner an imagined action related to that mathematical concept (Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004). For example, the idea of division might be introduced with the example of sharing cookies between friends. Children can draw on their real world experiences to understand the idea of sharing equally and, when asked to divide in future problems, they are more likely to remember the procedure via the imagined action of sharing cookies. In addition to these cognitive advantages it has been suggested that concrete examples have a motivational benefit. For example, LeFevre and Dixon (1986) demonstrated that students prefer working with concrete examples. Students were presented with conflicting abstract instructions and a concrete example and they predominately chose to follow the concrete example.

Despite the early support for concrete representations, in recent years there has been a growing body of research suggesting that concrete materials may not be universally beneficial when learning mathematics. Several costs associated with the use of concrete representations have been highlighted (Brown, McNeil, & Glenberg, 2009). First, concrete representations contain extraneous details that can distract the learner from the relevant information (DeLoache, 2000; Uttal, O'Doherty, Newland, Hand, & DeLoache, 2009), a phenomenon that has been referred to as the 'seductive details effect' (Garner, Gillingham, & White, 1989; Harp & Mayer, 1998). Second, it has been shown that learning through concrete examples can constrain transfer of knowledge to other problems (Day, Motz, & Goldstone, 2015; Sloutsky, Kaminski, & Heckler, 2005). For example, introducing fractions with concrete examples of cutting a cake can constrain children's transfer of knowledge to other problems, particularly if asked to multiply or divide fractions, as it no longer makes sense to think about fractions as portions of cake.

In support of this view, a high-profile paper by Kaminski, Sloutsky, and Heckler (2008) suggested that learning from abstract examples results in a more sophisticated level of understanding of mathematical concepts than learning from concrete examples. Kaminski et al. (2008) compared groups of undergraduates who learned about an algebraic concept - the group of order 3 - using abstract or concrete examples. Following a learning phase, participants' understanding was assessed using a multiple-choice test. Participants who learned the mathematical concept with an abstract example performed better on the post-test than those who learned with concrete examples. Kaminski et al. (2008) concluded that abstract examples are more effective than concrete examples when learning mathematical concepts. However, critics have argued that the questions in the post-test were more similar to the learning phase stimuli seen by the abstract learning group than the those seen by the concrete learning group and that the two groups could have learned somewhat different concepts (group of order 3 vs. addition modulo 3) from the examples they were given (Jones, 2009a; Jones, 2009b; Mourrat, 2008; De Bock, Deprez, Van Dooren, Roelens, & Verschaffel, 2011).

Although these criticisms limit the conclusions which can be drawn from Kaminski et al.'s (2008) study, other researchers have also suggested that there can be advantages to using abstract representations (Koedinger, Alibali, & Nathan, 2008; McNeil & Uttal, 2009). Abstract representations eliminate extraneous information, are more generalizable for transfer to other problems and focus attention on the to be learned information (Uttal et al., 2009; Son, Smith, & Goldstone, 2008). On the other hand, there may also be costs to using abstract representations. For example, they may lead to inefficient solution strategies (Koedinger & Nathan, 2004), an inflexible application of knowledge (McNeil & Alibali, 2005) and can lead to logical errors (Carraher & Schliemann, 1985).

Given the potential advantages and disadvantages of both concrete and abstract representations, a third approach, known as concreteness fading, has been proposed (Bruner, 1966; Fyfe et al., 2014; Goldstone & Son, 2005). The aim of the concreteness fading technique is to combine the advantages of both abstract and concrete representations. In this approach, learning begins with concrete examples, which allow the learner to access real-world concepts to help understand the key idea, and then gradually fades to more abstract examples, which have the advantage that the information learnt can be transferred to other problems. Bruner described an expanded version of this with three forms: an enactive form, which is a physical concrete model of the concept; an iconic form, which is a graphic pictorial model of the concept and a symbolic form, which is an abstract model of the concept. For example, the quantity 'two' could first be represented by two physical items e.g. apples, then a picture of two dots representing the apples and then the Arabic numeral for two. Several researchers have adopted the concreteness fading approach, finding positive effects (see Fyfe et al., 2014 for review). For example, Goldstone and Son (2005) investigated students' learning and transfer of a scientific principle by presenting them with a concrete representation (using pictures of ants and leaves), an abstract representation (using dots and patches) and a concreteness fading representation (fading from the ants to the dots). They found that students' transfer was better when they used a concreteness fading technique compared to concrete or abstract representations alone. Similarly McNeil and Fyfe (2012) found that transfer of mathematical knowledge was also better when presented using a concreteness fading technique compared to abstract or concrete representations alone.

A particular feature of the concreteness fading approach is that it involves two (or more) different representations of the same concept. This is typically also true of concrete approaches, where different concrete representations tend to be used together. It has been widely claimed that the use of multiple representations may lead to better learning outcomes in comparison to the use of a single representation (e.g., Ainsworth, 1999; Brenner et al., 1997; Jong et al., 1998; van der Meij & de Jong, 2006). Three advantages to the use of multiple representations have been proposed: two or more representations can provide complementary information; information in one representation can constrain the interpretation of information in another representation; and multiple representations can lead to the construction of deeper understanding by allowing learners to abstract information across different examples or representations (Ainsworth, 2006).

However, using multiple representations is not universally beneficial, and may come with costs (Ainsworth, 2006). Multiple representations can result in split attention and typically involve extraneous cognitive activities which may interfere with learning (Chandler & Sweller, 1992). Furthermore learners may fail to see how the representations are linked to each other and fail to extract the key concept. For example, Ainsworth, Bibby, and Wood (2002) explored children's computational estimation using a computerised intervention that involved either pictorial representations, mathematical representations or both. Across two experiments they found that children learning with either pictorial or mathematical representations improved their estimation accuracy, but children working with multiple representations did not. Furthermore Flack and Horst (2017) also found that multiple representations (i.e. illustrations on both pages) in children's storybooks hindered children's word learning when compared to storybooks with single representations (i.e. illustrations only on one page). The disadvantages of multiple representations are often interpreted within the framework of cognitive load theory: it is suggested that learners fail to benefit from multiple representations because they do not have sufficient cognitive resources (e.g. working memory capacity) to process the available information (Ainsworth, 2006; Ainsworth et al., 2002). This perspective raises a question over whether the multiple representations involved in the concrete or concreteness-fading approaches are beneficial or not.

8.3 Early number learning

These debates are pertinent to the question of what types of representations may best support children's early number learning. Number books are often used to teach child number symbols but it is important to understand how children learn number symbols before investigating the representations used by number books.

Current models of early number learning propose that, as children learn the meaning of number words and symbols, they connect these with magnitude information (see Fazio et al., 2014 for review). Although debates surround the precise nature of the quantity information which underpins the meaning of number words and symbols (Leibovich, Katzin, Harel, & Henik, 2016; Le Corre & Carey, 2007), the predominant model has focused on the role of the Approximate Number System (ANS), a cognitive system that allows individuals to represent the approximate quantity of a set of items (Halberda, Mazzocco, & Feigenson, 2008; Feigenson, Dehaene, & Spelke, 2004). ANS representations are thought to be generated whenever an individual perceives a set of items. When children learn number words and symbols, it has been proposed that they become mapped onto these internal ANS representations (to avoid confusion, I use the term 'ANS representation' to refer to an internal cognitive representation formed by the Approximate Number System, and 'representation' to refer to an external representation of a mathematical idea, such as an array of dots or a number symbol). Evidence to support this comes from studies showing that the accuracy of both nonsymbolic and symbolic number comparisons are subject to ratio effects. That is to say that when asked to select the more numerous of two stimuli (either arrays of dots, or Arabic numerals) participants' accuracy decreases as the ratio between the stimuli approaches 1 (Dehaene, 2011). Furthermore, children's performance on typical ANS acuity measures (i.e. dot comparison tasks) has been found to correlate with their school-level mathematics achievement (e.g., Halberda et al., 2008 but see Gilmore et al., 2013).

Huang et al.'s (2010) study of children's number learning focused on the learning of verbally presented number words, while much of the research exploring how numbers acquire their meaning has considered Arabic symbols. However, it has been suggested that number words and symbols are attached to meaning in the same way. Dehaene's (1992) Triple Code model incorporates separate codes for verbal and symbolic representations of number, but proposes that these are both mapped onto abstract magnitude representations (although leaves open the possibility that one of these mappings is dominant). Indeed, many theorists fail to draw a distinction, referring only to number symbols (words or digits) (e.g. Reynvoet & Sasanguie, 2016; Leibovich & Ansari, 2016; Piazza, 2010). The small number of studies to have explicitly distinguished words and digits have tended to find that children associate magnitude information with number words prior to Arabic digits (Bialystok, 1992; Knudsen, Fischer, Henning, & Aschersleben, 2015; von Aster & Shalev, 2007). When parents read numbers, children will typically receive both verbally-presented number words and visually-presented number symbols together, along with examples of the appropriate quantity.

8.4 Aims for this part of the thesis

In the remaining chapters of this part of the thesis, I explore whether the nature of the representations provided to a learner has an impact on their ability to learn symbolic number representations. In other words, might children learn number words and symbols more effectively if they are matched with abstract representations, concrete representations, or a combination of both? Number books typically provide concrete representations (Powell & Nurnberger-Haag, 2015; Ward et al., 2017) but previous evidence about learning from representations provides mixed predictions about whether abstract or concrete examples will be more beneficial. Concrete examples may support learning because they allow children to draw upon their real-world experiences of dealing with object sets and motivate children to learn. On the other hand, abstract examples may support learning because they do not include extraneous details and make it easier to identify and abstract the key information (i.e. numerosity). A concreteness fading approach incorporating both concrete and abstract examples may provide the benefits of both approaches, or alternatively may be less effective because the use of multiple representations leads to the overload of working memory.

Children begin to learn number words in their second year and are exposed to number words in combination with a wide range of different examples. Consequently, it is challenging to test this question in young children because they will have different prior experiences with symbolic representations. To study these processes in a more controlled fashion, researchers have, therefore, begun to use an artificial symbol learning paradigm, in which participants learn novel symbolic representations after being exposed to these symbols in combination with nonsymbolic quantity representations, for example dot arrays (Lyons & Beilock, 2009; Lyons & Ansari, 2009; Merkley & Scerif, 2015; Merkley & Ansari, 2016; Zhao et al., 2012). Using this paradigm, it has been shown that adults can learn the meaning of novel symbols and subsequent performance on symbolic comparison tasks show characteristic ratio effects, suggesting that adults are learning these symbols in a similar way to how children learn Arabic number symbols (Merkley & Scerif, 2015). However, to my knowledge no research has investigated whether children can learn novel symbols using non-symbolic representations, nor whether learning is affected by the nature of the examples with which novel symbols are paired in the training phase. In this part of the thesis, I present three studies that investigated the effects of abstract and concrete representations on children's accuracy in learning novel number symbols.

Chapter 9

Learning novel symbols through abstract and concrete representations (Study 3)

The aim of Study 3 was to investigate if children could successfully learn novel symbols using an artificial learning paradigm and to test whether children were more successful when presented with abstract or concrete non-symbolic representations.

9.1 Method

9.1.1 Participants

Seventy-four children ranging from 6 to 10 years old (M = 7.83 years, SD = 1.238, 35 boys) participated in the study. This study was powered to have 90% chance of detecting a small effect size ($\eta_p^2 = .03$, based on a correlation of 0.6 between performance in each condition). The children participated at 'Spring Scientist Week' at the University of Nottingham, an annual event run during the school holidays. Parents and children are invited to the university for half a day to take part in research activities and games. All studies were approved by the University of Nottingham School of Psychology Ethics Committee and all parents provided written consent for their child to participate. Children received a goody bag to thank them for taking part.

9.1.2 Procedure

I used a within-subjects design where participants completed two training phases, in each of which they learnt the meaning of five novel numerical symbols. In one training condition, the novel symbols were paired with abstract representations while in the other training condition the novel symbols were paired with concrete representations (see Figure 9.1). The symbols used in each training condition were different. Each training phase was immediately followed by a symbolic comparison task. The order of the training conditions was counterbalanced across participants. Between the two training phases, participants took a pencil and paper arithmetic test. The arithmetic task was included in order to evaluate if there was a relationship between the magnitude comparison task performance and mathematics skills, similar to the relationship found when using arabic digits in a magnitude comparison task (Halberda et al., 2008). The training and comparison tasks were presented on a laptop computer using PsychoPy software (Peirce & Peirce, 2009) and the entire experiment took approximately 20 minutes.

9.1.3 Training

In each of 100 training trials per condition, participants saw a symbol at the top of the screen with an array of dots/pictures (depending on whether they were in the abstract/concrete condition respectively) underneath, as shown in Figure 9.1. Children were asked to remember how many dots/pictures were associated with each symbol. Each trial was displayed for 1000ms with a blank screen for 200ms between each trial. The trials were presented in random order and participants received breaks every 25 trials.

Ten symbols were selected from the LaTeX amsmath package, so that they would be unfamiliar to children of this age. Five symbols were used for the concrete condition and a different set of five symbols used for the abstract condition (see Figure 9.2). Each symbol was associated with either 5, 10, 15, 20, or 25. Children saw each symbol 20 times, each time paired with a numerically equivalent but spatially different array.

In the abstract condition the nonsymbolic stimuli were arrays of dots. The dots, which were the same size in all trials, were randomly placed within

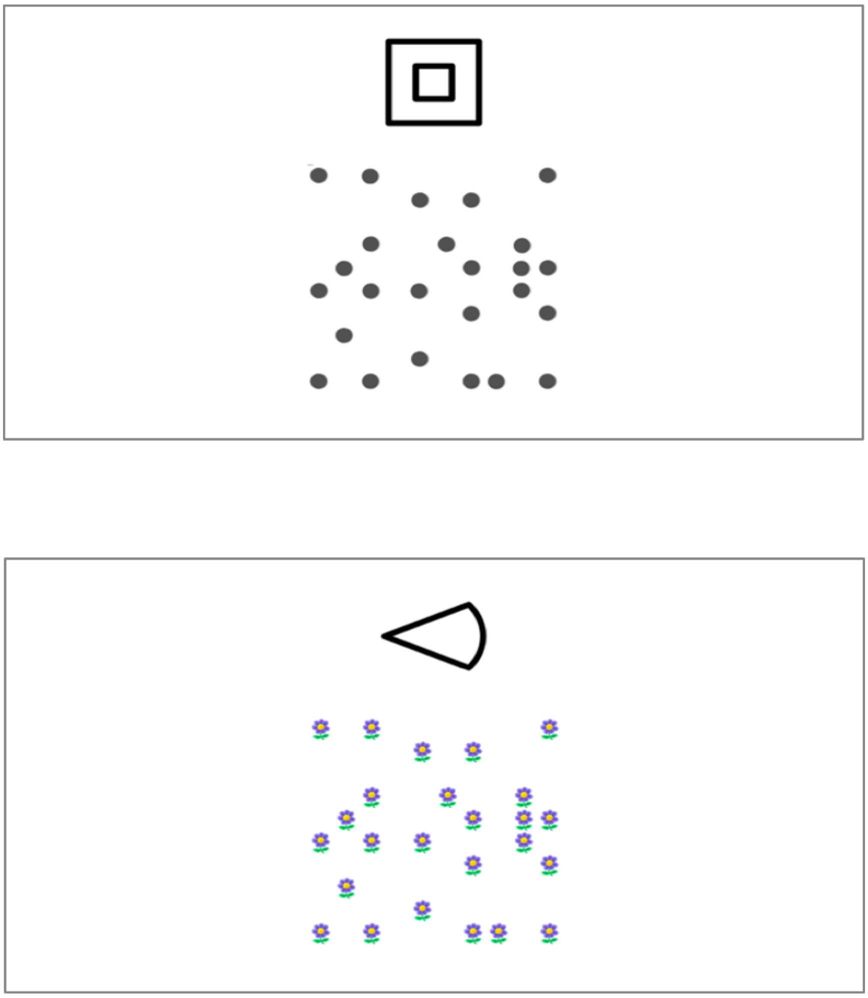

Figure 9.1: Examples of a) abstract and b) concrete stimuli used in training phase.

a 10 × 10 grid to create 20 different displays per numerosity. Each display was only presented once in the experiment. In the concrete condition, the stimuli were arrays of identical pictures. The pictures were the same size as the dots and were placed in exactly the same places as the dots in the abstract condition. A total of 20 different pictures that would be familiar to children of this age were selected (see Figure 9.3). Each picture was only used once for each number. For example, the novel symbol for 5 was displayed once each with a set of 5 frogs, 5 pizzas, 5 cars etc. Figure 9.4 shows examples of the screens the participant saw in each condition.

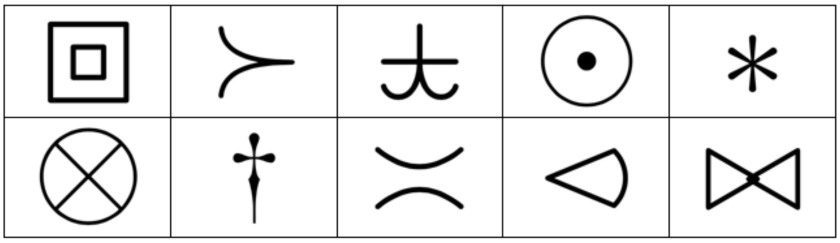

Figure 9.2: Symbols used in abstract (top row) and concrete (bottom row) conditions in training phase.

Figure 9.3: Pictures used in the concrete training phase.

9.1.4 Symbolic magnitude comparison task

Immediately following each training phase participants completed a symbolic comparison task to assess whether they had learnt the numerical meaning of the novel symbols. On each of 40 trials they were presented with two symbols, and asked to select the one that represented the larger number (see Figure 9.4). The symbols were presented until the participant responded by pressing a key on the keyboard. Every combination of the 5 symbols in each condition was presented four times with display side counterbalanced. I calculated a mean accuracy score for the abstract and concrete conditions as well as an overall mean accuracy.

9.1.5 Arithmetic test

Between the training conditions participants completed the Woodcock Johnson arithmetic fluency test, which requires participants to answer as many one- and two-digit sums as possible in 3 minutes. Raw scores (total correct)

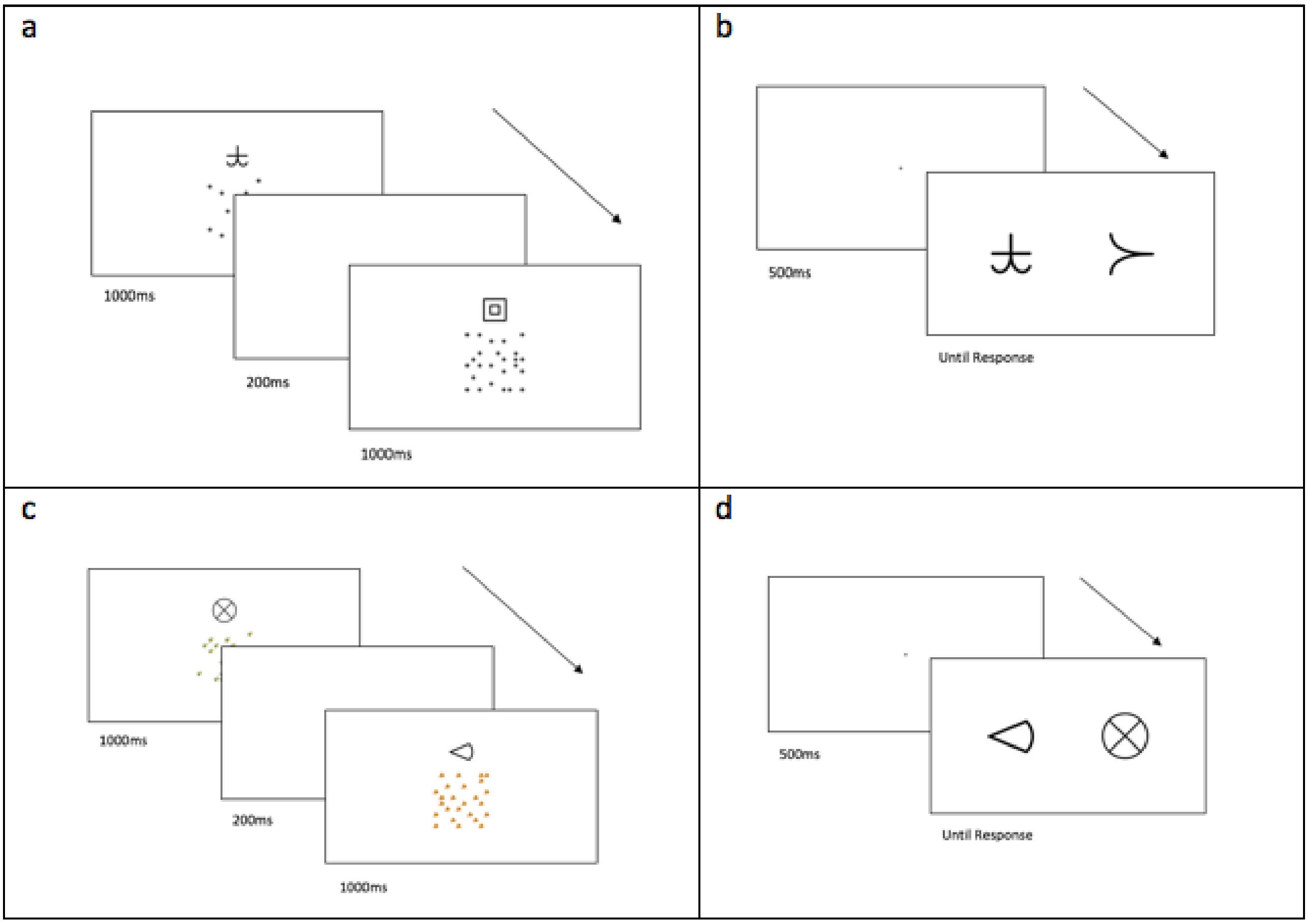

Figure 9.4: Examples of the screens seen by the participant in the a) abstract training b) abstract comparison task, c) concrete training and d) concrete comparison task.

were used in the analysis.

9.2 Results and discussion

I first assessed whether children were able to learn the meaning of the novel symbols and whether performance on the symbolic magnitude task showed characteristics typically observed on Arabic symbolic comparison tasks. I then compared performance for the two different training conditions.

Children were significantly more accurate than chance (0.5) in both the abstract, M = .72, SD = .166, $t(73) = 11.41$, $p < .001$, and concrete conditions, M = .66, SD = .195, $t(73) = 7.25$, $p < .001$, indicating that they were able to learn the meaning of the novel numerical stimuli.

To test whether children had learned the meanings of the full range of symbols or had simply learnt the symbols for the smallest and largest quantities, I examined children's performance on the symbolic magnitude comparison test after removing trials including the smallest and largest symbols. The accuracy for both the abstract (M = .57, SD = .221) and concrete (M = .61, SD = .227) conditions were still significantly above chance, $t(73)$ = 2.63, $p = .010$, $t(73) = 4.02$, $p < .001$, respectively, indicating that children had not just learned the meaning of the two extreme values.

The ratio between the two comparison symbols in this task ranged from .2 to .8. If children had mapped the novel symbols onto nonsymbolic representations of quantity then I would expect a significant effect of ratio on performance. I evaluated this by conducting a by-items linear regression, predicting the proportion of participants correctly responding to each trial by the trial's ratio (calculated as smaller: larger). This revealed a significant effect of ratio, $\beta = -.744$, $p < .001$, $R^2 = .554$ (see Figure 9.5). Overall accuracy was correlated with performance on the Woodcock Johnson arithmetic fluency test, $r = .391$, $p = .001$, which remained significant after controlling for age, $pr = .241$, $p = .040$. Therefore, performance on the symbolic comparison task showed performance characteristics that are typically observed on Arabic symbolic comparison tasks.

To explore the differing effects of abstract and concrete representations on symbol learning, I conducted a 2 × 2 ANOVA with condition (abstract or concrete) as a within-subjects factor and order (abstract condition first or concrete condition first) as a between-subjects factor. This revealed a significant main effect of condition, $F(1, 72) = 5.81$, $p = .019$, $\eta_p = .075$, with higher accuracy in the abstract condition than the concrete condition (M = .72, SD = .166; M = .66 SD = .195, respectively). There was no significant main effect of order, F < 1, however the condition by order interaction effect did reach significance, $F(1, 72) = 7.38$, $p = .008$, $\eta_p = .008$, as shown in Figure 9.6. For the group who completed the abstract condition first there was no significant difference between accuracy in the abstract and concrete conditions, $t(37)$=.235, $p = .815$. However, for the group who

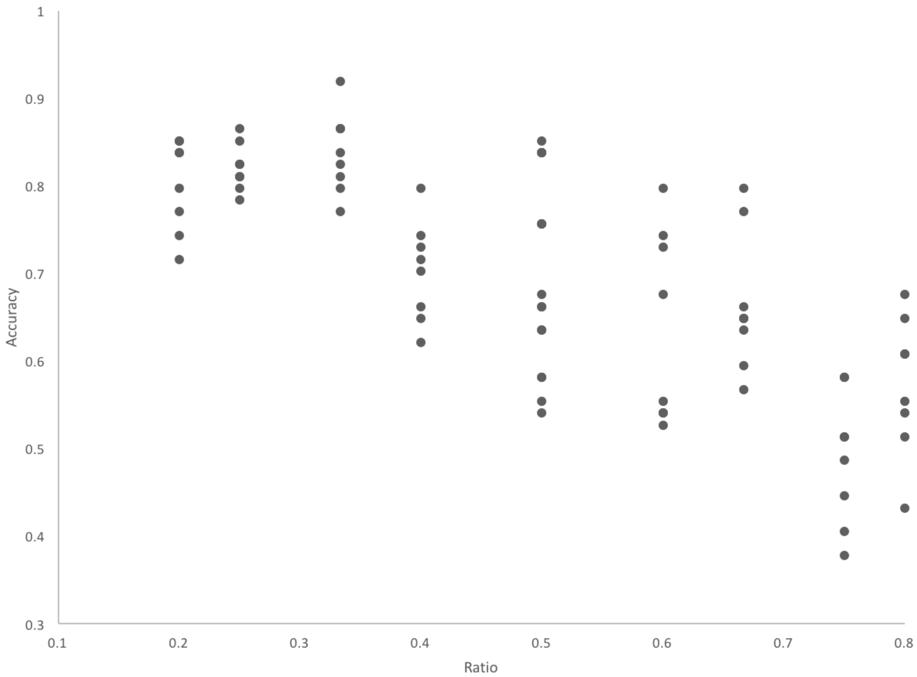

Figure 9.5: Mean accuracies on the symbolic comparison task by ratio (smaller:larger).

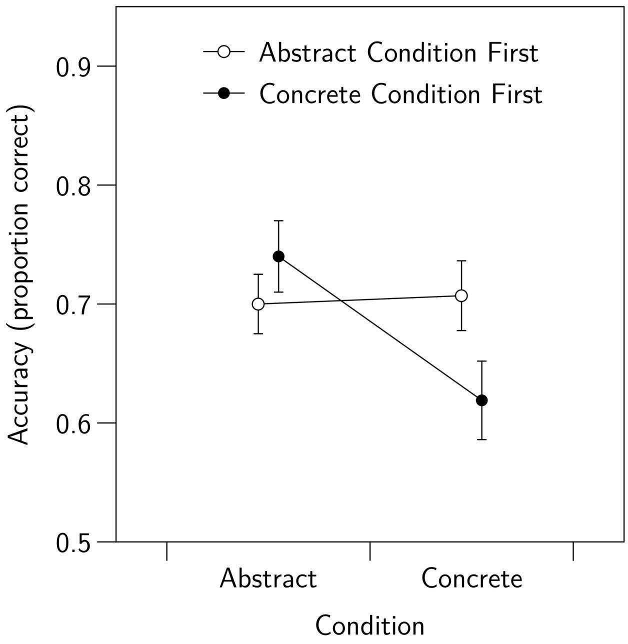

Figure 9.6: Mean accuracies on the symbolic comparison task by condition and order. Error bars show ±1 SE of the mean.

completed the concrete representations first, scores in the abstract condition were significantly higher than those in the concrete condition, $t(35) = 3.361$, $p = .002$. Finally, in a between-subjects comparison, I compared scores for only the condition each participant completed first. This revealed that those who learned from abstract stimuli scored higher than those who learned from concrete stimuli (*Ms* = .62, .70; *SDs* = .20, .16, for the abstract and concrete stimuli respectively), but this difference did not reach significance, $t(71) = 1.878$, $p = .065$; although I note that this study was powered for a within-subjects comparison not a between-subjects comparison.

In sum, I found three main results. First, children did seem able to learn the meaning of novel number symbols following a short training session: the subsequent symbolic comparison task showed above-chance performance. Second, children's performance when comparing the newly learned symbols showed similar ratio effects to those found in typical Arabic symbolic comparison tasks. This pattern of results is consistent with the suggestion that the children mapped the novel symbols to ANS representations in a similar fashion to that proposed by the dominant account of Arabic symbol acquisition (e.g., Fazio et al., 2014).

Finally, I found that children were more successful when they learned symbols using abstract representations than when they learned using concrete representations. There was a particularly low performance for children who learned from the concrete condition first. This finding is perhaps surprising in view of the suggestion that concreteness fading – using concrete representations first followed by abstract representations – is an effective method. However, children in this condition did not experience true concreteness fading, where concrete representations are gradually withdrawn in favour of abstract representations. To overcome this limitation, I conducted a further study.

Chapter 10

Learning novel symbols through concreteness fading (Study 4)

The main aim of Study 4 was to investigate whether concreteness fading would lead to more effective learning of new number symbols, compared to learning with abstract representations alone or concrete representations alone. I also included an abstractness-fading condition (concreteness fading in reverse, sometimes referred to as 'concreteness introduction') in a between-subjects experiment.

10.1 Method

10.1.1 Pre-Registration

Before data collection commenced, the study hypotheses, design, sample size, exclusion criteria and analysis plan was pre-registered at AsPredicted.com. The pre-registered protocol is available at `http://aspredicted.org/blind.php/?x=d8us8v`. Pre-registration of studies is important as it gives other researchers confidence in the results produced by that study. HARKing (hypothesising after the results are known) is a major problems in disciplines that use null hypothesis significance testing. John, Loewenstein, and Prelec (2012) estimated 90% of psychologists HARK. The purpose of preregistration is to give readers of the study confidence that HARKing has not taken place. With pre-registration researchers state their hypothesis and analysis plans in advance of data collection. Therefore making clear the distinction between the stage of generating the hypothesis and the stage of testing the hypothesis (Nuzzo, 2015). This does not prevent exploratory analyses but it does prevent researchers presenting exploratory findings as confirmatory findings (Wagenmakers & Dutilh, 2016).

10.1.2 Participants

Participants were a new group of 216 children (mean age 8 years, range: 7 years 1 month to 9 years 1 month, 114 boys). This sample size was chosen to give 90% power to detect a medium effect of $\eta_p^2 = 0.06$. Children were recruited from three primary schools in Nottinghamshire, UK, which had varying socio-economic status (SES), based on free school meals eligibility: 157 children came from two low-SES schools and 59 children came from a high-SES school. Children received stickers to thank them for taking part. Ethical approval for the study was received from the Loughborough University Ethics Approvals (Human Participants) Sub-Committee.

10.1.3 Procedure

I used a between-subjects design where participants were randomly assigned to one of four conditions: abstract only, concrete only, concreteness fading and abstractness fading. In each condition participants completed a training session, followed by a symbolic comparison task. The training and comparison tasks were presented on a laptop computer using PsychoPy software (Peirce & Peirce, 2009).

10.1.4 Training

The training phase consisted of 200 trials in which participants saw a symbol at the top of the screen with an array of dots/pictures (depending on condition) underneath, as in Study 3. Children were asked to learn how many dots/pictures were associated with each symbol. Each trial was displayed for 1000ms with a blank screen for 200ms between each trial. The trials were presented in random order and participants received breaks every 20 trials.

Five of the same novel numerical symbol stimuli from Study 3 were used (top row of Figure 9.2). Again, these symbols represented the numerosities 5, 10, 15, 20 and 25. To ensure that there was no inherent magnitude information included in the symbols, the association between the symbols and numerosities was counterbalanced across participants.

The dot/picture arrays were the same as those used in Study 3. Children in all conditions saw 200 training trials. In the abstract and concrete conditions children saw the same 100 trials as Study 3 but each trial was presented twice, in a random order. In the concreteness fading and abstractness fading conditions participants saw 100 dot arrays and 100 picture arrays. The combination of dot and picture arrays used in each condition is shown in Table 10.1.

	Blocks of 20 trials				
	1& 2	3 & 4	5 & 6	7 & 8	9 & 10
Abstract	100% Abstract	100% Abstract	100% Abstract	100% Abstract	100% Abstract
Concrete	100% Concrete	100% Concrete	100% Concrete	100% Concrete	100% Concrete
Concreteness Fading	100% Concrete	75% Concrete	50% Concrete	25% Concrete	0% Concrete
	0% Abstract	25% Abstract	50% Abstract	75% Abstract	100% Abstract
Abstract Fading	0% Concrete	25% Concrete	50% Concrete	75% Concrete	100% Concrete
	100% Abstract	75% Abstract	50% Abstract	25% Abstract	0% Abstract

Table 10.1: Combinations of abstract and concrete arrays used for each block in all four conditions.

10.1.5 Symbolic magnitude comparison task

Immediately following the training phase participants completed a symbolic magnitude comparison task, which was identical across conditions and identical to that used in Study 3, other than each stimuli pair was presented 8 times giving a total of 80 trials. Participants were given a short break halfway through the trials. Mean accuracy was calculated for each participant.

10.2 Results and discussion

All children completed the full experiment and no-one met our pre-registered exclusion criteria of performance more than 3 SDs above or below the mean. Therefore all analyses were performed on the full pre-registered sample of 216.

Mean accuracy in the symbolic magnitude comparison task was significantly above chance for all conditions, abstract only $t(53) = 7.34$, $p < .001$; concrete only, $t(53)=4.34$, $p < .001$; abstractness fading, $t(53) = 2.69$, $p = .009$; concreteness fading, $t(53) = 3.12$, $p = .003$. This replicates and extends Experiment 1 by demonstrating that children can learn the meaning of novel numerical symbols from a training session with either abstract, concrete or a mixture of both abstract and concrete representations. Next, I examined children's performance on the symbolic magnitude comparison test after removing trials including the smallest and largest symbols. The accuracy for the abstract only (M = .61, SD = .235), concrete only (M = .59, SD = .203) and concreteness fading conditions (M = .56, SD = .209) were still significantly above chance, $t(53) = 3.70$, $p = .001$, $t(53) = 3.26$, $p = .002$,$t(53) = 2.25$, $p = .029$, respectively. However for the accuracy for the abstractness-fading condition (M = .54, SD = .216) was not significantly above chance, $t(53) = 1.21$, $p = .232$. This indicates that children had not just learned the meaning of the two extreme values in most of the conditions. Next I checked whether children's performance on the symbolic comparison task showed the canonical ratio effect by conducting a by-items linear regression predicting the proportion of children answering each problem correctly by the problem's ratio. This revealed a significant effect of ratio, $\beta = -.839$, $p < .001$, $R^2 = .703$ (see Figure 10.1).

A one-way between-subjects ANOVA was conducted to compare mean accuracies from the four training conditions. This revealed a significant

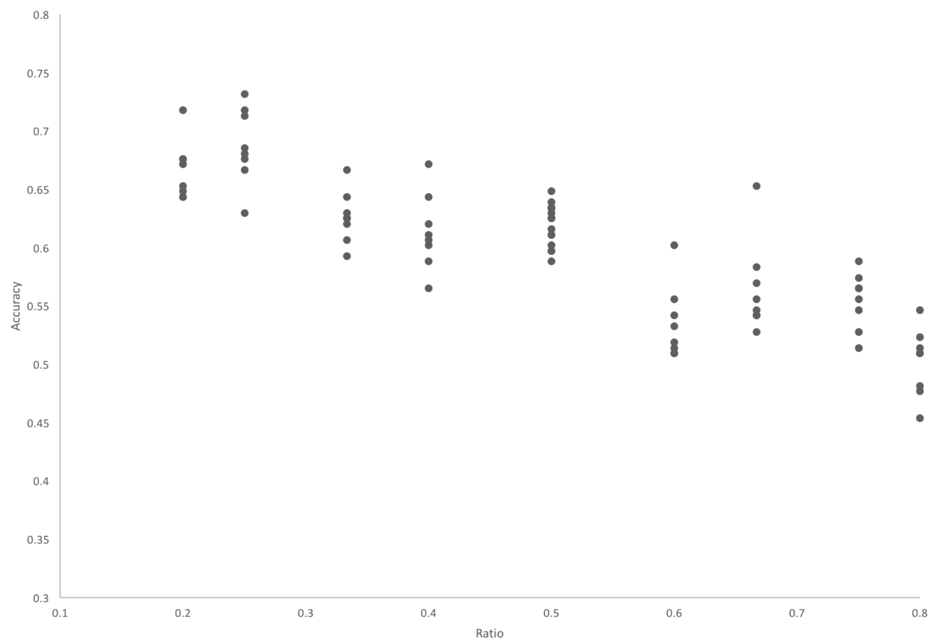

Figure 10.1: Mean accuracies on the symbolic comparison task by ratio (smaller:larger).

effect of training condition on accuracy, $F(3, 212) = 8.05$, $p < .001$, $\eta_p = .102$, shown in Figure 10.2. Post hoc comparisons using Tukey HSD tests indicated that the mean score for the abstract condition ($M = .69$, $SD = .19$) was significantly higher than the mean score for the concrete condition ($M = .58$, $SD = .14$, $p = .004$), the abstractness fading condition ($M = .56$, $SD = .16$, $p < .001$) and the concreteness fading condition ($M = .56$, $SD = .14$, $p < .001$). There were no significant differences between any other conditions.

Although I pre-registered a sample size of 216, I in fact tested 259 due to the number of children in the schools who wanted to take part. Running the analysis on this larger sample of N = 259 yielded essentially identical results. There was a significant effect of training condition on accuracy, $F(3, 255) = 9.42$, $p < .001$, $\eta_p = .100$. Post hoc comparisons indicated that the mean score for the abstract condition ($M = .69$, $SD = .19$) was significantly higher than the mean score for the concrete condition ($M = .58$, $SD = .14$, $p = .002$), the abstractness fading condition ($M = .55$, $SD = .15$, $p < .001$) and the concreteness fading condition ($M = .56$, $SD = .15$, $p < .001$). There

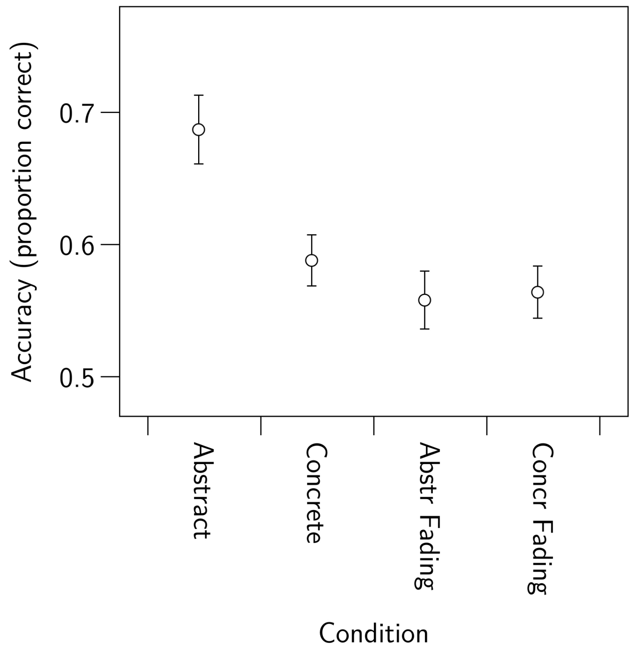

Figure 10.2: Mean accuracies on the symbolic comparison task, by condition. Error bars show ±1 SE of the mean.

were no significant differences between any other conditions.

In sum, as in Study 3, I found that children learned novel symbols more effectively when they were presented with abstract representations alone than when they were presented with concrete representations alone. However, here I also found that abstract representations alone were more effective than a mixture of abstract and concrete representations, using both the concreteness fading and the abstractness fading techniques. Performance did not differ between the concreteness fading and abstractness fading conditions.

In both Studies 3 and 4 I found that abstract representations appeared to help children acquire the meaning of novel number symbols. But why might this be the case? In both these studies the presence of concrete representations was confounded with the presence of multiple representations: as in traditional children's number books, I presented multiple concrete representations in the concrete conditions. In other words, children learned to associate the new symbol for 5 with five frogs, five pizzas, five cars and so on. In contrast, in the abstract condition children learned to associate this symbol with many arrays of five dots. This observation leaves open the possibility that the abstract advantage I found across Studies 3 and 4 was the result of using a single representation for numerosities – dots – rather than multiple representations. To disentangle the effect of concrete representations from the effect of multiple representations, I ran a third study in this area.

Chapter 11

Learning novel symbols through multiple concrete representations (Study 5)

The goal of Study 5 was to compare the effectiveness of learning novel number symbols from a single concrete representation compared to multiple concrete representations. In a between-subjects experiment I compared three learning conditions, where children learned new symbols with abstract representations, single concrete representations (five fish, five fish, ten fish, ten fish, etc.) and multiple concrete representations (five fish, five cakes, ten rockets, ten cars, etc.). If concrete representations per se disadvantage children, then I would expect to see the same abstract advantage found in Studies 4 and 5. If, however, the abstract advantage was the result of the concrete condition using multiple representations, I would expect to see children in the abstract and single-concrete conditions outperform those in the multiple-concrete condition.

11.1 Method

The study hypotheses, design, sample size, exclusion criteria and analysis plan was pre-registered at AsPredicted.com. The pre-registered protocol is available at `http://aspredicted.org/blind.php/?x=4hi37g` .

11.1.1 Participants

A new sample of 120 children took part in this study. This sample size gave 90% power to detect an effect of $\eta_p^2 = .1$ (based on the effect size found in Study 4). The children were recruited from two primary schools in Nottinghamshire, UK, which were both of low socio-economic status (SES), based on free school meal eligibility. Children's ages ranged from 7 years 4 months to 9 years 3 months ($M = 8$ years 3 months, 52 boys). Children received stickers to thank them for taking part. Ethical approval for the study was received from the Loughborough University Ethics Approvals (Human Participants) Sub-Committee.

11.1.2 Procedure

Participants were randomly assigned to one of three conditions: abstract, single-concrete and multiple-concrete. Participants in the single-concrete condition were then randomly assigned to one of four sets of stimuli (fish, cake, rocket or cars). In each condition participants completed a training session followed by a symbolic comparison task. The training and comparison tasks were presented on a laptop computer using PsychoPy software (Peirce & Peirce, 2009).

11.1.3 Training

Participants completed a similar training phase to that in Study 4. In each of 200 trials participants saw a symbol at the top of the screen with an array of dots/pictures (depending on condition) underneath. Children were asked to learn how many dots/pictures were associated with each symbol. Each trial was displayed for 1000ms with a blank screen displayed for 200ms between each trial. The trials were presented in random order and participants received breaks every 20 trials.

The same five novel numerical symbol stimuli from Study 4 were used. Again, these symbols represented the numerosities 5, 10, 15, 20 and 25. To ensure that there was no inherent magnitude information in the symbols, the association between the symbols and numerosities was counterbalanced across participants.

New nonsymbolic arrays were created. In the abstract condition the nonsymbolic stimuli were arrays of dots. The dots, which were the same size in all trials, were randomly placed within a 10 × 10 grid to create 20 different displays per numerosity. The multiple-concrete condition was identical to the concrete conditions in Studies 3 & 4. The stimuli were arrays of pictures, which were the same within an array but varied across the arrays. Twenty arrays of different pictures were created for each of the 5 numerosities by placing pictures of the same size as the dots in the exactly the same position as the dots stimuli. In the single-concrete condition the pictures were the same both within and across arrays. Four different sets of the stimuli were created for the concrete stable condition and participants were randomly assigned to one. Each set of stimuli used a different picture (fish, cake, rocket or cars) and contained 20 arrays of pictures for each of the five numerosities. These were created by placing pictures of the same size as the dots in the exactly the same position as the dots stimuli. In all conditions each trial was presented twice, in random order, resulting in 200 training trials. Examples of screens presented to the participant are shown in Figure 11.1.

11.1.4 Symbolic magnitude comparison task

Children were presented with the same symbolic magnitude comparison task used in Study 4.

11.2 Results and discussion

All children completed the full experiment and no-one met the pre-registered exclusion criteria of performance more than 3 SDs above or below the mean. Therefore, all analyses were performed on the full pre-registered sample of 120.

Participants performed at above chance levels in all three conditions (abstract, $t(39) = 7.92$, $p < .001$; single-concrete $t(39) = 6.39$, $p < .001$; multiple-concrete $t(39) = 2.96$, $p = .005$). Thus I again replicated the finding that children can accurately learn the meaning of novel symbols by associating them with the magnitude of nonsymbolic abstract or concrete representations. Next, I examined children's performance on the symbolic magnitude comparison test after removing trials including the smallest and largest symbols. The accuracy for the abstract ($M = .57$, $SD = .173$) and single-concrete conditions ($M = .60$, $SD = .180$) were still significantly

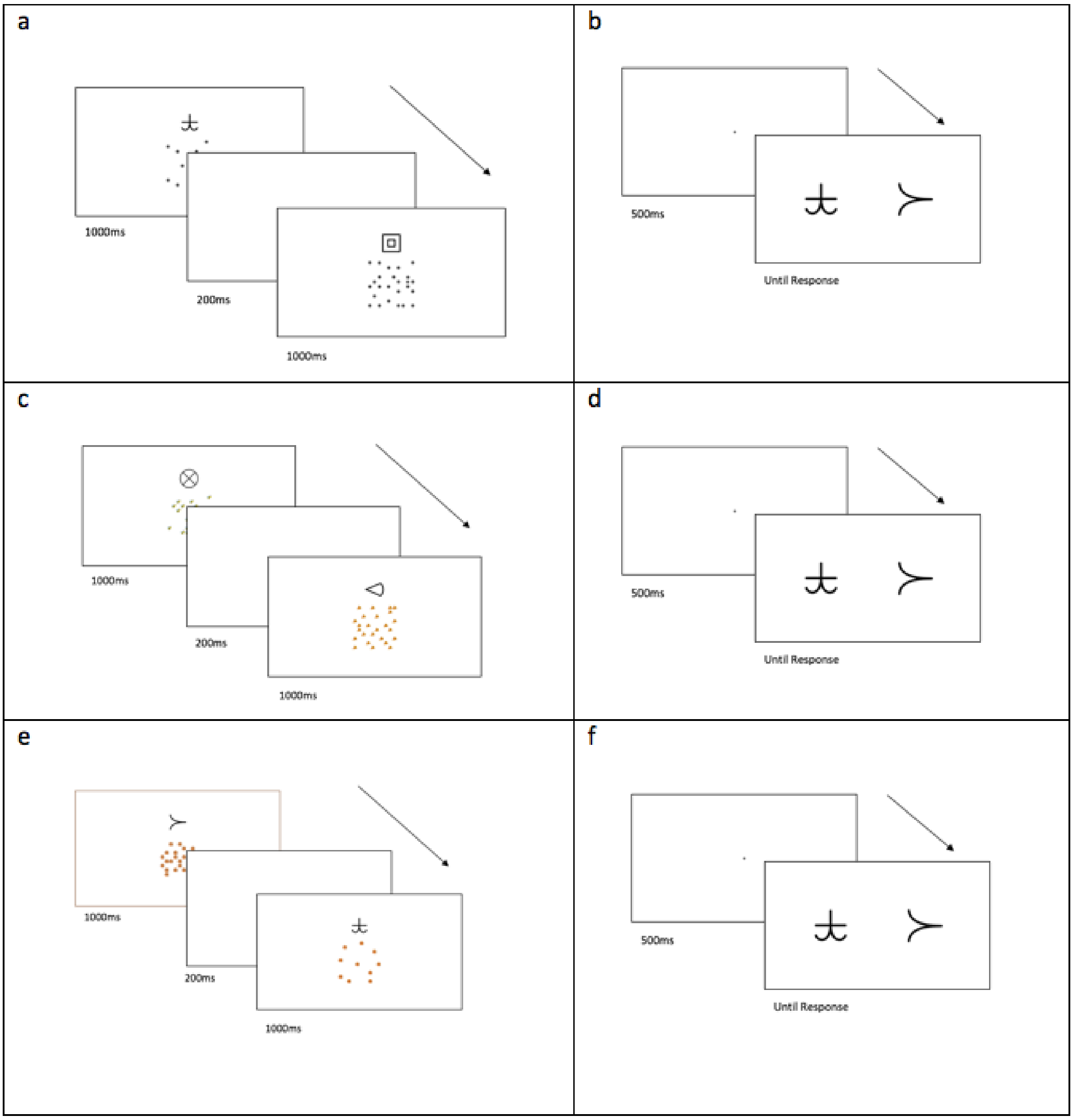

Figure 11.1: Examples of the screens seen by the participant in the a) abstract training b) abstract comparison task, c) multiple-concrete training, d) multiple-concrete comparison task, e) single-concrete training and f) single-concrete comparison task

above chance, $t(39) = 2.56$, $p = .015$, $t(39) = 3.37$, $p = .002$, respectively. However, the accuracy for the multiple-concrete condition ($M = .50$, *SD*

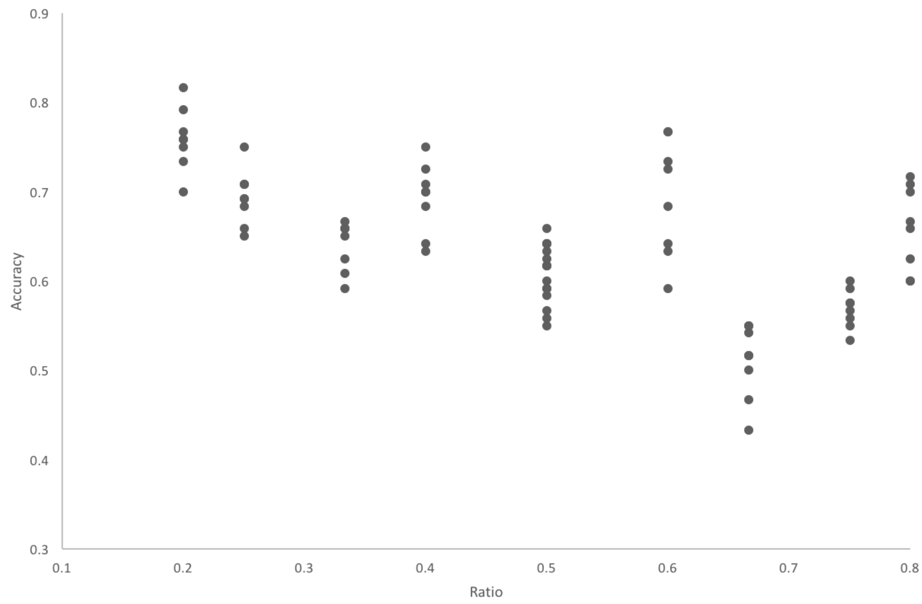

Figure 11.2: Mean accuracies on the symbolic comparison task by ratio (smaller:larger).

= .186) was not significantly above chance, $t(39) = 0.00$, $p = 1.00$. This indicates that children had not just learned the meaning of the two extreme values in most of the conditions. As before I conducted a by-items linear regression to assess whether there was a canonical ratio effect and, again as before, I found a significant effect of ratio, $\beta = -.545$, $p < .001$, $R^2 = .296$ (see Figure 11.2).

A one-way between-subjects ANOVA was conducted to compare accuracy on the symbolic comparison task following the three training conditions. There was a significant effect of condition on accuracy, $F(2, 117) = 8.66$, $p < .001$, $\eta_p = .129$, shown in Figure 11.3. Post hoc Tukey HSD tests indicated that the mean score for the multiple-concrete group ($M = .56$, $SD = .135$) was significantly lower than for the abstract ($M = .70$, $SD = .160$, $p < .001$) or single-concrete groups ($M = .67$, $SD = .164$, $p = .010$). These latter two groups did not differ significantly ($p = .566$).

In sum, I again found that children who learned the novel symbols using abstract representations outperformed those who learned using multiple-concrete representations. However, I found that there was no benefit to learning from abstract representations compared to single-concrete repre-

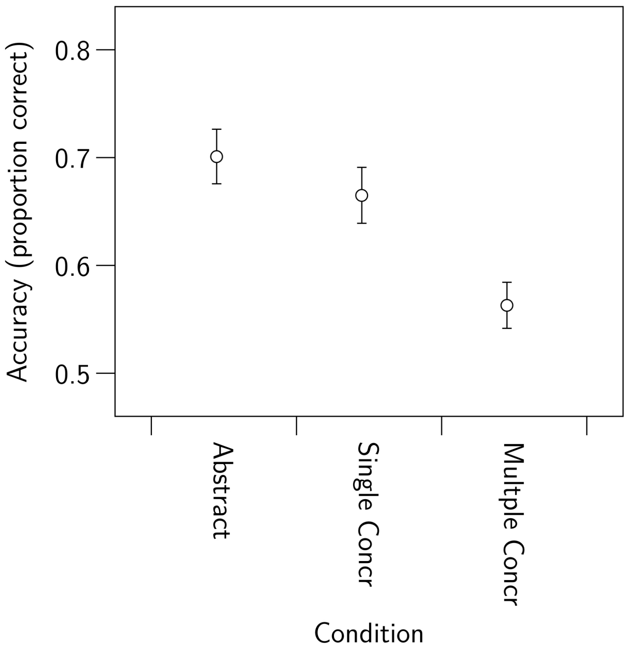

Figure 11.3: Mean accuracies on the symbolic comparison task, by condition. Error bars show ± 1 SE of the mean.

sentations, or vice versa. This pattern of results suggests that children's difficulty with the concrete stimuli used in Studies 3 & 4 stemmed not from their concreteness per se, but rather from the difficulty of dealing with multiple representations across trials. This will be discussed further in the next chapter.

Chapter 12

Learning number symbols through concrete and abstract representations

12.1 Summary of main findings from Chapters 9, 10 and 11

The first part of this thesis showed that there was need for more in-depth analysis of specific areas of the home numeracy environment. Therefore Part III of this thesis has focused on an in-depth analysis of how children can learn number symbols from number books. In particular, how can we help children to attach numerical meaning to number words and symbols? I focused on the relative merits of learning novel number symbols using abstract and concrete representations.

In Study 3, I found an abstract advantage: children who learned the meaning of novel symbols by pairing them with numerosities represented by an array of dots performed better on a subsequent symbolic comparison task than those who paired them with equivalent concrete representations. Study 4 replicated this result and extended it by also demonstrating an abstract advantage over both concreteness fading and abstractness fading approaches, each of which involved a mixture of abstract and concrete representations. Finally, Study 5 demonstrated that the abstract advantage found in Study 3 and 4 was not due to abstract representations being inherently superior to concrete representations, but rather due to the use of multiple concrete representations in the concrete conditions. In Study 5 I found no difference in symbolic comparison performance between those who learned from dots and those who learned from a single concrete representation. But both these groups outperformed those who learned with multiple concrete representations.

Overall, I found that there is a cost to context. Learning number symbols from multiple concrete representations – the approach adopted by the majority of children's number books (Ward et al., 2017) – seems to be less effective than learning from consistent concrete representation or an abstract representation. My discussion of these findings falls into three main sections. First, I discuss possible cognitive mechanisms for these results, focusing on the 'seductive details' effect. Second, I discuss implications for the wider debate about whether abstract or concrete representations should be favoured when teaching mathematics. Finally, I draw out implications for early number learning, and particularly discuss how my experimental setting differed from that in which children typically encounter Arabic numerals for the first time and conclude with outlining future directions for this research area.

12.2 Mechanisms

Why did those children who learned from abstract representations outperform those who learned from multiple concrete representations? The so-called 'seductive details' effect provides a natural account. Prior research has found that seductive details – the provision of information unconnected to the learning goal – can harm learning by activating irrelevant prior knowledge that the learner may try to integrate with the to-be-learned knowledge (e.g., Harp & Mayer, 1998). For instance, showing a child an array of five frogs may bring to mind knowledge about frogs that is irrelevant to the fiveness of the representation. If the child is to be successful, then this irrelevant prior knowledge must be inhibited, and only the numerosity of the array associated with the novel symbol. The seductive details account also seems to explain the difference in performance between the multiple- and single-concrete conditions seen in Study 5. It is likely that inhibiting prior knowledge is easier when the same knowledge is activated on every trial than when new knowledge is activated from trial to trial. For example, performance on trial n of a Stroop task is facilitated when the to-be-inhibited text is identical to that presented on trial $n-1$ (Lowe, 1979; MacLeod, 1991).

Many researchers have found that the failure to inhibit irrelevant prior knowledge can damage learning by consuming limited working memory capacity (e.g., Harnishfeger & Bjorklund, 1993; Sanchez & Wiley, 2006). These factors, therefore, suggest that one plausible account of the lower performance of children in the multiple-concrete conditions was that they failed to inhibit prior knowledge automatically activated by the concrete representations, that this increased the load on their working memory, and that this therefore damaged their ability to map the novel symbols onto their ANS representations. Another possibility is that children's ANS representations themselves were less precise in the multiple-concrete conditions, due to an increase in working memory load caused by a failure to successfully inhibit irrelevant prior knowledge activated by the concrete representations. The literature offers conflicting evidence about the plausibility of this latter account. Some researchers have found that performance on nonsymbolic comparison tasks is correlated with measures of working memory capacity, suggesting that working memory resources are implicated in the ability to form precise ANS representations (Xenidou-Dervou, De Smedt, van der Schoot, & van Lieshout, 2013; Xenidou-Dervou, Molenaar, Ansari, van der Schoot, & van Lieshout, 2017). This conclusion is also consistent with the finding that working memory loads damage participants' ability to perform nonsymbolic arithmetic (Xenidou-Dervou, van Lieshout, & van der Schoot, 2014). On the other hand, Fuhs, McNeil, Kelley, O'Rear, and Villano (2016) asked children to tackle a nonsymbolic comparison using stimuli similar to those used in my abstract and multiple-concrete conditions. They found that nonsymbolic comparison performance did not significantly differ between the abstract and multiple-concrete conditions, suggesting that children are able to form similarly precise ANS representations from both stimuli types. This would be surprising if the lower performance in my studies was primarily due to imprecise ANS representations. A third possibility is that both these accounts – less precise ANS representations and lower quality ANS-to-symbol mapping – played a role in the lower performance exhibited by the children in the multiple-concrete conditions.

12.3 Abstract and concrete representations

There is a longstanding debate about whether instructional materials in mathematics should favour abstract or concrete representations. While many teachers and researchers have argued in favour of using concrete representations (e.g., Bruner, 1966; Piaget, 1971), others have pointed out that there are reasons to prefer abstract representations (e.g., Kaminski et al., 2008; Uttal et al., 2009). Still others have proposed combining both abstract and concrete representations using a concreteness fading technique (e.g., Fyfe et al., 2014).

My results are clear. In the context of associating numerosities with novel symbols, I found that children who learned from abstract representations outperformed those who learned from either concrete representations, or from a sequence of representations that faded from concrete to abstract (or vice versa). These results, combined with those from the wider literature, perhaps suggest that looking for a universal answer to the abstract versus concrete debate may be misguided. For instance, Day et al. (2015) found that abstract representations were more effective than concrete representations when teaching beginning psychology undergraduates about measures of central tendency. Koedinger and Nathan (2004) found that high school student's algebra problem solving performance was improved when using concrete story problems opposed to abstract mathematical equations. McNeil and Fyfe (2012) found that concreteness fading improved undergraduates' learning of modulo arithmetic compared to the use of concrete and abstract representations alone. One way of making sense of these disparate findings is to propose that there is no universal answer to the question of what type of representations are better for learning in general. It may be that different answers will emerge for young children learning number symbols to high school students solving algebra problems, to undergraduates learning mathematical concepts. If this suggestion is correct, then it would be beneficial for researchers to move beyond the question of whether abstract or concrete representations are better, and instead to ask when they are better.

Consistent with this suggestion, the result of Study 5 demonstrated that where concrete representations have been found to be less effective than abstract representations, this may not be down to the concrete nature of the representations per se, but rather to the use of multiple concrete representations rather a single abstract representation. This finding echoes Ainsworth's (2006) warning that although multiple representations can sometimes be beneficial for learning, this is not always the case.

12.4 Implications for early number learning and number books design

Parents are commonly encouraged to help their children acquire number words and symbols by reading number books. What implications do my findings have for the design of such resources? I highlight two important differences between the context of the experiments reported here, and children's first introduction to Arabic number symbols.

First, typically children first encounter Arabic number symbols at a much younger age than the participants in my studies (who ranged from 6 to 10 years old). Clearly some care is needed before I can generalize the lessons learned from how older children performed on my artificial symbol learning paradigm to the learning of Arabic symbols by younger children. However, if the mechanism behind my results is as I have suggested, then there are two reasons to suppose that the abstract advantage would be even greater with younger children. Earlier research has found that the harmful effects of seductive details are greater for participants with lower working memory capacity (Sanchez & Wiley, 2006). Since working memory capacity is developmental (e.g., Gathercole, Pickering, Ambridge, & Wearing, 2004), it is reasonable to suppose that younger children would be more distracted by irrelevant knowledge when reading number books than older children, and therefore that multiple concrete representations would have a more deleterious effect on their number symbol acquisition.

Second, the older participants in my study were all familiar with the notion of representing numerosities with symbols, a fact which allowed us to simply tell them that the novel symbols were related to the number of items in the display. In contrast, younger children encountering Arabic symbols or number words for the first time must first infer that the concept the symbol represents is the number of items in the display, and not some property of the objects represented in the concrete representations. Indeed, as discussed above, Huang et al. (2010) found that children sometimes find it difficult to generalise number words from the real-world contexts in which they were introduced.

Both these factors suggest that the abstract advantage I found here may be even greater in the context of young children learning Arabic symbols or number words for the first time. However, directly investigating this issue would be a valuable goal for future research.

These considerations highlight a further issue. In all three experiments I used children's symbolic comparison performance as a measure of the extent to which they had learned the novel symbol system. However, it is also important that children are able to map between number symbols and non-symbolic quantities. In other words, although I would certainly like children to understand that 5 is greater than 3 – the skill I tested – I would also like them to know that the symbol '5' and number word 'five' should be attached to a picture of five cars rather than a picture of three cars. Would the abstract advantage I found here generalize to alternate measures of numerical performance such as mapping tasks? I cannot directly answer this question, but do note that there is evidence that Arabic symbolic comparison seems to be a more important skill for formal mathematics than symbolic-to-nonsymbolic mapping. Mundy and Gilmore (2009) found that 6-7 year old children's symbolic comparison performance was strongly correlated with a test of school mathematics achievement ($r = .53$), whereas performance on a mapping task was not significantly correlated with the same test ($r = .167$). In other words, assigning numerical meaning to number symbols in such a way that permits fluent symbolic comparison appears to be a more important for children's future mathematical development than mapping, and one that I found to be advantaged by abstract representations.

Therefore in order to overcome some of the limitations of these studies and be confident that the way number books are designed should be changed, future studies should design a number book using a single concrete representations and compare learning through this book to learning through a number book using multiple concrete representations. Future studies should be conducted with young children and could be designed using novel number symbols to avoid children's prior experiences with Arabic number symbols. This type of study could be run as an intervention in which parents are encouraged to share these books at home with their child for a number of weeks. The children could then be tested on their understanding of the novel symbols using a range of comparison and mapping tasks. If the single representation number book is more effective than the multiple representations book then the design of future number books should be considered. Furthermore a scheme, similar to BookTrust (2017), could be developed in which all children are provided with some single representation number books to aid children's early number learning in the home numeracy environment.

12.5 Summary

To summarise the work presented in Part III of this thesis focused on children learning number symbols from number books, specifically the benefits and costs of using concrete or abstract representations to support children's learning of novel number symbols. Three studies investigated children's learning of novel number symbols when using abstract and/or concrete representations. Initially in Study 3 and 4 it was thought that concrete representations were less effective for learning than abstract representations but Study 5 importantly highlighted that it was not the concrete representations per se that were causing the difficultly but the multiple representations used within the concrete representations conditions. These studies highlight that multiple representations come at a cost for early number learning and the design of the majority of number books is not the most effective design for learning number symbols. In conclusion, this section highlights that when designing resources for early number learning the cost of using multiple representations should be taken into account.

Part IV

General Discussion

Chapter 13

Conclusions

This chapter aims to bring together the findings from both parts of this thesis. I will start by reviewing the aims of this thesis outlined in Chapter 1. Next, I will highlight the key findings from the research presented in this thesis followed by my thoughts for future research into the home numeracy environment. Finally, I will finish on my concluding remarks for this thesis.

13.1 Overview of thesis aims

Number skills have shown to be important not just for job prospects but also for personal health, and it was shown that children's individual differences in number skills when starting formal schooling predict later mathematical ability. Four factors have been found to influence young children's individual differences in mathematics performance, domain-general cognitive abilities, domanin specific cognitive abilities, dispositional factors and environmnental factors. The environment factor and in particular the home numeracy environment section had received a lot less attention in the literature than the other factors. Therefore, the aim of this thesis was to address this gap by investigating the effects of individual differences in the home numeracy environment.

Four mains aims were highlighted for the home numeracy environment research. Firstly, to investigate the relationship between the home numeracy environment and mathematics achievement in the UK. Secondly, to investigate the reliability and validity of a novel text message method to measure the home numeracy environment. Thirdly, to investigate how three different measures of the home numeracy environment relate to each other, mathematics achievement and demographic variables. Finally, to evaluate if it is appropriate to classify home numeracy activities as formal and informal activities.

The findings from Part II of the thesis lead to a more in-depth analysis of one aspect of the home numeracy environment in Part III. The main aim of this part was to investigate if using concrete images in number books is beneficial for teaching children number symbols.

13.2 Research findings and future directions

The main finding from Part II of this thesis was that the home numeracy environment's relationship with mathematics and other variables (such as demographic variables, parents' expectations and attitudes) is inconsistent. In my studies, I showed that while a new text message measure of the home numeracy environment appeared to be valid and reliable, when three different measures of the home numeracy environment were compared, the self-report measures were not related to the observation measures and none of the measures, apart from child number talk, related to mathematics performance, even when dividing the questionnaire into formal and informal activities.

A number of factors could explain the lack of relationship between the home numeracy environment and mathematics: 1) this could be due to the mathematics measure used, 2) another demographic variable that we did not measure (such as parent's IQ) might be more strongly related to childrens' mathematics, 3) the home numeracy environment could be too complex to be captured accurately by these measures or 4) my findings could show the true relationship between the home numeracy environment and mathematics achievement. It was decided that it would not be possible to explain these results using the current measures of the home numeracy environment because correlational studies do not tell us about the direction or causality of the relationship. Therefore, this research led me to suggest two ways to develop the home numeracy environment research for the future: intervention studies or a more in-depth analysis. Intervention studies would only change one aspect of the home numeracy environment and therefore we would be able to see which aspects of the home numeracy environment have an impact on mathematics skills. Alternatively, future research could break down the home numeracy environment into specific activities parents do with their child at home, for example reading number books, and conduct an in-depth analysis of these activities to find out how they are effective in teaching number skills. Analyses of a range of specific activities put together could help us understand the whole home numeracy environment.

For Part III of this thesis, I conducted an in-depth analysis of how children learn number symbols from number books, particularly looking into abstract and concrete representations. The main finding for the three studies investigating the use of abstract and concrete representations in number symbol learning was that it is not important if the representation is abstract or concrete, but it is important that the representation is consistent throughout. This has important implications for the design of number books where the majority of number books have different pictures on each page (e.g. one cow, two pigs, three sheep etc.). Future studies should develop a number book using the same representation throughout (e.g. one cow, two cows, three cows etc.) and evaluate its effectiveness.

In a similar way, future research should investigate the effectiveness of other specific home numeracy environment activities in learning and teaching mathematics skills.

13.3 Concluding remarks

It is well-known that young children's mathematics skills are important for later life. This thesis highlights that the home numeracy environment could be the cause of some of the individual differences in children's mathematics performance, but the home numeracy environment is too broad of a concept that could be influenced by so many other factors to be measured accurately. This has key implications for the future of home numeracy environment research. In education research, it is accepted that just by teaching children more does not mean that children will learn more and we should look at the type and quality of the teaching and the same ideas should apply to the home numeracy environment, Just doing more number activities will not result in better mathematics skills, we should be investigating the type and quality of number activities in more detail. This thesis starts this process by investigating number books, but there are still many more home numeracy activities to investigate.

Part V

References & Appendices

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*, 131–152. doi: 10.1016/S0360-1315(99)00029-9

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*, 183–198. doi: 10.1016/j.learninstruc.2006.03.001

Ainsworth, S., Bibby, P., & Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences*, *11*, 25–61. doi: 10.1207/S15327809JLS1101_2

Alfvn, G. (2010). SMS pain diary: a method for real-time data capture of recurrent pain in childhood. *Acta Pdiatrica*, *99*, 1047–1053. doi: 10.1111/j.1651-2227.2010.01735.x

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. doi: 10.1016/j.ecresq.2011.08.003

Anderson, A., Anderson, J., & Shapiro, J. (2004). Mathematical discourse in shared storybook reading. *Journal for Research in Mathematics Education*, *35*, 5–33. doi: 10.1007/1-4020-7914-1_2

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. doi: 10.1037/0022-0663.96.4.699

Axn, I., Bodin, L., Bergstrm, G., Halasz, L., Lange, F., Lvgren, P. W., … Jensen, I. (2012). The use of weekly text messaging over 6 months was a feasible method for monitoring the clinical course of low back pain in patients seeking chiropractic care. *Journal of Clinical Epidemiology*, *65*, 454–461. doi: 10.1016/j.jclinepi.2011.07.012

Axn, I., Bodin, L., Kongsted, A., Wedderkopp, N., Jensen, I., & Bergstrm, G. (2012). Analyzing repeated data collected by mobile phones and frequent text messages. An example of low back pain measured weekly for 18 weeks. *BMC Medical Research Methodology*, *12*, 105. doi: 10.1186/1471-2288-12-105

Batchelor, S. (2014). *Dispositional factors affecting children's early numerical development* (Thesis). `https://dspace.lboro.ac.uk/2134/17474`.

Bialystok, E. (1992). Symbolic representation of letters and numbers. *Cognitive Development*, *7*, 301–316. doi: 10.1016/0885-2014(92)90018-M

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. doi: 10.1080/0300443001650104

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 and Parenting*, *5*, 35–45. doi: 10.1002/(SICI)1099-0917(199603)5:1⟨35::AID-EDP113⟩3.0.CO;2-0

BookTrust. (2017). *BookStart.* Retrieved from `http://www.bookstart.org.uk/d/bookstart/home/`

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to Meta-Analysis.* Sussex, United Kingdom: John Wiley & Sons.

Bornstein, M., & Bradley, R. (2014). *Socioeconomic status, parenting, and child development.* Abingdon, Oxen: Taylor & Francis.

Brenner, M. E., Mayer, R. E., Moseley, B., Brar, T., Durn, R., Reed, B. S., & Webb, D. (1997). Learning by understanding: The role of multiple representations in learning algebra. *American Educational Research Journal*, *34*, 663–689. doi: 10.3102/00028312034004663

Brown, M. C., McNeil, N. M., & Glenberg, A. M. (2009). Using concreteness in education: Real problems, potential solutions. *Child Development Perspectives*, *3*, 160–164. doi: 10.1111/j.1750-8606.2009.00098.x

Bruner, J. S. (1966). *Toward a Theory of Instruction.* Cambridge, Massachusetts: Harvard University Press.

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford, UK: Oxford University Press.

Bull, R., Espy, K. A., & Wiebe, S. A. (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. doi: 10.1080/87565640801982312

Carraher, T. N., & Schliemann, A. D. (1985). Computation routines prescribed by schools: Help or hindrance? *Journal for Research in Mathematics Education*, *16*, 37–44. doi: 10.2307/748971

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, *62*, 233–246. doi: 10.1111/j.2044-8279.1992.tb01017.x

Christie, A., Dagfinrud, H., Dale, ., Schulz, T., & Hagen, K. B. (2014). Collection of patient-reported outcomes; - text messages on mobile phones provide valid scores and high response rates. *BMC Medical Research Methodology*, *14*, 52. doi: 10.1186/1471-2288-14-52

Ciping, D., Silinskas, G., Wei, W., & Georgiou, G. K. (2015). Cross-lagged relationships between home learning environment and academic achievement in Chinese. *Early Childhood Research Quarterly*, *33*, 12–20. doi: 10.1016/j.ecresq.2015.05.001

Clayton, S., & Gilmore, C. (2014). Inhibition in dot comparison tasks. *ZDM*, 1–12. doi: 10.1007/s11858-014-0655-2

Cunha, F., & Heckman, J. (2007). *The technology of skill formation* (Working Paper). National Bureau of Economic Research. Retrieved from `http://www.nber.org/papers/w12840`

Day, S. B., Motz, B. A., & Goldstone, R. L. (2015). The cognitive costs of context: The effects of concreteness and immersiveness in instructional examples. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.01876

Dearing, E., Casey, B. M., Ganley, C. M., Tillinger, M., Laski, E., & Montecillo, C. (2012). Young girls arithmetic and spatial skills: The distal and proximal roles of family socioeconomics and home learning experiences. *Early Childhood Research Quarterly*, *27*, 458–470. doi: 10.1016/j.ecresq.2012.01.002

De Bock, D., Deprez, J., Van Dooren, W. V., Roelens, M., & Verschaffel, L. (2011). Abstract or concrete examples in learning mathematics? a replication and elaboration of Kaminski, Sloutsky, and Heckler's

Study. *Journal for Research in Mathematics Education*, *42*, 109–126. doi: 10.5951/jresematheduc.42.2.0109

DeFlorio, L. (2013). *The influence of the home learning environment on preschool children's informal mathematical development: Variation by age and socioeconomic status.* (Doctoral dissertation, US). Retrieved from `http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2013-99010-509&site=ehost-live`

DeFlorio, L., & Beliakoff, A. (2014). Socioeconomic status and preschoolers' mathematical knowledge: The contribution of home activities and parent beliefs. *Early Education and Development*, *26*, 1–23. doi: 10.1080/10409289.2015.968239

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42. doi: 10.1016/0010-0277(92)90049-N

Dehaene, S. (2011). *The Number Sense: How the Mind Creates Mathematics, Revised and Updated Edition.* Oxford University Press, USA.

DeLoache, J. S. (2000). Dual representation and young children's use of scale models. *Child Development*, *71*, 329–338. doi: 10.1111/1467-8624.00148

Department for Education. (2014). Statutory framework for the early years foundation stage. *Runcorn: Department for Education*.

De Smedt, B., Nol, 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. doi: 10.1016/j.tine.2013.06.001

Dilworth-Bart, J. E. (2012). Does executive function mediate SES and home quality associations with academic readiness? *Early Childhood Research Quarterly*, *27*, 416–425. doi: 10.1016/j.ecresq.2012.02.002

Doig, B. A. (1989). *Links: A Guide to Maths in Children's Literature.* Melbourne: Nelson.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428–1446. doi: 10.1037/0012-1649.43.6.1428

Dunst, C. J., Hamby, D. W., Wilkie, H., & Dunst, K. S. (2017). Meta-analysis of the relationship between home and family experiences and young childrens early numeracy learning. In S. Phillipson, A. Gervasoni, & P. Sullivan (Eds.), *Engaging families as children's first mathematics educators* (pp. 105–125). Singapore: Springer.

Esplin, J. A., Neilson, B. G., Austin, A. B., Blevins-Knabe, B., Hendershot, S. M., & Loesch, L. A. (2016). Number line skills and home numeracy activities for preschoolers in center-based and family-based child care. In B. Blevins-Knabe & A. M. B. Austin (Eds.), *Early childhood mathematics skill development in the home environment.* (pp. 105–126). Cham, Switzerland: Springer International Publishing. doi: 10.1007/978-3-319-43974-7_7

Evans, & Shaw, D. (2008). Home grown for reading: Parental contributions to young children's emergent literacy and word recognition. *Canadian Psychology*, *49*, 89. doi: 10.1037/0708-5591.49.2.89

Evans, J. S. B. T. (2005). *How to do research: A psychologist's guide.* Sussex, United Kingdom: Taylor & Francis.

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, *123*, 53–72. doi: 10.1016/j.jecp.2014.01.013

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314. doi: 10.1016/j.tics.2004.05.002

Flack, Z. M., & Horst, J. S. (2017). Two sides to every story: Children learn words better from one storybook page at a time. *Infant and Child Development*. doi: 10.1002/icd.2047

Fuhs, M. W., McNeil, N. M., Kelley, K., O'Rear, C., & Villano, M. (2016). The role of non-numerical stimulus features in approximate number system training in preschoolers from low-income homes. *Journal of Cognition and Development*, *17*, 1–28. doi: 10.1080/15248372.2015.1105228

Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: a systematic review. *Educational Psychology Review*, *26*, 9–25. doi: 10.1007/s10648-014-9249-3

Gailey, S. K. (1993). The mathematics - children's literature connection. *The Arithmetic Teacher*, *40*, 258–261.

Galindo, C., & Sheldon, S. B. (2012). School and home connections and children's kindergarten achievement gains: The mediating role of family involvement. *Early Childhood Research Quarterly*, *27*, 90–103. doi: 10.1016/j.ecresq.2011.05.004

Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of 'seductive details' on macroprocessing and microprocessing in adults and children. *Cognition and Instruction*, *6*, 41–57. doi: 10.1207/s1532690xci0601_2

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177–190. doi: 10.1037/0012-1649.40.2.177

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., ... Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLOS ONE*, *8*, e67374. doi: 10.1371/journal.pone.0067374

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). *Mathematics education for young children: What it is and how to promote it.* (Vol. 22). Society for Research in Child Development.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology*, *96*, 424–436. doi: 10.1037/0022-0663.96.3.424

Goldstone, R. L., & Son, J. Y. (2005). The transfer of scientific principles using concrete and idealized simulations. *Journal of the Learning Sciences*, *14*, 69–110. doi: 10.1207/s15327809jls1401_4

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–668. doi: 10.1038/nature07246

Halsey, P. (2005). Assessing mathematics tradebooks: do they measure up? *Reading Improvement*, *42*, 158–164.

Hannula, M. M., & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, *15*, 237–256. doi: 10.1016/j.learninstruc.2005.04.005

Hannula, M. 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. doi: 10.1016/

j.jecp.2010.06.004

Harnishfeger, K. K., & Bjorklund, D. F. (1993). The ontogeny of inhibition mechanisms: A renewed approach to cognitive development. In *Emerging Themes in Cognitive Development* (pp. 28–49). Springer, New York, NY. doi: 10.1007/978-1-4613-9220-0_2

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*, 414–434. doi: 10.1037/0022-0663.90.3.414

Hassinger-Das, B., Jordan, N. C., & Dyson, N. (2015). Reading stories to learn math: Mathematics vocabulary instruction for children with early numeracy difficulties. *The Elementary School Journal*, *116*, 242–246. doi: 10.1086/683986

Hattie, J. (2009). Visible learning: A synthesis of 800+ meta-analyses on achievement. *Abingdon: Routledge*.

Heckman, J. J. (2013). *Giving kids a fair chance.* Cambridge, Mass: The MIT Press.

Hellwig, S. J., Monroe, E. E., & Jacobs, J. S. (2000). Making informed choices: Selecting children's trade books for mathematics instruction. *Teaching Children Mathematics; Reston*, *7*, 138.

Higgins, S., Katsipataki, M., Kokotsaki, D., Coe, R., Major, L. E., & Coleman, R. (2013). The sutton trust-education endowment foundation teaching and learning toolkit: Technical appendices. *London: Education Endowment Foundation*.

Hojnoski, R. L., Columba, H. L., & Polignano, J. (2014). Embedding mathematical dialogue in parent–child shared book reading: A preliminary investigation. *Early Education and Development*, *25*, 469–492. doi: 10.1080/10409289.2013.810481

Hong, H. (1996). Effects of mathematics learning through children's literature on math achievement and dispositional outcomes. *Early Childhood Research Quarterly*, *11*, 477–494. doi: 10.1016/S0885-2006(96)90018-6

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1-55. doi: 10.1080/10705519909540118

Huang, Y. T., Spelke, E., & Snedeker, J. (2010). When is four far more than three? Childrens generalization of newly-acquired number words. *Psychological science*, *21*, 600–606. doi: 10.1177/0956797610363552

Hunsader, P. D. (2004). Mathematics trade books: Establishing their value and assessing their quality. *The Reading Teacher*, *57*, 618–629.

Huntsinger, C., Jose, P., & Luo, Z. (2016). Parental facilitation of early mathematics and reading skills and knowledge through encouragement of home-based activities. *Early Childhood Research Quarterly*, *37*, 1–15. doi: 10.1016/j.ecresq.2016.02.005

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. doi: 10.1037/0012-1649.27.2.236

Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*, 128–141. doi: 10.1207/S15328007SEM1001_6

Jenner, D. M. (2002). Experiencing and understanding mathematics in the midst of a story. *Teaching Children Mathematics*, *9*, 167–71.

Jennings, C. M., Jennings, J. E., Richey, J., & Dixon-Krauss, L. (1992). Increasing interest and achievement in mathematics through children's literature. *Early Childhood Research Quarterly*, *7*, 263–276. doi: 10.1016/0885-2006(92)90008-M

Johansen, B., & Wedderkopp, N. (2010). Comparison between data obtained through real-time data capture by SMS and a retrospective telephone interview. *Chiropractic & Manual Therapies*, *18*, 10. doi: 10.1186/1746-1340-18-10

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*, 524-532. doi: 10.1177/0956797611430953

Jones, M. G. (2009a). Examining surface features in context. *Journal for Research in Mathematics Education*, *40*, 94–96.

Jones, M. G. (2009b). Transfer, abstraction, and context. *Journal for Research in Mathematics Education*, *40*, 80–89.

Jong, T. d., Ainsworth, S., Dobson, M., Hulst, A. v. d., Levonen, J., Reimann, P., ... Swaak, J. (1998). Acquiring knowledge in science and mathematics: the use of multiple representations in technology based learning environments. In M. W. Someren (Ed.), *Learning with multiple representations* (pp. 9–41). Oxford: Elsevier Science.

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice*, *22*, 36–46. doi: 10.1111/j.1540-5826.2007.00229.x

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*, 850–867. doi: 10.1037/a0014939

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science*, *320*, 454–455. doi: 10.1126/science.1154659

Kleemans, T., Peeters, M., Segers, E., & Verhoeven, L. (2012). Child and home predictors of early numeracy skills in kindergarten. *Early Childhood Research Quarterly*, *27*, 471–477. doi: 10.1016/j.ecresq.2011.12.004

Kleemans, T., Segers, E., & Verhoeven, L. (2013). Relations between home numeracy experiences and basic calculation skills of children with and without specific language impairment. *Early Childhood Research Quarterly*, *28*, 415–423. doi: 10.1016/j.ecresq.2012.10.004

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. doi: 10.1037/0012-1649.42.1.59

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling.* New York: Guilford Publications.

Kluczniok, K., Lehrl, S., Kuger, S., & Rossbach, H.-G. (2013). Quality of the home learning environment during preschool age – Domains and contextual conditions. *European Early Childhood Education Research Journal*, *21*, 420–438. doi: 10.1080/1350293X.2013.814356

Knudsen, B., Fischer, M. H., Henning, A., & Aschersleben, G. (2015). The development of arabic digit knowledge in 4- to 7-year-old children. *Journal of Numerical Cognition*, *1*, 21–37. doi: 10.5964/jnc.v1i1.4

Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: evidence from algebra problem solving. *Cognitive Science*, *32*, 366–397. doi: 10.1080/03640210701863933

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences*, *13*, 129–164. doi: 10.1207/s15327809jls1302_1

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, *17*, 248–294. doi: 10.1016/0010-0285(85)90009-X

Kuntsche, E., & Robert, B. (2009). Short Message Service (SMS) Technology in Alcohol ResearchA Feasibility Study. *Alcohol and Alcoholism*, *44*, 423–428. doi: 10.1093/alcalc/agp033

Lakens, D. (2017). *On the reproducibility of meta-analyses.* `http://daniellakens.blogspot.co.uk/2014/08/on-reproducibility-of-meta-analyses.html`.

Lakens, D., & Etz, A. J. (2017). Too true to be bad: When sets of studies with significant and nonsignificant findings are probably true. *Social Psychological and Personality Science*, 1948550617693058. doi: 10.1177/1948550617693058

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. doi: 10.1016/j.cognition.2006.10.005

LeFevre, J., Polyzoi, E., Skwarchuk, S., 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. doi: 10.1080/09669761003693926

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 Englishspeaking Canadian children. *Early Child Development and Care*, *172*, 283–300. doi: 10.1080/03004430212127

LeFevre, J.-A., & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction*, *3*, 1–30. doi: 10.1207/s1532690xci0301_1

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. doi: 10.1037/a0014532

Leibovich, T., & Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology*, *70*, 12–23. doi: 10.1037/cep0000070

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2016). From sense of number to 'sense of magnitude' – The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, 1–62. doi: 10.1017/S0140525X16000960

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. doi: 10.1037/a0019671

Lewis, B. A., Long, R., & Mackay, M. (1993). Fostering communication in mathematics using children's literature. *The Arithmetic Teacher*, *40*, 470–473.

Liberatos, P., Link, B. G., & Kelsey, J. L. (1987). The measurement of social class in epidemiology. *Epidemiologic reviews*, *10*, 87–121.

Lowe, D. G. (1979). Strategies, context, and the mechanism of response inhibition. *Memory & Cognition*, *7*, 382–389. doi: 10.3758/BF03196943

Lukie, I. K., Skwarchuk, S.-L., LeFevre, J.-A., & Sowinski, C. (2013). The role of child interests and collaborative parentchild interactions in fostering numeracy and literacy development in canadian homes. *Early Childhood Education Journal*, *42*, 251–259. doi: 10.1007/s10643-013-0604-7

Lyons, I. M., & Ansari, D. (2009). The cerebral basis of mapping nonsymbolic numerical quantities onto abstract symbols: an fMRI training study. *Journal of Cognitive Neuroscience*, *21*, 1720–1735. doi: 10.1162/jocn.2009.21124

Lyons, I. M., & Beilock, S. L. (2009). Beyond quantity: individual differences in working memory and the ordinal understanding of numerical symbols. *Cognition*, *113*, 189–204. doi: 10.1016/j.cognition.2009.08.003

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, *109*, 163–203. doi: 10.1037/0033-2909.109.2.163

Manolitsis, G., Georgiou, G. K., & Tziraki, N. (2013). Examining the effects of home literacy and numeracy environment on early reading and math acquisition. *Early Childhood Research Quarterly*, *28*, 692–703. doi: 10.1016/j.ecresq.2013.05.004

McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? knowledge of operational patterns hinders learning and performance on equations. *Child Development*, *76*, 883–899. doi: 10.1111/j.1467-8624.2005.00884.x

McNeil, N. M., & Fyfe, E. R. (2012). Concreteness fading promotes transfer of mathematical knowledge. *Learning and Instruction*, *22*, 440–448. doi: 10.1016/j.learninstruc.2012.05.001

McNeil, N. M., & Uttal, D. H. (2009). Rethinking the use of concrete materials in learning: Perspectives from development and education. *Child Development Perspectives*, *3*, 137–139. doi: 10.1111/j.1750-8606.2009.00093.x

Melhuish, E. C., 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. doi: 10.1111/j.1540-4560.2008.00550.x

Melhuish, E. C., Sylva, K., Sammons, P., Siraj-Blatchford, I., Taggart, B., Phan, M. B., & Malin, A. (2008). Preschool influences on mathematics achievement. *Science*, *321*, 1161–1162. doi: 10.1126/science.1158808

Merkley, R., & Ansari, D. (2016). Why numerical symbols count in the development of mathematical skills: evidence from brain and behavior. *Current Opinion in Behavioral Sciences*, *10*, 14–20. doi: 10.1016/j.cobeha.2016.04.006

Merkley, R., & Scerif, G. (2015). Continuous visual properties of number influence the formation of novel symbolic representations. *The Quarterly Journal of Experimental Psychology*, *68*, 1860–1870. doi: 10.1080/17470218.2014.994538

Missall, Hojnoski, R. L., Caskie, G. I. L., & Repasky, P. (2014). Home numeracy environments of preschoolers: Examining relations among mathematical activities, parent mathematical beliefs, and early mathematical skills. *Early Education and Development*, *26*, 1–21. doi: 10.1080/10409289.2015.968243

Missall, Hojnoski, R. L., & Moreano, G. (2016). Parent-child mathematical interactions: examining self-report and direct observation. *Early*

Child Development and Care, *26*, 1–13. doi: 10.1080/03004430.2016.1193731

Montessori, M. (1917). *The advanced montessori method.* Michigan: Frederick A. Stokes Company.

Mourrat, J.-C. (2008). Testing the limits of "concrete" and "generic". *Science*, *322*, 1632a–1632a. doi: 10.1126/science.322.5908.1632a

Mudkiff, B., & Cramer, M. M. (1993). Stepping stones to mathematical understanding. *Arithmetic Teacher*, *40*, 303–5.

Mundy, E., & Gilmore, C. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, *103*, 490–502. doi: 10.1016/j.jecp.2009.02.003

Musun-Miller, L., & Blevins-Knabe, B. (1998). Adults' beliefs about children and mathematics: How important is it and how do children learn about it? *Early Development & Parenting*, *7*, 191–202. doi: 10.1002/(SICI)1099-0917(199812)7:4⟨191::AID-EDP181⟩3.0.CO;2-I

Mutaf, B., Sasanguie, D., De Smedt, B., & Reynvoet, B. (2016). *The association between home numeracy and childrens calculation skills is mediated by basic number processing.* Leuven, Belgium.

National Numeracy. (2015). *National numeracy.* Retrieved from `http://www.nationalnumeracy.org.uk`

Nesmith, S., & Cooper, S. (2010). Trade books in the mathematics classroom: The impact of many, varied perspectives on determinations of quality. *Journal of Research in Childhood Education*, *24*, 279–297. doi: 10.1080/02568543.2010.510086

Niklas, F., Cohrssen, C., & Tayler, C. (2016). Parents supporting learning: A non-intensive intervention supporting literacy and numeracy in the home learning environment. *International Journal of Early Years Education*, *24*, 121–142. doi: 10.1080/09669760.2016.1155147

Niklas, F., & Schneider, W. (2013a). Casting the die before the die is cast: the importance of the home numeracy environment for preschool children. *European Journal of Psychology of Education*, *29*, 327–345. doi: 10.1007/s10212-013-0201-6

Niklas, F., & Schneider, W. (2013b). Home Literacy Environment and the beginning of reading and spelling. *Contemporary Educational Psychology*, *38*, 40–50. doi: 10.1016/j.cedpsych.2012.10.001

Nuzzo, R. (2015). How scientists fool themselves and how they can stop.

Nature News, *526*, 182. doi: 10.1038/526182a

OpenScienceCollaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*. doi: 10.1126/science.aac4716

Pan, Y., Gauvain, M., Liu, Z., & Cheng, L. (2006). American and Chinese parental involvement in young children's mathematics learning. *Cognitive Development*, *21*, 17–35. doi: 10.1016/j.cogdev.2005.08.001

Parsons, S., & Bynner, J. (2005). Does numeracy matter more? *National Research and Development Centre for adult literacy and numeracy*. Retrieved from http://www.nrdc.org.uk

Peirce, J. W., & Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. doi: 10.3389/neuro.11.010.2008

Petrill, S. A., Pike, A., Price, T., & Plomin, R. (2004). Chaos in the home and socioeconomic status are associated with cognitive development in early childhood: Environmental mediators identified in a genetic design. *Intelligence*, *32*, 445 - 460. doi: http://dx.doi.org/10.1016/j.intell.2004.06.010

Piaget, J. (1971). The theory of stages in cognitive development. In *Measurement and Piaget.* New York, NY, US: McGraw-Hill.

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*, 542–551. doi: 10.1016/j.tics.2010.09.008

Plewis, I., Mooney, A., & Creeser, R. (1990). Time on educational activities at home and educational progress in infant school. *British Journal of Educational Psychology*, *60*, 330–337. doi: 10.1111/j.2044-8279.1990.tb00949.x

Powell, S. R., & Nurnberger-Haag, J. (2015). Everybody counts, but usually just to 10! A systematic analysis of number representations in childrens books. *Early Education and Development*, *26*, 377–398. doi: 10.1080/10409289.2015.994466

Ramani, G. B., & Eason, S. H. (2015). It all adds up learning early math through play and games. *Phi Delta Kappan*, *96*, 27–32. doi: 10.1177/0031721715583959

Ramani, G. B., Rowe, M. L., Eason, S. H., & Leech, K. A. (2015). Math talk during informal learning activities in Head Start families. *Cognitive Development*, *35*, 15–33. doi: 10.1016/j.cogdev.2014.11.002

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development*, *79*, 375–394. doi: 10.1111/j.1467-8624.2007.01131.x

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. doi: 10.1016/j.appdev.2011.02.005

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. doi: 10.1080/15248372.2012.664593

Reading Agency. (2013). *The reading agency.* Retrieved from `http://readingagency.org.uk`

Reynvoet, B., & Sasanguie, D. (2016). The symbol grounding problem revisited: A thorough evaluation of the ANS mapping account and the proposal of an alternative account based on symbol-symbol associations. *Frontiers in Psychology*, *7*, 1581. doi: 10.3389/fpsyg.2016.01581

Ribner, D., Andrew. (2014). *What matters most? Examining socioeconomic disparities and predictors of early math ability.* (Unpublished doctoral dissertation). Wesleyan University, Connecticut.

Rodriguez, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first 5 years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child Development*, *82*, 1058–1075. doi: 10.1111/j.1467-8624.2011.01614.x

Roth McDuffie, A. M., & Young, T. A. (2003). Promoting mathematical discourse through children's literature. *Teaching Children Mathematics; Reston*, *9*, 385–389.

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, *34*, 344–355. doi: 10.3758/BF03193412

Sarama, J., & Clements, D. H. (2009). Concrete computer manipulatives in mathematics education. *Child Development Perspectives*, *3*, 145–150. doi: 10.1111/j.1750-8606.2009.00095.x

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online*, *8*, 23–74.

Schiro, M. (1997). *Integrating children's literature and mathematics in the classroom: Children as meaning makers, problem solvers, and literary critics*. New York: Teachers College Press, Columbia University.

Schliemann, A. D., & Carraher, D. W. (2002). The evolution of mathematical reasoning: Everyday versus idealized understandings. *Developmental Review*, *22*, 242–266. doi: 10.1006/drev.2002.0547

Segers, E., Kleemans, T., & Verhoeven, L. (2015). Role of parent literacy and numeracy expectations and activities in predicting early numeracy skills. *Mathematical Thinking and Learning*, *17*, 219–236. doi: 10.1080/10986065.2015.1016819

Sénéchal, M., & LeFevre, J.-A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child development*, *73*, 445–460. doi: 10.1111/1467-8624.00417

Sénéchal, M., & LeFevre, J.-A. (2014). Continuity and change in the home literacy environment as predictors of growth in vocabulary and reading. *Child development*, *85*, 1552–1568. doi: 10.1111/cdev.12222

Shanahan, T., & Lonigan, C. J. (2010). The national early literacy panel: A summary of the process and the report. *Educational Researcher*, *39*, 279–285. doi: 10.3102/0013189X10369172

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*, *4*, 1–32. doi: 10.1146/annurev.clinpsy.3.022806.091415

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science*, *11*, 655–661. doi: 10.1111/j.1467-7687.2008.00714.x

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. doi: 10.1037/a0014239

Simonsohn, U. (2015). *The effect size does not exist* (Blog No. 33). `http://datacolada.org/33`.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534–547. doi: 10.1037/a0033242

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better P-curves: Making P-curve analysis more robust to errors, fraud, and ambitious P-hacking, a Reply to Ulrich and Miller (2015). *Journal of Experimental Psychology: General*, *144*, 1146–1152. doi: 10.1037/xge0000104

Simpson, A. (2017). The misdirection of public policy: comparing and combining standardised effect sizes. *Journal of Education Policy*, *32*, 450-466. doi: 10.1080/02680939.2017.1280183

Skwarchuk, S.-L. (2009). How do parents support preschoolers' numeracy learning experiences at home? *Early Childhood Education Journal*, *37*, 189–197. doi: 10.1007/s10643-009-0340-1

Skwarchuk, S.-L., Sowinski, C., & LeFevre, J.-A. (2014). Formal and informal home learning activities in relation to childrens early numeracy and literacy skills: The development of a home numeracy model. *Journal of Experimental Child Psychology*, *121*, 63–84. doi: 10.1016/j.jecp.2013.11.006

Sloutsky, V. M., Kaminski, J. A., & Heckler, A. F. (2005). The advantage of simple symbols for learning and transfer. *Psychonomic Bulletin & Review*, *12*, 508–513. doi: 10.3758/BF03193796

Smith, T. E., & Graham, P. B. (1995). Socioeconomic stratification in family research. *Journal of Marriage and the Family*, 930–940. doi: 10.2307/353413

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: Short-cutting abstraction in childrens object categorizations. *Cognition*, *108*, 626–638. doi: 10.1016/j.cognition.2008.05.002

Sonnenschein, S., Galindo, C., Metzger, S. R., Thompson, J. A., Huang, H. C., & Lewis, H. (2012). Parents' beliefs about children's math development and children's participation in math activities. *Child Development Research*, *2012*, e851657. doi: 10.1155/2012/851657

Sowinski, C., LeFevre, J.-A., Skwarchuk, S.-L., Kamawar, D., Bisanz, J., & Smith-Chant, B. (2015). Refining the quantitative pathway of the Pathways to Mathematics model. *Journal of Experimental Child Psychology*, *131*, 73–93. doi: 10.1016/j.jecp.2014.11.004

Swick, D. C. (2007). *The effects of parental involvement on children's school readiness skills.* (Doctoral dissertation, US). Retrieved from `http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2007-99230-157&site=ehost-live`

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004). *The effective provision of pre-school education (eppe) project: Findings from pre-school to end of key stage 1.* Department for Education and Skills.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B., Smees, R., ... others (2004). The effective provision of pre-school education (eppe) project. *Department for Education and Skills*.

Taggart, B., Sylva, K., Melhuish, E., Sammons, P., & Siraj, I. (2015). Effective pre-school, primary and secondary education project (eppse 3-16+): how pre-school influences children and young people's attainment and developmental outcomes over time.

Texttank. (2017). *Texttank.* Retrieved from `http://www.texttank.co.uk`

Thatcher, D. H. (2001). Reading in the math class: Selecting and using picture books for math investigations. *Young Children*, *56*, 20–26.

Trochim, K. W., M. (2006). *The research methods knowledge base.* Retrieved 2014-10-15, from `http://www.socialresearchmethods.net/kb/index.php`

Trust, L. (2017). *The national literacy trust.* Retrieved from `http://www.literacytrust.org.uk`

Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., & DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives*, *3*, 156–159. doi: 10.1111/j.1750-8606.2009.00097.x

Vandermaas-Peeler, M., Nelson, J., Bumpass, C., & Sassine, B. (2009). Numeracy-related exchanges in joint storybook reading and play. *International Journal of Early Years Education*, *17*, 67–84. doi: 10.1080/09669760802699910

Vandermaas-Peeler, M., & Pittard, C. (2014). Influences of social context on parent guidance and low-income preschoolers' independent and guided math performance. *Early Child Development and Care*, *184*, 500–521. doi: 10.1080/03004430.2013.799155

van der Meij, J., & de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and Instruction*, *16*. doi: 10.1016/j.learninstruc.2006.03.007

von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine and Child Neurology*, *49*, 868–873. doi: 10.1111/j.1469-8749.2007.00868.x

Wachs, T. D. (1996). Known and potential processes underlying developmental trajectories in childhood and adolescence. *Developmental Psychology*, *32*, 796. doi: 10.1186/1471-244X-13-303

Wade, B., & Moore, M. (2000). A sure start with books. *Early Years*, *20*, 39–46. doi: 10.1080/0957514000200205

Wagenmakers, E.-J., & Dutilh, G. (2016). *Why, what, where of preregistration.* Retrieved from `http://psyres.uva.nl/scientific-integrity/preregistration/preregistration.html`

Ward, J. M., Mazzocco, M. M., Bock, A. M., & Prokes, N. A. (2017). Are content and structural features of counting books aligned with research on numeracy development? *Early Childhood Research Quarterly*, *39*, 47–63. doi: 10.1016/j.ecresq.2016.10.002

Wilensky, U. (1991). Abstract meditations on the concrete and concrete implications for mathematics education. In I. Harel & S. Papert (Eds.), *Constructionism* (pp. 193–204). Evanston: Ablex.

Xenidou-Dervou, I., De Smedt, B., van der Schoot, M., & van Lieshout, E. C. D. M. (2013). Individual differences in kindergarten math achievement: The integrative roles of approximation skills and working memory. *Learning and Individual Differences*, *28*, 119–129. doi: 10.1016/j.lindif.2013.09.012

Xenidou-Dervou, I., Molenaar, D., Ansari, D., van der Schoot, M., & van Lieshout, E. C. D. M. (2017). Nonsymbolic and symbolic magnitude comparison skills as longitudinal predictors of mathematical achievement. *Learning and Instruction*, *50*, 1–13. doi: 10.1016/j.learninstruc.2016.11.001

Xenidou-Dervou, I., van Lieshout, E. C. D. M., & van der Schoot, M. (2014). Working memory in nonsymbolic approximate arithmetic processing: A dual-task study with preschoolers. *Cognitive Science*, *38*, 101–127. doi: 10.1111/cogs.12053

Young-Loveridge, J. M. (2004). Effects on early numeracy of a program using number books and games. *Early Childhood Research Quarterly*, 82–98. doi: 10.1016/j.ecresq.2004.01.001

Zhao, H., Chen, C., Zhang, H., Zhou, X., Mei, L., Chen, C., ... Dong, Q. (2012). Is order the defining feature of magnitude representation?

an ERP study on learning numerical magnitude and spatial order of artificial symbols. *PLOS ONE*, *7*, e49565. doi: 10.1371/journal.pone.0049565

Zippert, E. L., & Ramani, G. B. (2016). Parents' estimations of preschoolers' number skills relate to at-home number-related activity engagement. *Infant and Child Development*, *26*, e1968. doi: 10.1002/icd.1968

Appendices

Appendix A

P-Curve Disclosure Table

Original Paper	1) Quoted text from original paper indicating prediction of interest to researchers	2) Study design	(3) Key statistical result (looking up column 2 in table 3)	4) Quoted text from original paper with statistical results	(5) Results			(6) Robustness results
					Combined	Formal	Informal	
Anders et al., 2012	we examine the influence of different aspects of the home learning environment (quality of stimulation in numeracy and [pre]reading literacy) on development.	Latent Growth Curve		The results showed that the quality of the home learning environment already explained substantial variance in numeracy at the first assessment, when children were on average 3 years old. There was no significant effect of HLE on the slope, indicating that the early advantages of children with a high-quality HLE were maintained over the next two years.	Latent Growth Curve Statistic – contacted for correlation but no response			
Blevins-Knabe & Musun-Miller, 1996	Study 2 was designed to examine the relationship between reported number activities in the home and children's performance in mathematics	Correlational	Pearson correlation	There were no significant correlations between the TEMA-2 score and the mean activities of the activities parents reported they engaged in with their children.	$r(47) = .09$ (reported via email)			
Blevins-Knabe, Austin, Musun, Eddy, & Jones, 2000	We addressed the question of whether the beliefs and reported activities of parents and providers would predict children's	Correlational	Pearson correlation	The frequency of math activities reported by parents or family day care providers was not significantly correlated with children's mathematics achievement scores in either age group.	*r* not reported – contacted but data not available			

	performance on the Test of Early Mathematics Ability-2							
Ciping, Silinskas, Wei, & Georgiou, 2015	to examine the cross-lagged relationships between home literacy/numeracy activities and children's reading/mathematics ability in Chinese	Correlational	Pearson correlation	Table 3 presents the correlations between the measures used in the study. […] mathematics ability in Grade 1 correlated negatively with formal numeracy activities in Grades 1.		*r*(175) = -.18	*r*(175) = .05	
Dearing et al., 2012	We predict that numerical and spatial activities will be the most proximally related predictors of math and spatial achievement at school entry for girls.	Correlational	Pearson correlation	girls' home learning environments and, in turn, their engagement in math-related activities were proximal predictors of arithmetic performance (r = .19 and .29, respectively addition and subtraction)	*r (125)* = .19			*r (125)* = .29
L. DeFlorio, 2011	In this correlational study, I employed a mixed methods approach to examine the ways in which aspects of the home environment may contribute to the SES gap present in early mathematics […] A subset of families	Correlational	Pearson correlation	There were no significant correlations between the general characteristics of the activities demonstrated, including activity type, length, minutes including math, total number of math occurrences, and mean number of math occurrences per minute and children's CMA scores.	*r* not reported – contacted for r value but no response			

	were observed in the home as well. The purpose of the home observations was to capture potential qualitative differences by age and SES in children's home activities perhaps not captured by the parent questionnaire.							
L. DeFlorio & Beliakoff, 2014	Is there a correlational relationship between the frequency, range, or type of activities children do in the home and their early mathematical knowledge?	Correlational	Pearson correlation	significant correlations with the CMA included the activity frequency score (r. .17, p. .027)	*r* (176) = .17			
Esplin et al., 2016	What is the relationship for children in both child care types between the stimulation parents provide via the home numeracy environment and the child's number line performance?	Correlational	Pearson correlation	Number line scores were not significantly related to any of the HNAS factors. Table 7.1 shows the correlation between working with numbers factor and TEMA and playing with numbers factor and TEMA.		*r (87)* = .39	*r* (87) = .23	
Huntsinger,	Explore the links	Correlational	Pearson	See Table 7 for correlations among		*r (*198)	*r (*198*)* =	Formal T2 – *r*

Jose, & Luo, 2016	between reports of parent-provided mathematics and literacy experiences and children's performance on tests of early mathematics and reading concurrently and longitudinally.		correlation	key study variables.		= .40	.00	(95) = .21 Informal T2 - *r* (95) = -.14
Kleemans, Peeters, Segers, & Verhoeven, 2012	We expected home numeracy experiences to be related to early numeracy skills	Correlational	Pearson correlation	As can be seen in Table 3, all child and home measures correlated significantly with early numeracy skills.	*r* (87) = .47			
Kleemans, Segers, & Verhoeven, 2013	To what extent can the variation in basic calculation skills of children with SLI and children with NLA in grade one be related to their home numeracy experiences, measured at kindergarten, while taking into account their cognitive and linguistic capacities?	Correlational	Pearson correlation	The correlations are presented in Table 3.	*r* (148) = .635			*r* (148) = .475 (subtraction)
J.-A. LeFevre et al., 2009	In the present research we sought to explore not only activities related to specific number skills	Correlational	Pearson correlation	For math knowledge, only the frequency of participation in games correlated with performance. However, for math fluency, three of the four activity factors correlated		*r* (144) = -.06	*r* (144) = .27	Math fluency Formal – *r* (135) = .07 Informal - *r* (135) = .26

	but also to assess the frequency of a variety of situations that might involve children in quantitative activities, but where the focus is not necessarily on direct learning of number skills.			with performance.				
J. LeFevre, Polyzoi, Skwarchuk, Fast, & Sowinski, 2010	expect a relation between home experiences and children's numeracy performance	Correlational	Pearson correlation	the frequency of direct numeracy activities was correlated with numeracy scores, r(98) = .38 (Greek) and r(102) = .37 (Canadian), ps < .05, whereas the frequency of indirect activities and speeded activities were not significantly correlated with numeracy in either country, all rs < .07, ps > .50.		*r* (98) = .38 (reported via email)	*r* (98) = .023 (reported via email)	Canadian *r* (102) = .374 (formal) *r* (102) = .052 (informal)
Levine, Suriyakham, Rowe, Huttenlocher, & Gunderson, 2010	We examined variation in parent talk about numbers during naturalistic interactions with their 14- to 30-month-olds and the relation of this variation to children's subsequent	Correlational	Pearson correlation	Children's point-to-x task performance was positively related to parent cumulative number talk (r = .47, p < .01).	*r* (42) = .47			

	numerical understanding.							
Manolitsis, Georgiou, & Tziraki, 2013	Formal numeracy activities will predict math fluency in grade 1 through the effects of early math concepts and verbal counting.	Correlational	Pearson correlation	Table 3 presents the zero-order correlations among all the variables involved in the present study. Both parent literacy and numeracy teaching correlated with counting at the beginning of kindergarten.		*r* (80) = -.02 (Grade 1)		*r* (80) = .09 (Beginning Kindergarten) *r* (80) = .14 (End Kindergarten)
Ramani, Rowe, Eason, & Leech, 2015	The third goal was to understand how the two different types of number experiences related to children's current numerical knowledge	Correlational	Pearson correlation	Caregiver report of frequency of engaging in activities that directly teach children about math was positively associated with children's performance on both the foundational and advanced number measures composites, r = 55, p < .001 and r = 46, p < .01, respectively. […] Caregiver report of engaging in board games, card games, and computer games was correlated with children's performance on the foundational number skills composite, r = 35, p < .05, but not the advanced number skills composite.		*r*(31) = .46	*r*(31) = .24	Formal *r*(31) = .55 Informal *r*(31) = .35
Missall,	To what extent does	Correlational	Pearson	Table 4 presents Pearson product–	*r*(70)=-			

Hojnoski, Caskie, & Repasky, 2014	the home numeracy environment (i.e., activities and beliefs) predict children's performance on measures of early mathematics?		correlation	moment correlations between math activities, math beliefs, and early math performance. […]However, only low and non–statistically significant correlations were found between the three EMQ subscales and the measures of early math and early school readiness.	.03			
Niklas & Schneider, 2013	A second assumption was that HNE should be associated with subsequent mathematical competencies. More specifically, children who experienced a more favourable HNE should achieve greater mathematical performance.	Correlational	Pearson correlation	Table 2 presents the results of the correlational analyses of HNE with the control variables and the mathematical measures as well as the descriptive statistics. […] We found small but significant correlations between HNE and the mathematical competencies from t2 onwards (HNE was obtained at t2).	*r*(491) = .09			
Niklas, Cohrssen, & Tayler, 2016	Are the home literacy and numeracy environments associated with child and family characteristics and children's cognitive outcomes?	Correlational	Pearson correlation	Table 3 provides an over- view of the correlation between all variables. […] the home numeracy environment showed close interrelations with the mathematical tasks.	*r*(111) = .23			

Segers, Kleemans, & Verhoeven, 2015	In the present study, we therefore examined to what extent the home numeracy environment adds to the prediction in children's early numeracy skills, after controlling for cognitive and linguistic child factors and aspects of the home literacy environment	Correlational	Pearson correlation	An exploration of Table 4 shows relationships between child and home factors on the one hand and early numeracy on the other hand. Correlations between early numeracy and cognitive and linguistic child factors were all medium to high. This was also the case for the relationship of early numeracy with the home numeracy environment	*r*(55) = .41			
Skwarchuk, 2009	The purpose of this project was to examine the opportunities that exist for preschoolers to explore numeracy concepts at home, and the extent to which numeracy exposure predicts mathematical achievement.	Correlational	Pearson correlation	Achievement scores correlated with […] complex numeracy (r = .520, p < .05)		*r*(23) = .52	*r* not significant, not reported – contacted for r statistic but data not available	
Skwarchuk, Sowinski, & LeFevre, 2014	We hypothesized that formal numeracy practices would uniquely predict children's	Correlational	Pearson correlation	Table 3 contains the correlations, raw means, and standard deviations among all variables and composite scores. […] Informal numeracy practices were correlated with non-		*r*(119) = .27	*r*(119) = .12	Non-symbolic arithmetic *r*(118) = .14 (formal) *r*(118) = .30

	knowledge of the number system but not their non-symbolic arithmetic performance.			symbolic arithmetic but not symbolic number knowledge. […] advanced formal numeracy practices were correlated with symbolic number knowledge				
Swick, 2007	Are there specific sub- groups of parental practices that predict specific academic skills and social competence?	Correlational	Pearson correlation	Tables 3 and 4 present the descriptive statistics of and correlations among predictors and dependent variables.	*r*(177) = .03			
Vandermaas-Peeler & Pittard, 2014	Parental reports of literacy and numeracy practices at home were expected to be positively correlated with children's mathematics achievement	Correlational	Pearson correlation	The home numeracy practices composite score was significantly and positively correlated with the TEMA-3 percentile rank, r(14) = .57, p < .05	*r*(14) = .57			
Zippert & Ramani, 2016	to examine the types of number-related experiences that parents share with their preschoolers at home, and how their engagement is associated with children's number	Correlational	Pearson correlation	We then correlated children's performance on the number measures and their at-home number-related activity engagement As shown in Table 2, the only significant correlation found was between children's advanced numeracy skills and their engagement in advanced at-home		*r*(40) = .41	*r*(40) = .12	Conventional *r*(40) = .22 (formal) *r*(40) = .25 (informal) Non-symbolic *r*(40) = .10 (formal) *r*(40) = -.10