Research Report DCSF-RR118

Development of Maths Capabilities and Confidence in Primary School

Terezinha Nunes, Peter Bryant, Kathy Sylva and Rossana Barros
Department of Education,
University of Oxford

In collaboration with ALSPAC, University of Bristol

department for
children, schools and families

Research Report No
DCSF-RR118

Development of Maths Capabilities and Confidence in Primary School

Terezinha Nunes, Peter Bryant, Kathy Sylva and Rossana Barros

Department of Education, University of Oxford

In collaboration with ALSPAC, University of Bristol

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Children, Schools and Families.

© University of Oxford 2009

ISBN 978 1 84775 457 8

June 2009

Contents

Tables and Figures

Acknowledgements

We are very grateful to the DCSF team of Stephen Witt, Deborah Wilson and Ali Bushell for the careful and sensitive support and also for the excellent advice that they have given us throughout this project, and to the DCSF for the financial support that made the project possible. We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council, the Wellcome Trust and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors who will serve as guarantors for the contents of this paper. This research was specifically funded by the DCSF. We would also like to thank Kerry Humphries and Jon Heron for providing all the links that we needed with the ALSPAC treasure-trove and for the help that they have given us directly or indirectly. We would also like to thank Dr Kathryn Duckworth from the Institute of Education, University of London for her contribution in developing the original research questions. We are also grateful to Lars Malberg for his statistical advice

Executive Summary

Introduction

The project looked at the development of competence in different aspects of maths and the effect of this on young people's key stage results. It used data from a large longitudinal survey of young people (ALSPAC - the Avon Longitudinal Survey of Parents and Children) which has unique information on levels and patterns of understanding in mathematics at different stages of young people's progress through primary and early secondary school.

Key Findings

- Mathematical reasoning, even more so than children's knowledge of arithmetic, is important for children's later achievement in mathematics

    - Mathematical reasoning and knowledge of arithmetic (as assessed in year 4) make independent contributions to children's achievement in mathematics in KS2 and 3. While both are important, mathematical reasoning is more important than knowledge of arithmetic for achievement in KS2 and 3.

- Spatial skills are important for later attainment in mathematics, but not as important as mathematical reasoning or arithmetic

    - The influence of spatial skills (being able to rotate and manipulate shapes and imagine the results of these actions) play a role in achievement that is independent both of mathematical reasoning and arithmetic. This role increases from KS2 to KS3 as the teaching of geometry gains in importance and spatial competence is being tested.

- Children's attention and memory also plays a small but consistent part in their mathematical achievement

    - The ability to focus and maintain attention and to remember relevant information are important factors in mathematics achievement, even after controlling for individual differences in arithmetic and mathematical reasoning.

- Children from high socio-economic status backgrounds are generally better at mathematical reasoning than their peers

- Streaming, or ability grouping, in Primary school improves the mathematical reasoning of children in the top ability group, but the effect is small. It hinders the progress of children in the other groups.

- Children's self confidence in maths is predicted most strongly by their own competence, but also by gender (girls are less confident than boys) and by the ability group in which the child is placed. Children's attainment, although largely determined by cognitive and social factors, is also influenced by their self-confidence.

Background

Quite striking differences in mathematical achievement exist among children in the UK and in other countries. We need to be particularly concerned about pupils who do not succeed in mathematics at school and about the general level of mathematical skills in the population and the UK's ability to compete in the world market with other technologically advanced societies.

The more we can find out about differences in mathematical achievement and their causes, the more likely it is that we will also find ways of helping those who have difficulties in learning about mathematics.

There are several possible reasons for differences among children in mathematics achievement. Some are cognitive: it is quite possible that differences in some underlying cognitive skill, such as the ability to reason about mathematical relations, or to calculate, or to pay sustained attention in the classroom and remember information, might determine how well and how badly children do in mathematics at school. Other possible reasons are social, such as the child's own social background or the social composition of the school that the child goes to, or the influence of their experience at home or school.

Still other reasons for differences in mathematics learning could be categorised as affective: among these are the children's self-confidence in mathematics and in their ability to learn in the future. Each of these ideas is plausible enough, and yet up to now very little progress has been made in determining why children differ so widely in mathematics.

Methodology

There were four main research questions:

1. Are there differences in the development of specific types of mathematical understanding and if so, how do these influence pupils' progression across the Key Stage 1-3 period?

2. Is the development of some mathematical skills more important for progression than others?

3. Are there differences in the development of skills by gender, ethnicity, or socio-economic factors, such as parents' education or SES?

4. How is development of understanding affected by children's self perception of their own ability in maths?

To answer these questions we used the large and impressive data bank collected in the ALSPAC study. This contains more than 14000 children born in Avon in the West of England in 1991- 92 and it has followed them over the whole of their childhood. The data include plentiful information on possible underlying cognitive differences, on the wide range of the children's social backgrounds and on relevant affective variables, such as their self confidence in their own abilities.

The number of children whose results we analysed depended on how many of them had been given the various cognitive, social and emotional measures that we included in each of the analyses. These numbers were always large: they varied from roughly 800 to roughly 4000 children in the different analyses.

Very nearly all our analyses were about how well a particular predictor was related to an outcome measure that was given some time later. Our main outcome measures were the mathematics assessments at Key Stage 2, when the children were 11, and at Key Stage 3, when they were just 14. The predictor variables were cognitive, social and emotional measures, many of which were administered three years before the Key Stage 2 and six years before the Key Stage 3 measures.

The most successful of the cognitive predictors were two mathematical reasoning tasks, one of which was given when the children were 8 and the other when they were 10 and again at 12. All the other cognitive predictors were sub-tests of the WISC (Wechsler Intelligence Scale for Children). Our social measures were of the children's self-perception and their parents' SES: we used measures of the mother's and the father's social economic status, as defined by their occupations, and of the mother's educational level.

Our self-perception and self-confidence scores were based on a measure of how much the children liked maths in Year 3 and their self-perception as learners of maths in Year 4. These two measures were highly correlated (r=.83) even though they were given to the children in different school years. We integrated the information from these two measures into a single factor and treated this as a measure of a single concept, children's self-confidence in mathematics.

We analysed the relations between the predictor and the outcome variables in three main ways:

1. Through multiple regressions in which we usually looked at the effect of different variables on the outcome measure. These regressions provide a measure of the contribution that each predictor variable makes to the outcome measure independent of all the other predictor variables.

2. Through Structural Equation Models (SEMs) in which we measured the strength of the pathways between the predictors and the outcome measure and considered whether these pathways went directly from the predictor to the outcome measure or indirectly through another predictor variable.

3. Through multi-level modelling in which we looked at the level at which social variables operated. One was at the level of the individual, and the other at the level of the school. Thus, our aim was to see how much of the difference in children's mathematical abilities was due to differences in the children's own SES status and how much to the SES composition of the pupils in the schools that they go to.

Findings and Implications

Mathematical reasoning, even more so than children's knowledge of arithmetic, is important for children's later achievement in mathematics

Arithmetic here is defined as *'learning how to do sums and using this knowledge to solve problems*', mathematical reasoning is 'learning to reason about the underlying relations in mathematical problems they have to solve' and can include both additive and multiplicative reasoning.

Finding: Children's ability to reason about mathematical relations was easily the most powerful predictor of children's mathematical achievement, out of all the relevant cognitive measures in the ALSPAC data bank. It strongly predicted their mathematics achievement in Key Stage 2 and 3 assessments even after controls for the effects of differences in intelligence and calculation ability.

Implication: It is important to promote reasoning about mathematical relations in primary school. Children at risk for difficulties in mathematics can be identified early through mathematics reasoning assessments and receive early intervention, which would translate into better mathematics achievement later on.

Finding: The contribution of calculation skills to mathematics achievement was modest but it was independent of reasoning.

Implication: In the context of time pressures, more time should be allocated to developing children's reasoning than to practising calculation skills.

Finding: Mathematical reasoning scores were strongly related to the Key Stage mathematics assessments, less strongly but quite well related to the science assessments, and only weakly related to the English assessments. The relation between mathematical reasoning and the science assessments was stronger with the Key Stage 3 than with the Key Stage 2 assessments.

Implication: The teaching of both science and mathematics could benefit from an analysis of the forms of mathematical reasoning used in science and from the close coordination of the teaching mathematics and science at school. This coordination should be particularly beneficial in the education of older primary school children.

Spatial skills are important for later attainment in mathematics, but not as important as mathematical reasoning or arithmetic

Finding: Children's spatial skills also predicted their Key Stage mathematics results, but to a lesser extent. Spatial skills were more important for Key Stage 3 than Key Stage 2 achievement. Children's achievement in the mathematics reasoning questions about space was very modest.

Implication: Greater attention to children's spatial reasoning in mathematics lessons should improve children's achievements in geometry.

Children's attention and memory also plays a small but consistent part in their mathematical achievement

Finding: Memory and attention also made a modest, but independent, contribution to children's mathematical achievement in Key Stage 2 and 3. Some children may be at risk for difficulties in learning mathematics due to low performance in these cognitive factors.

Implication: Recent developments in research that show that it is possible to improve children's attention and memory through training. This should be considered as part of a personalised programme for children at risk.

Children from high SES backgrounds are generally better at mathematical reasoning than their peers

Finding: Individual children whose SES status is high are on the whole better at mathematical reasoning than those from lower SES homes, even if they are at the same school. The influence of SES at the individual level is similar to an effective educational intervention: it raises the average level of performance and reduces the variation between children, and consequently the proportion of children who show difficulty in mathematics.

SES also operates at the level of the school. Children, who go to Primary schools in which the SES composition is high, reason about mathematics more successfully than children at schools with a lower SES composition, whatever their own social background.

Implication: The identification of a mechanism that mediates the connection between SES and mathematics achievements should help us to explain the effects of social background on mathematics achievement. In addition, improving children's mathematical reasoning through instruction in school can make an important contribution to reducing SES inequalities in mathematics achievement.

Streaming, or ability grouping, in primary school improves the mathematical reasoning of children in the top ability group, but the effect is small. It hinders the progress of children in the other groups.

When young people were in Year 3, teachers were asked in which stream or ability group the target children were placed for maths, but this did not distinguish between putting children into different classes according to ability (streaming), or grouping pupils by ability within classes (ability grouping).

Finding: Streaming only improves mathematical achievement of children in the top ability group, but this effect is small. It actually hinders the progress of children in the other groups. It also influences children's self-confidence in maths: children of the same level of ability placed in a lower ability group are less confident in their ability than those placed in a higher group.

Implication: Schools need to consider alternatives to grouping children by ability levels if they want to promote higher achievement for children with all levels of ability.

Children's self confidence in maths is predicted most strongly by their own competence, but also by gender and by the ability group in which the child is placed.

Finding: Children's self-confidence in mathematics is predicted most strongly by their own mathematical competence but also, independently by their gender (girls are less confident than boys) and by ability group. Children's attainment in Key Stage 2 and 3 mathematics, although largely determined by cognitive and social factors, is also influenced by their self-confidence.

Implication: It is important to pay attention to the affective aspects of children's mathematics learning as well as to differences in their cognitive abilities and their social background.

1. Introduction

The least controversial proposition that one could make about children's success in learning mathematics is that it varies a great deal from child to child. Some children learn mathematics quickly and well; others make only slow progress and abandon the subject as soon as they are allowed to do so. These differences operate at different levels. They occur at the individual level, since even children at the same schools and from much the same social backgrounds vary a great deal from each other in their mathematical development (Dowker, 2005), but there are also striking and consistent differences between children in different countries (TIMS: PISA : Stevenson *et al.*, 1985, 1986; Stigler *et al.*, 1987), from different social backgrounds (Reyes & Stanic, 1988; Sacker, Schoon & Bartley, 2002) and, as we shall show later in this report, in different schools. Children's success in mathematics, therefore, depends not just on their own abilities and motivations but also on their family, their school and their nationality.

There are two good reasons to study why these differences occur. The more obvious of the two is that they are a serious cause for concern. Adults with low mathematics skills are more likely than those with low literacy skills to be unemployed; among those employed, low numeracy skills are a stronger predictor than low literacy skills of low levels of employment (Bynner & Parson, 2000). So, those who do poorly in the subject and have learned very little about it by the time that they leave school are likely to suffer at work as a result, especially in a technological society in which mathematical skills are increasingly needed (Noss & Hoyles, 1992, 1996). For the same reason, society at large may suffer if there is a shortage of mathematical skills in the community. The generally declining number of pupils who pursue mathematics after the age of 16 in the UK suggests that this is a particular danger for this country. We urgently need to find ways of removing the difficulties that so many children encounter in mathematics lessons and thus to decrease the spread and improve the general level of mathematical abilities among schoolchildren.

The second reason, a related one, for paying close attention to differences in mathematical success is as important as the first. It is that research on this issue could give us information that we badly need about the basis for children's success or lack of it in mathematics. By studying the factors that are related to how well children do in mathematics, we should be able to identify the kind of experiences that children should have, and the kind of teaching that they should be given, to help them most effectively. We can consider, for example, what underlying skills determine how well children do in mathematics: how important is it for children to be able to take in and remember mathematical information well, to calculate efficiently and to reason about mathematical relations logically? We can study, too, the social and emotional bases of children's mathematical success, such as their liking of mathematics and their self-confidence in their own mathematical ability and also the effect on their mathematical learning of their family background in socio-economic terms and of the school that they happen to be in. Convincing answers to any of these questions should make it easier to devise a better system for teaching mathematics than we have already.

Many researchers have recognised the value of studying the causes of differences between children for this purpose, but the questions that they set out to answer are still open ones. The reason for this general lack of progress probably lies in the nature of the research that has been done so far. Much of it has dealt with children at one time only and this has made it hard to disentangle cause from effect among the many different variables that have been considered. Also, many of the studies have been of relatively small numbers of children in a restricted set of environments, with the inevitable consequence that we cannot be sure how representative their results are or even that they are relevant to the mathematical learning of other children brought up in different kinds of homes and educated in different kinds of schools. It has also been quite impossible in these small-scale studies to distinguish effects at the individual level from the effects at the social level, the effects for example of the child's own social background from the effects of the social composition of the school that he or she is attending.

The Avon Longitudinal Study of Parents and Children (ALSPAC) offers an almost unique opportunity to surmount these research obstacles. This impressive project has accumulated a wealth of data on a representative group of slightly more than 14,000 children. The large body of data in the project includes information about the children's educational progress, including their progress in mathematics, and their performance in several relevant psychological tests such as a well-known intelligence test and a test of mathematical reasoning that was given to a sub-sample of the children when they were 8-, 10- and 12-years old. The size of the ALSPAC sample and the longitudinal nature of the data about them remove the difficulties that have hindered the progress of so much previous research on differences in children's mathematical learning.

In this report we shall describe the results of a set of analyses in which we were able to take advantage of the welcome opportunities offered by the ALSPAC data bank. All the analyses are about the ALSPAC children's mathematics learning and the underlying variables that predict how successful they have been.

2. Background

2.1 Cognitive issues

Two kinds of cognitive issue will be analysed in this investigation: issues that are specific to mathematics learning, and issues that are related to children's cognitive processes and skills in general and are not specifically mathematical.

2.1.1 Learning mathematics: the nature of the task

Learning mathematics in primary school involves learning about numbers, quantities and relations, and about the connections and distinctions between these three concepts.

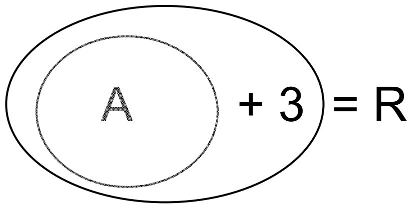
Numbers and quantities are not the same. A quantity can be represented by a number but we do not always need to measure a quantity and represent it by a number: we can compare one person's height to another's, for example, without any measurements or numbers.

Thompson (1993) provided a framework which we shall use throughout this report when he distinguished between numbers, quantities and relations. According to him "Quantitative reasoning is the analysis of a situation into a quantitative structure - a network of quantities and quantitative relationships. A prominent characteristic of reasoning quantitatively is that numbers and numeric relationships are of secondary importance, and do not enter into the primary analysis of a situation. What is important is relationships among quantities" (Thompson, 1993, p. 165). Elsewhere, Thompson (1994) emphasised that "a quantitative operation is nonnumerical; it has to do with the *comprehension* (italics in the original) of a situation" (p. 187-188).

The differences and links between quantities and relations are best understood if we consider an example. Figure 1 presents two problems and identifies the quantities and relations in each one. Both problems describe a quantity, the total number of books that Rob and Anne have, and the relation between two quantities, the number of books that Rob has and the number of books that Anne has. The relations between the quantities in Problem 1 are described in terms of a part-whole structure, as illustrated in the diagram. Part-whole relations are additive relations. The relations between the quantities in Problem 2 are described in terms of one-to-many correspondence, as illustrated in the diagram; these are multiplicative relations.

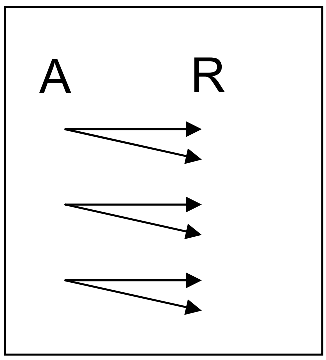
Figure 1 - Two problems illustrating two types of relations between quantities

(1)
Together Rob and Anne have 15 books (quantity)
Rob has 3 more books than Anne (or Anne has 3 books fewer than Rob) (relation)
How many books does each one have? (quantity)



(2)

Together Rob and Anne have 15 books (quantity)
Rob has twice the number of books that Anne has (or Anne has half the number of books that Rob has) (relation)
How many books does each one have? (quantity)



One characteristic of relations, which distinguishes them from quantities, is that they have a converse: if A is greater than B, then B is smaller than A. So understanding relations involves understanding the connections between these two ways of thinking about the same relation.

In order to learn mathematics, children must be able to coordinate their understanding of quantities with their understanding of relations, and must also distinguish between them. We can think about relations between quantities and represent the relation by a number even if we do not know what the quantities are. For example, if Paul plays a game of marbles and wins 7, then plays a second game and loses 1, and then plays a third game and loses 2, we know that at the end of the three games he has 4 more marbles than before. However, we do know not know how many marbles he has nor how many he started with.

Numbers are signs, spoken or written. Each number sign is part of a network of signs (i.e. all other numbers), the meanings of which are interconnected. Children can be said to truly understand the meaning of numbers only when they understand, for example, that all sets with the same number of objects are equivalent and that if two sets are equivalent they necessarily have the same number of objects (Piaget, 1952). They should also understand, for example, that 4 is one less than 5 and one more than 3, and that you only change the number in a set if you add or subtract elements. These meanings are part of the network of meanings in a number system.

Numbers can be used to represent both quantities and relations. Quantities and relations are meanings for numerical signs and children must learn how to interpret numbers by connecting them to quantities and relations and must also learn how to represent quantities and relations by numbers.

2.1.2 Theories about children's mathematics learning

Current hypotheses about the cognitive processes involved in children's mathematical learning fall into two broad camps. One stresses the importance of numbers and procedures related to determining the number in a set, such as counting, remembering number facts and calculating. The other emphasises the significance of children being able to understand quantities and relations and being able to reason logically and imaginatively about them.

The first type of theory, which focuses on numbers without considering quantities and relations, is best exemplified at pre-school age by the ideas of Gelman and Gallistel (1978). They argued that children are born in possession of innate, psychological mechanisms that eventually make it possible for them to learn, and to understand, the counting system with very little difficulty. With the help of these innate mechanisms children, who are learning to count, immediately apply a set of essential principles, such as the principle of cardinality, which Gelman and Gallistel define as the knowledge that the number of items in a set that they have counted corresponds to the last number that they produced in the count.

Other researchers (Dehaene, 1997; Butterworth, 2003; Durand, Hulme, Larkin, & Snowling, 20005) agree on the central importance of pre-school children's knowledge of the counting system but, when considering school-age children, have turned to children's success in comparing the magnitude of different numbers as a crucial index of mathematical ability, independent of or in coordination with knowledge of counting (Gelman & Butterworth, 2005). Still others (Geary & Brown, 1991) have argued that the ability to carry out simple calculations and to remember number facts lies at the heart of children's mathematical successes and failures. Children first learn how to use numbers and carry out calculations and later, through their use, come to understand why the procedures work (i.e. the concepts on which they are based).

All these theories agree in stressing the importance of the numerical procedures that children have to conquer at the start of their mathematical career at school without any reference to their understanding of quantities or relations. However, there is evidence that children can use these procedures without understanding their connection with quantities, and this shows that learning about numbers by themselves is not sufficient for learning mathematics in school. Fuson (1988), for example, showed that 3-years old children who satisfied Gelman's criterion for using the cardinality principle continue to use the last number word in the counting sequence to say how many items are in a set even if the counting started from two, rather than from one. Freeman, Antonuccia and Lewis (2000) also assessed 3- and 5-year-olds' rejection of the last word after counting if there had been a mistake in counting. The children watched a puppet count an array with either 3 or 5 items, but the puppet miscounted, either by counting an item twice or by skipping an item. The children were asked whether the puppet had counted right, and if they said that the puppet had not, they were asked to evaluate the puppet's knowledge of the cardinal for the set. Although all children had shown that they could count 5 items accurately, only about one third of the children were able to say that the answer was not right *after* they had detected the error. These results suggest that at least some children learn the counting words and a procedure for answering the question "how many?" relatively independently of their understanding of the connection between counting and the quantity represented. After observing a puppet counting wrongly, about two thirds still believed that the cardinal for the set is the last word used in the counting string.

There is also evidence that later, at school age, children continue to learn how to carry out operations with numbers - i.e. they learn number facts and computation - but may not know when to use them. In Thompson's terms, they may not understand the relations between quantities in problem situations well enough to know how to use numerical operations to solve a problem. Teachers are well aware of this discrepancy between knowing how and when to do sums, and there is research that supports their practical knowledge: two problems that require the same computation may have radically different levels of difficulty because they differ in the demands that they make of children's understanding of relations between quantities (e.g. Brown, 1981; Bryant, 1985; Carpenter & Moser, 1982; Thompson, 1993).

The fact that children often know about numerical symbols and procedures and yet fail to use them to solve problems about quantities and relations can be interpreted in different ways. One possibility is that children's counting and calculation skills develop independently of their understanding of quantities and relations between quantities. According to this hypothesis, this number knowledge should be a predictor of children's mathematics achievement independently of their ability to reason about quantities and relations. The latter should play no role in predicting mathematics achievement. An alternative hypothesis has been proposed by some researchers (e.g. Baroody & Gannon, 1984; Siegler & Crowley, 1991) who acknowledge that children often learn numerical procedures without understanding their conceptual basis. They suggest that, by learning procedures and practising them, children eventually understand the principles behind them.

These distinct possibilities are worth testing in a longitudinal study because there is so far no evidence for the idea that children's knowledge of number facts and calculations forms the basis for the development of their mathematical reasoning. A prediction from it about school children is that their knowledge of number signs and computation at an early age predicts their later understanding of quantities and relations and also their mathematics achievement. The present study will provide a test of this hypothesis.

The second type of theory focuses instead on the importance of children's reasoning about quantities and relations for mathematical learning, and assigns numerical procedures only a secondary and sometimes a very minor role. Piaget's theory, of course, is the best known example (Piaget, 1952; Piaget & Inhelder, 1974; 1975) but many other researchers have since stressed the significance of understanding quantities and relations for children's mathematics learning (e.g. Brown, 1981; Carpenter & Moser, 1983; Clements, Copple & Hyson, 2005 ; Ginsburg, Klein, & Starkey, 1998; Thompson, 1994; Vergnaud, 1979, to name only a few). This theoretical approach is at odds with the view that what matters is children's knowledge of numerical procedures on every important point.

Piaget's central idea was that children's understanding of what number really means depends crucially on their understanding of quantities and relations and on their ability to reason about these logically. This idea was not the reason why he was so heavily criticised in the 1970s and 80s. The aspect of his theory that caused all that controversy was his claim that this kind of reasoning is for the most part quite impossible for young pre-school children. There can be little doubt now that Piaget did to some extent underestimate young children's reasoning abilities, but on the whole the controversy which surrounded Piaget's work then, and which was the main reason for his relative fall from grace, left his central claim about the crucial importance of reasoning about quantities and relations for learning mathematics quite intact.

A report by Stern (2005) on some of the results of the impressive Munich longitudinal study of development did provide some interesting support for the importance of reasoning about relations and quantities for later mathematics learning. She showed that a measure of children's understanding of the inverse relation between addition and subtraction when they were eight years old was a good predictor of their performance in an algebra test many years later in university, even after allowing for differences in the participants' intelligence. In fact, the measure of the children's understanding of the inverse relation between addition and subtraction correlated as highly with their performance in algebra as the measure of the children's general intelligence that had been obtained at the same time.

Another longitudinal study, this time in the U.K., shows that differences between children in their understanding of quantities and relations in a measure obtained when they start school predict their mathematics achievement later (Nunes *et al*., 2007). The interval between the predictor (a measure of children's understanding of relations) and the outcome (KS1 Maths) was certainly more modest than that in the Munich study because the children's mathematics achievement was measured through their Key Stage 1 results, about a year and a half after the measure of mathematical reasoning had been administered. Nevertheless this is quite an important result because the study was longitudinal and appropriate controls were used in the analysis: the contribution of children's reasoning to the prediction of mathematics achievement was still considerable even after controlling for the children's specific knowledge of number and for general cognitive skills. The study was carried out with a small sample (59 children); evidence from a larger sample is urgently needed before one could draw implications for practice with great confidence. The present study will test this hypothesis longitudinally with a very large sample and over much longer intervals between the assessment of the children's reasoning and the measures of their mathematics achievement.

In summary, the evidence for both sets of theories regarding the prediction of mathematics achievement is sketchy. Longitudinal studies, which are so vital in this context, are so far very limited. The opportunities, therefore, that are provided by the ALSPAC study, are welcome and timely.

2.1.3 The importance of general cognitive processes

In the previous sections, we considered what children need to learn to be successful mathematics learners and the theories about the development of mathematics knowledge. In this section we consider general cognitive processes. No one would doubt that children's intelligence plays a role in their mathematics learning. The question is which aspects of children's intelligence have a specific connection with their mathematics learning.

Measures of children's intelligence assess a verbal and a non-verbal component. The connection between the verbal component and children's mathematics achievement is usually high: children's understanding of the teacher's verbal explanations in any context affects what they learn from instruction. Their reading ability also affects their performance in problem solving when the problems are presented in written form. However, this connection is not specific and does not help us understand children's mathematics achievement. In order to understand the connection between children's intelligence and their mathematics achievement, we need to take the concept of intelligence apart and to identify specific components that influence mathematics achievement. After accomplishing this, we could ask whether it is possible for education to help children at risk for difficulties in mathematics due to low performance in these specific measures.

Most people readily recognise that learning and using mathematics must draw on some general cognitive capacities. In order to use even the simple mathematics that we learn in primary school, we must be able to do several things:

- pay attention to information in whatever form it is given to us (for example, in a story about a situation);

- pick out and remember the relevant parts of this information;

- remember number facts and procedures that help us answer the question we want to answer.

Researchers have proposed some hypotheses regarding the importance of these general cognitive processes for children's mathematics learning. One of these hypotheses is about memory and attention. When children use a computation procedure in order to solve a subtraction or a division sum with four digits, for example, or when they solve an applied problem, they need to keep in mind the information in the problem and the steps to implement the solution, while monitoring what they have done and what still needs to be done. Working memory is the ability to keep information in mind and at the same time operate on it. It is therefore expected that working memory should affect how well they can keep numerical information in mind and thus their success with mathematical procedures and problem solving. The hypothesis that working memory determines children's mathematics achievement has been proposed by many researchers both in the U.K. and in other countries (e.g. Hitch & McAuley, 1991; Towse & Hitch, 1995; Adams & Hitch, 1997; McLean & Hitch, 1999; Siegel & Linder, 1984; Siegel & Ryan, 1989) but so far there is little evidence for it from large scale longitudinal studies.

Attention is most certainly involved in working memory but it can be measured relatively independently of memory. Most teachers would agree that some children make mistakes in solving problems that they should solve correctly because they do not pay sufficient attention to what they are doing. No specific theory links attention to mathematics learning but teachers' practical knowledge can be analysed in the context of this study. Contemporary measures of children's intelligence, such as the Wechsler Intelligence Scale for Children (WISC), show that measures of attention and memory are strongly related and seem to assess a single component, called "freedom from distractibility".

It has also been suggested that we must be able to use our visual and spatial skills in the process of attending to and interpreting information (see, for instance, a collection of papers edited by Zimmermann & Cunningham, 1991, which discusses the role of visualisation in algebra, calculus and computer programming). For example, if we are asked a question like, "The head teacher has 36 books to distribute to classes Pink and Yellow. There are more children in class Pink so she decides to give 6 more books to class Pink than to the Yellow class. How many books will each class receive?", it should help to imagine from the start a pile of books for class Pink that is taller than the one for the Yellow class, and has 6 books more. The problem then becomes trivial: all we need to do is separate out the 6 extra books for class Pink and share the remaining 30 books equally between the two classes. Imagination in this case helps us solve the problem and we do not need to know anything about the order in which the operations of subtraction and division should be carried when solving problems. Thus, although spatial and visualisation skills are general cognitive skills, they might play a specific role in predicting children's mathematics achievement.

Our study includes analyses on whether differences between children in attention and memory influence their mathematics achievement later on. It also investigates whether children's spatial and visualisation skills are specific predictors of mathematics attainment at a later age. The importance of testing these hypotheses through long term and large scale longitudinal studies cannot be over-emphasised.

2.2 Social and emotional issues

Cognitive factors are arguably central to children's mathematics learning but cognition is not the only factor that influences children's learning, which is not free from social and affective influences. In this section we discuss social influences connected to children's own characteristics (social background, ethnicity and gender), connected to school practices (streaming) and affective factors in mathematics learning (children's self-confidence in mathematics).

2.2.1. Socioeconomic status and mathematics learning

There is a strong relationship between school children's social background and what they achieve in mathematics at school, both at early stages (Ginsburg & Russell, 1981) and later on (Lee & Bryk, 1989; Reyes & Stanic, 1988; Sacker, Schoon, & Bartley, 2002; Tate, 1997). As in other school subjects, so in mathematics, children from more prosperous families do better than children with less prosperous backgrounds in the classroom and in exams. This is a consistent and also a familiar correlation, but it raises some unanswered questions.

One is about mediators: there must be a channel of some sort between children's social background and their mathematical success, some factor or factors associated with social background that lead to children from high SES families having relatively more success in mathematics than others from lower SES backgrounds.

Several related possibilities have been discussed in the literature. One is that children from different backgrounds come to school with different sorts of mathematical knowledge. Children from lower SES know some things about mathematics, those from higher SES know other things, and these different forms of knowledge affect what they learn in school and how they perform in school tasks. Ginsburg and Russell (1981) have indeed observed these different levels of performance between children from different social backgrounds on various mathematical tasks at pre-school level. However, their study did not allow for conclusions on the influence of these different skills in relation to the children's mathematics learning in school because it was not a longitudinal study.

A second possibility is that the link between SES and mathematical achievement happens as a result of associated differences in what is often called "cultural capital". There might be differences in the material that supports mathematical learning at home, such as access to a computer, or the presence of a clock, books, comics and games (e.g. Snakes and Ladders) at home, and the activities that are associated with these objects. For example, learning to use a computer involves learning to follow rules in a rigorous order, a sort of knowledge that could prepare children for learning computation algorithms in school. Children who have a watch or a clock in their bedrooms might learn more about time than those who do not, and thus have an advantage in school. Previous research (Siegler & Ramani, 2009) has shown that children's performance in some number tasks was improved by an intervention in which an experimenter played a board game like Snakes and Ladders with the children: having board games at home, therefore, might give children a head start in number tasks. These differences in the cultural capital might be the underlying reason why children from more prosperous families do relatively well in mathematics.

A third possibility, which has received great attention from sociologists, is the relationship between the overt and hidden curricula of schools - i.e. the stated and un-stated, or explicit and implicit goals of mathematics education. Apple (1979), for example, argued that a school curriculum always represents a selection from a much broader knowledge base, and that students from higher SES backgrounds may benefit more from the choices that are made about the form and content of school curricula than those from lower SES backgrounds. This benefit can operate in two ways. First, achievement tests may draw more on the sorts of knowledge developed by higher SES students, and thus they would perform better. This is the argument formulated when one asks whose knowledge gets into tests. Second, the unstated, implicit curriculum may include more of the sort of knowledge developed outside school by students from higher SES backgrounds, leaving students from lower SES backgrounds to fend for themselves in achieving this sort of knowledge, exactly because it remains implicit.

Finally, social background might have a direct effect on some underlying ability which in turn influences how well children do in mathematics at school. This kind of mediation seems quite plausible if one accepts that relatively well educated parents may be readier than others to discuss mathematical ideas and to promote mathematical thinking with their children at home. This kind of informal interaction, if it has any effect at all, is likely to encourage children to think about quantities and mathematical relations and therefore to reason mathematically. Thus, children's mathematical reasoning might be the mediator between social background and children's mathematics.

These are not alternative hypotheses and it is quite likely that they operate together. For example, if children from higher SES backgrounds have more opportunities to engage in informal interactions that promote mathematical reasoning in school *and* if schools leave much of the learning about mathematical reasoning implicit in the curriculum, then children from lower SES background would start school at a disadvantage and also have fewer opportunities to develop the reasoning skills which they need to develop in order to learn better the mathematics that they are taught explicitly.

One must also consider that SES levels influence children's mathematics achievement at the level of the school as well as at the individual level. The average social class of the school has an effect on students' achievement above and beyond the effect of the individual's social class (Alexander, McDill, Fenessy, & D'Amico, 1979; Melhuish, Romaniuk, Sammons, Sylva, Siraj-Blatchford, & Taggart, 2006; Opdenakker & Van Damme, 2007). So it is quite possible that the social composition of the pupils in a school might influence the underlying abilities that form the basis for learning mathematics in school. In other words, whatever their own social background, children might be more likely to develop mathematical reasoning well in a school in which the average SES level is high than in one with a lower average SES composition. As far as we know, there is evidence for the effect of schools' different social compositions on mathematics achievement in standardised assessments (e.g. Lee & Bryk, 1989; Reyes & Stanic, 1988; Sacker, Schoon, & Bartley, 2002; Tate, 1997) but not in assessments of mathematical reasoning that would explain differences in the achievement measures.

This study will consider SES effects on children's mathematics reasoning both at the individual and at the school level. School level influences can be such that all children in a school where the mean SES is high benefit equally from being in the school but it is also possible that the social composition of the school interacts with the individual's own SES background, and that some children benefit more than others from being in a school where the mean SES is high. If the effects of school composition differ for children from different individual SES backgrounds, there will be different slopes for the linear functions that describe the relation between individual SES and achievement across schools with different social compositions. A multilevel approach to the analysis of SES effects on mathematics reasoning is thus important and will be used in this study.

2.2.2 Ethnic background

Different studies have documented that there is a gap between students from minority ethnic groups and those from the dominant group in the society in school achievement in general and also specifically in mathematics (e.g. Secada, 1992; Tate, 1997). Although SES and ethnic background are confounded in many studies, ethnic differences in mathematics attainment persist even when SES is controlled for (Green, Dugoni, Ingels, & Cambum, 1995).

There is only a reduced percentage of children in the ALSPAC sample with minority ethnic backgrounds. This study will investigate effects of ethnicity on mathematics learning as far as possible within the sample and using controls for SES background.

2.2.3 Gender

The study of gender differences in mathematics achievement captured much interest from researchers for about two decades but is has attracted less attention in this last decade because it has become progressively clear that, of the groups that define inequalities in mathematics learning - gender, ethnic origin and social background - gender has the least impact. Lubienski and Bowen (2000) carried out a survey of the studies that analyse group differences in mathematics attainment and report that gender differences were the focus of the largest number of publications about inequalities in education during the 80s and 90s; they located 323 articles about gender effects on mathematics attainment, 112 about effects of ethnicity and 52 about SES effects. After some speculation on the importance of gender differences in brain structure and how the age of adolescent maturation could influence mathematics learning (see Leder, 1992), a sort of speculation that made gender into a biological variable, gender has become firmly placed in the realm of social influences on mathematics learning.

In his review of research on gender, Tate (1997) concluded that, although males tend to perform better than females on standardised tests, effects tend to be small and generally non-significant. Lockheed, Thorpe, Brooks-Gunn, Casserly, and McAloon (1985) observed that the interaction between gender and ethnicity was more important than gender effects *per se*, and criticised previous research in which this interaction had not been analysed.

In this study, we will analyse both gender effects and the interaction between gender and ethnicity. The variation in ethnic background is restricted in the ASLPAC sample, but there is still a sufficiently large sample for an analysis to be carried out.

2.2.4 The effects of streaming on mathematical reasoning

Slavin (1990, p. 472) defines streaming as between-class ability groups, an arrangement in which students are placed in different abilities across all subjects. This form of ability grouping is found more commonly in secondary schools; primary schools tend to use within-class ability grouping, which allows for the assignment of the same student to different ability groups in different subjects. In spite of his recognition of this difference, in his review of the effects of streaming, he included both between-class and within-class ability grouping in his analysis (p. 475). Slavin (1990) and Boaler (1997) carried out excellent reviews of the impact of ability grouping on mathematics learning. Both reviews led to the same conclusion: that streaming only benefits students placed in the top ability group and, even so, just marginally so.

In the studies reviewed, the evidence is based on achievement tests, without the benefit of a measure obtained independently which could be used to assess whether students with the same level of measured ability but placed on different streams for mathematics teaching perform differently.

In this study, it will be possible to carry out an analysis of the effects of streaming on students' achievement controlling for an independent measure of their mathematical reasoning. When the participants were in Year 3, the teachers were asked in which stream or ability group the target children were placed for maths and for English. Their response with respect to maths could either indicate that the school did not use streaming or the stream or ability group in which the child was placed. This information will be used in subsequent analyses. The expressions "stream" and "ability group" will be used interchangeably, as they were used in the questionnaire, but streaming will be used more often for brevity.

2.2.5. Affective factors: children's self-confidence in mathematics

The DCSF publication "Making Good Progress in Key Stage 3 Mathematics" reports that pupils who struggle to progress from Level 3 to 5 in mathematics have a low self-concept as mathematics learners: they feel that they were never able to do maths and have become used to not understanding mathematics. This makes it important for research to describe what influences children's self-perception as mathematics learners and also how it relates to mathematical ability. It is possible that low self-concept as a mathematics learner is a result from low achievement. It is just as plausible that the reverse is true: low self-confidence results in low achievement. These are not necessarily alternative hypotheses: both may be true, with achievement and self-confidence reinforcing each other over time.

Past research has shown that at least four factors influence individual differences in children's self-confidence in maths:

- their mathematics competence (Marsh, 1986; Pretzlik, Olsson, Nabuco, & Cruz, 2003);

- others' (i.e. their teachers' perception and their peers') perception of their ability (Crocker & Cheeseman, 1991; Dermitzaki & Efklides, 2000; Pretzlik, Olsson, Nabuco, & Cruz, 2003);

- gender (girls are reported to have less confidence in their mathematical ability than boys and to attribute their success more to luck than to ability, whereas boys do the opposite; Fennema, 1977; Fennema & Peterson, 1985);

- their verbal ability (past research shows a small negative correlation between verbal ability and self-perception as a learner of maths; Marsh, 1986; Marsh & Yeung, 1997; Skaalvik & Rankin, 1995).

There are conflicting results concerning the importance of these affective factors in the prediction of later achievement. Mortimore, Sammons, Stoll, Lewis and Ecob (1988) found that attitudes and achievement were almost independent of each other whereas Pajares and Miller (1994) found that pupils' attitudes were related to achievement. It is quite possible that these discrepancies result from the use of different measures: in the study by Mortimore *et al.*, the measures were of attitudes towards the subject and in the study by Pajares they included a component of self-perception.

In this study, we investigate the relative importance of the four factors listed earlier on for predicting children's self-confidence in maths. We also investigate whether children's self-confidence makes a contribution to explaining later achievement after controlling for cognitive measures that predict achievement. The ALSPAC data base offers a unique opportunity to study these factors longitudinally: at approximately the same time, the children were given measures of mathematical reasoning, self-perception as learners of mathematics, and their liking of mathematics. These measures preceded the Key Stage 2 and 3 assessments, thus providing the opportunity of an analysis of their power in a longitudinal prediction.

2.3 Summary

Children vary considerably in what they achieve in mathematics at school. The consequences of this wide variation are significant both for the children themselves and for society, given both the importance of mathematical ability for progress in education and employment and also the competitiveness of contemporary societies in the global economy.

The variation in children's mathematical abilities is related to personal factors as well as to group membership. Personal factors include cognitive, social and affective differences. We distinguished between specific cognitive abilities, related to the learning of mathematics in itself, and more general information processing abilities that might influence mathematics learning. Social factors are mostly determined by group membership (SES background, gender, ethnic group, membership in a specific school). Affective aspects of learning mathematics take account of children's self-confidence as learners of mathematics and its relation to achievement: it is hypothesised that ability and self-confidence influence each other, reinforcing each other over time. The ALSPAC data base offers a unique opportunity for explaining differences between children in mathematics achievement, considering all of these aspects in a single study.

3. Method

3.1 The ALSPAC Sample

The Avon Longitudinal Study of Parents and Children (ALSPAC) project has accumulated a wealth of data on slightly more than 14,000 children, born between April 1991 and December 1992 in the geographical region in the West of England that was covered at the time by the Avon Health Authority. The sample is not only a large but also a representative one, since it includes over 80% of the children born in the area during the 21-month recruitment period. The large body of data in the project includes a great deal of information about the children's educational progress, including of course their performance in mathematics and in several relevant psychological tests, such as a well-known intelligence test, the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1992), a test of mathematical reasoning that was given to a sub-sample of the children when they were 8-, 10- and 12-years old, and a measure of children's self-perception as learners of mathematics. The data bank also provides information about the children's social background and about their schools

3.2 The cognitive measures

3.2.1 Mathematical Reasoning - the Year 4 and the Year 6/8 tasks

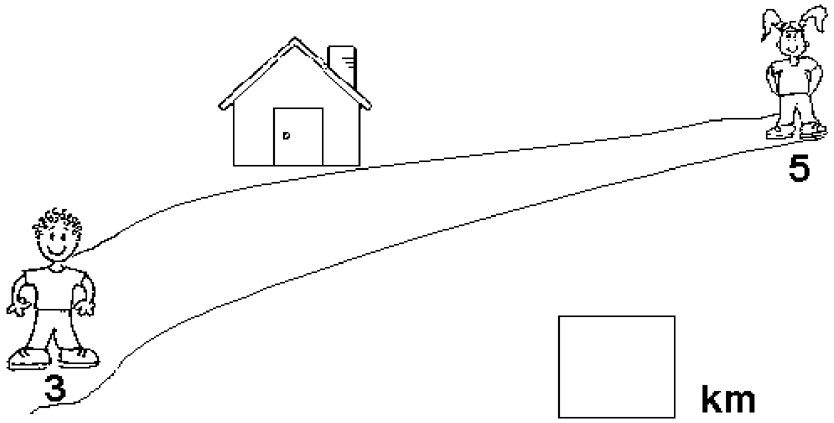
The aim of the Mathematical Reasoning tasks was to assess children's understanding and use of the quantitative relations, which we discussed in the introduction, in order to solve mathematical problems. The tasks were originally devised by Nunes and Bryant for the ALSPAC study. Our view, when we designed the items in these tasks was that it was highly likely though not certain that the importance of reasoning about mathematical relations in mathematical learning is seriously underestimated in mathematical education. However, we recognised that there was no decisive evidence to support that view, and that the ALSPAC study could establish, one way or the other, the part that mathematical reasoning plays in children's learning.

We designed two Mathematical Reasoning tasks. One, containing 17 items, was given to school-children in Year 4 (N= 5275, mean age 8 years 9m). The other, containing 35 items, was given to children in Year 6 (N= 7981, mean age 11 years 2m) and in Year 8 (N= 2755, mean age 12 years 8m).

The aim of both tasks was to assess children's reasoning about quantities and the relations between quantities in mathematical problems independently of their computation skills. None of the items in these tests contained difficult calculations; the children had to reflect on the relations between quantities in each item in order to decide how to solve the problem. All the items were presented with the support of drawings; the children could actually use counting to solve many of the problems if they did not know the number facts that might be used in the solution. All the problems were presented orally by the teachers in order to avoid an undue influence of reading difficulties on the children's performance.

Thus the calculation that is needed in the problem in Figure 2, which was an item in the Mathematical Reasoning test given in Year 4, is not at all difficult and well within the range of calculations that 8-year-olds can do. The taxing part of this item is in reasoning that an addition is needed to solve the problem.

Figure 2 - An additive reasoning problem in the Year 4 Mathematical Reasoning Task



Two friends were walking on the beach. They started walking from the house. The boy walked 3km one way. The girl walked 5km the other way. How far apart are the two friends?

Write your answer in the empty box.

A similar item was included, in which the two friends started walking from the house and in the same direction. In this second item, in spite of the similarity in the language of the problem, the relation between the quantities (how far the friends are from each other) is different and the problem is solved by a subtraction.

Three types of item were included in the Year 4 Mathematics Reasoning Task: additive reasoning items about quantities, additive reasoning items about relations, and multiplicative reasoning items about quantities. The assessments used in Years 6 and 8 included six types of item: additive reasoning items about quantities; additive reasoning items about relations; multiplicative reasoning items about quantities; multiplicative reasoning items involving relations (i.e. proportions); items about spatial reasoning and items about fractional quantities. Our hypothesis is that, although these different items can be distinguished in terms of their content, they measure the same concept, Mathematics Reasoning. Examples of items in these tasks are presented in Appendix A.

Analyses of their internal consistency using the Cronbach's α showed that on all three occasions the mathematics reasoning tasks had good levels of inter-item reliability: .74 at Year 4 (N=5275), .89 at Year 6 (N=7881) and .91 at Year 8 (N=2755).[1] This high internal consistency justifies the addition of all the items into a single score in some of the analyses.

Exploratory factor analyses, followed by confirmatory analyses in structural equation models, were carried out by entering scores for each of the sets of items described previously in this section. In these analyses, a single factor was identified, again justifying the use of a total

[1] The Year 8 Mathematics Reasoning Task was only given to the cohort of children born between 1st September 1991 and 31st August 1992, so the number of participants in this assessment is considerably smaller although it is still a large sample.

score. The details of these analyses will be presented later, when the structural equation models are described.

3.2.2 The Wechsler Intelligence Scale for Children (WISCIII)

The WISC is very well-known, and is probably the most widely used general intelligence test for children. It was administered to 7354 children in the ALSPAC sample, whose mean age at the time was 8 years 7m. (SD 3.92). It consists of 12 subtests, one of which is a measure of children's knowledge of arithmetic.

This arithmetic subtest contains a series of very simple word problems, which pose no conceptual difficulties to children aged about 8 years. For example, "John had 4 pence and his mother gave him 2 more. How many pence did he have altogether?"; "A shop had 25 cartons of milk and sold 14 of them. How many cartons were left?"; "At £8 each, how much will 3 T-shirts cost?" Problems about these types of situation demand little conceptual analysis; when small numbers are used, pre-school children show high rates of success in these problems (for studies showing pre-school children's success in addition and subtraction problems, see Carpenter & Moser, 1982 ; for pre-school children's success in simple multiplication problems, see Becker, 1993; Carpenter, Ansell, Franke, Fennema, & Weisbeck, 1993). Thus this sub-test of the WISC is a simple measure of children's knowledge of arithmetic, as indicated in the task's label. This measure will be used in analyses carried out in this study on its own, as an index of the children's knowledge of number facts and computation skill.

The WISC was originally conceived as a measure of two aspects of intelligence, verbal (measured by six subtests) and non-verbal (measured by the remaining six subtests). Over time, the conception of general intelligence has developed and so did the analyses of the different tasks included in the WISC. Recent analyses of the WISC tasks led to the identification of three, rather then two, aspects of intelligence: in the new analyses the Verbal Comprehension Index is much the same as the original Verbal Scale, but the non-verbal tasks are regrouped in two indices, a Perceptual Organisation Index (which contains items about spatial reasoning and ordering pictures to form an appropriate time sequence) and a Freedom from Distractibility Index (which contains tasks that measure attention and memory; the arithmetic subtest also relates to this factor). We ran a factor analysis on the raw scores of these tasks and confirmed that these three indices could be identified for the ALSPAC sample.

In some of our analyses, we will use an estimate of the children's full IQ score as a control in order to see whether differences in the children's mathematics reasoning allow us to predict their mathematics attainment after controlling for differences in their intelligence. In other analyses we separated out and looked at particular sub-test scores in detail. In particular we were interested in the WISC Arithmetic task, two spatial sub-tasks from the Perceptual Organisation Index, which were the Block Design and the Object Assembly tasks, and two tasks from the Freedom from Distractibility Index, the Digit Span task which gave us a measure of Working Memory and the Coding task which provided us with a measure of each child's attentional powers. Details of these sub-tests are given in Appendix A.

3.3 The social and affective measures

3.3.1 Socioeconomic status

In the UK, and in many other countries, there are sharp socio-economic differences between children. The children in the ALSPAC sample reflect this social diversity well. The ALSPAC data set contains three measures that were used in combination to assess the children's SES: the mother's occupation, the father's occupation, and the highest level of education attained by the mother. One of the SES categories (army) cannot be placed in the ordinal scale formed by other occupations. There are only a few cases in the data set of mothers (n=4) and fathers (n=28) whose SES is described by this category: because this is such a small number, it was decided to exclude this category from the analyses in order to avoid noise in the data. The three measures are highly correlated with each other and so we integrated them into a single factor through a principal components analysis. In most analyses this combined index will be used to measure the children's socio-economic background.

3.3.2 Streaming

Teachers of the children participating in the ASLPAC study were asked whether the school used streaming by ability level for mathematics teaching. Their responses will be treated in two ways in the analyses of effects of streaming. First, a variable will be created to test whether streaming per se affects children's outcomes. Second, the mathematics achievement of children in each ability group (top, middle, and bottom) will be compared with that of children in schools that do not stream children by ability levels for mathematics instruction.

3.3.3 Ethnic group

Information about the children's ethnic group is included in the ASLPAC data set. The groups represented in the data file were classified as White, Black Caribbean, Black African, Other Black, Indian, Pakistani, Bangladeshi, Chinese and Other. Of these groups, two (Black African and Pakistani) had numbers smaller than 5, and so these were excluded from our analyses. When there was a loss of participants due to missing values, groups that numbered less than 5 were systematically excluded for the specific analysis.

The sample of children with scores in mathematics reasoning and results in KS mathematics is mostly (98%) white. Thus the possibilities for investigating effects of ethnicity on mathematics achievement are limited but the data will be analysed as far as possible.

3.3.4 Self-confidence in maths

The children were given a measure of how much they liked maths in Year 3 and a measure of self-perception as learners of maths in Year 4. These two measures were highly correlated (r=.83) even though they were given to the children in different school years. They can be treated as a measure of a single concept, *children's self-confidence in maths*. We integrated the information from these two measures into a single factor, through a principal components analysis. This factor will be used in different statistical analysis.

3.3.5 Others' perception of the child's ability and further influences on self-confidence in maths

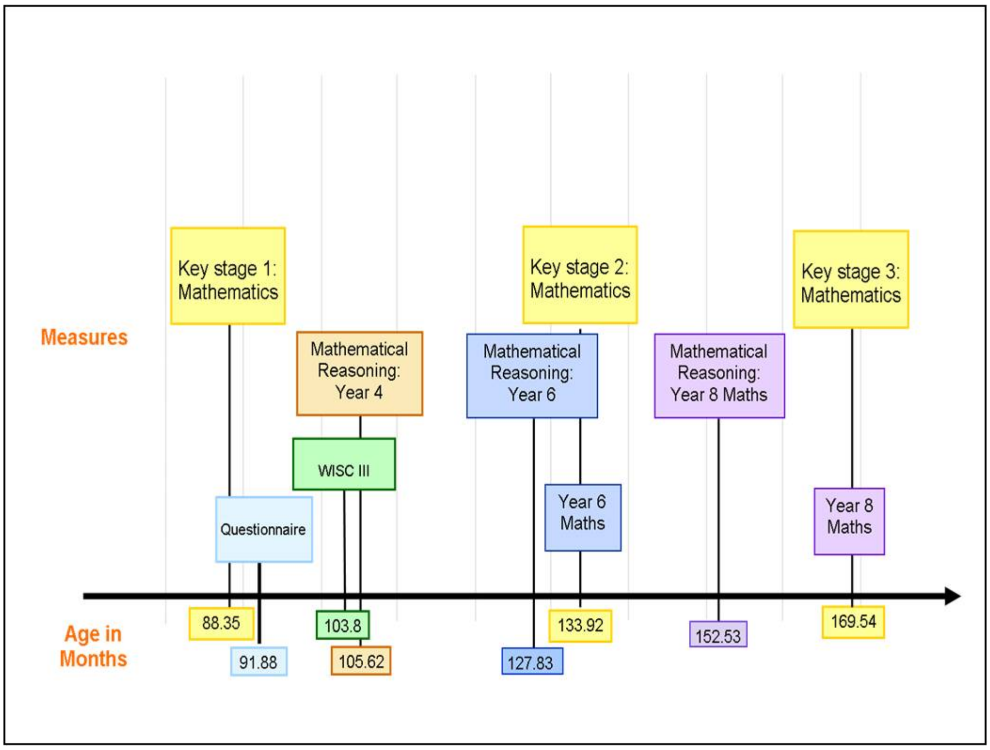
The teachers' perception of the children was assessed in Year 3 with questions that do not refer specifically to the children's mathematical ability but to their general ability and knowledge. We analysed the relation between teachers' perceptions and their pupils' reading and mathematical competence in order to test whether this factor, as measured in the ALSPAC data, might be relevant to the children's self-confidence in maths.

Boaler (1997) observed that students are well aware of which stream they themselves and their peers are placed in for mathematics. Although there is no measure of how peers perceive the children's mathematics ability, when we analyse the effects of streaming on the children's self-confidence in maths we should remain aware that this effect might be mediated by peers' perception of the children.

3.4 The sequence of psychological and educational tests in the ALSPAC data

In order to make the relative timing of the different measures clear, we end this section with the time-line in Figure 3.

Figure 3 - Timeline for ALSPAC cognitive and affective measures



Notice that the WISC III test, the first Mathematical Reasoning task and the questionnaire given to the teachers were all administered in roughly the same time period and over a year after the KS 1 assessments. This made it possible for us to measure how well these centrally important variables predicted the KS2 and the KS3 mathematics assessments over time. The intervals between the time when the children took the Year 4 Mathematical Reasoning task and when the KS2 and 3 assessments were administered were respectively 2 years 4m and 5 years 4m. The interval between the administration of the Year 6 mathematical reasoning task and of the KS3 assessments was 3years 5m. These are long intervals and any variable that predicts children’s achievement over such periods can be counted as powerful and important.

4. Results

We have divided our analyses of the ALSPAC data into two main sections. One deals with the relation between the children's performance in the cognitive tasks and the national Key Stage assessments of Mathematics and, to a certain extent, of Science and English as well, which we take as a good measure of children's achievement in these subjects. The other section presents our analyses of the impact of social and affective factors on children's mathematical achievement, again as measured by the level of their success in the national Key Stage assessments.

Our aim in both sections will be to give a clear and simple summary of a large number of analyses, some of which are quite complex. With each set of results we shall say what the analyses were and what variables were involved and we shall present the essential and important results. However, the reader, who would like to know more about the details of the methods and of the full results of our analyses, should also look at Appendices A-D.

In our analyses we took seriously the proposition, mentioned in the introduction, that individual differences in children's mathematical learning could provide us with the information that we need about the bases for success in mathematics. With hardly an exception our aim in each analysis was to find which variables account for children's mathematics achievement. Each analysis, therefore, had an outcome measure, and this was usually children's results in the national assessments at KS2 and 3 Maths. Each analysis also included predictor variables, our aim being to find how well different variables and different combinations of variables predict success in these Key Stage assessments.

We used mainly three kinds of statistical analysis to look at these predictive relationships. One was the fixed-order multiple regression. The aim of this kind of regression is to see how strongly a set of predictor variables is related to a particular outcome measure. The analysis provides a standardised co-efficient, called β, for the relationship between each predictor variable and the outcome measure, when the influence of all the other variables has been taken into account. The higher the β coefficient, the stronger is the relationship between the predictor variable and the outcome measure. These regressions also tell you how much of the variance in the outcome measure is explained by a particular predictor variable after the effect of all other predictor variables previously entered into the equation has been controlled. The additional variance that each predictor accounts for when it is entered into the equation is the R^2 change figure for that variable. Thus, regressions tell us whether there is a specific connection between a measure used as a predictor and the children's mathematics achievement, after taking into account what both the predictor and mathematics achievement have in common with other predictor variables in the analysis.

The second kind of analysis that we use throughout this report was Structural Equation Models (SEMs). A SEM tells you about the strength of the pathways connecting a set of variables to an outcome measure. These pathways can be direct: for example you can look at the direct path between socio-economic status (SES) and Key Stage Maths performance. They can also be indirect: an example would be a pathway that starts with a path between SES and mathematical reasoning, and then continues on from mathematical reasoning to the outcome measure of mathematical achievement. If this indirect pathway turns out to be stronger than the direct path from SES to mathematical achievement, you will have established that socio-economic status influences mathematical achievement through the effects that it has on children's mathematical reasoning. SEMs usually also involve 'latent variables' which are factors formed from two or more directly measured variables. So, one can form a latent variable of mathematical reasoning on the basis of different reasoning scores (additive reasoning, multiplicative reasoning etc.). Similarly, one can form a measure of children's self-confidence by considering their answers about how much they like maths and also how good they think they are in the subject.

In reporting our results of these two types of analyses, we shall present β figures in the main text and also figures for the percentage of variance in the outcome measures that is explained by some of the predictor variables. The detailed tables for each of these multiple regressions can all be found in Appendix B. We shall present the figures for some SEMs in the main part of the report.

The third kind of analysis that we used in this report was multilevel analysis. Multilevel models are necessary, for example, when the same variable might affect mathematics achievement at the individual level and at the school level. SES is a most common example of this sort of influence: the social composition of the school (i.e. the mean SES of students in a school) may affect their achievement above and beyond their individual SES. In these analyses, we first assessed how much of the variation between individuals could be accounted for by variation between schools. When the proportion of variation accounted for by between-school differences turns out to be important, it is then necessary to analyse how the school composition affects children's achievement. Variation among schools can be of two types. First, it is possible that, in schools with a higher mean SES composition, students from all SES levels perform better than their counterparts in schools with a lower mean SES. Second, it is possible that the benefits of attending a school with a higher SES composition vary between children depending on their own individual SES. We analysed both possibilities using hierarchical linear (multilevel) models.

4.1 Cognitive factors

4.1.1 The effects of cognitive measures on mathematics achievement

The main cognitive measures were the three Mathematical Reasoning tasks, and the various sub-tests of the WISC intelligence test that we listed in the previous section. The main question that we asked was how well each of these variables predicted children's mathematical achievement after we had taken into account certain essential controls.

(a) Mathematical Reasoning in Year 4 and WISC Arithmetic

Our first question was about the relative importance of children's numerical and computation skills, on the one hand, and their understanding of quantities and mathematical relations, on the other. Both kinds of ability have been suggested as the basis for children's mathematics achievement, but we knew of no decisive comparison of these suggestions.

To answer this question we compared the predictive power of the Mathematical Reasoning tasks given in Year 4 and of the WISC Arithmetic sub-test which was administered at much the same time (see timeline in Figure 3) and which we took to be a measure of the children's computation ability. We wanted to find out how well these two variables predicted the children's performance in the Key Stage assessments of mathematics and which predicted these assessments better. Table 1 gives the correlations between these measures of mathematical ability and the children's achievement in KS 1, 2 and 3.

Table 1 - Correlations between the children's Mathematical Reasoning in Year 4, their WISC Arithmetic scores and the mathematics achievement at KS2 and 3

	WISC Arithmetic	KS2 mathematics	KS3 mathematics
Maths Reasoning Year 4	.49	.66	.68
WISC Arithmetic		.57	.58
KS2 mathematics			.88

All the correlations were highly significant statistically. The lowest correlation was between mathematical reasoning and calculation, which suggests that there was some independence between these two cognitive measures. The highest (.88) was between the two Key Stage mathematics assessments, and it establishes that the relative level of the children's mathematical achievement during this period was remarkably stable (for a detailed analysis of progression and stability of Key Stage attainments, see Duckworth, 2007).

Both cognitive measures correlated well with the two Key Stage assessments, but the correlations were noticeably higher for the Mathematical Reasoning tasks than for the WISC Arithmetic task. Thus, these correlations provided us with initial evidence for a strong link between both abilities (reasoning and calculation) and the children's success in mathematics over a five year period, but they also showed that this link was probably stronger in the case of children's mathematical reasoning than in the case of their ability to calculate.

There is a danger with any simple correlation that it might be the result not of a direct link between the two variables concerned but of a common relationship that both of them have with some other, third variable. For example, children's mathematical reasoning might correlate with their success in mathematics at school, not because one of the influence of mathematical reasoning on mathematical learning but because both variables are determined by differences in the children's IQ. One way to make sure that this is not the case is to enter the suspect third variable into a regression equation, as we explained at the beginning of the Results section.

In the multiple regressions that we are about to describe we routinely controlled for the influence of differences in the children's ages, their IQ and their working memory, when we looked at the effects of mathematical reasoning and of children's calculation abilities on children's KS Mathematics achievement. We also took into account the effects of each of the two principal predictors in these analyses when we were measuring the contribution of the other: thus, we controlled for differences in WISC arithmetic when we measured how much of the variance in children's achievement was explained by mathematical reasoning and we controlled for differences in mathematical reasoning when we measured how much variance in the children's achievement was explained by WISC arithmetic.

We carried out these multiple regressions to see how well the children's WISC Arithmetic scores and their Year 4 mathematical reasoning scores predicted the assessment of their mathematical achievement later at KS2 and at KS3. The regressions (which are summarised in detail in Tables 3 and 4 in Appendix B), showed that both measures predicted the children's performance in both of the KS mathematical assessments very well, and independently of each other. Thus, the children's ability both to calculate and to reason about mathematical relations when they were 8-years old played a key role in their mathematical achievement over the next five years.

A further point to be made about this impressive result is that the contribution made by the children's ability to reason mathematically was far greater than the contribution made by their ability to calculate. The β coefficients were consistently higher for Year 4 Mathematical Reasoning than for WISC Arithmetic. The β coefficients for Year 4 Mathematical Reasoning stood at .35 with the KS2 and .34 with the KS3 assessments. For WISC Arithmetic the figures were .21 with KS2, and .18 with KS3, as the outcome measures.

Another way of comparing these two predictor variables was to look at the proportion of variance in the outcome measure that each predictor accounted for when it was entered as the last step in the regression and therefore after the effects of the other four variables in the equation had been taken into account. We found that the WISC Arithmetic task explained 3.0% and Mathematical Reasoning 7.6% of the variance in the KS2 mathematics assessments when each was entered as the last step in the regression. The equivalent figures for the KS3 assessments were 2.3% and 7.5%.

(b) The sustained effect of mathematical reasoning

The data on mathematical reasoning that we have analysed so far were restricted to the test given in Year 4. However, when an ability that has been measured at one particular age turns out to be strongly related to children's achievement at school, one needs to find out whether this ability is only important around the age at which the measure was given or whether it continues to be important when measured at different ages.

In the ALSPAC project the children were given mathematical reasoning tasks at three times. They were given one such task at Year 4, and another mathematical reasoning task (the same task) at Years 6 and 8.

In order to find out whether the task given in Year 6 and 8 also predicted KS Maths achievement, we ran three new multiple regressions on the relations between Year 6 and Year 8 Mathematical Reasoning tasks and the KS2 and 3 assessments (Tables 5, 6 and 7 in Appendix B give the details of these regressions). These regressions took the same form as the regressions that we have already described in which the Year 4 Mathematical Reasoning task was the final step (the regressions reported in Tables 3 and 4). In the three new regressions we took into account the effect of differences in the children's age, their IQ, their Working Memory, and their WISC Arithmetic scores, when we measured the effect of differences in mathematical reasoning.

The new regressions showed that the predictive power of the Mathematical Reasoning scores, already strong in the Year 4 task, was even stronger in Year 6 and stronger still at Year 8 Mathematical reasoning tasks. When the KS2 assessment was the outcome measure, the β coefficient was .39 for the Year 6 Mathematical Reasoning scores, which of course is more than the .35 figure for the Year 4 scores that we have reported already. With the KS3 mathematics achievement as the outcome measure, the β figure for the relationship between mathematical reasoning was .40 for the Year 6 scores and .52 for the Year 8 scores, again much greater than the .35 β figure for the Year 4 scores. Of course, one good reason for the progressively stronger relationship over time between the Mathematical Reasoning tasks and the Key Stage assessments must have been that the interval in time between when the children took these tasks and when they went through the Key Stage assessments was shorter for the Mathematical Reasoning tasks given later than for the Year 4 Mathematical Reasoning task.

(c) Specific and general predictions made by the mathematical reasoning tasks

We have established a strong relation between children's ability to reason mathematically and their achievement in mathematics, but it is possible that their performance in the mathematical reasoning tasks may predict their attainment in other subjects as well. The question is an interesting and important one, because the answer to it will tell us more about the reason why these tasks predict mathematics so strongly. This may be because the relations that children have to reason about in these tasks are mathematical ones, in which case it would be unlikely that the tasks would predict the children's achievement in, for example, English, a subject that does not involve mathematical reasoning. However, children certainly have to reason in English and, if mathematical reasoning tasks predict mathematics achievement well because they measure reasoning in general, these tasks should also predict children's achievement in English.

We decided, therefore, to look at the relations between mathematical reasoning and the Key Stage assessments in English, and also in Science. We chose Science as well because it seemed to us to lie somewhere between mathematics and English in terms of its use of mathematics. Science, as it is taught at school, does involve mathematical relations. Children have to reason about proportions and fractions in physics and chemistry; both proportions and fractions involve reasoning about multiplicative relations (see, for example, Vergnaud, 1983). Therefore, if the reason for the predictive power of the mathematical reasoning tasks is that they specifically deal with mathematical relations, we would expect these tasks to predict attainment in Mathematics and Science a great deal better than in English.

The ALSPAC data set includes records of Key Stage assessments in Science and in English as well as in Mathematics. We were able, therefore, to run identical multiple regressions to the five-step regressions that we have already reported on mathematics achievement, except that in these new regressions the outcome measures were the children's scores in the KS2 and 3 assessments of Science and of English. These analyses are summarised in detail in Table 8 in Appendix B. They produced two main results.

The first was that the children's ability to reason mathematically was more strongly related to their achievement in mathematics than in the other two subjects. The second was that the mathematical reasoning scores predicted the children's achievement in Science better than in English. The β coefficient for the relationship between the Year 4 Mathematical Reasoning task and the KS2 Mathematics achievement was .35; in the case of the Science it was .19; and in the case of English it was .13. The equivalent β figures for the Year 6 Mathematical Reasoning task predictions of the KS2 assessments were .39 for Mathematics, .28 for Science and .17 for English. The discrepancy in the power of the mathematical reasoning tasks to predict Science and English was even more striking with the Year 8 task than with the other two tasks. The β coefficient for the relationship between the Year 8 Mathematical Reasoning task and KS3 assessments was .52 with Mathematics, .40 with Science and .22 with English.

Thus, there was a modest but consistent relationship between the mathematical reasoning tasks and the KS English assessments, even when the effects of differences in IQ and working memory had been controlled. This suggests that, to a small degree, the mathematical reasoning tasks were acting as a measure of children's reasoning in general. However, Table 8 in Appendix B shows that the mathematical reasoning tasks never accounted for more than 3.1% of the variance in the English assessments. This should serve as a reminder that this general function of the mathematical reasoning measures is a highly limited one.

The much stronger relationship between mathematical reasoning and the Science assessments is important and it entails a serious educational implication, which is that teachers of science should pay attention to children's awareness and understanding of the mathematical relationships that are involved in scientific learning.

Finally, the fact that the mathematical reasoning tasks consistently predicted children's mathematical achievement much better than in the other two subjects confirms the power and importance of mathematical reasoning in learning mathematics.

Key findings on mathematical reasoning

- Children's ability to reason about mathematical relations was easily the most powerful predictor of children's mathematical achievement, out of all the relevant cognitive measures in the ALSPAC data bank. It strongly predicted their mathematics achievement in KS2 and 3 assessments even after controls for differences in age, IQ and associated skills.

- The contribution of calculation skills to mathematics achievement was independent of mathematical reasoning but modest.

- Mathematical reasoning scores were strongly related to the KS mathematics assessments, less strongly but quite well related to the science assessments, and only weakly related to the English assessments. The relation between mathematical reasoning and the science assessments was stronger with the KS3 than with the KS2 assessments.

4.1.2 General cognitive measures

(a) Spatial abilities

The WISC intelligence test includes two well-established standardised spatial tasks, Block Design and Object Assembly, which are described in Appendix A. Here we will simply note that in the Block Design test children are asked to construct a series of abstract geometric patterns; in the Object assembly test they are given a series of jigsaw type tasks in which they have to construct some familiar figures (e.g. a face). Both tasks require the child to rotate and manipulate shapes and to imagine the results of these manipulations and rotations. Thus, we were able to consider whether the children's performance in two purely spatial tasks predicts their mathematical achievement at school.

We carried out four multiple regressions. In two of these the outcome measure was the KS2 Mathematics achievement and in the other two the KS3 Mathematics achievement. Block Design was entered as the last step in two of the multiple regressions and Object Assembly in the other two. In all four multiple regressions the first four steps were the children's ages, their IQ, their working memory and their mathematical reasoning.

The two analyses in which Block Design was the last step (Tables 9 and 10, Appendix B) suggested that children's spatial abilities, as measured by Block Design, have a small but consistent effect on their mathematical achievement. Performance in the Block Design task accounted for 1.2% (β= .12) of the variance in children's mathematical achievement at KS2 and 2.3 % (β= .18) at KS3 after controls for the other four variables. Thus, the amount of variance in the mathematics assessments accounted for by the Block Design scores was almost twice as strong five years after, as it was two years after this task was administered. The fact that this link between spatial intelligence and mathematics apparently increased in strength over time is intriguing and needs further investigation. The increase suggests that the demands on children's spatial abilities made in mathematics classes, particularly in classes on geometry, increase between Years 6 and 8 at school.

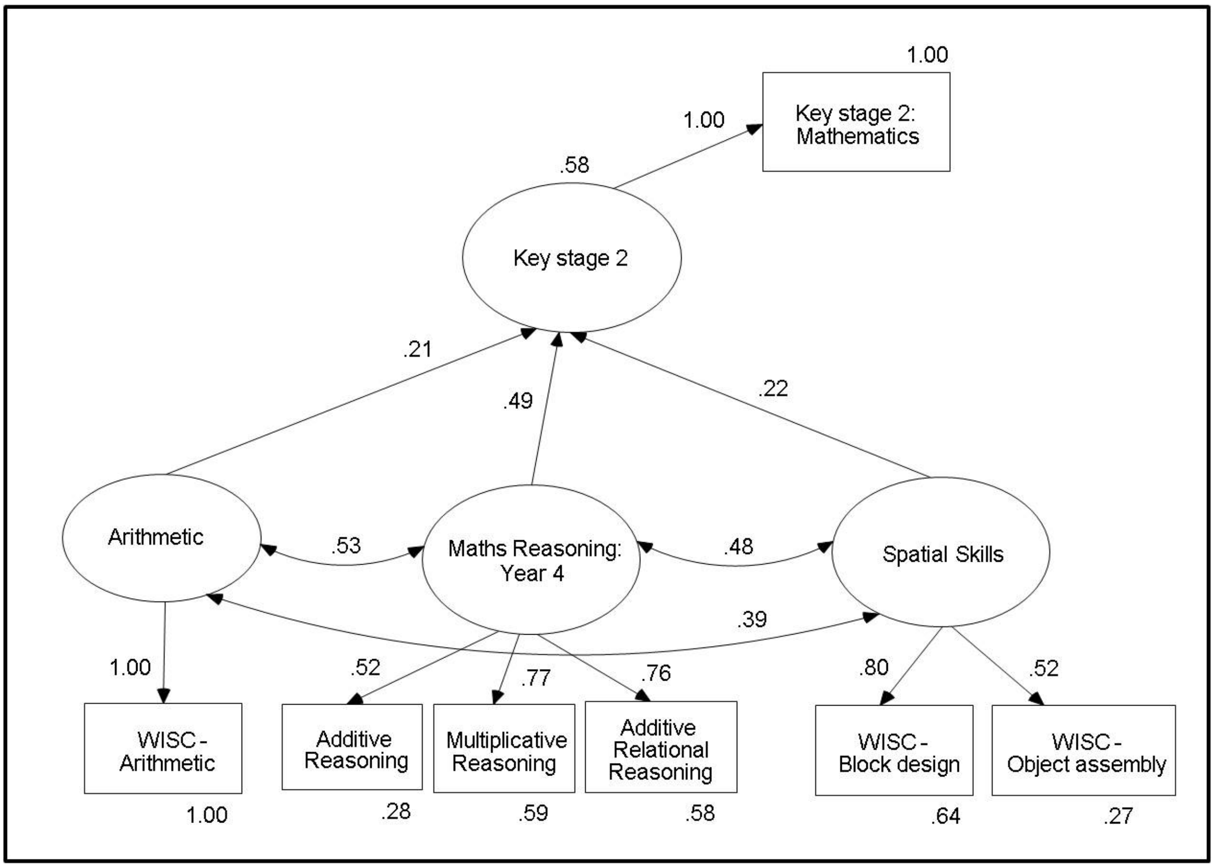
We ran two further multiple regressions (Table 11 and 12, Appendix B) to analyse the links between the Object Assembly scores and children's success in mathematics. These took the same form and contained the same controls as the regressions in which we looked at the contribution of Block Design. The analyses revealed a significant connection between the Object Assembly scores and both Key Stage assessments, but the contribution of this spatial task were very modest indeed. The task accounted for only 0.3% of the variance in the KS2 and for only 0.2% in the KS3 mathematics assessments. This result confirms the connection between children's spatial ability and their mathematics achievement but it also shows that the strength of the connection varies a great deal with the kind of spatial task employed.

We can only speculate why Block Design turned out to be a much better predictor than Object Assembly. The reason may lie in the nature of the two tasks. The Object Assembly task is a jigsaw task in which the children have to put together a set of irregular but familiar shapes. In the Block Design task they have to manipulate and rotate regular geometric shapes (plain squares and squares divided diagonally into two equilateral triangles) in order to copy complex symmetrical geometrical figures. The requirements of learning geometry may account for much of the connection between children's spatial skills and their mathematical achievement. The fact that a task that involves geometric shapes and standard geometric transformations like rotation and symmetry predicts mathematical achievement relatively well fits this idea.

(b) A model of the joint effects of children's mathematical reasoning and spatial abilities

We have established, mainly with the use of multiple regressions, three consistent facts about children's mathematical achievement at school. First, mathematical reasoning is a very powerful predictor, over several years, of this achievement. Second, children's ability to do sums, as measured by WISC Arithmetic, is also a strong predictor, though not as strong as mathematical reasoning. Third, children's spatial abilities also predict mathematical achievement, but to a much smaller extent than the other two measures. The multiple regressions show that each of these three variables predicted achievement after the effects of the other two had been controlled. This suggests that each variable - mathematical reasoning, WISC arithmetic and spatial ability - makes an independent contribution to children's mathematical learning. We used a structural equation model, summarised in Figures 4 and 5, to test this conclusion.

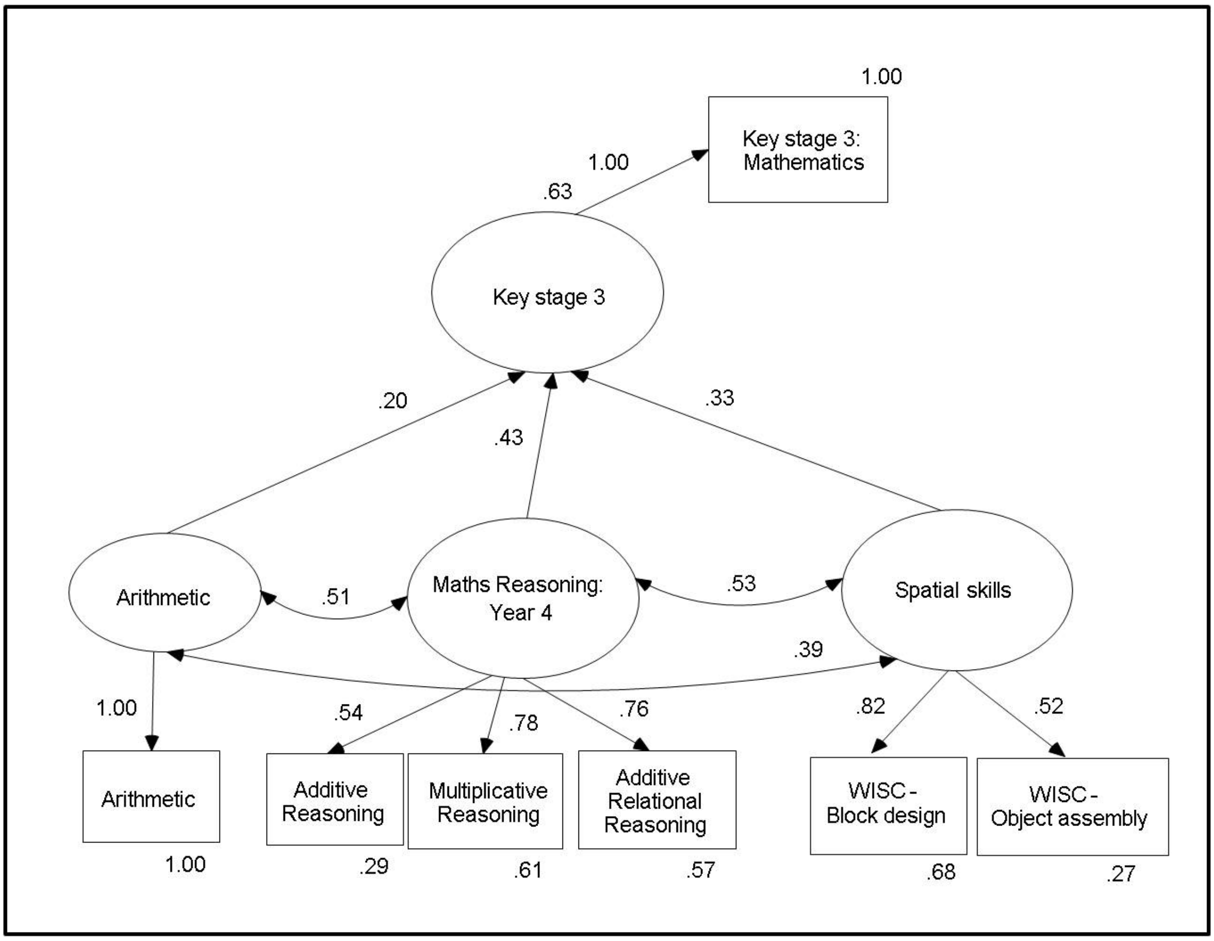
Figure 4 - Structural equation model of the relationships between Arithmetic, Maths Reasoning: Year 4, Spatial Skills, and KS2: Mathematics achievement (N=2488)



In this model we assigned separate scores to the three kinds of item (additive reasoning about quantities, additive reasoning about relations and multiplicative reasoning) in the Year 4 Mathematical Reasoning task and formed a latent variable based on these three scores. The arrows connecting each group of items to the latent variable mathematical reasoning show that, in Year 4, reasoning about additive relations and about multiplicative relations had a stronger connection with the overall mathematical reasoning measure than additive reasoning about quantities.

We also formed another latent variable for spatial ability based on the scores in the Block Design and the Object Assembly tasks. These two latent variables were treated as predictors in the model and so was the manifest variable, Arithmetic, which was entirely based on the WISC Arithmetic subtest.

Figure 5 - Structural equation model of the relationships between Arithmetic, Maths Reasoning: Year 4, Spatial Skills and KS3: Mathematics achievement (N=1630)



The numbers by each of the straight arrows from these three predictors to the outcome measure, KS2 (Figure 4) and KS3 (Figure 5) mathematics, represent the β coefficients for the predictive strength of these three variables. All three variables made direct and independent contributions to the outcome measure, but the influence of mathematical reasoning (β=.49) was far greater than that of the other two variables.

Key Findings

- When mathematical reasoning, arithmetic, and spatial skills are considered together in a model designed for predicting KS2 and 3 Mathematics achievement, mathematical reasoning is found to be the strongest of the three predictors. The addition of spatial skills, as a general cognitive skill, does not detract from the importance of the specific ability in mathematical reasoning.

- Children’s spatial skills predicted their KS mathematics results, but to a lesser extent. Spatial skills were more important for KS3 than KS2 achievement, although the gap in time between the assessment of spatial reasoning and KS achievement was greater for KS3 than KS2.

4.1.3 Memory and attention

Another possible influence on children's mathematics learning is the degree to which they can remember information and can focus their attention in a sustained way in order to learn more generally and to solve problems in the mathematics classroom. There are many approaches to measuring children's attention and memory. The Freedom from Distractibility Index of the WISC contains three tasks: Arithmetic is part of this index along with Digit Span, which assesses memory, and Coding, which assesses attention. The Digit Span task has two components. Forward-digit span assesses short-term memory: the child hears a series of digits and attempts to repeat them in the same order. Backward-digit span assesses working memory: the child hears a series of digits and attempts to repeat them from the last heard to the first. Backward-digit span is considered to measure working memory because the child has to work on the input - i.e. the series of digits - and reverse its sequence, and at the same time recall the input.

In the analyses that we have just described on the effects of the two spatial tasks on mathematics learning, we also included the backward-digit span measure of working memory as a control; this was the third step in each of the multiple regressions (see Tables 9-12 in Appendix B). In these analyses, the IQ estimate excluded the subtest Working Memory. We found that Working Memory accounted independently for a small but significant amount of variance in KS Maths achievement (about 2% in KS2 and 1% in KS3) and its weight in the regression equations (β coefficient) was approximately .1 when the outcome measure was the KS2 assessment and .06 when it was the KS3 assessment. Thus, working memory plays a role in mathematics achievement that is distinct from general intelligence, even though it makes a small contribution.

Another possible influence on children's mathematical learning is the degree to which they can attend in a sustained way to what is going on in the classroom. It is quite hard to measure children's attentional powers, but the WISC battery does contain one task which amounts to a plausible attempt to do so. This is the Coding subtest in which children have in front of them the digits from 1 to 9; under each digit, is a symbol (e.g. a half circle, a plus sign). Below this key there are four rows of digits in random order. The children's task is to draw in the appropriate pattern under each digit. The task is a timed one and the score is the number of items filled in correctly in the fixed period. Any child who attends meticulously and uninterruptedly to the task at hand should do well in this particular task, and it is hard to see any other possible constraint on children's performance in WISC Coding than failures in attention.

Therefore, we carried out two further multiple regressions which took exactly the same form as the ones in which we analysed the role of spatial intelligence, except that the last step in the analyses was now the WISC Coding scores. Again the outcome measure was the KS2, in one analysis, and the KS3 mathematics results in the other analysis. The results for these two multiple regressions are presented in Tables 13 and 14 in Appendix B.

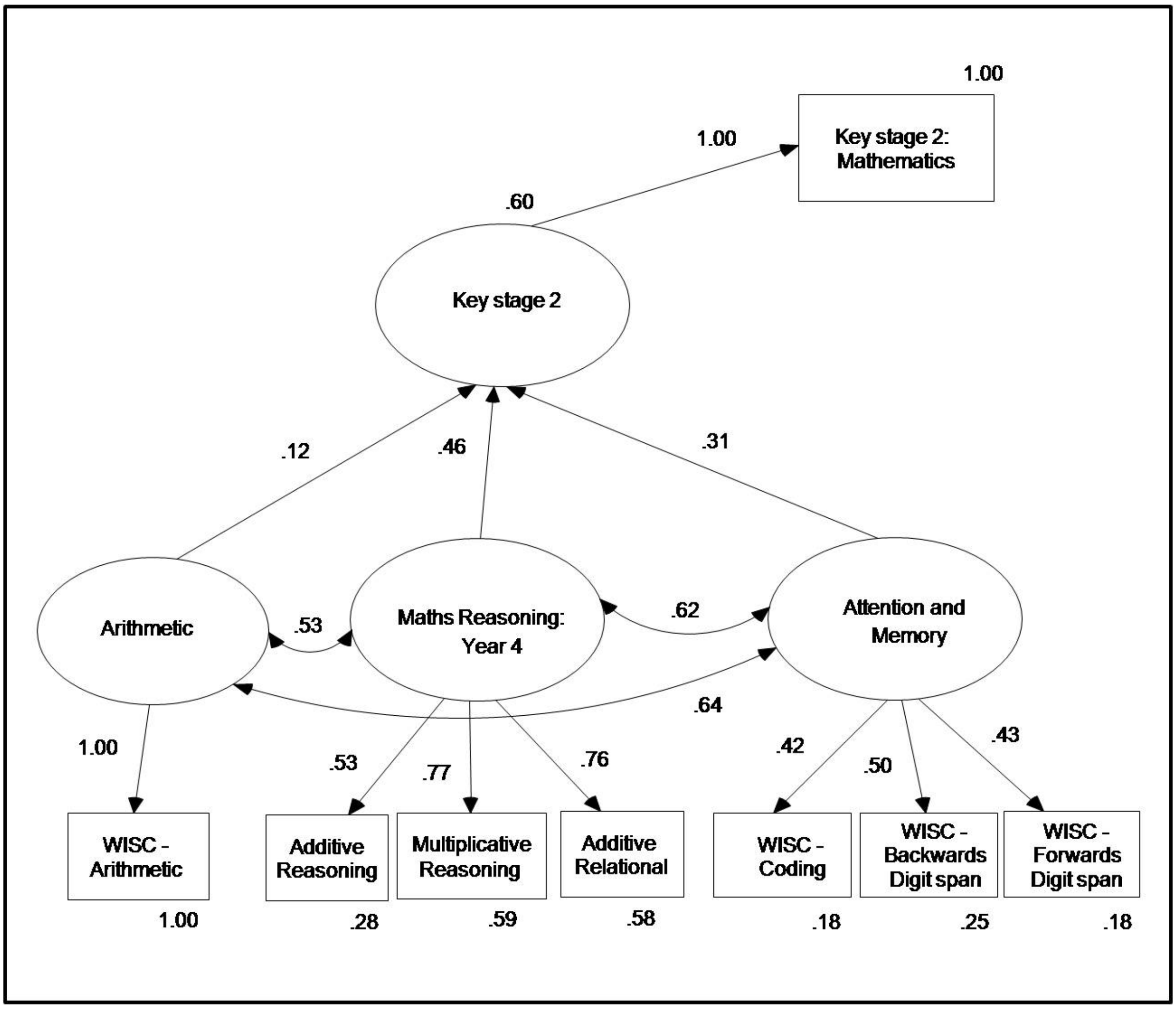
The children's scores in the WISC Coding sub-test did significantly predict the children's mathematical success in the two Key Stage assessments, even after controls for the effects of differences in the children's ages, IQ, working memory and their mathematical reasoning. The amount of variance that the Coding scores accounted for in these assessments was small in both analyses,1.3% (β coefficient = .12) and 1.4% (β coefficient = .13) respectively, but it was consistent and it could be important.

There are established ways of improving children's attentional and working memory strategies (e.g. Klingberg, Fernell, Olesen, Johnson, Gustafsson, Dahlström *et al*. 2005; Nunes, Bell, & Evans, 2008), and the results of these two analyses imply that we now need to investigate whether attention and working memory training would also benefit children's mathematical learning.

(a) A model of the joint effects of children's mathematical reasoning, memory and attention

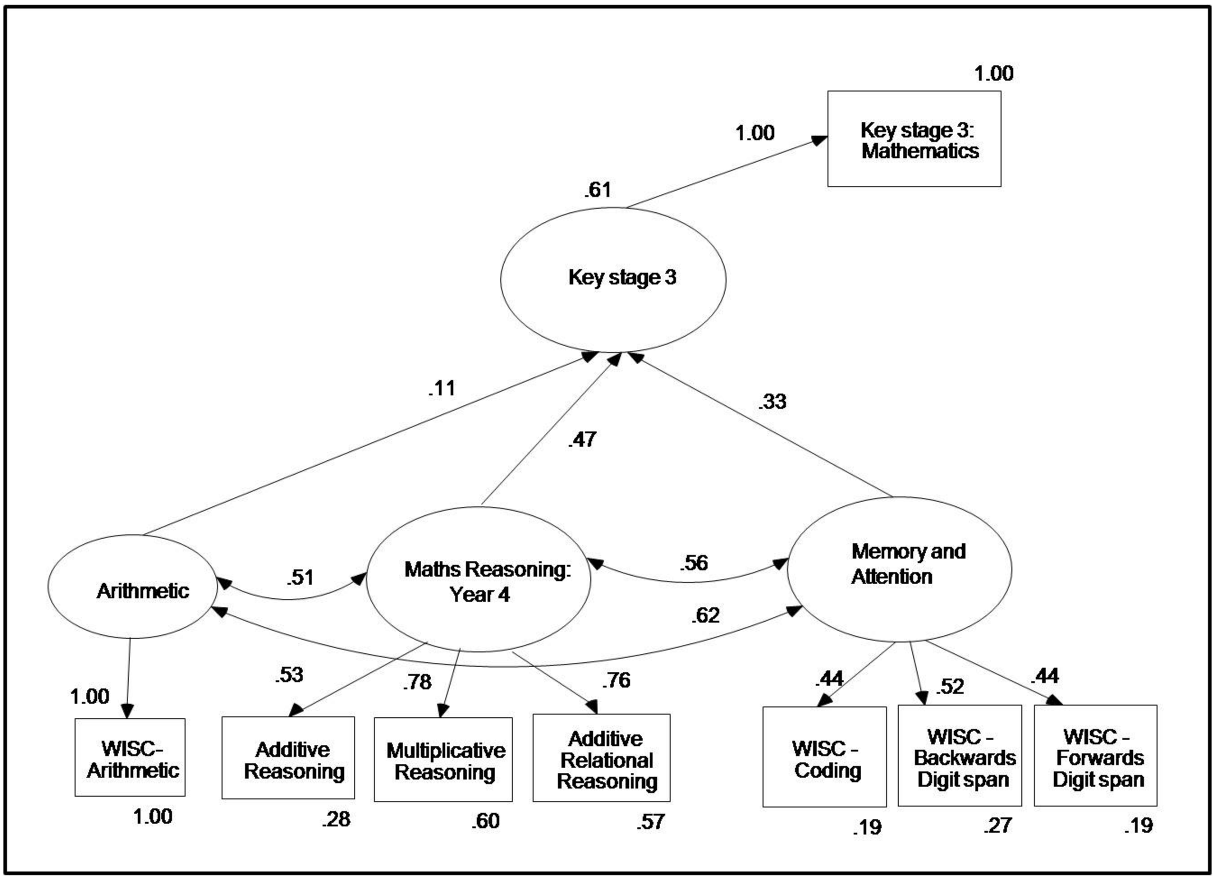
We have reached much the same position with the measures of memory and attention as with the spatial measures. The memory and attention measures made an independent contribution to children's mathematical attainment, but not as strong as mathematical reasoning. For this reason we ran a structural equation model that was similar to the one that we described in Figures 4 and 5 except that instead of spatial abilities we formed a new latent variable, which we called Memory and Attention, based on the memory and attention measures. We present this model in Figures 6 and 7.

Figure 6 - Structural equation model of the relationships between Arithmetic, Maths Reasoning: Year 4, Attention and Memory Tasks, and Key Stage 2: Mathematics achievement (N=2579)



The most surprising feature of this model is the predictive strength of the latent variable, Attention and Memory, which contrasts with the relatively small predictive contributions of Coding and of Digit Span on their own in the multiple regressions just described. The success of this latent variable confirms the importance of attention and memory, and adds some urgency to our plea for research on the effect of training on mathematical learning. We would like to remind readers that, despite the success of the attention and memory variable (β=.31 = in the prediction of KS2 results and β=.33 in the prediction of KS3), mathematical reasoning predicted KS achievement even better (β=.46 and β=.47, respectively, for KS2 and 3).

Figure 7 - Structural equation model of the relationships between Arithmetic, Maths Reasoning: Year 4, Attention and Memory Tasks, and Key Stage 3: Mathematics achievement (N=1680)



Another finding to be considered is the relatively low value of β for Arithmetic: β=.11 and β=.12 for KS2 and 3, respectively). We would like to remind the readers that the subtest Arithmetic is actually part of the Freedom from Distractibility Index, together with Attention and Memory. This connection between Arithmetic, Attention, and Memory raises the possibility that learning arithmetic may be one instance of how attention and memory affect mathematics learning: arithmetic facts and calculation rules may be committed to memory with a relatively small contribution of other general cognitive factors, which in the WISC are verbal comprehension and perceptual integration. For this reason, the importance of arithmetic for predicting KS Maths achievement is not affected significantly when spatial skills are added to the model but it is affected when attention and memory are added to the model.

Key Findings

- Mathematical reasoning makes the most important contribution to the prediction of KS2 and 3 achievement when its importance is tested together with arithmetic, attention and memory.

- Memory and attention made a modest, but independent, contribution to children's mathematical achievement in KS2 and 3. Some children may be at risk for difficulties in learning mathematics due to low performance in these cognitive factors.

- The importance of arithmetic as a predictor of KS2 and 3 achievement is reduced by adding attention and memory to the same model. This suggests that attention and memory play a role in children's learning of arithmetic facts and calculation.

4.2 Social and emotional factors

4.2.1 The effect of mother's education and socio-economic status (SES)

The sharp socio-economic differences between children in the UK are known to play a part in their educational achievements. The variations among the children in the ALSPAC sample reflect the extent of the diversity in socio-economic status that exists in the UK well, but there was a loss of participants over time and more participants from higher SES and higher levels of mother's education remained in the sample at later ages. Although there is slightly less variability in the sample at later ages, there is considerable variation and the sample can still provide a good picture of variation in SES. The sample of children with scores in mathematics reasoning and results in KS2 and 3 mathematics is mostly (98%) white. The analyses in this section, therefore, exclude children from other ethnic backgrounds to avoid noise in the data. Ethnic background is considered later on in a special section.

Mother's highest educational qualification (which will be referred to as mother's education for greater conciseness) and mother's and father's SES are highly correlated. Separate analyses were carried out with each of these measures initially in order to seek replication of the findings. There were no discrepancies in the results when each of measures was treated independently so we report here initially only the analyses with mother's education. Later on, a latent variable, SES background, will be formed using the information from all three measures.

The initial analyses are Analyses of Variance (ANOVAs) with mother's education and gender treated as independent variables; this allows for reporting each of these effects and interactions between social background and gender. This section focuses on the children's social background and the one that follows on the social composition of the schools. The subsequent section focuses on gender and its interaction with social background.[2] These exploratory analyses were used to develop a more precise hypothesis about the form that the relation between social background and mathematics achievement might take. We were interested in observing the trends described by the means and the spread of the scores within each level of mother's education.

Figures 8 - 10 present the means for children's KS1 - 3 Maths achievement, mathematics reasoning and general cognitive measures by mother's education. The results for KS attainment are presented in National Curriculum Levels. The figures show that the children's scores in all the measures increase in an almost linear fashion with the mother's level of education. Smaller differences tend to be observed between mothers with CSE and

---

[2] In a set of preliminary analyses we established that there were no age differences between boys and girls nor between children from different social backgrounds. Age was therefore not entered in these analyses.

Vocational qualifications than between other educational levels. The overall effect of social background on all the measures was significant; the effect size was large and consistent across all the measures.

Figure 8 - Mothers' Educational Levels and the children's attainment scores in the KS 1, 2 and 3 mathematical assessments

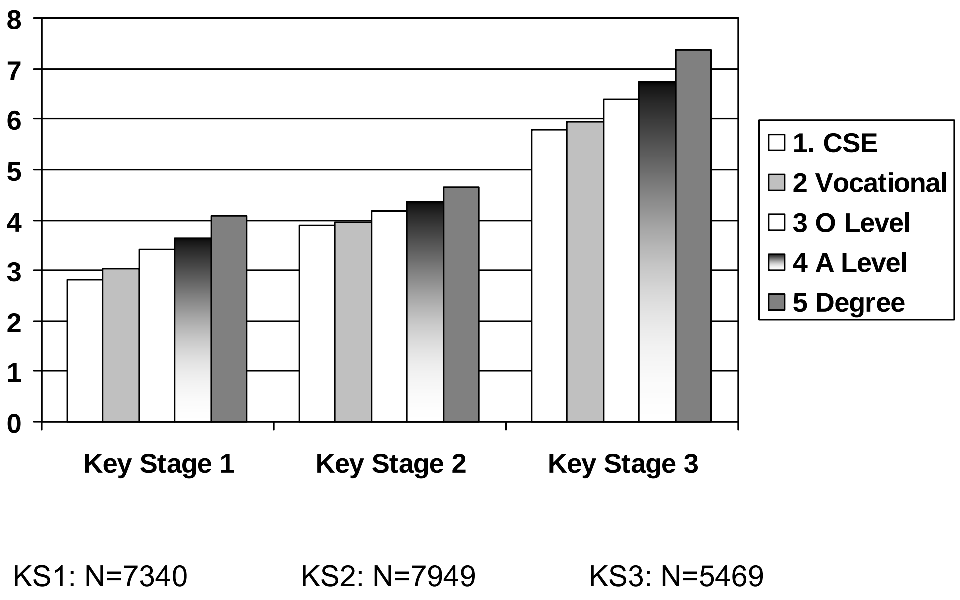


Figure 9 - Mothers' Educational Levels and the children's scores in the three mathematical reasoning tasks

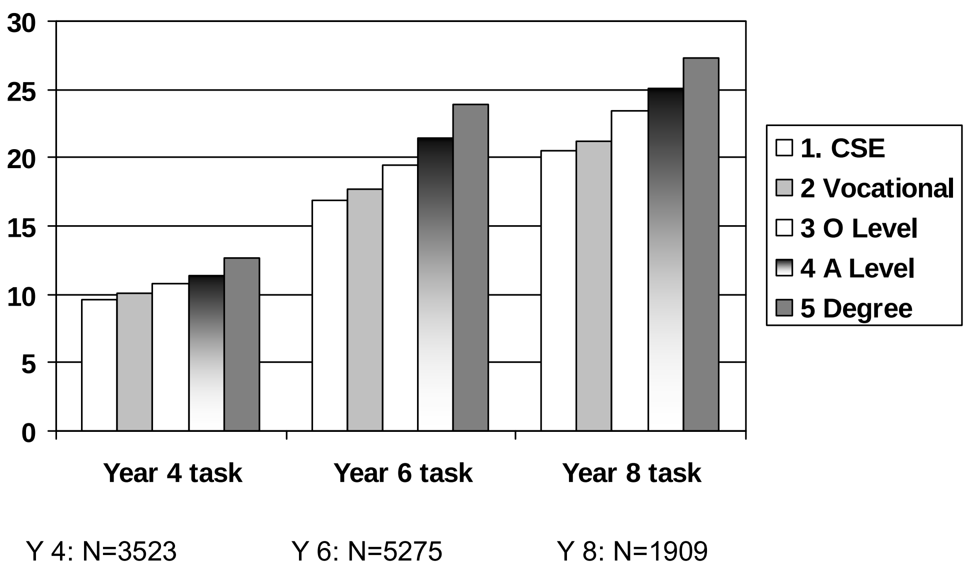


Figure 10 - Mothers' Educational Levels and the children's scores in the WISC Arithmetic and Block Design tasks (N=5564)

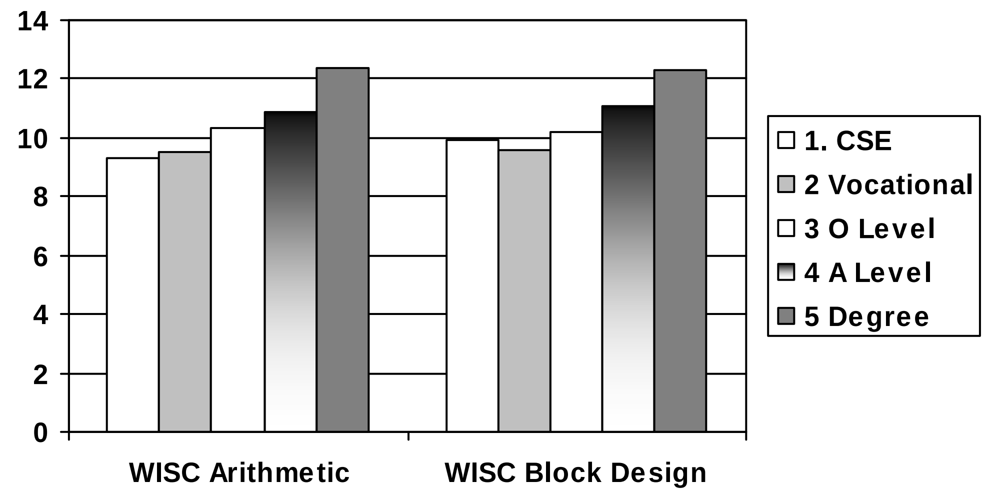
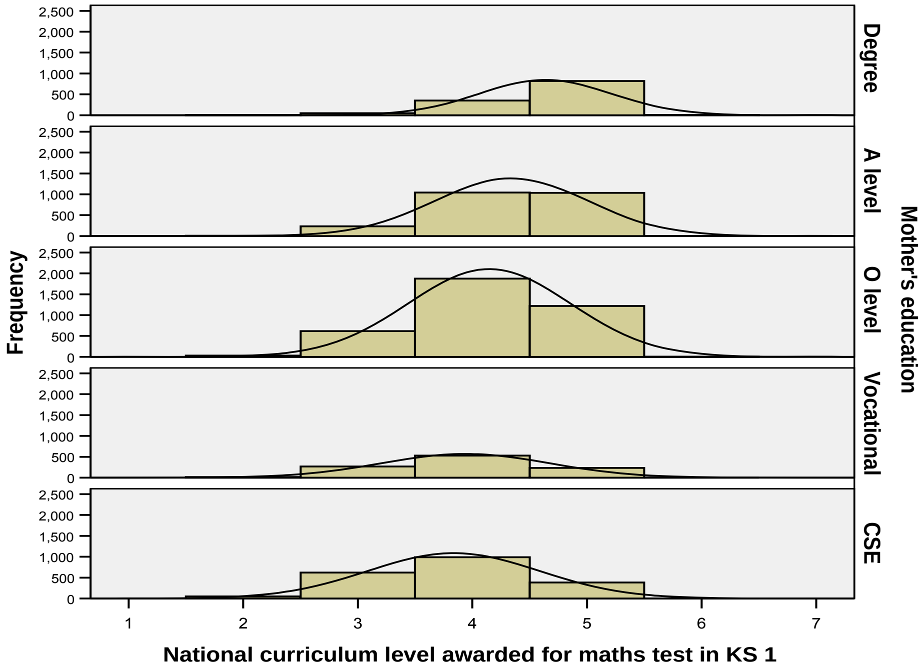
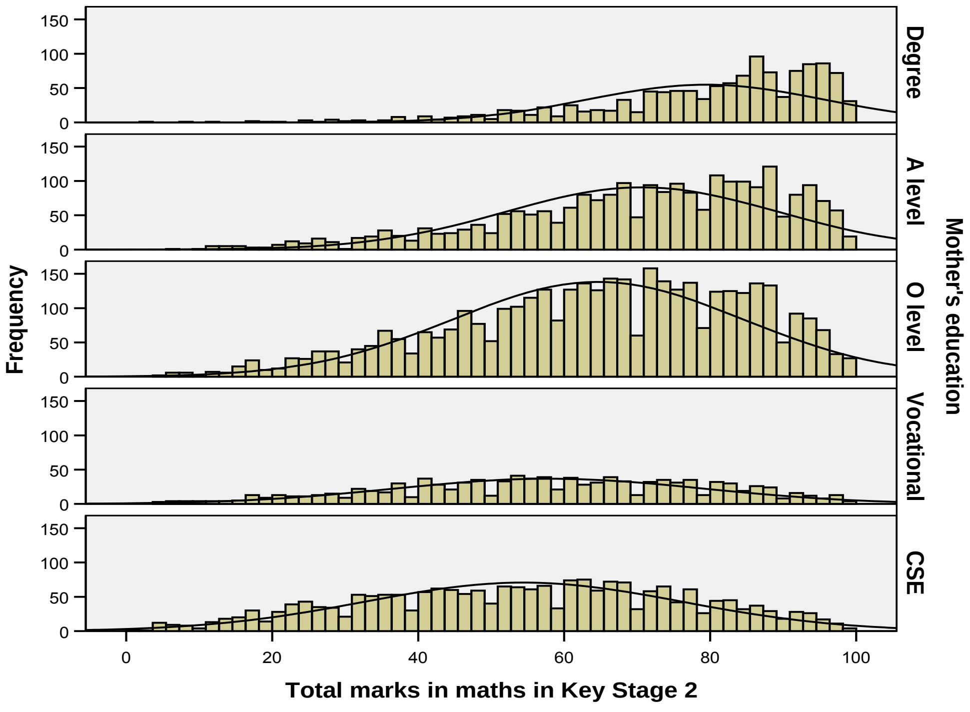


Figure 11 shows the distribution of KS1 and 2 Maths attainment for the different levels of mother's education. We selected these two assessments because KS1 is reported in National Curriculum levels and KS2 is reported in points attainment so they offer different degrees of refinement of the measures. The two graphs demonstrate that social background (in these cases, measured by mother's education) operates by a reduction in the variability of scores at the higher levels of mother's education. Children whose mother's education is classified as vocational or CSE display a greater variation in attainment, with results ranging from the lowest to the highest levels. For the sake of brevity, we did not include the graph for KS3 results, but the distributions are highly similar in this graph also.

These results are compatible with the idea that children from different social backgrounds would in principle show the same amount of variation. However, social background operates as an intervention: the higher the parents' SES and education, the better is the "home intervention" that they provide. Effective treatments improve results and reduce variability in educational outcomes. Thus, understanding the nature of these natural "home interventions" is crucial to promoting equality in education.

Figure 11 - Distribution of Key Stage results by different levels of mother's education





SES and mother's education effects are, to some extent, a black box. It is important in educational research to try to look into how these effects are produced in the children's daily lives in and out of school. We carried out two further analyses in order to understand better what mediates the effects of the children's social background on their attainments. Two possible explanations could be tested through the analysis of the ALSPAC data: they are that (1) "cultural capital" and (2) mathematical reasoning are important mediators of the social background effects.

4.2.2 The mediation between social background and mathematics attainment by the cultural capital

The ALSPAC data set contains information about what we might call the material aspects of the cultural capital that children from more affluent homes have access to: one of the questionnaires records whether the child has a computer, a clock, a radio, a television, books, comics and games (e.g. Snakes and Ladders) in the bedroom. The questionnaire was given to the children when they were about 9 years old. As discussed in the introduction, these objects might create the opportunity for parents to carry out activities that impact on concepts that will provide a basis for their later mathematics achievement. Therefore, we analysed how these measures correlated with each other and with the social background measures. By providing clues about the activities in more affluent homes that lead to better mathematics results, this information could help us understand what mediates social background effects on children's mathematics learning.

An exploratory analysis through non-parametric correlations showed that there was a positive but weak correlation between the social background measures and the child having a TV or a computer in the bedroom; these correlations varied between r=.2 and r=.3. There was also a weak but positive correlation between the child's ownership of a TV and a computer in the bedroom (r=.27). All other correlations were very low; the items could not be added together to provide a single indicator of how the cultural capital was being transmitted in the home because they did not form a reliable scale (the Cronbach's α index of internal consistency was .3 and an index of .7 or higher is expected in a reliable scale). There was no correlation between the child having games in the room or a computer in the bedroom, on the one hand, and the KS1 or 2 Mathematics results and the children's performance in the mathematics reasoning tests, on the other. These material indicators, therefore, did not provide information on what mediates the relation between social background and mathematics attainment.

4.2.3 The mediation between social background and mathematics attainment by mathematics reasoning

So far we have established (1) a strong relationship between children's mathematics reasoning and their achievement in the mathematics assessments at KS2 and 3; (2) a strong relationship between measures of the children's social background and their performance in these Key Stage assessments; and (3) a moderate relationship between the same social background measures and the children's success in the mathematical reasoning tasks. It is highly likely that these three relationships are connected to each other, and one possible way in which they might be linked is through the children's mathematical reasoning ability playing a "mediating" role in the connection between social background and success in the KS mathematics assessments. *The idea here is that the children's social background has an effect on their ability to reason about mathematical relations, and the extent to which they can reason mathematically then determines how well they learn mathematics at schoo*l. This seemed to us a most plausible hypothesis and we carried out a series of analyses to test it.

First we formed one unified measure of the children's social background. The three social background measures, mother's and father's SES and mother's education, are significantly correlated. They can be considered as assessing a broader construct, social background, which was identified by means of a principal components analysis. A single factor emerged from the analysis; the factor loadings were .79 for mothers' SES, .72 for father's SES and .82 for mother's education.[3]

We ran a series of structural equation models (SEMs) as a direct test of the hypothesis that mathematics reasoning mediates the relation between social background and KS2 and 3 Mathematics achievement. An SEM establishes that one variable (*B*) mediates the effect of another variable (*A*) on the outcome measure (*O*) when the indirect pathway from *A* to the *O* via *B* is as strong as or stronger than the direct pathway from *A* to *O*. A strong indirect pathway like this actually shows that *A* influences the outcome through its effect on *B*.

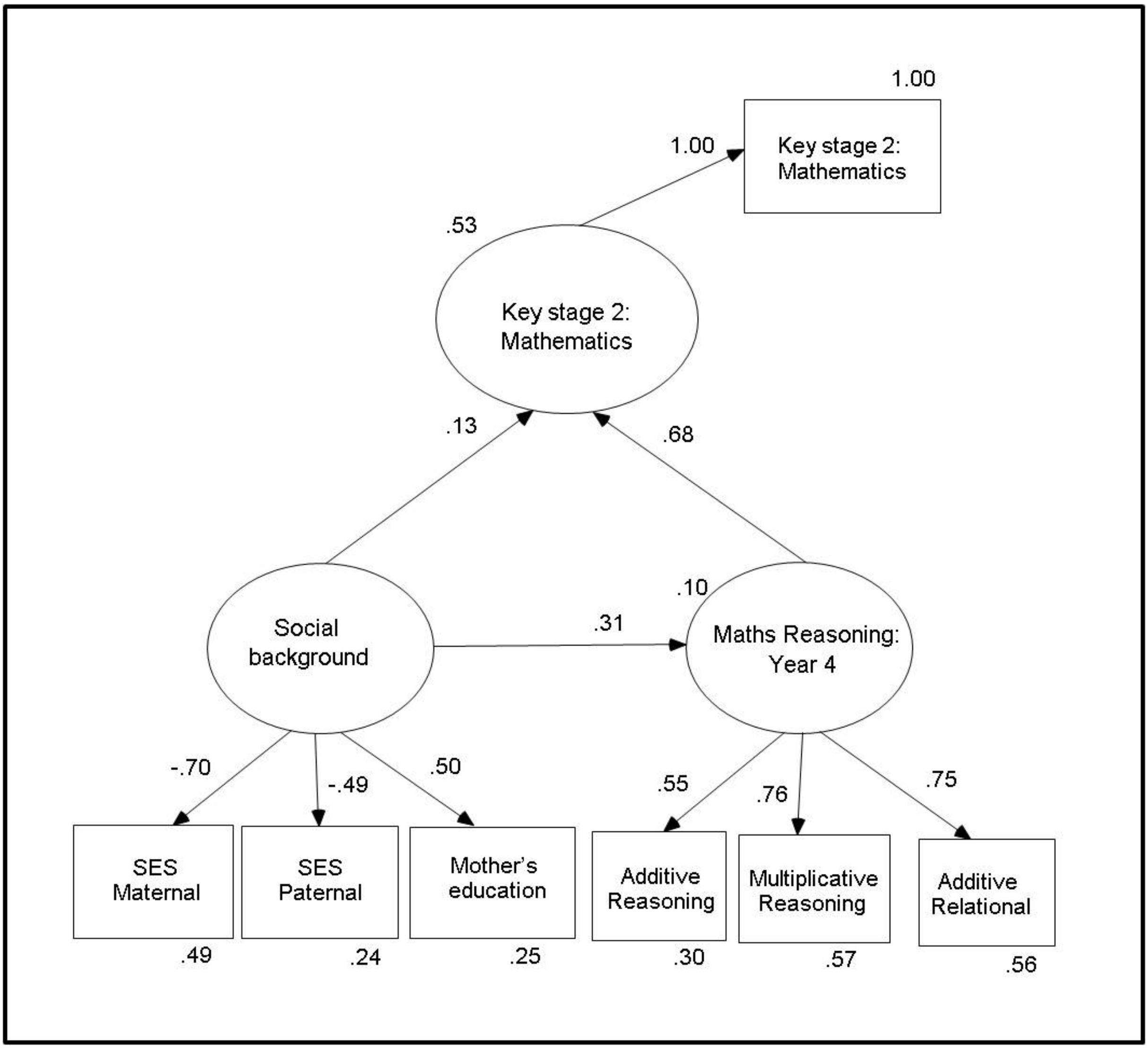
Figures 12 - 15 all provide good evidence for a strong indirect pathway from Social Background to the Key Stage achievement through mathematical reasoning. The path from Social Background to the outcome measure in each model (which was either the KS2 or KS3 mathematics attainment levels) was via mathematical reasoning.[4] This means that mathematics reasoning mediates a great deal of the effect of social background on children's mathematical achievement at school. In other words one of the main reasons why children's mathematical attainment varies so strongly with the children's SES is that children's social background has a powerful effect on children's ability to reason mathematically and their mathematical reasoning in turn plays an important role in their mathematical progress at school.

In our view these structural equation models suggest that the home interventions may operate through some form of teaching of mathematical reasoning in the more affluent homes. As a result, children from such home environments develop their mathematics reasoning to significantly higher levels than children from less privileged homes, and are at an educational advantage when they are taught mathematics in school.

[3] In some analyses a negative sign will appear because the SES scale in the data set has 1 as the highest value and the mother's education has 1 as the lowest value.

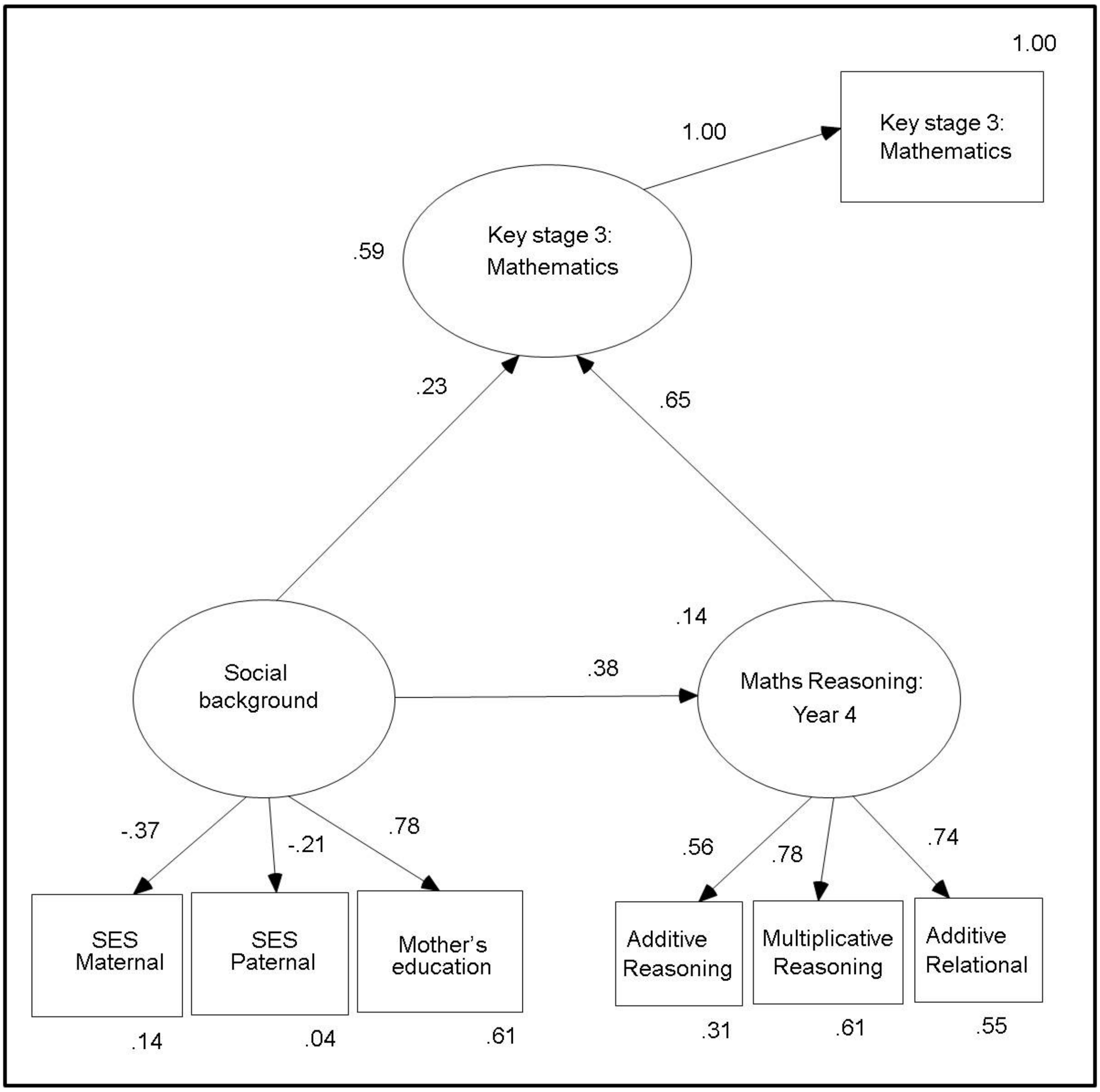
[4] Similar results were obtained using the assessments of mathematics reasoning in Year 8 as mediator and KS3 Maths as outcome. We did not include these figures here for the sake of brevity.

Figure 12 - Path analysis showing the coefficients in a model where mathematics reasoning measured in Year 4 is treated as a mediator between social background and KS2 Maths attainment (N=3183)



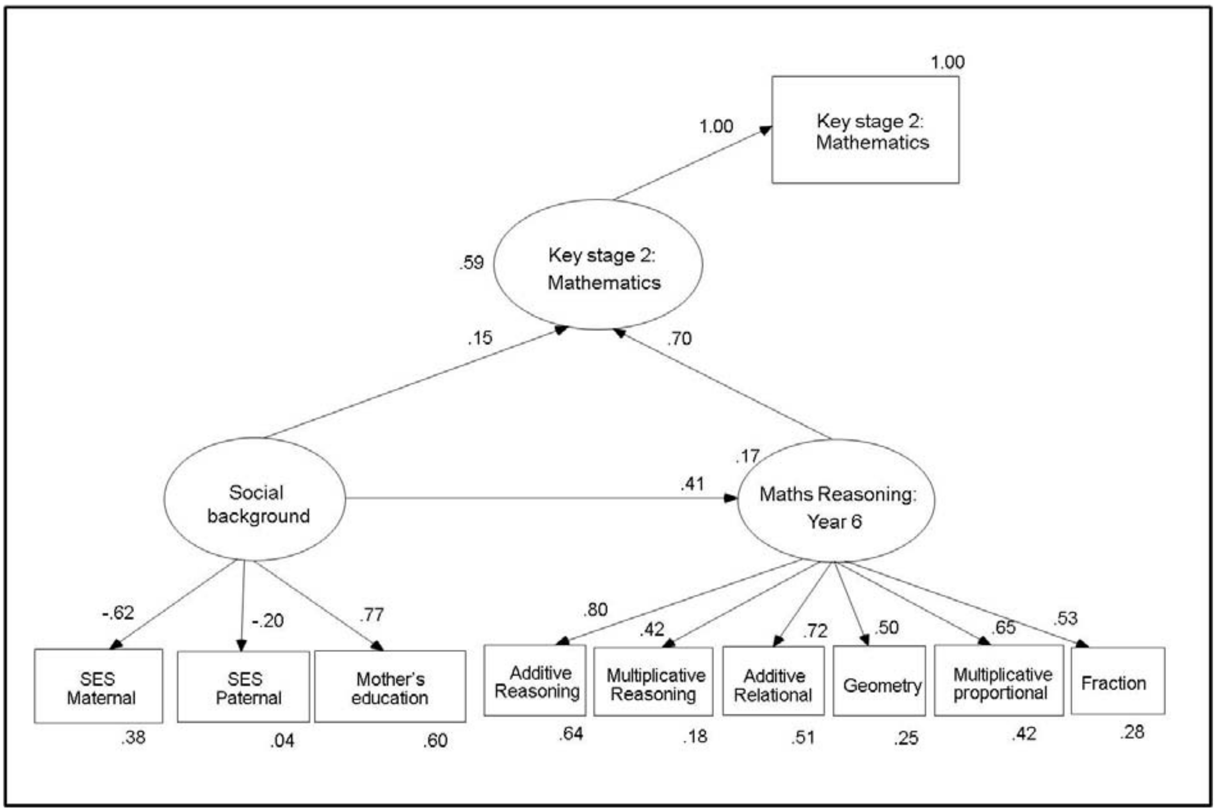
The value next to the arrow connecting Social Background to Key Stage 2 indicates the importance of the direct path from Social Background. The value next to the arrow connecting Mathematical Reasoning to Key Stage 2 indicates the importance of the indirect path. The model shows that the indirect path is considerably stronger than the direct path.

Figure 13 - Path analysis showing the coefficients in a model where mathematics reasoning measured in Year 4 is treated as a mediator between social background and KS3 Maths attainment (N=2083)



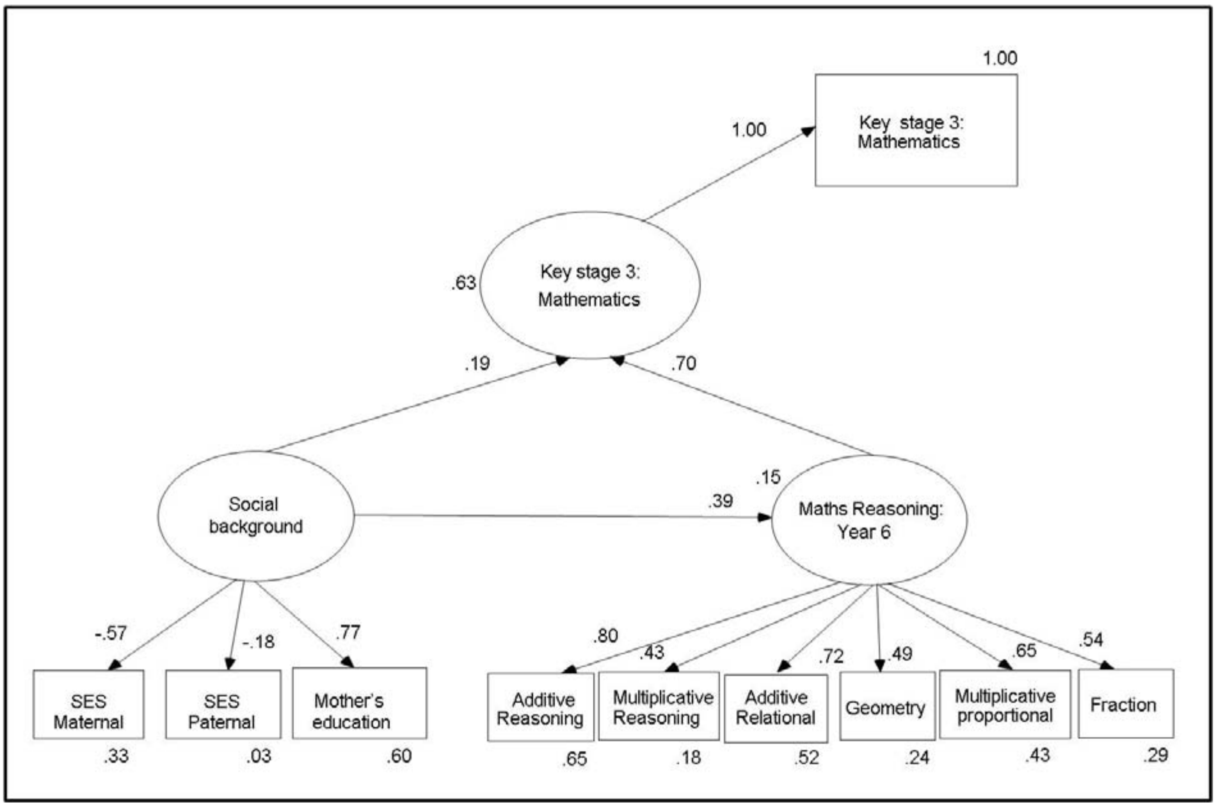
The value next to the arrow connecting Social Background to Key Stage 3 indicates the importance of the direct path from Social Background. The value next to the arrow connecting Mathematical Reasoning to Key Stage 3 indicates the importance of the indirect path. The model replicates the finding described in Figure 12: the indirect path is considerably stronger than the direct path.

Figure 14 - Path analysis showing the coefficients in a model where mathematics reasoning measured in Year 6 is treated as a mediator between social background and KS2 Maths attainment (N=4953)



The value next to the arrow connecting Social Background to Key Stage 2 indicates the importance of the direct path from Social Background. The value next to the arrow connecting Mathematical Reasoning to Key Stage 2 indicates the importance of the indirect path. This analysis provides a third replication of the same finding. The measure of mathematical reasoning here is the Year 6 Maths Reasoning measure and the outcome KS2 Maths: the indirect path is clearly stronger than the direct path between SES and KS2 Maths achievement.

Figure 15 - Path analysis showing the coefficients in a model where mathematics reasoning measured in Year 6 is treated as a mediator between social background and KS3 Maths attainment (N=3428)



Once again, the indirect path between Social Background to Key Stage 3 Maths, through Mathematical Reasoning, is stronger (.70) than the direct path (.19). Thus this set of SEMs shows that analyses with two different measures of mathematical reasoning (Year 4 and Year 6 tasks) and two different outcome measures (KS2 and 3 Maths) converge and demonstrate that the indirect path between SES and KS Maths achievement, through mathematical reasoning, is stronger than the direct path.

These results are stimulating for educators. They suggest that the negative effects of coming from less privileged homes can be offset by an approach to education that offers teaching which improves children's mathematics reasoning. The results converge with evidence from a quasi-experimental study (Nunes *et al*., 2007) in which one cohort of children at risk for difficulties in learning mathematics was assigned to a control group and another cohort, sampled from the same schools, was assigned to an experimental group. The children in both groups were considered at risk for difficulties in learning mathematics because they had performed below the 20th percentile in their schools in mathematical reasoning. After the intervention, the children who received teaching on mathematics reasoning significantly improved their results in KS1 Maths in comparison to the control group. They also showed a level of attainment slightly above the average (i.e. the 50th percentile) for British children.

4.2.4 The social composition of schools and its effect on mathematics reasoning

Educational researchers are naturally interested in understanding differences between schools and attempting to clarify the nature of these differences even when no information about the leadership, pedagogy and other school practices is available in a data set. In the ALSPAC data set, there is no information about school practices that could have an impact on mathematics learning, except for whether the children are in a school that uses streams for mathematics teaching or not, and in the former case which stream each child was in.

However, it is still possible to find out whether the demographic composition of the school can explain some of the variation observed in the children's performance. The first step in the analysis of school effects is to compare between-school differences with within-school differences and estimate the amount of variance between individual children that can be accounted for by between-school differences. Between-school differences accounted for approximately 10% of the individual differences in mathematics reasoning measured in Year 4. This is a substantial amount of variance and justifies further investigation regarding the nature of this effect.

In the UK, and in many other countries, there are sharp socio-economic differences between schools as well as between individual children. Some schools are largely composed of children whose parents are well qualified educationally and have relatively high incomes; others are not. The schools in the ALSPAC sample reflect this social diversity in schools well. The average score for the children's social background in each school varies a great deal between the different schools in the sample, and at the same time most of schools contain children from a relatively wide range of socio-economic levels.

The existence of these differences between schools and also between children within schools raises an interesting and potentially important question. It is possible that individual children's mathematical understanding might be affected not only by their own socio-economic levels, but also by the average socio-economic level of the pupils in the school that they happen to be in (i.e. the social composition of the school). In the introduction, we referred to research that shows that social background affects children's mathematics attainment both at the individual level and through the social composition of schools; we now propose to analyse if the same is true of social background influences on mathematical reasoning.

We already know that children from higher socio-economic backgrounds tend to reason about mathematics more effectively than children from lower socio-economic backgrounds. Now, we have to consider the possibility that the socio-economic composition of the schools also plays a role. Individual children in schools with pupils from predominantly high social backgrounds might do better in mathematical reasoning tasks than children in schools with a lower socio-economic composition, and this effect might be independent of the effect on individual children of their own social background. In other words, a child from a poor social background might be better at reasoning mathematically if he or she attends a school in which the average socio-economic level of the pupils is relatively high than one in which it is relatively low.

The most efficient way to examine the effect of the differing social composition of schools on children's mathematical learning is *multilevel modelling*. This type of analysis deals with "nested" variables. In the ALSPAC study the children are "nested" in schools: there are several schools, but at any particular time each child belonged to one school and not to any other school. Multilevel modelling allows investigators to examine whether the nests (in our case, the schools) that individual participants happen to be in do have an independent effect on these participants' scores.

There are at least two levels of analysis in every multilevel model. Some have two levels, others three or four or even more. The multilevel analyses that we shall describe all had two levels only. One level was the individual child, and here we examined, or rather re-examined, the relation between the individual children's mathematical performance and their social background. The other level was the school (or in other words "the nest"): at this level we looked at the overall success of the children in the different schools in relation to the average socio-economic level of the pupils in these schools. The number of schools in these analyses varied from 27 to 133. The number of the participants in each school ranged from 6 to 168.

We shall describe three multilevel analyses of the connection between socio-economic status, both at the level of the individual and at the level of the school, and mathematical reasoning. The children's scores in the Year 4 Mathematics Reasoning task was the outcome measure in the first analysis (Table 15, Appendix C): their scores in the Year 6 and the Year 8 Mathematics Reasoning tasks were the outcome measures in the second and third analyses respectively (Tables 16 and 17, Appendix C). Each analysis measured three possible relationships. The first was between the children's social background at the individual level and their ability to reason mathematically, independently of the schools they attend. The second was between the socio-economic composition of the schools that the children attended at the time and their mathematical reasoning. The third, which we shall describe in more detail below, is of an interaction between these two levels.

All three analyses confirmed the relation between socio-economic status and the individual children's mathematical reasoning, and showed that this effect was significant quite independently of the schools that the children attended.

The three analyses also showed that the socio-economic composition of the schools influenced the children's performance in the mathematical reasoning tasks. This significant effect was quite independent of the relation between the individual children's social background and their scores in the three reasoning tasks. Thus, these analyses established that socio-economic factors influence children's mathematical reasoning at two levels, the level of the individual and the level of the school.

The interaction term was not significant in the analyses of the Year 4 and 6 mathematical reasoning tasks (Tables 15 and 16, Appendix C), but it was just significant in the analysis of the Year 8 mathematical reasoning task (Table 17, Appendix C). This is a modest effect but it is worth considering what it means.

A significant interaction indicates that socio-economic differences at one level, the level of the school, influence the pattern of socio-economic effects at the other level, the level of the individual. Different schools could have different effects on the relation that undoubtedly exists between individual children's socio-economic background and their success in mathematical reasoning. In some schools, this relation might be a powerful one, so that the children with advantageous backgrounds would do far better in mathematics than children from less prosperous families. In other schools, the differences might be less radical, and disadvantaged children would do almost as well as advantaged ones. A significant interaction indicates that the socio-economic composition of the school influences how powerful the relation between individual children's social background and their mathematical reasoning becomes.

The small, but significant, interaction in the multilevel model of the Year 8 Mathematical Reasoning task, therefore, shows that the social composition of the schools in the sample influenced the relationship between the individual pupils' social background and their mathematical reasoning scores. We need to know exactly how this relationship varied.

Appendix D presents a detailed analysis of this interesting interaction. The analysis showed that the relationship between the children's SES and their mathematical reasoning was quite strong in schools whose social composition was low in SES terms or around average, but that it was not so strong in schools with a high SES composition. It appears therefore that schools with a relatively high SES social composition managed to dampen down the otherwise pervasive effects of social background on children's mathematical reasoning. How they manage to do so is an important matter. A note of caution about these results is necessary. The number of children from lower SES in schools with a high average SES is relatively small and there are no children from higher SES background in the schools in the lowest quartile of average SES. It is possible that the attenuation of the connection between individual children's SES and mathematics reasoning in the schools with the highest average SES is not due to the schools themselves but to a selection of children from lower SES in the admission process.

In summary, social background affects children's mathematical reasoning at two different levels. (1) There is a strong relationship between individual children's social background and the Mathematical Reasoning tasks in Years 4, 6 and 8. The higher the children's SES, the better those children tend to reason mathematically. Multilevel analyses established that this effect at the individual level is a significant one, even after controlling for the influence of differences between schools that the children attended. (2) There is also a strong relationship between the socio-economic composition of the different schools and the mathematical reasoning scores of the children in them. Children in schools with a high average socio-economic level do better in the mathematical reasoning tasks than children in schools with a lower overall socio-economic composition. This effect at the level of schools is significant even after the relationship between individual children's socio-economic background and their mathematical reasoning scores is controlled.

By the time children reach Year 8 at school, the effects of their individual social background on their mathematical reasoning appear to differ between schools. A moderate, but significant, interaction led us to conclude that the usual differences between pupils with high and low socio-economic backgrounds in schools do not apply to children in schools in which the social composition is generally high (the top quartile), even though they are clearly evident in schools with lower social compositions.

Key Findings on the effects of social background

- Individual children whose SES status is high are on the whole better in mathematical reasoning than those from lower SES homes, even if they are at the same school. The influence of SES at the individual level is similar to an effective educational intervention: it raises the average level of performance and reduces the variation between children, and consequently the proportion of children who show difficulty in mathematics.

- Children's SES influences how well they reason mathematically which in turn affects their Key Stage mathematics attainment. Mathematical reasoning is a significant mediator of the link between SES and achievement.

- SES also operates at the level of the school. Children, who go to schools in which the SES composition is high, reason about mathematics more successfully than children at schools with a lower SES composition, whatever their own social background.

4.2.6 The effect of gender on KS 1 - 3 attainment, mathematics reasoning and cognitive skills related to mathematics learning

We found no consistent differences between boys' and girls' mathematical achievement as measured by the Key Stage results (Table 18 in Appendix C). For the most part, the mean for the Key Stage levels in mathematics were identical for the two sexes. Boys did out-perform girls in KS2 Mathematics when the number of points obtained rather than KS level was used as measure of achievement, but the effect size (Cohen's d) indicated that this difference was not important: it was less than 0.2 of a standard deviation. These results are entirely consistent with the reviews of gender effects on mathematics achievement reported in the introduction (e.g. Tate, 1997).

In contrast, there were definite gender differences in mathematical reasoning (Table 19 in Appendix C). In all three tasks there was a significant difference in favour of boys. However, the effect sizes (Cohen's d) indicated consistently that this is not an important difference. In the Year 4 Mathematical Reasoning task, the difference was equal to 0.08 standard deviations; in the Year 6 task, the difference was equal to 0.23 standard deviations, and in the Year 8 task it was equal to 0.17. Thus, although the difference is not due to chance, it is certainly not a cause for concern at Year 4. It is puzzling that the difference seems to increase over time, even though it remains small.

We also looked at the cognitive skills measured by the WISC which are relevant to mathematics attainment. The mean results for these comparisons are presented in Table 20 in Appendix C.

Boys out-performed girls in Arithmetic, Block Design and Object Assembly, but the difference was not significant for Object Assembly. The effect sizes (Cohen's d) for Arithmetic and Block Design were small as they were less than 0.2 of a standard deviation. Girls significantly out-performed boys in Coding and Digit Span; the effect size for Digit Span was small (0.2 of a standard deviation) and moderate for Coding (0.45 of a standard deviation). This latter result does give some cause for concern in view of the finding that the standard deviation for boys was also larger than that for girls. This means that more boys than girls might be at risk for learning due to attention problems.

In only one of these analyses the interaction between gender and social background was significant (father's SES and gender interacted significantly when the dependent variable was Object Assembly). In the context of so many analyses and a non-replication when the other measures of social background were used in the analyses, we view this result with caution and perhaps a product of chance because so many analyses were carried out.

In conclusion, the analyses of gender effects showed only one important difference, which might call for further consideration and perhaps intervention. Boys showed a lower performance than girls in attention (Coding Test), which could certainly place some boys at risk for learning mathematics.

4.2.7 The effect of ethnicity on KS 1 - 3 attainment, mathematics reasoning and cognitive skills related to mathematics learning

The possibility of analysing effects of ethnicity on attainment, mathematics reasoning and cognitive skills in the ALSPAC data is very limited due to the small number of participants from other ethnic backgrounds than white. The groups represented in the data file which had more than 5 participants were classified as White, Black Caribbean, Other Black, Indian, Bangladeshi, and Chinese.

Before running the analyses to assess effects of ethnicity, we identified in the sample all the schools that had at least one participant from an ethnic minority group; only these schools were included in the analyses. This allowed us to control for school differences and still retain a large enough number of participants from the majority group in the analyses. Thirty-four schools had at least one participant from a minority ethnic background.

Next we ran analyses of co-variance, in which we controlled for mother's education,[5] and investigated the means and standard error for each of the ethnic groups in KS Maths attainment, mathematics reasoning and the subtests in the WISC which relate to mathematics attainment. These analyses indicated a significant effect of ethnicity at the .05 level on the children's performance in the mathematical reasoning task given in Year 4 and the measure of Spatial Skills (which included Block Design and Object Assembly), in the WISC.

In order to understand the nature of this effect better, we separated out the results for boys and girls and in both cases the difference remained significant for boys but not for girls. Some power is lost when the analysis is split by gender, because the number of participants decreases, but the number of participants in the analyses remained high enough for meaningful effects to be identified.

These analyses identified two significant ethnic differences among boys but none for girls as a function of ethnicity. The results for boys are presented in Table 2. The group of white boys was the comparison group in the analyses because it has the largest number of participants and thus allows for the identification of significant differences where they exist. Boys from the Black Caribbean and the Black Other groups had significantly lower scores than the boys in the White comparison group in the Year 4 Mathematical Reasoning task.

The significance of this difference in mathematics reasoning measured in Year 4 is that this measure was a strong predictor of KS2 and 3 Mathematics and that this analysis controls for mother's education. It is quite possible that the effects would be stronger for this group without such controls, as the children might be exposed to two risk factors at the same time. The educational implication of this finding is that schools should be particularly aware of the importance of promoting mathematical reasoning among these two groups of black children early on in their educational lives.

[5] Mother's education rather than the factor scores for social background was used as the covariate in these analyses because of loss of data when the factor scores were used as the covariate. Note the remark by Lockheed et al. (1985), referred to in the introduction, about the need to control for SES in the analysis of effects of ethnicity on mathematics achievement.

Table 2 - Adjusted means (controlling for mother's education), number of participants, and standard error of the means for the different ethnic groups in Mathematics Reasoning measured in Year 4 [6]

Ethnic Group	Mathematics Reasoning – Year 4		
Boys only	**n**	**Adjusted means**	**Standard error of the mean**
White	541	11.20	0.13
Black Caribbean	16	9.50*	0.74
Black other	9	9.17*	0.99
Indian	6	12.17	1.21
Chinese	7	11.28	1.13

* Differences significant at the .05 level.

4.2.8 The effect of streaming on mathematics reasoning and on cognitive skills related to mathematics learning

It has already been found in other studies that streaming affects children's attainment in Mathematics (see Slavin, 1990, for a summary of previous research). In our analyses we considered two other outcome measures, mathematics reasoning and the subtests of the WISC that measure cognitive skills related to mathematics learning, namely Arithmetic, Spatial Skills (measured by the sub-tests Block Design and Object Assembly), and Attention and Memory (measured by Coding and Digit Span).

Information about whether the school uses ability grouping for mathematics or not is available for 11,409 participants in the ALSPAC study; 7% of the teachers answered that the schools do not stream for mathematics teaching. In the sub-sample that we analyse here, which comprises participants who took the relevant measures, the same percentage of answers indicates that the schools do not stream. Thus in this respect the sample analysed here reflects the larger ALSPAC sample. Approximately 82% of the children in schools that do not stream for maths do not stream for English either and 77% of the children were placed in the same ability group for maths and English. This indicates that at least some of the schools use within-class ability grouping but we do not know which arrangement each school used. In all the analyses carried out here, we consider only the teachers' answers with respect to whether school streams for mathematics teaching and the ability group in which the child is placed in mathematics.

Preliminary analyses showed that children's attainment in KS1 Maths accounted for 47% of the variance in children's assignment to the bottom, middle or top level stream. *Gender and Social Background did not make a contribution to the assignment to a stream after controlling for KS1 Maths results*. This is an interesting finding in view of Boaler's (1997) hypothesis that children from lower SES and girls were more likely to be assigned to the lower stream. We did not find evidence to support this bias in assignment in the ALSPAC data. Therefore, all the analyses reported here control for the children's KS1 Mathematics attainment but not for gender or social background. If an effect of streaming is found, this cannot be attributed to the children's mathematical ability as measured by KS1 attainment.

[6] When the number of participants was smaller than 5, the group was excluded from the analysis.

We compared children in schools without streaming to children in the different streams (Top, Middle, Bottom) in schools in which there was streaming in mathematics. We looked at the mathematical reasoning scores of these four groups in the Year 4 and Year 6 Mathematical Reasoning tasks. The clear result in both comparisons (see Table 21 in Appendix C) is that schools that do not stream the children for mathematics are able to maintain the average attainment of the children in Mathematics Reasoning *above* that obtained by the middle group in schools that do stream the children for mathematics teaching. This result is important because it cannot be explained by the children's KS1 attainment as this was used as a control in the analyses. Thus, in schools that streamed, the children, both in the bottom and in the middle group, seem to be offered fewer opportunities to develop their mathematics reasoning than children in schools that do not stream. Children placed in the top group in schools that stream performed better in the Mathematics Reasoning Task in Year 4 than those in schools that do not stream but this effect was small (Cohen's d = 0.18 of a standard deviation).

The pattern of the four groups' scores in the WISC cognitive skills that are related to mathematics attainment was much the same (see Table 22 in Appendix C). Children in schools that do not stream performed significantly *better* than those in the middle and lower group in schools that stream for mathematics teaching, and *as well as* those in the top group. Schools that do not stream are able, therefore, to bring the children to a level of attainment above that of the middle group in schools that stream for mathematics attainment.

Thus the pattern of results is slightly different for the skills that are not explicitly taught in school, namely spatial skills, attention and memory. The children in schools that do not stream perform as well as those in the top group in schools that do stream for mathematics teaching, and significantly better than the middle and the lower group. It is likely that, although these cognitive skills are not explicit aims of instruction, schools that stream for mathematics teaching may make different demands on the children across the streaming groups.

In summary, the analyses suggest that the schools that do not stream are able to bring the average performance of all children to a level above that of the middle group in schools that stream. The only children who seem to benefit from streaming are the children assigned to the top level but only in academic skills; this is not observed in the analysis of general cognitive skills that support this academic learning. These results are completely in line with previous research about mathematics attainment and its relation to streaming (see: Boaler, 1997; Slavin, 1990). Their implication is that schools that wish to promote further mathematics learning among their strongest pupils should seriously consider measures alternative to streaming.

Key Finding

- Streaming only improves mathematical achievement of children in the top stream, but this effect is small. It actually hinders the progress of children in the middle and bottom stream in terms of the development of cognitive abilities that support mathematics learning.

4.2.10 Self-confidence in learning mathematics

In the introduction we discussed the contribution that large scale longitudinal studies can offer to clarify the nature of the connection between self-confidence in maths and competence. On the one hand, it is possible that low self-concept as a mathematics learner is a result of low achievement. On the other hand, it is just as plausible that the reverse is true: low self-confidence results in low achievement. In fact, both hypotheses may be true: achievement and self-confidence may reinforce each other over time. Longitudinal studies can shed light on which of these alternatives is consistent with children's development.

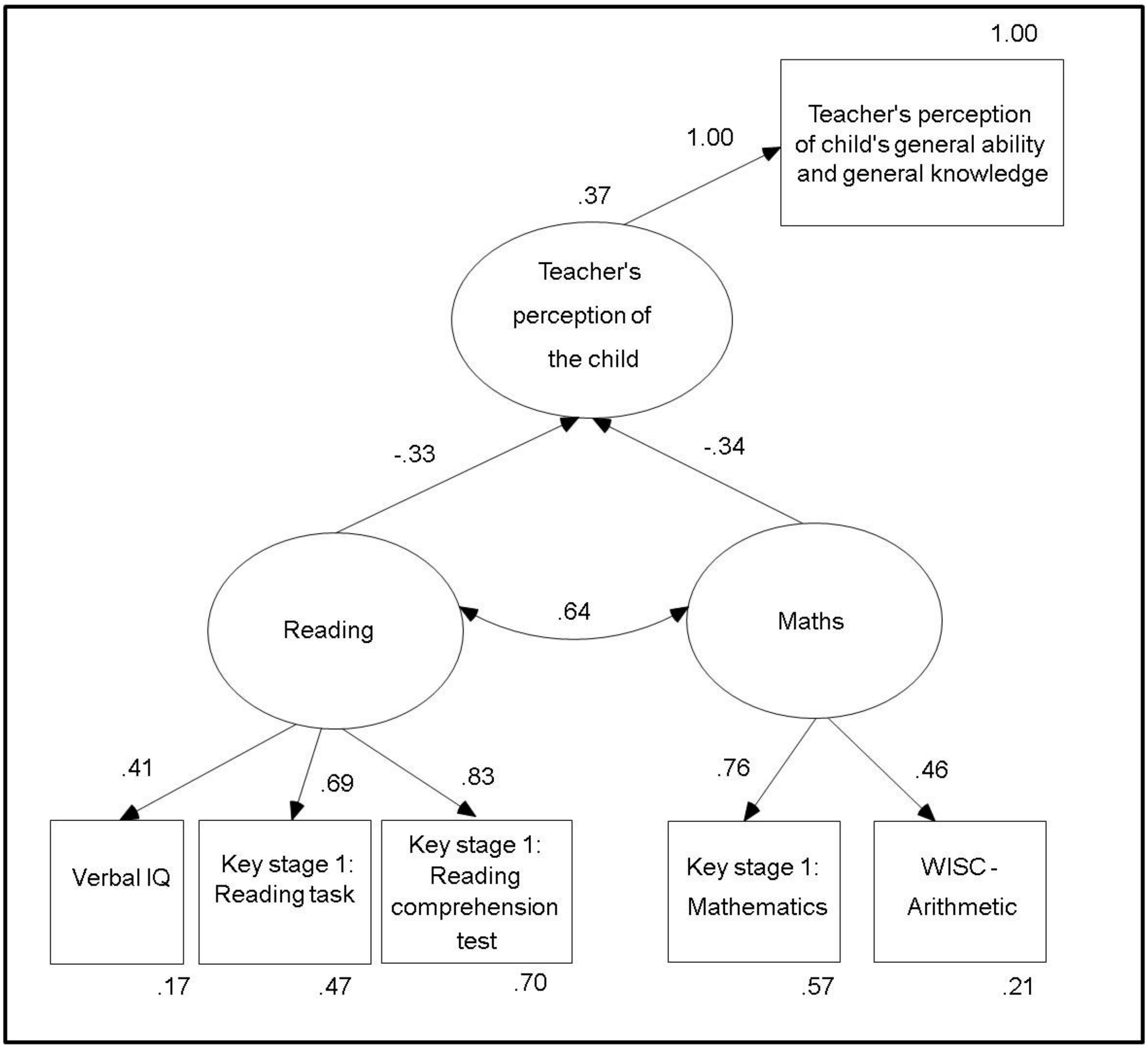
We investigated in the analyses regarding children's self-confidence in maths the factors that predict self-confidence and the long-term effects of self-confidence on mathematics achievement.

As described in the methods section, self-confidence was treated as a broader concept, measured through the children's liking of maths and their self-perception as learners of maths. A single factor, based on a principal components analysis, was used in the regression analyses and analyses of variance.

The children's competence in mathematics was also treated as a latent variable, measured through KS1 Maths achievement assessed in Year 2, performance in the WISC Arithmetic assessed in Year 3, and the scores in the Year 4 mathematics reasoning. These three measures show significant inter-correlations; a principal components analysis shows that they form a single factor and can be treated as measuring the same concept, which we will call mathematical competence. The factor loadings of these measures were all high: .85, .80, and .86, respectively. So they can be seen as assessing the same construct, mathematics competence.

Past research has shown that others' perception of the children's ability has an impact on their self-perception. The ALSPAC data set includes some measures of the teachers' perception of the children, obtained in Year 3, but the questions relate to the children's general ability and knowledge. We analysed whether the teachers' perception of the children's general ability was related to children's mathematical competence to test whether this factor might be relevant to the children's self-confidence in maths. Because past research (Pretzlik *et al*., 2003) has shown that teachers' judgements of children's general ability are more influenced by their verbal than their mathematical ability, we used a broad measure of the children's verbal ability in this analysis too. We obtained an estimate of the children's verbal competence by combining three measures, their achievement in the two reading tasks in KS1 and their verbal IQ. The estimate of the children's mathematical competence for the analysis of the teachers' judgements was based on their KS1 attainment and the WISC Arithmetic because the Year 4 Mathematics Reasoning assessment was administered after the teachers had answered the questions about the children's general ability. A regression analysis showed that both types of competence, verbal and mathematical, contributed to explaining variance in the teachers' judgements of the children's ability and had similar importance. Figure 16 displays this information graphically. So we decided to include the teachers' judgement obtained in Year 3 as a predictor of the children's self-confidence.

Figure 16 - Structural equation model showing the influence of maths and reading/oral ability on teachers' perceptions of children's general ability (1355)



There were no direct measures of peer's perception of the children's ability but past research (Boaler, 1997) shows that children are aware of ability groupings in the classroom. Streaming can influence the children's self-confidence directly and also indirectly, through the perception that their peers have of them. Because these two possible impacts of streaming cannot be separated in the ALSPAC data, we analyse the effects of streaming *in lieu* of peer's perception. We first analysed whether the practice of streaming had an influence on the children's self-perception or not; then, in those schools that stream children for mathematics teaching, we analysed whether the stream they were assigned to had an influence on their self-confidence.

It has been argued that girls have lower self-confidence in maths even if they have the same level of measured competence as boys. It was found in the previous analyses that there were no important gender effects in KS1 Maths achievement and mathematics reasoning, even though some statistically significant differences were observed. We investigated here whether gender explains individual differences in maths self-confidence as well as liking of other subjects in school. The children were asked how much they liked different things that they did in school in Year 3. A factor analysis showed that three factors described how the different subjects were seen by the children. First, children's liking of Science and Natural

History, Geography, and History were highly inter-correlated. Second, their liking of English, another language, Arts and Music were highly correlated with each other. Third, children's liking of maths was correlated with their liking of sports. This latter finding did suggest that maths and sports might be seen as "boys' territory". So we explored whether gender explained individual differences in self-confidence in maths and also in how much the children liked other subjects.

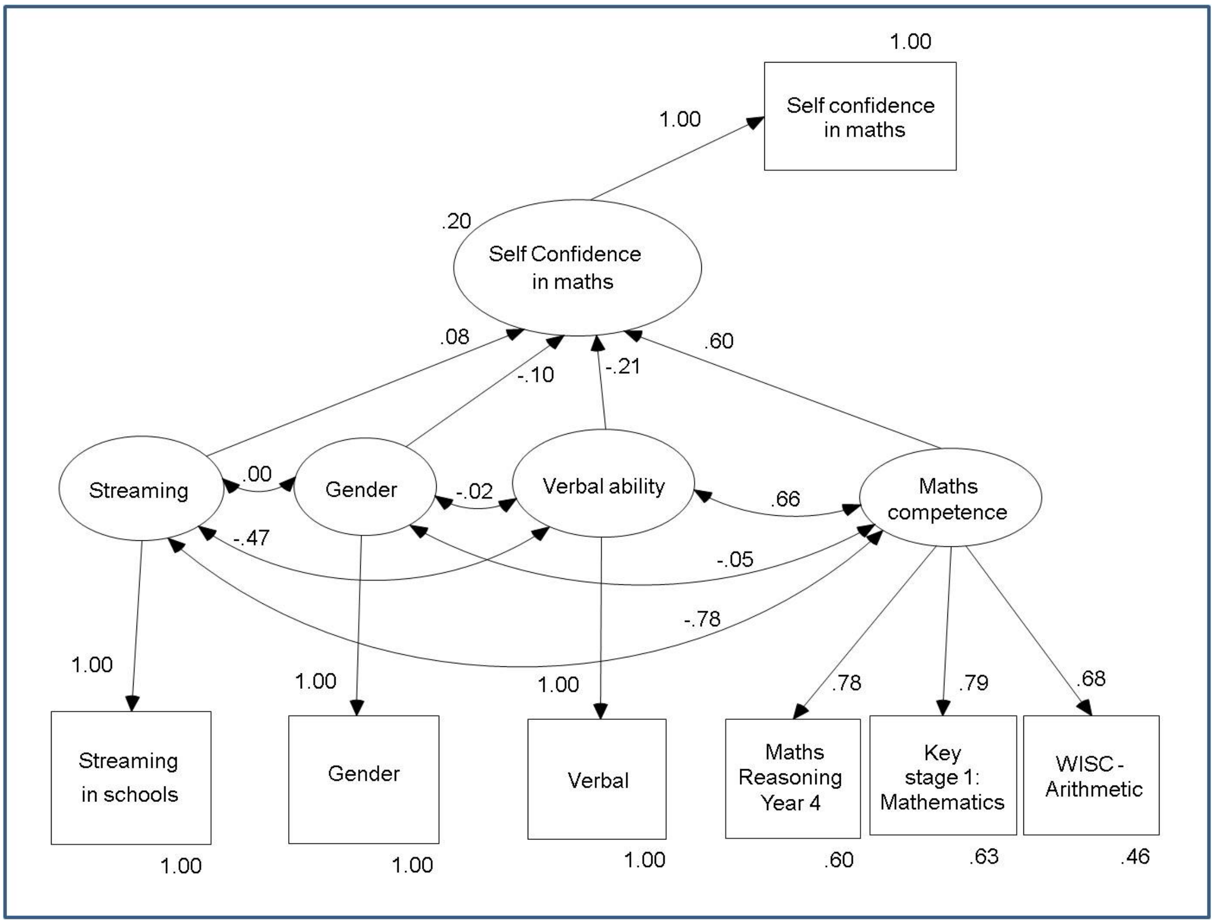
Finally, past research, reviewed in the introduction, suggests that, although verbal and mathematical ability are correlated, people seem to judge themselves in relative terms in these two domains (e.g. Marsh, 1986). Those whose verbal ability is higher than their mathematical ability think of their mathematical ability as low, even if they do well in mathematics assessments. So, we also explored the connection between children's verbal ability and self-confidence in maths.

The preliminary analyses showed that streaming *per se* did not have an effect on the children's self-confidence: the level of self-confidence expressed by the children in schools where they are streamed for maths teaching did not differ on average from that expressed by children where they are not. Therefore, in the analyses of effects of streaming, only schools that do stream for mathematics teaching were included.[7]

We investigated whether the five factors listed here (maths competence, verbal ability[8], teacher's perception, gender and stream to which child is assigned) contributed to explaining individual differences in self-confidence in maths by performing an exploratory stepwise regression analysis in order to identify the model that explained the largest amount of variance in children's self-confidence with the smallest number of variables from this set of five. This analysis led us to discard the teachers' perception of the children's general ability from subsequent analyses because it does not explain any variance in the children's self-confidence in maths independently of the child's mathematical competence. The other four variables explain some variation independently of each other. Figure 17 shows an overview of the findings, using a structural equation model. This model takes into account the interactions between the four factors as well as the importance of the factors by themselves. The amount of variance in the children's self-confidence that was explained by these four factors was modest: a total of 20%.

[7] The number of participants in the analysis is still large (N=1044) after this exclusion, so the loss of participants is not important for the results.
[8] Verbal ability was measured by five sub-tests of the WISC (Information, Similarities, Vocabulary, Comprehension and Digit Span), and excluded Arithmetic, which is one of the original subtests in the Verbal Ability Scale.

Figure 17 - A structural equation model that shows the relation between the children's self-confidence in maths and the different factors that explain independent variance in their self-confidence (N=1022)

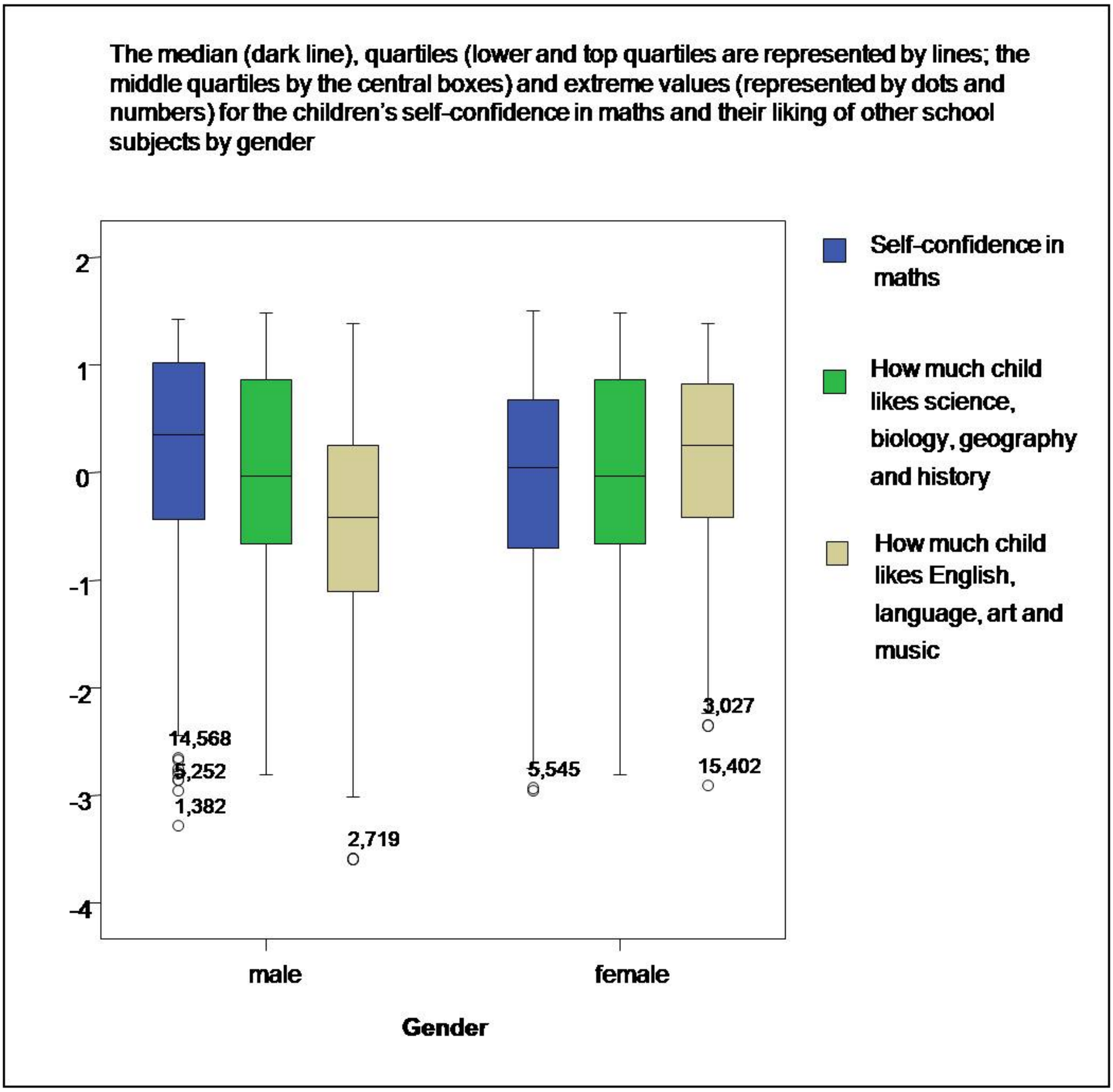


Inspection of the standardised regression coefficients, which are represented by the numbers next to the arrows that connect the factors to the measure of children's self-confidence, shows that the most important factor is the children's competence in mathematics. This has a positive value of .60, and indicates a strong effect. So the better children actually are in maths, the more confident they are.

The next factor in order of importance was the children's verbal IQ. It showed a negative but modest impact on the children's self-confidence in maths: the coefficient was -.21. This result replicates findings from the literature and suggests that children rate themselves relatively: the better they do in verbal tasks, the less confidence they have in their mathematical ability.

Gender also had a significant impact on the children's self-confidence: girls were less confident than boys. Although this impact was small (the coefficient is -.10), it was significant. Figure 18 displays the median and the distribution of scores for boys and girls in the measures of their self-confidence in maths and their liking of other school subjects.

Figure 18 - Children's self confidence in maths and liking of other subjects

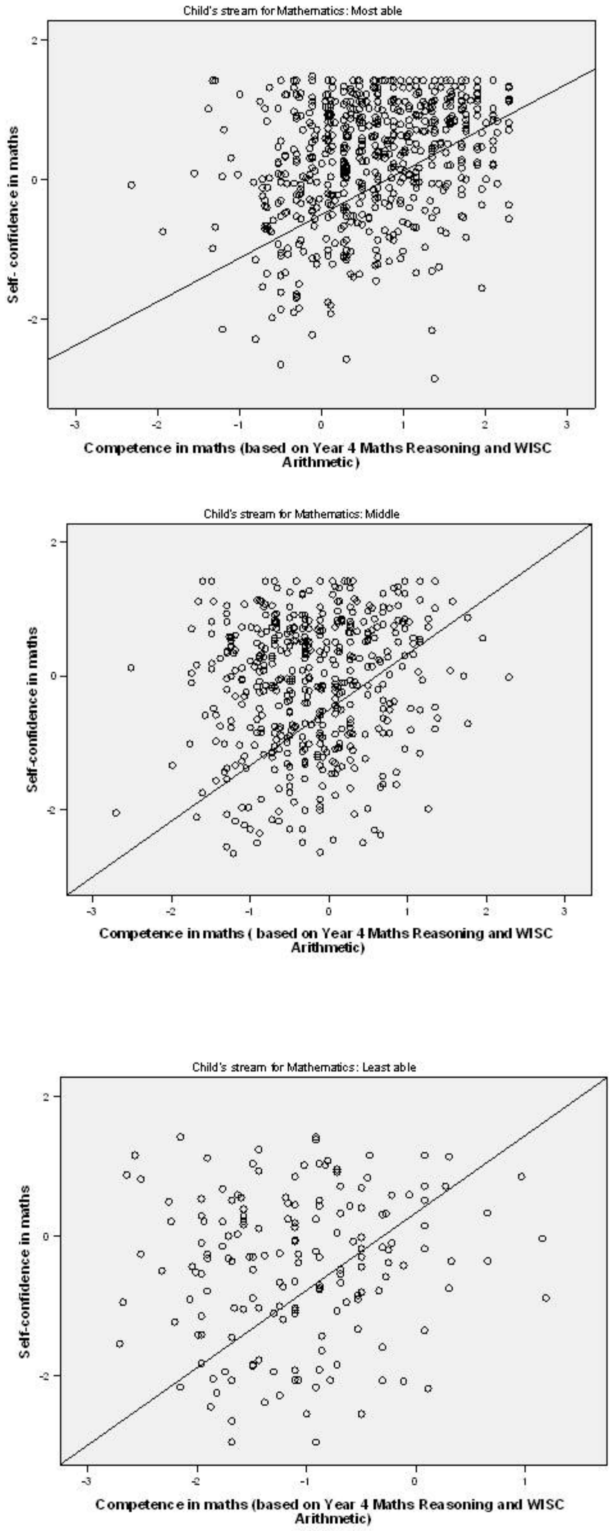


School subjects that were strongly correlated are considered together in one measure. All three measures are displayed on the same scale, where 0 is the median, positive values indicate a positive attitude and negative values a negative attitude. The graphs show that girls and boys differ on all three measures. Boys show greater self-confidence in maths; girls show greater liking of the other school subjects. The difference is particularly large between boys and girls in liking English, a second language, arts and music. These analyses were followed up with analyses of co-variance, in which we controlled for the children's ability and compared their self-confidence and liking of the school subjects. All analyses showed significant differences between boys and girls (at the .001 level). Thus the differences in self-confidence in maths and liking of the other subjects could not be explained by differences in ability, as they persisted in the analyses of covariance.

The Cohen's d effect sizes indicated that the difference between boys and girls in liking of English and the related subjects was quite large: 0.72 standard deviations. The other differences were small (0.23 standard deviations for liking of science and related subjects and 0.21 for self-confidence in maths). These results suggest that the most important differences between boys' and girls' attitudes towards school subjects are in the negative attitudes that boys seem to have towards English, foreign languages and arts. The differences between girls and boys in the other subjects are small, even though significant statistically. None of these differences can be explained by differences in ability. We can only speculate that they may reflect cultural stereotypes and family attitudes but we have no evidence for this.

Finally, the child's stream also impacted on the children's self-confidence: the effect was small but significant. It indicated that the children in the lower stream were less confident than those in the higher streams. The effect of streaming may operate through peers' perception but streaming may also have a direct influence on the child's self-confidence. Figure 19 shows the self-confidence of children assigned to the top, middle and bottom streams as a function of their competence. Inspection of the graphs shows that there were children with average levels of ability in all three streams. If we concentrate on those whose ability was at the mean, which is zero in this scale, or up to one standard deviation above the mean (i.e. between 0 and 1 on the graph), it is quite easy to see that most of the children with this level of ability in the bottom stream are below the regression line; in the middle stream they are distributed below and above the regression line; and in the top stream there are many more above than below the regression line. This indicates a negative effect of streaming on children's self-perception if they are in the bottom stream and a positive effect for the children in the top stream.

Figure 19 - Graphs showing the relation between the children's competence and their self-confidence in maths for each of the streams in mathematics teaching.



It should be remembered that there is a high correlation between the children's competence in mathematics as defined by their KS1 results and their assignment to different streams in Year 3. So the differences in self-confidence could stem largely from a direct effect of ability on self-confidence. In order to assess whether streaming has an effect beyond ability, we created a score for the children's competence that does not include the KS1 Maths achievement, based on the children's mathematical reasoning assessed in Year 4 and on their performance in the WISC Arithmetic sub-test. This factor is highly correlated with KS1 results but is an independent assessment of the children's mathematical competence. It was important to have an independent assessment because the KS1 Maths results might have been used to assign the children to a stream. We then ran an analysis of covariance, in which we controlled for the children's competence as measured by this factor, and compared children in different streams. This analysis showed a significant effect of streaming (at the .001 level) on the children's self-confidence in maths. This analysis confirms what is suggested by inspection of Figure 19: children in the bottom stream are considerably less self-confident than those in the middle and top streams even if they show the same ability on the measures used in ASLPAC.

In conclusion, of the five factors that were hypothesised to explain variance in individual children's self-confidence in maths, four had a significant and independent impact. The children's measured competence had the highest impact; this was followed by verbal ability and gender, both of which had a negative and similar impact, and finally by streaming. Verbal ability had a low negative correlation with self-confidence in maths. Girls displayed lower self-confidence than boys, even after controlling for their ability. Streaming had a negative impact on children who, despite showing an average level of competence, had been assigned to the bottom stream for maths.

4.2.11 School effects on the children's self-confidence

Past research (e.g. Marsh, 1991) indicates that high-ability students in schools where the overall ability is high have less positive views of themselves as learners than students of comparable levels of ability in schools with a lower average ability. This effect is known as the "big-fish-little-pond" effect and has been replicated in the UK by Tymms (2001), who found it to be a small effect. We analysed whether this effect could be replicated in the analysis of the ALSPAC data.

A preliminary analysis of variance was carried out on the proportion of variance in children's self-confidence that is explained by between-school differences. This showed that only 3% of the variation in children's self-confidence could be attributed to between-school differences, which is a small amount.

We followed up this analysis by creating four quartiles of school-level mathematical ability: i.e. we estimated the average level of mathematical ability in each school and placed each school within a quartile, bottom, middle-low, middle-high, and top. We then carried out an analysis of covariance to see whether students with the same level of ability would show different levels of self-confidence depending on the average level of mathematical ability in their school. In this analysis, we controlled for mathematical ability and tested whether children in the different school-quartiles differed significantly from each other. The analysis produced a negative result: i.e., we did not find the predicted "big-fish-little-pond" effect. There was no sign that children in the schools in the top quartile had lower self-confidence in maths than those children in schools in any of the other quartiles, when competence was controlled for.

Thus, there are modest school effects on the children's self-confidence in maths, but we have no evidence that these could be attributed to differences between the schools in their average level of ability.

4.2.12 Self-confidence in maths as a predictor of maths achievement

The aim of the previous analyses was to understand what explains differences between children in their self-confidence in maths. The major factor in explaining differences between children in self-confidence was their estimated competence.

The analyses carried out in this section seek to understand whether self-confidence in maths, in return, also affects children's achievement. If we find that self-confidence predicts the children's later attainment, providing information that goes above and beyond what we know from the children's competence, we can conclude that self-confidence and competence exert mutual influences on each other. Thus improving self-confidence should lead to improved attainment for the same level of ability.

In the previous analyses, we established that schools explained a small amount of variance in the children's self-confidence in maths. For this reason, we will use simple regressions in order to test whether self-confidence and competence in maths explain independent variance in children's attainment.

The measures used as predictors in the analyses will be:

- the factor created to describe the children's self-confidence (composed by the degree to which the child likes maths, measured in year 3, and self-confidence in maths, measured in Year 4); and

- the factor created to describe the children's competence (composed by their KS1 Maths achievement, their performance in the WISC Arithmetic sub-test and in the Year 4 assessment of Mathematical Reasoning).

The outcome measure is attainment in KS2 and KS3 Maths [9]. KS1 Maths cannot be used as an outcome because the predictors were measured after the children had already been through the KS1 assessments.

We tested whether self-confidence and competence explain independent portions of variance in predicting KS2 Maths by using two fixed-order regressions. In the first, we entered self-confidence first and then competence; in the second, we reversed the order of these predictors.

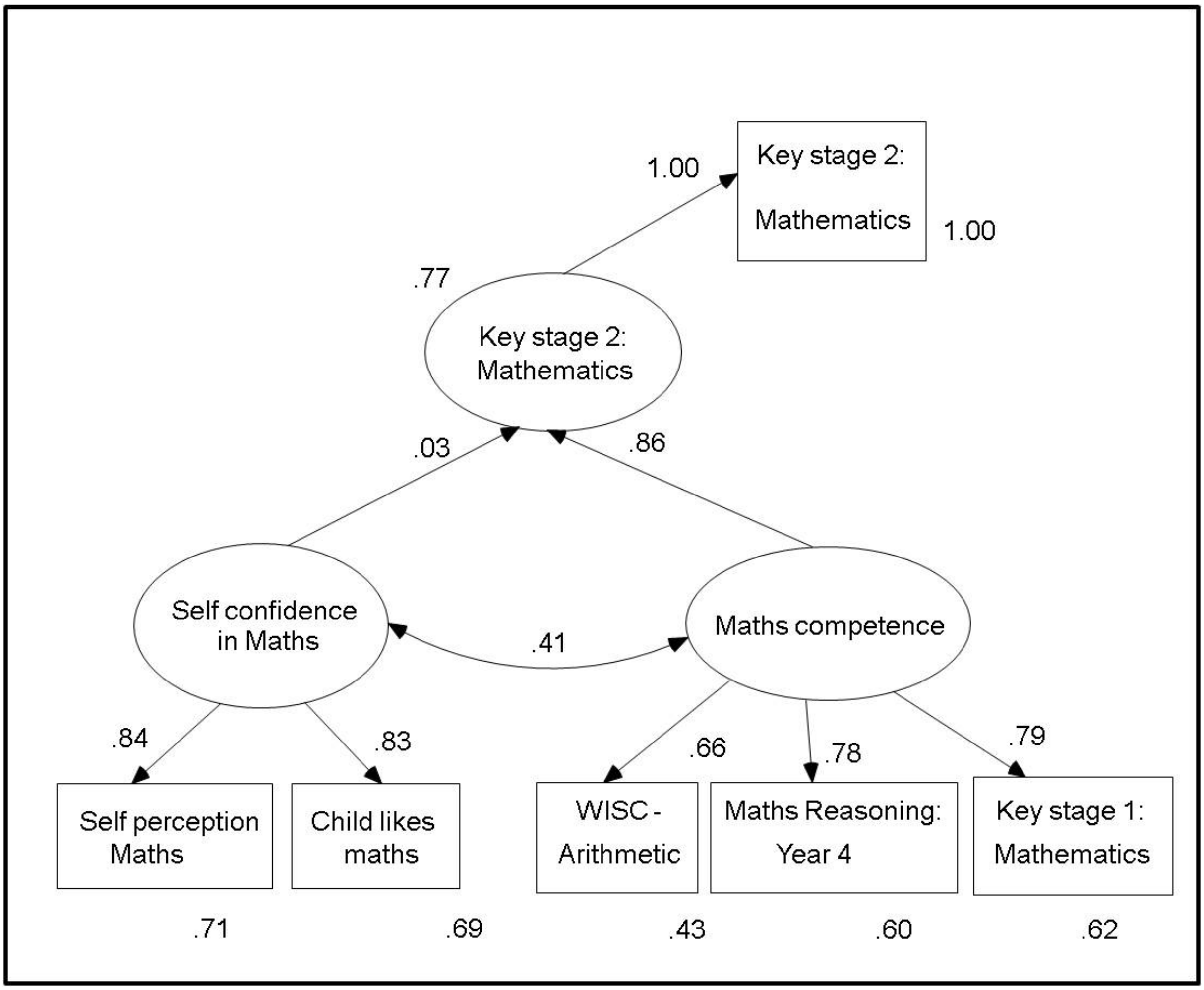
When self-confidence was entered first in the equation, it explained 13.5% of the differences between the children in KS2 Maths attainment; the children's competence explained a further 42.5%. When the predictors were entered in the reverse order, the children's competence explained 53% of the differences between the children and their self-confidence explained a further 3%. This may seem a small amount of variance but it is independent of the children's competence and statistically significant (at the .001 level).

The coefficients β provide an indication of the relative importance of these two factors, competence and confidence, for later achievement. They show that competence is certainly the more important of the two: the coefficient for the regression equation for competence was .678 and for self-confidence it was .180.

[9] We carried out the analyses with National Curriculum Levels attained in KS2 and also with the finer measure of total marks in maths. The results did not differ so we report here only the analyses using National Curriculum Levels.

Figure 20 shows these results graphically for KS2 Maths. The figure was obtained through a structural equation model using the same measures. The standardised coefficients, displayed next to the arrows connecting the predictors to KS2 Mathematics attainment, indicate that the children's competence has the strongest effect of KS2 Maths achievement. It should be pointed out that in this model the mutual influences between self-confidence and competence are taken into account, whereas this is not the case in a simple regression. Thus the coefficients differ in the structural equation model and in the simple regression: the importance of maths competence is actually greater (the coefficient goes up to .86) when these mutual influences are taken into account. The effect of self-confidence decreases correspondingly (it is .03 in this model). Nevertheless, the amount of variance explained in KS2 Maths increases to 77% when self-confidence is considered: this contrasts with 62% of variance (see Figure 4) explained in the same outcome measure when only cognitive factors were used as predictors.

Figure 20 - A structural equation model that shows the relation between the children's attainment in KS2 Maths, their self-confidence and competence in maths (N=1648)



Parallel analyses were carried out with KS3 Maths attainment as the outcome measure with similar results. When self-confidence was entered first in the analysis, it explained 13% of the variance and competence explained a further 42%. When competence was entered first in the analysis, it explained 53% of variance and self-confidence a further 2%. This addition was significant at the .001 level. It must be noted that KS3 attainment is measured about six years after the measure of self-confidence had been given to the children (see the timeline in Figure 3). So self-confidence has a long-lasting effect on children's attainment. An analysis of the coefficients β in the regression equations showed, once again, the robustness of this finding: the coefficient for competence was .682 and for self-confidence was .146. Due to the similarity between the two results for KS2 and KS3, we did not include a figure to show the results for KS3 graphically.

We started the analysis of children's self-confidence with a question about the link between achievement and self-confidence in maths. Does achievement lead to greater self-confidence in children or does self-confidence have an effect on the children's later achievement?

This set of analyses allowed us to find, without a doubt, that the link between self-confidence and competence is a two-way street. Children's self-confidence is influenced by their competence but their attainment in KS2 and 3 Maths, although largely determined by their competence, is also influenced by their self-confidence. This result is robust, as it was replicated across key stages, and also persistent, because the measure of the children's self-confidence was obtained about six years before they were given the KS3 Maths assessments. The implication is that it is very important to pay attention to the social aspects of children's mathematics learning: not everything can be explained by their cognitive abilities and their mathematical competence.

Key Findings

- Children's self-confidence in mathematics is predicted most strongly by their own mathematical competence but also, independently by their gender (girls are less confident than boys) and by streaming.

- Children's attainment in KS2 and 3 mathematics, although largely determined by cognitive and social factors, is also influenced by their self-confidence.

- So children's self-confidence and competence seem to reinforce each other over time.

5. Discussion

Our analyses of the ALSPAC data led to a large number of interesting and, in our view, important conclusions about the factors that determine how well children learn mathematics at school. Our first aim in this brief discussion, therefore, will be to summarise our different findings and clarify the relations between them. Our other main aim will be to explore the implication of these results for teaching children mathematics and also for future research on mathematics education.

5.1 The importance of mathematical reasoning

We start with the question, raised in the opening sections of the introduction, about the role of mathematical reasoning. There, we argued that there are good reasons for thinking that it is at least as important for children to be able to reason logically and flexibly about quantities and relations as to learn how to calculate and to do simple sums. We acknowledged, however, the absence of any firm evidence that mathematical reasoning is as important as that, and gave this as a reason for analysing the relevant variables in the ALSPAC data bank.

We have already explained the unique opportunity that ALSPAC offers to researchers interested in studying the bases for children's development longitudinally. Here, we will only add the comment that one of the great advantages of this data bank is that it includes two reliable mathematical reasoning tasks, in which all the items test children's ability to work out the relevant quantitative relations and all include only very simple calculations or no calculations at all. Thus, we could test the relation between children's successes and failures in these tasks and their progress in learning mathematics at school.

The children's scores in the mathematical reasoning tasks were consistently and strongly related to their achievement in mathematics. Not only were these the best of all the relevant predictors in the ALSPAC data bank: the strength of their predictions, also, was remarkably constant over time. The children's scores in the mathematical reasoning task given to them when they were 8-years-old accounted for nearly exactly the same amount of variance in the mathematics assessments at KS2 two years later and at KS3 five years later (Tables 3 and 4 in Appendix B) and the β coefficient for these two relations was virtually the same as well (.35 and .34 respectively). The scores in the mathematical reasoning task given to the children at 11-years actually accounted for more of the variance in the KS3 (3½ years later) than in the KS2 (only 6 months later) assessments.

The consistency with which the predictions made by the mathematical reasoning scores outstripped those made by any other available set of scores was also impressive, but we would like to dwell here on the comparison between the predictive power of mathematical reasoning and one other measure, the measure of Arithmetic. Our reason for concentrating on this comparison is that the two measures are in genuine competition with each other in the classroom as well as in our analyses. Teachers who provide less explicit instruction on mathematical relations usually concentrate on teaching how to calculate instead, and when teachers do devote more time to fostering mathematical reasoning, they undoubtedly do so at the expense of time spent on teaching calculation. Our analyses certainly showed that teaching children about calculation is not wasted time, since the ability to calculate proved to be a good predictor of mathematical success. However, the far stronger relation between mathematical reasoning and mathematical achievement is evidence that there is a genuine need for teachers to spend ample time ensuring that their pupils know what quantitative relations are and how to reason about them logically and enterprisingly.

Mathematical reasoning emerged as an important conduit between children's social backgrounds and their mathematical achievement. Children from prosperous social backgrounds did well in the key stage mathematics assessments because they could reason well mathematically. Because the mathematical reasoning by children from less prosperous backgrounds was less effective, they were less successful in the key stage assessments too.

The pattern of relations between the mathematical reasoning scores and the Mathematics, Science and English key stage assessments confirms the central importance of these tasks. By showing that the children's mathematical reasoning scores were more strongly related to the children's key stage assessments in Mathematics than in Science and English, we have also established that the mathematical reasoning tasks are valid, as well as reliable, instruments for measuring children's mathematical progress. Our discovery that the reasoning tasks predicted Science assessments much better than English assessments also seems to us be a result of prime educational importance. Its strong implication is that mathematical reasoning plays an important part in learning science; this is most likely to result from the ubiquitous need for children to reason about quantitatively scientific concepts that involve relations (such as density) and about relations between variables. It follows that the teaching of these two subjects should be a great deal more integrated than it is now.

5.2 Other cognitive variables

It has often been suggested that various other cognitive skills, not in themselves obviously mathematical, nevertheless play an important part in children's mathematical development, but the evidence for these hypotheses has on the whole been unsatisfactory for the reasons that we gave in our introduction. We had the opportunity to investigate some of these hypotheses with the help of the ALSPAC data bank.

One quite plausible suggestion is that children's mathematical progress is affected by how skilled they are in processing and thinking about spatial relations. Geometry, which is about mathematical representations of space and spatial relations, is an important part of most children's mathematical education, and this at least must make considerable demands on children's spatial knowledge. Since the ALSPAC data included two standardised spatial tasks, we were able to examine the contribution of this knowledge to mathematics achievement, and we found some evidence for a moderate but consistent connection between children's spatial ability, measured when they were 8-years-old, and their mathematical achievements later on.

One interesting point to emerge from these analyses was that one of the spatial tests predicted mathematical achievement more strongly and in a more sustained way than the other. The children were required to manipulate regular, abstract, geometric shapes in the more successful task and irregular shapes, which when combined correctly added up to a meaningful figure (a face, a ball), in the less successful task. The implication of this difference is that the connection between spatial scores and mathematics may be entirely due to the rather specific spatial demands made by school geometry, which concentrates on regular shapes rather than objects used in everyday life. This however is a hypothesis - one that badly needs testing. The hypothesis is supported by the fact that the successful spatial task predicted children's mathematics achievement at 14-years, five years after the task was administered, better than at 11 years, about two and a half years after spatial test. On the whole children spend more time on geometry at 14- than at 11-years, and this may be why a spatial test given to 8-year olds predicts their mathematics achievement better at 14- than at 11-years. However, this is only indirect evidence for a link between spatial skills and learning geometry: we need a lot more research on this possible connection.

The other two cognitive measures, of memory and of attention, are usually treated as information processing tasks. It is of course quite possible that children's ability to deal with incoming information does affect how well they learn mathematics, and could account for some of the variation in mathematical achievement, and this is an idea that has been expressed many times, particularly about working memory. In fact, our multiple regressions did produce some evidence of a quite modest connection between working memory and mathematics achievement and a slightly stronger connection in the case of Coding, our most direct measure of attention. However, when we formed a latent variable, which we called Attention and Memory, on the basis of the two memory measures, Forward and Backward Digit Span, and the Coding task, we found a fairly strong connection between this new variable and children's mathematical achievement in a Structural Equation Model. The latent variable, Mathematical Reasoning, was, as usual, an even better predictor, but nevertheless the Attention and Memory variable did make an impressive contribution. This result, therefore, supports the idea that children's ability to process information also affects their mathematical progress. It should prompt research on the effect on children's mathematical achievement of improving children's memory and attention with the help of existing tried and tested intervention methods.

We have a final comment to make about the cognitive measures. The clear pattern that has emerged is of several quite different cognitive abilities making separate and independent contributions to children's successes and failures in mathematics. There certainly is a pecking order in the relative strength of these contributions, with Mathematical Reasoning at the top and Attention and Memory near the bottom, but it would be wrong to dismiss any of the abilities that we have considered as too negligible to bother with. The pattern of multi-determination is in itself important, because it does not fit well with hypotheses that the serious problems that some children have with mathematics are due to the impairments in just one underlying ability, such as working memory or calculation skills. Our analyses suggest that there probably is no one golden bullet here: several factors determine children's successes and their failures in mathematics.

5.3 Social factors: social background

The transition from cognitive to social factors in our analyses was nearly seamless because we used our main cognitive variable, mathematical reasoning, as a tool to find out how children's social backgrounds affect their mathematical achievement. We had expected a relation between children's background and their mathematical achievement, and we did find an almost linear relation between these two variables, but in our view when you find such a relation, you must also explain the pathway between them. A person's social background affects many aspects of his or her life, and our question was which of these aspects in turn leads to the individual doing well or badly in mathematics. Our attempts to find particular items in the 'cultural capital' provided by the children's family environment, which might play this intervening role, came to nothing. In contrast, as we have mentioned already in this section, when we turned to our cognitive measures we were able to produce a convincing model of the relationships between social background, mathematical reasoning and mathematical achievement. In this model, mathematical reasoning provides the pathway, and is therefore the main reason why social background has such a strong effect on children's mathematical learning. This raises another question. The question is whether schools can themselves provide more focused instruction on mathematical reasoning and through their own intervention promote greater equality in mathematical attainment across SES groups. Hypotheses such as this one should be urgently investigated.

We pursued the relationship between children's social background and their mathematical reasoning in our attempts to disentangle the effects of children's social backgrounds at two different levels: at the level of the individual child and at the level of the school the child happens to be in. Of these two levels, the level of the school probably needs the most explanation. It is in principle possible that the overall SES composition of the different schools could affect the pupils' mathematical progress. Children who attend schools where most of the children come from prosperous homes might do better than those in schools where the majority of pupils are from less prosperous homes simply because of the different environments provided by those different schools.

The multilevel models that we constructed showed quite clearly that there were strong effects at both levels on the children's mathematical reasoning. Their own social status played an important part, but so did the social composition of their schools. As far as we know, this is a new result and in our view an intriguing one. Since SES differences in the children's school environment, as well as in their home environment, play a role in their mathematical reasoning, the question that we asked earlier about SES effects on mathematical reasoning becomes a more complicated one. We need to know how the environments provided by schools with high and low SES compositions differ, and which differences influence the level of children's mathematical reasoning. We also need to know whether these are genuine school effects or whether they may be a result of selection in the recruitment in different schools. Once again, intervention studies that aim to promote the development of mathematical reasoning would help distinguish between these two possible interpretations of the school effects.

5.4 Streaming and self-confidence

One difference between school environments is their policy on streaming. Some schools in the ALSPAC sample put children into different streams in their mathematics classes: others did not. When we compared children's mathematical achievements in these different kinds of school, we made sure to compare the children in the different streams with the children in the non-streamed schools separately. We found, as others found before us, that streaming only helps the children placed in the top stream. The rest are mostly hindered by streaming in the development of their mathematical reasoning.

We also found that the self confidence of children in the middle and bottom streams was lower than that of children of the same ability levels in the top stream. This is important because our analyses of the effects of self confidence show a relation between how self confident a child is and how well that child does several years later in the key stage assessments.

Self-confidence itself was consistently an important factor. We found that children's previous success in the Key Stage 1 mathematical assessments predicted their self confidence very well. Undoubtedly their confidence is largely built on past successes or failures in mathematics. However, we also showed that the children's self-confidence made a modest contribution to their success in later key stage assessments, even after controls for the effects of differences in their level of attainment in the previous key stage assessment. Thus, children's own feelings about how good they are at mathematics play an independent part in how well they do in mathematics. The strong general implication of this result is that affective factors are important too. The result also implies that teachers should find try to ways of boosting children's beliefs in the quality of their mathematical thinking.

5.5 Implications for education and for future research

We will conclude our report by summarising what we see as the best, next steps to take, given the results that we have presented. The educational implications of these results are clear. We have identified certain abilities that are essential to good progress in mathematical learning. The next step should be to find out how well schools already foster and encourage these abilities, and to develop ways for schools to do so where research does suggest ways forward.

We can take two such abilities, mathematical reasoning and calculation, as an example. It is highly likely, and the content of the national primary strategy for teaching numeracy supports the conclusion, that schools already do quite enough to teach children about calculation, but perhaps not nearly enough to show them how to reason about quantitative relations. If this is the case, steps should be taken to devote more teaching time to mathematical reasoning, which should not be difficult to do, since successful methods for teaching this form of reasoning have already been tested and documented (Nunes *et al*., 2007).

We advocate that other positive results in our analyses, such as the contributions of spatial and other information processing abilities to mathematical achievement, should be considered in the same way. Again, if needs be, there are tested and available methods of improving some of these underlying abilities through teaching, which could easily be considered in the development of personalised instruction for children who might be at risk or which could be integrated in the school curricula.

The data on the different effects of different social background might seem to raise a more intractable problem, but the relationship between social background and mathematical reasoning gives us a clear solution. The main pathway from social background to mathematical achievement is through the children's ability to reason mathematically. So, improving children's reasoning should decrease the effects of differences in social backgrounds. To some extent this seems to be happening in some schools already, at any rate with older children. The interesting and important interaction, which we described at some length in Appendix D, established that the effects of social background at the children's individual level are minimal with 14-year-old children who attend schools in which the majority of children come from relatively prosperous backgrounds. It is possible that these schools have found ways, which still need to be identified, of promoting mathematical reasoning to such an extent that there are no longer differences in the mathematical reasoning of children from different social backgrounds at these schools. However, we are cautious about this hypothesis because of the limited number of children from the lower SES quartile in schools with a high SES social composition.

The results and conclusions of our analyses are, we think, quite comprehensive and indisputable, given the size of the sample and the longitudinal nature of the ALSPAC data. However, they do raise important questions for further research. We have highlighted these already, and will only mention them briefly again here.

We certainly need to know more about the social effects and, in particular, about what families and schools can do to have an effect on children's reasoning. We should like to know more about the connection between children's spatial skills and the spatial demands of learning geometry. It would also be valuable to know more about children's mathematical reasoning itself, although the ALSPAC data have already provided a great deal of relevant information. There are, however, some important quantitative relations that we were not able to include in our original measures of mathematical reasoning, such as the inverse relation between addition and subtraction (Bryant, Christie & Rendu, 1999; Nunes, Bryant, Hallett, Bell, & Evans, 2009; Stern, 2005) and between multiplication and division, and it would be useful to know whether the reasoning measures, already extremely effective, would become even more powerful if these were included as well.

Such research projects are the natural consequence of an enquiry like ours which has established beyond doubt the importance of teaching and learning mathematical reasoning, and the central part that mathematical reasoning plays in determining the striking differences in children's mathematical achievements at school. Our most important claim is that improvements in teaching children about mathematical reasoning will radically reduce the extent of these worrying differences.

References

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology, 67*, 21-38.

Alexander, K. L., McDill, E. L., Fenessy, J., & D'Amico, R. J. (1979). School SES influences - Composition or context? *Sociology of Education, 52*, 222-237.

Apple, M. W. (1979). *Ideology and curriculum*. London: Routledge & Kegan Paul.

Baroody, A. J., & Gannon, K. E. (1984). The development of the commutativity principle and economical addition strategies. *Cognition and Instruction, 1*, 321-339.

Becker, J. (1993). Young children's numerical use of number words: counting in many-to-one situations. *Developmental Psychology, 19*, 458-465.

Boaler, J. (1997). Setting, Social Class and Survival of the Quickest. *British Educational Research Journal, 23*(5), 575-595.

Brown, M. (1981). Number operations. In K. Hart (Ed.), *Children's Understanding of Mathematics: 11-16* (pp. 23-47). Windsor, UK: NFER-Nelson.

Bryant, P. E. (1985). The distinction between knowing when to do a sum and knowing how to do it. *Educational Psychology, 5*, 207-215.

Bryant, P., Christie, C., & Rendu, A. (1999). Children's understanding of the relation between addition and subtraction: inversion, identity and decomposition. *Journal of Experimental Child Psychology, 74*, 194-212.

Butterworth, B. (2003). *Dyscalculia Screener*. London: Nelson.

Bynner, J., & Parsons, S. (2000). The impact of poor numeracy on employment and career progression. In C. Tikly & A. Wolf (Eds.), *The maths we need to know* (pp. 26-51). London: University of London Institute of Education (The Bedford Way Papers).

Carpenter, T. P., Ansell, E., Franke, M. L., Fennema, E., & Weisbeck, L. (1993). Models of Problem Solving: A Study of Kindergarten Children's Problem-Solving Processes. *Journal for Research in Mathematics Education, 24*, 428-441.

Carpenter, T. P., & Moser, J. M. (1982). The Development of Addition and Subtraction Problem-Solving Skills. In T. P. Carpenter, J. M. Moser & T. A. Romberg (Eds.), *Addition and Subtraction* (pp. 9-24). New York: Lawrence Erlbaum Associates.

Clements, D. H., Copple, C., & Hyson, M. (2005). *Early childhood mathematics: Promoting good beginnings. A joint position statement of the National Association for the Education of Young Children (NAEYC) and the National Council of Teachers of Mathematics (NCTM).* Reston, VA: NCTM.

Crocker, T., & Cheeseman, R. (1991). The ability of young children to rank themselves for academic ability. In M. Woodhead, P. Light & R. Carr (Eds.), *Growing up in a Changing Society* (pp. 156-162). London: Routledge and The Open University.

Dehaene, S. (1997). *The Number Sense*. London: Penguin.

Department for Children, S. a. F. (2008). *Making good progress in Kes Stage 3 Mathematics*: www.teachernet.gov.uk/publications, DCSF-00189-2008.

Dermitzaki, I., & Efklides, A. (2000). Aspects of Self-Concept and Their Relationship to Language Performance and Verbal Reasoning Ability. *The American Journal of Psychology, 113*(4), 621-637.

Dowker, A. (2005). *Individual differences in Arithmetic*. Hove: Psychology Press.

Duckworth, K. (2007). *What role for the three Rs? Progress and attainment during primary school*. London: Centre for Research on the Wider Benefits of Learning, Institute of Education.

Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7- to 10-year-olds. *Journal of Experimental Child Psychology, 91*, 113-136.

Fennema, E. (1977). Influences of selected cognitive, affective, and educational variables on sex-related differences in mathematics learning and studying. . In J. Shoemaker (Ed.), *Women and mathematics: Research perspectives for change. N.I.E. Papers in Education and Work, No. 8* (pp. 79-135). Washington, DC: Education and Work Group, National Institute of Education.

Fennema, E., & Peterson, P. L. (1985). Autonomous learning behavior: A possible explanation of gender-related differences in mathematics. In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 17-35). Orlando, FL: Academic Press.

Freeman, N. H., Antonuccia, C., & Lewis, C. (2000). Representation of the cardinality principle: early conception of error in a counterfactual test. *Cognition, 74*, 71-89.

Fuson, K. C. (1988). *Children's Counting and Concepts of Number*. New York: Springer Verlag.

Geary, D. C., & Brown, S. C. (1991). Cognitive addition: strategy choice and speed of processing differences in gifted, normal and mathematically disabled children. *Developmental Psychology, 27*, 398-406.

Gelman, R., & Butterworth, B. (2005). Number and language: how are they related? *Trends in Cognitive Science, 8*, 6-10.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, Mass: Harvard University Press.

Ginsburg, H. P., Klein, A., & Starkey, P. (1998). The Development of Children's Mathematical Thinking: Connecting Research with Practice. In W. Damon, I. E. Siegel & A. A. Renninger (Eds.), *Handbook of Child Psychology. Child Psychology in Practice* (Vol. 4, pp. 401-476). New York: John Wiley & Sons.

Ginsburg, H. P., & Russell, R. L. (1981). Social class and racial influences on early mathematical thinking. *Monographs of the Society for Research in Child Development, 46*(6), Serial No. 193.

Green, P. J., Dugoni, B. L., Ingels, S. J., & Cambum, E. (1995). *A profile of the American high school senior in 1992*. Washington, DC:: U.S. Department of Education.

Hitch, G. J., & McAuley, E. (1991). Working memory in children with specific arithmetic learning difficulties. *British Journal of Psychology, 82*, 375-386.

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., *et al*. (2005). Computerized training of working memory in children with ADHD - A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*, 177-186.

Leder, G. C. (1992). Mathematics and Gender: Changing Perspectives. In D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 597-622). New York: Macmillan Publishing Company.

Lee, V. E., & Bryk, A. S. (1989). A Multilevel Model of the Social Distribution of High School Achievement. *Sociology of Education, 62*(3), 172-192.

Lockheed, M. E., Thorpe, M., Brooks-Gunn, J., Casserly, P., & McAloon, A. (1985). *Sex and ethnic differences in middle school mathematics, science and computer science: What do we know?* Princeton, NJ: Educational Testing Service.

Lubienski, S. T., & Bowen, A. (2000). Who's Counting? A Survey of Mathematics Education Research 1982-1998. *Journal for Research in Mathematics Education, 31*(5), 626-633.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240-260.

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*, 129-149.

Marsh, H. W. (1991). Failure of High-Ability High Schools to Deliver Academic Benefits Commensurate With Their Students' Ability Levels. *American Educational Research Journal, 28*(2), 445-480.

Marsh, H. W., & Yeung, A. S. (1997). Coursework Selection: Relations to Academic Self-Concept and Achievement. *American Educational Research Journal, 34*, 691-720.

Melhuish, E., Romaniuk, H., Sammons, P., Sylva, K., Siraj-Blatchford, I., & Taggart, B. (2006). *The Effective Pre-School and Primary Education 3-11 Project (EPPE 3-11):The Effectiveness of Primary Schools in England in Key Stage 2 for 2002, 2003 and 2004*. London: DfES / Institute of Education, University of London.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters*. Wells Somerset: Open Books.

Noss, R., & Hoyles, C. (1992). Looking back and looking forward. In C. Hoyles & R. Noss (Eds.), *Learning mathematics and LOGO* (pp. 431-470). Cambridge (MA): The MIT Press.

Noss, R., & Hoyles, C. (1996). *Windows on mathematical meaning: Learning cultures and computers*. Dordrecht, The Netherlands: Kluwer.

Nunes, T., Bell, D., Evans, D. (2008). Improving children’s working memory through guided rehearsal. Paper presented at the American Educational Research Association. New York, April.

Nunes, T., & Bryant, P. (1996). *Children doing mathematics*. Oxford: Blackwell.

Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. N. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology, 25*, 147-166.

Nunes, T., Bryant, P., Hallett, D., Bell, D., & Evans, D. (2009). Teaching Children About the Inverse Relation Between Addition and Subtraction. *Mathematical Thinking and Learning, 11*(1), 61-78.

Opdenakker, M., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal, 33*(2), 179-206.

Pajares, F., & Miller, D. M. (1994). The role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal of Educational Psychology, 86*, 193-203.

Piaget, J. (1952). *The child's conception of number*. London: Routledge and Kegan Paul.

Piaget, J., & Inhelder, B. (1974). *The child's construction of quantities*. London: Routledge & Kegan Paul.

Piaget, J., & Inhelder, B. (1975). *The Origin of the Idea of Chance in Children*. London: Routledge and Kegan Paul.

Pretzlik, U., Olsson, J., Nabuco, M. E., & Cruz, I. (2003). Teachers' implicit views of intelligence predict pupils' self-perceptions as learners. *Cognitive Development, 18*, 579-600.

Reyes, L. H., & Stanic, G. M. A. (1988). Race, Sex, Socioeconomic Status, and Mathematics. *Journal for Research in Mathematics Education,, 19*(1), 26-43.

Sacker, A., Schoon, I., & Bartley. M. (2002). Social inequality in educational achievement and psychosocial adjustment throughout childhood. *Social Science and Medicine, 55*, 863-880.

Secada, W. G. (1992). Race, ethnicity, social class, language, and achievement in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 623-660). New York: Macmillan.

Siegel, L. S., & Linder, B. A. (1984). Short term memory processes in children with reading and arithmetic learning disabilities. *Developmental Psychology, 20*, 200-207.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973-980.

Siegler, R., & Crowley, K. (1991). The Microgenetic Method. A Direct Means for Studying Cognitive Development. *American Psychologist, 46*, 606-620.

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, in press*.

Skaalvik, E. M., & Rankin, R. J. (1995). A Test of the Internal/External Frame of Reference Model at Different Levels of Math and Verbal Self-Perception. *American Educational Research Journal, 32*(1), 161-184.

Slavin, R. E. (1990). Achievement effects of ability grouping in secondary schools: a best evidence synthesis. *Review of Educational Research, 60*, 471-499.

Stern, E. (2005). *Transitions in mathematics: From intuitive quantification to symbol-based reasoning.* Paper presented at the International Society for the Study of Behavioral Development (ISSBD), Melbourne, Australia.

Stevenson, H. W., Stigler, J. W., Lee, S.-Y., Kitamura, S., & Hsu, C.-C. (1985). Cognitive performance and academic achievement of Japanese, Chinese and American children. *Child Development, 56*, 718-734.

Stevenson, H. W., Stigler, J. W., Lucker, G. W., Lee, S.-Y., Hsu, C. C., & Kitamura, S. (1986). Mathematics achievement of Chinese, Japanese and American children. *Science, 231*, 693-699.

Stigler, J. W., Lee, S.-Y., & Stevenson, H. W. (1987). Mathematics classrooms in Japan, Taiwan and the United States. *Child Development, 58*, 1272-1285.

Tate, W. F. (1997). Race-Ethnicity, SES, Gender, and Language Proficiency Trends in Mathematics Achievement: An Update. *Journal for Research in Mathematics Education, 28*(6), 652-679.

Thompson, P. (1994). The Development of the Concept of Speed and Its Relationship to Concepts of Rate. In G. Harel & J. Confrey (Eds.), *The Development of Multiplicative Reasoning in the Learning of Mathematics* (pp. 181-236). Albany, New York: State University of New York Press.

Thompson, P. W. (1993). Quantitative Reasoning, Complexity, and Additive Structures. *Educational Studies in Mathematics, 25*, 165-208.

Towse, J. N., & Hitch, G. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? Quarterly Journal of Experimental Psychology, 48A, 108–124.

Tymms, P. (2001). A Test of the Big Fish in a Little Pond Hypothesis: An Investigation Into the Feelings of Seven-Year-Old Pupils in School. *School Effectiveness and School Improvement. An International Journal of Research, Policy and Practice, 12*(2), 161-181.

Vergnaud, G. (1979). The Acquisition of Arithmetical Concepts. *Educational Studies in Mathematics, 10*, 263-274.

Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 128-175). London: Academic Press.

Wechsler, D. (1992). *WISC-III UK Manual*. Sidcup, Kent: The Psychological Corporation, Harcourt Brace Jovanovich.

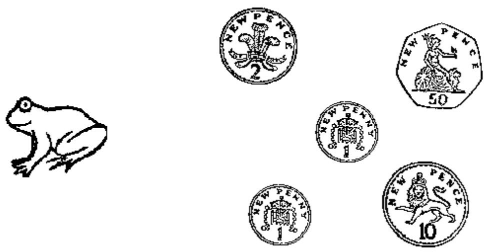
Zimmermann, W. & Cunningham, S (1991). Visualization in teaching and learning mathematics. Washington (DC): Mathematical Association of America

Appendix A - Details of the ALSPAC tasks analysed in this report

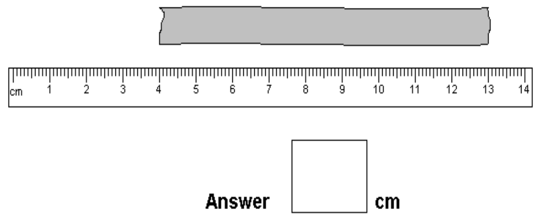
A. Cognitive tasks

The Mathematical Reasoning Tasks were presented orally, with the support of pictures. The instructions are written here for the readers' information. Figure 21 presents some examples of items in the Year 4 Mathematical Reasoning Tasks. The items on top are examples of additive reasoning problems and those in the bottom illustrate multiplicative reasoning.

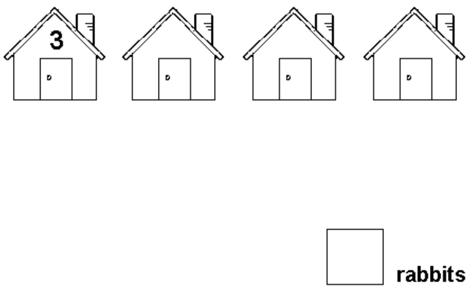
Figure 21 - Examples of Items in Y4 Mathematic reasoning tests



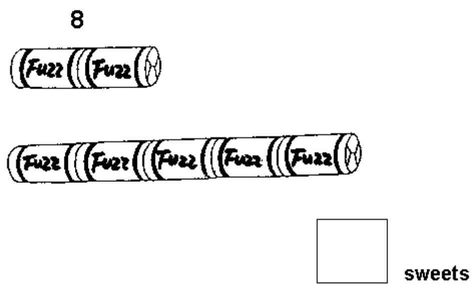
The frog costs 11p. Tick the coins you need to give the exact money.



(Make sure there are no rulers on the table before this question is asked) Here is a picture of a ribbon and a ruler. How long is the ribbon? Use the ruler in the picture to help you find out. Write your answer in the empty box.



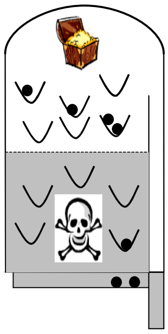
3 rabbits live in each house. How many rabbits live altogether in the 4 houses? Write your answer in the box.



The roll on top has 8 sweets. How many sweets do you think there are in the big roll below? Write you answer in the empty box.

Figure 22 presents some examples of items in the Year 6/8 Mathematical Reasoning Tasks. The items on top are examples of additive reasoning problems and those in the bottom illustrate multiplicative reasoning.

Figure 22 - Examples of items in Y6/8 Mathematical Reasoning Tasks

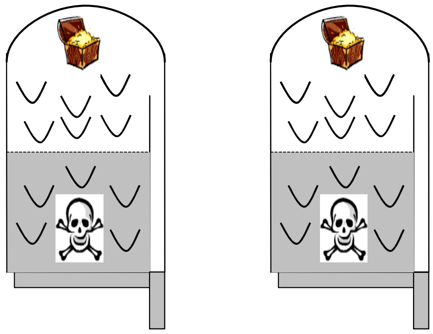


In the pinball game, you win one point for each pinball that falls into the top area, where there is a treasure. Look at game 1. There are 4 pinballs in the treasure area. This means: won 4.

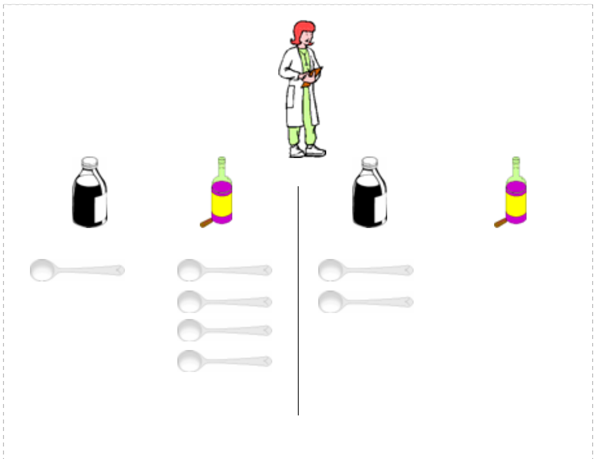
You loose one point for each pinball that falls into the bottom area, where there is a skull. Look at game 1. There is one pinball in the skull area. This means lost 1.

If your pinballs fall into the tube, you do not score. Look at game 1. There are 2 pinballs in the tube. No points there.

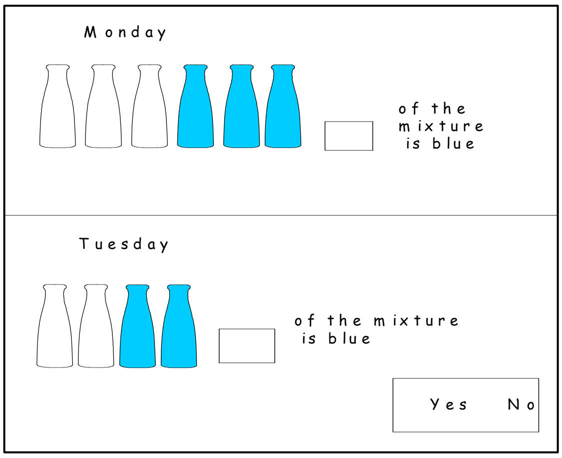
What was the score for this game?



Ali played 2 games. When he played Game 1, he lost 3 points. Draw in the 7 pinballs to make him end with a loosing score of 3 points. When he played Game 2 he won 4 points. Draw in the 7 pinballs to make him end with a winning score of 4.



There is a medicine that is very bitter and the chemist mixes it with syrup for the children to make it taste better.
Yesterday she mixed 1 spoon of medicine with 4 spoons of syrup.
Today she had to make more mixture and she will have to use 2 spoons of medicine.
How many spoons of syrup will she need for the mixture to taste the same as yesterday?
Draw the spoons or write the number under the syrup bottle.



Imagine you are mixing paint. On Monday you mix 3 bottles of white and 3 of blue paint. The blue paint appears grey in the picture. On Tuesday you mix 2 bottles of white and 2 of blue. Will the colour of the mixed paint look the same on Monday and Tuesday? Circle 'yes' or 'no' in the box at the bottom of the page.
What fraction of the paint is blue on Monday? Write your answer in the box.
What fraction of the paint is blue on Tuesday? Write your answer in the box.

The WISC III subtests treated as separate measures in the analyses

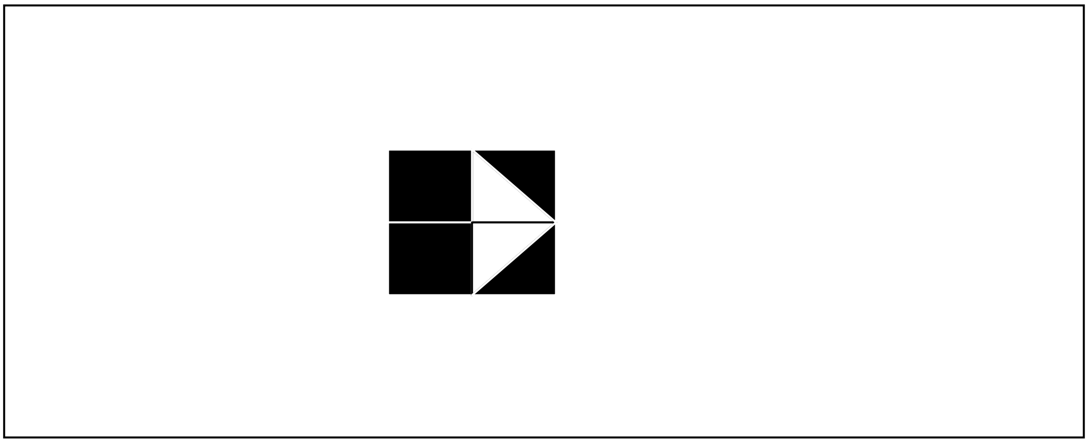
Arithmetic

WISC Arithmetic contains a series of sums and arithmetical problems, in each of which it is quite clear what calculation the child has to do. Thus the sub-test is a pure measure of the ability to calculate, and it involves no mathematical reasoning because there is no need in any of the items in WISC arithmetic to work out what calculation is needed. That is explicit from the start.

Block Design

The children are shown a card that depicts a pattern in two colours (Figure 23 displays an example of a possible item; for copyright reasons, it is not possible to include an actual item). The child is offered a number of square blocks which have faces that are painted in a single colour or in two colours with a division along the diagonal. The child is asked to reproduce the figure on the card using the blocks. The task is timed: there is a time limit for each shape.

Figure 23 - An item that illustrates the Block Design task



Object Assembly

The child is given a collection of cut-out shapes, which, if put together correctly, will represent something familiar to the child (e.g. a ball, a hand). The child is not told what the resulting figure will be. The task is timed: there is a time limit for each object.

Digit Span - forward and backward

Forward-digit span assesses short-term memory: the child hears a series of digits and attempts to repeat them in the same order. Backward-digit span assesses working memory: the child hears a series of digits and attempts to repeat them from the last heard to the first. Backward-digit span is considered to be a measure of working memory because the child has to work on the input - i.e. the series of digits - and at the same time recall the input.

Coding

Children have in front of them, as a key, the digits from one to nine; under each digit, there is a symbol (e.g. a half circle, a plus sign). Below this key there are several rows of digits in random order. The children's task is to draw in the appropriate pattern under each digit. The task is a timed one and the score is the number of items filled in correctly in the fixed period, with penalties for mistakes. Any child who attends meticulously and uninterruptedly to the task at hand should do well in this particular task, and it is hard to see any other possible constraint on children's performance in WISC Coding than failures in attention.

Appendix B - Regression analyses used to test whether different cognitive measures made independent contributions to the predictions of Key Stage Maths attainment

Table 3
Prediction of Key Stage 2 mathematics assessments by WISC Arithmetic & Mathematical Reasoning in Year 4

Two multiple regressions in which the outcome measure was the mathematics assessments at Key Stage 2. The first three predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. WISC arithmetic was the 4th and Maths reasoning the final step in one regressions and vice versa in the other regression (N=2413)

Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.028	.12	.023	.003
2nd step WISC IQ Full scale minus Arithmetic & Working Memory	.369	.32	.015	.001
3rd step WISC Working Memory	.030	.09	.076	.013
4th step in 1st regression WISC Arithmetic	.073	.21	.046	.004
5th step in 1st regression Year 4 Maths Reasoning Task	.076	.35	.087	.004
4th step in 2nd regression Year 4 Maths Reasoning Task	.119	.35	.087	.004
5th step in 2nd regression WISC Arithmetic	.030	.21	.046	.004

Table 4
Prediction of Key Stage 3 mathematics assessments by WISC Arithmetic & Mathematical Reasoning in Year 4

Two multiple regressions in which the outcome measure was the mathematics assessments at Key Stage 3. The first three predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. WISC arithmetic was the 4th and Year 4 Maths Reasoning the final step in one regression and vice versa in the other regression (N= 1595)

Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.011	.09	.028	.005
2nd step WISC IQ Full scale minus Arithmetic & Working Memory	.463	.40	.030	.001
3rd step WISC Working Memory	.019	.06	.092	.024
4th step in 1st regression WISC Arithmetic	.059	.18	.065	.007
5th step in 1st regression Year 4 Maths Reasoning Task	.075	.34	.138	.008
4th step in 2nd regression Year 4 Maths Reasoning Task	.111	.34	.138	.008
5th step in 2nd regression WISC Arithmetic	.023	.18	.065	.007

Table 5
Prediction of Key Stage 2 mathematics assessments by Mathematical Reasoning in Year 6

Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 2. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) WISC Arithmetic. The final variable was Year 6 Mathematical Reasoning (N= 3752)

Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.017	.13	.027	.002
2nd step WISC IQ Full scale minus Arithmetic & Working Memory	.393	.28	.014	.001
3rd step WISC Working Memory	.033	.11	.100	.010
4th step in 1st regression WISC Arithmetic	.068	.21	.047	.003
5th step in 1st regression Year 6 Maths Reasoning Task	.099	.39	.044	.001
4th step in 2nd regression Year 6 Maths Reasoning Task	.137	.39	.044	.001
5th step in 2nd regression WISC Arithmetic	.031	.21	.047	.003

Table 6
Prediction of Key Stage 3 mathematics assessments by Mathematical Reasoning in Year 6

Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 3. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) WISC Arithmetic. The final variable was Year 6 Mathematical Reasoning (N= 2590)

Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.007	.06	.020	.004
2nd step WISC IQ Full scale minus Arithmetic & Working Memory	.469	.35	.026	.001
3rd step WISC Working Memory	.026	.09	.135	.018
4th step in 1st regression WISC Arithmetic	.061	.19	.068	.005
5th step in 1st regression Year 6 Maths Reasoning Task	.098	.40	.070	.003
4th step in 2nd regression Year 6 Maths Reasoning Task	.135	.40	.070	.003
5th step in 2nd regression WISC Arithmetic	.025	.19	.068	.005

Table 7
Prediction of Key Stage 3 mathematics assessments by Mathematical Reasoning in Year 8

Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 3. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) WISC Arithmetic. The final variable was Year 8 Mathematical Reasoning (N= 1169)

Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.005	.06	.020	.005
2nd step WISC IQ Full scale minus Arithmetic & Working Memory	.436	.29	.021	.001
3rd step WISC Working Memory	.019	.06	.087	.024
4th step in 1st regression WISC Arithmetic	.054	.17	.057	.006
5th step in 1st regression Year 8 Maths Reasoning Task	.178	.52	.085	.003
4th step in 2nd regression Year 8 Maths Reasoning Task	.211	.52	.085	.003
5th step in 2nd regression WISC Arithmetic	.020	.17	.057	.006

Table 8
Prediction of Key Stage Mathematics, English and Science by Mathematical Reasoning tasks

The R^2 change and β figures for the relation between the Mathematics Reasoning Task scores and the children's achievement in Mathematics, English and Science.

Outcome Measure							
		Mathematics		English		Science	
Year in which Mathematics Reasoning Task was given	Key Stage	R^2 change	B	R^2 change	β	R^2 change	B
4	2	.076	.35	.011	.130	.022	.19
6	2	.099	.39	.019	.172	.050	.28
4	3	.075	.34	.014	.144	.028	.21
6	3	.098	.40	.016	.158	.067	.33
8	3	.178	.52	.031	.218	.108	.40

Table 9
Prediction of Key Stage 2 mathematics assessments WISC Block Design

Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 2. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) Year 4 Mathematical Reasoning. The final variable was WISC Block Design (N= 2232)

Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.035	.12	.024	.003
2nd step WISC IQ Full scale minus Block Design and Working Memory	.369	.36	.020	.001
3rd step WISC Working Memory	.024	.10	.095	.014
4th step in 1st regression Year 4 Maths Reasoning Task	.105	.36	.092	.004
5th step in 1st regression WISC Block Design	.012	.12	.007	.001
4th step in 2nd regression WISC Block Design	.025	.12	.007	.001
5th step in 2nd regression Year 4 Maths Reasoning Task	.091	.36	.092	.004

Table 10
Prediction of Key Stage 3 mathematics assessments by WISC Block Design

Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 3. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) Year 4 Mathematical Reasoning. The final variable was WISC Block Design (N= 1485)

Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.013	.08	.027	.005
2nd step WISC IQ Full scale minus Block design & Working Memory	.440	.39	.034	.002
3rd step WISC Working Memory	.015	.08	.108	.025
4th step in 1st regression Year 4 Maths Reasoning Task	.104	.35	.141	.008
5th step in 1st regression WISC Block Design	.023	.18	.017	.002
4th step in 2nd regression WISC Block Design	.044	.18	.017	.002
5th step in 2nd regression Year 4 Maths Reasoning Task	.084	.35	.141	.008

Table 11
Prediction of Key Stage 2 mathematics assessments by WISC Object Assembly

Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 2. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) Year 4 Mathematical Reasoning. The final variable was WISC Object assembly (N= 2221)

Step in regression	R^2 change	β coefficient	B	Standard error of B
1st step Age at outcome	.038	.12	.023	.003
2nd step WISC IQ Full scale minus Object assembly & Working Memory	.380	.40	.022	.001
3rd step WISC Working Memory	.020	.10	.093	.014
4th step in 1st regression Year 4 Maths Reasoning Task	.096	.36	.092	.004
5th step in 1st regression WISC Object assembly	.003	.06	.005	.001
4th step in 2nd regression WISC Object assembly	.006	.06	.005	.001
5th step in 2nd regression Year 4 Maths Reasoning Task	.093	.36	.092	.004

Table 12 **Prediction of Key Stage 3 mathematics assessments by WISC Object Assembly**				
Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 3. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) Year 4 Mathematical Reasoning. The final variable was WISC Object assembly (N= 1487)				
Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.016	.09	.028	.006
2nd step WISC IQ Full scale minus Object Assembly and Working Memory	.463	.46	.039	.002
3rd step WISC Working Memory	.013	.08	.118	.025
4th step in 1st regression Year 4 Maths Reasoning Task	.093	.36	.145	.008
5th step in 1st regression WISC Object assembly	.002	.05	.006	.002
4th step in 2nd regression WISC Object Assembly	.005	.05	.006	.002
5th step in 2nd regression Year 4 Maths Reasoning Task	.090	.36	.145	.008

Table 13 **Prediction of Key Stage 2 mathematics assessments by WISC Coding subtest**				
Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 2. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) Year 4 Mathematical Reasoning. The final variable was WISC Coding (N= 2225)				
Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.037	.13	.024	.003
2nd step WISC IQ Full scale minus Coding and Working Memory	.359	.37	.020	.001
3rd step WISC Working Memory	.024	.10	.090	.014
4th step in 1st regression Year 4 Maths Reasoning Task	.107	.36	.092	.004
5th step in 1st regression WISC Coding	.013	.12	.013	.002
4th step in 2nd regression WISC Coding	.027	.12	.013	.002
5th step in 2nd regression Year 4 Maths Reasoning Task	.092	.36	.092	.004

Table 14
Prediction of Key Stage 3 mathematics assessments by WISC Coding subtest

Multiple regression in which the outcome measure was the mathematics assessments at Key Stage 3. The first four predictor variables entered were: (1) Age at key stage assessment. (2) IQ, (3) Working memory. (4) Year 4 Mathematical Reasoning. The final variable was WISC Coding (N= 1482)

Step in regression	R^2 change	B coefficient	B	Standard error of B
1st step Age at outcome	.014	.09	.030	.006
2nd step WISC IQ Full scale minus Coding and Working Memory	.443	.43	.036	.002
3rd step WISC Working Memory	.016	.08	.112	.025
4th step in 1st regression Year 4 Maths Reasoning Task	.100	.35	.143	.008
5th step in 1st regression WISC Coding	.014	.13	.021	.003
4th step in 2nd regression WISC Coding	.026	.13	.021	.003
5th step in 2nd regression Year 4 Maths Reasoning Task	.088	.35	.143	.008

Appendix C - Different types of analysis investigating the effects of social factors on Key Stage Mathematics attainment and on children's self confidence in maths

Table 15
Multilevel analysis of the effects of socio-economic status on the Year 4 Mathematical Reasoning scores at two levels: the level of individual child and the level of the school *

	Estimate	t	Significance
Intercept	10.87		
Effect of SES at the individual level	-0.73	-9.78	<.001
Effect of SES at the school level	-0.70	-4.03	<.001
Interaction between the two levels	0.16	1.30	.192
Residual	7.66		

* Number of children: 2213
* Number of schools: 86
* Mean no. of children analysed in each school: 25.73 (range 6-65)

Table 16
Multilevel analysis of the effects of socio-economic status on the Year 6 Mathematical Reasoning scores at two levels: the level of individual child and the level of the school *

	Estimate	t	Significance
Intercept	19.82		
Effect of SES at the individual level	-1.70	-9.78	<.001
Effect of SES at the school level	-1.88	-4.03	<.001
Interaction between the two levels	0.08	1.30	.760
Residual	39.42		

* Number of children: 3440
* Number of schools: 133
* Mean no. of children analysed in each school: 25.87 (range 7-132)

Table 17 **Multilevel analysis of the effects of socio-economic status on the Year 8 Mathematical Reasoning scores at two levels: the level of individual child and the level of the school ***			
	Estimate	t	Significance
Intercept	23.77		
Effect of SES at the individual level	-1.93	-9.78	<.001
Effect of SES at the school level	-2.11	-4.03	<.001
Interaction between the two levels	-0.66	1.30	.042
Residual	41.31		

* Number of children: 1894
* Number of schools: 27
* Mean no. of children analysed in each school: 70.15 (range 15-168)

Table 18 **Mean, Standard Deviation, Number of Participants for each Key Stage Mathematical Results by Gender**						
Gender	KS1	Standard Deviation and n	KS2	Standard Deviation and n	KS3	Standard Deviation and n
Male	3.38	1.36 n=3775	4.15	0.72 n= 4022	6.44	1.33 n=2810
Female	3.44	1.27 n=3583	4.14	0.72 n=3944	6.44	1.22 n=2674

Table 19
Mean, Standard Deviation, Number of Participants for each Mathematical Reasoning Test by Gender

Gender	Year 4 Maths Reasoning Max: 17	Standard Deviation and n	Year 6 Maths Reasoning Max: 35	Standard Deviation and	Year 8 Maths Reasoning Max: 35	Standard Deviation and n
Male	11.12	3.07 N=1761	20.86	6.96 N=2659	24.54	6.90 N=1026
Female	10.88	3.00 N=1768	19.26	6.69 N=2631	23.39	6.86 N=887

Table 20
Mean Scaled Score[10] and Number of Participants for Mathematically Relevant Cognitive Skills Measured in the WISC by Gender

Gender	Arithmetic	Block Design	Object Assembly	Coding	Digit Span	N
Male	10.85	10.98	10.30	9.94	10.14	2774
Female	10.49	10.49	9.91	11.29	10.74	2795

[10] Scaled scores are adjusted for age differences. Their expected mean is 10 and Standard Deviation is 3. The N is approximately the same across tests because these are subtests of a measured administered on a single occasion.

Table 21
Means (controlling for KS1 results) in the Year 4 Mathematics Reasoning Test, standard error of the mean, and number of participants by group defined for streaming

	Mathematics Reasoning Year 4		Mathematics Reasoning Year 6	
Streaming Group	Mean (n)	Standard Error	Mean (n)	Standard Error
School does not stream	10.80 (n=205)	.16	19.77 (n=403)	.27
Top group	11.36 (n=995)	.08	20.76 (n=2353)	.12
Middle group	10.32 (n=1065)	.07	17.96 (n=1892)	.12
Lower group	9.22 (584)	.12	15.98 (n=1106)	.19

Table 22
Means (controlling for KS 1 results) in the WISC cognitive skills measured in Year 4, standard error of the mean, and number of participants by group defined for streaming in Year 3

	Arithmetic		Spatial Skills		Attention and Memory	
Streaming Group	Mean (n)	Standard Error	Mean (n)	Standard Error	Mean (n)	Standard Error
School does not stream	10.72 (n=189)	.25	21.86 (n=177)	.43	21.49 (n=180)	.31
Top group	11.52* (n=1180)	.12	21.66 (n=1118)	.20	22.13 (n=1161)	.14
Middle group	10.06* (n=1099)	.10	20.37* (n=1020)	.18	20.51* (n=1063)	.13
Lower group	8.91* (n=519)	.18	18.68* (n=481)	.31	18.56* (n=502)	.22

Appendix D - The cross-level interaction in the multilevel model of the Year 8 Mathematical Reasoning task

We had already measured each school's social composition by calculating its mean score on the Social Background Factor. Now, we divided the schools into four groups in terms of their social composition. We called the four social composition school groups Top, Medium High, Medium Low and Bottom.

Then we looked at the relationship between the individual children's social background and their mathematical reasoning in the Year 8 Mathematical Reasoning task separately within each of the school groups. The significant interaction suggested that this relationship should be different in the high and low social composition schools. So, forming these groups allowed us to look at such differences.

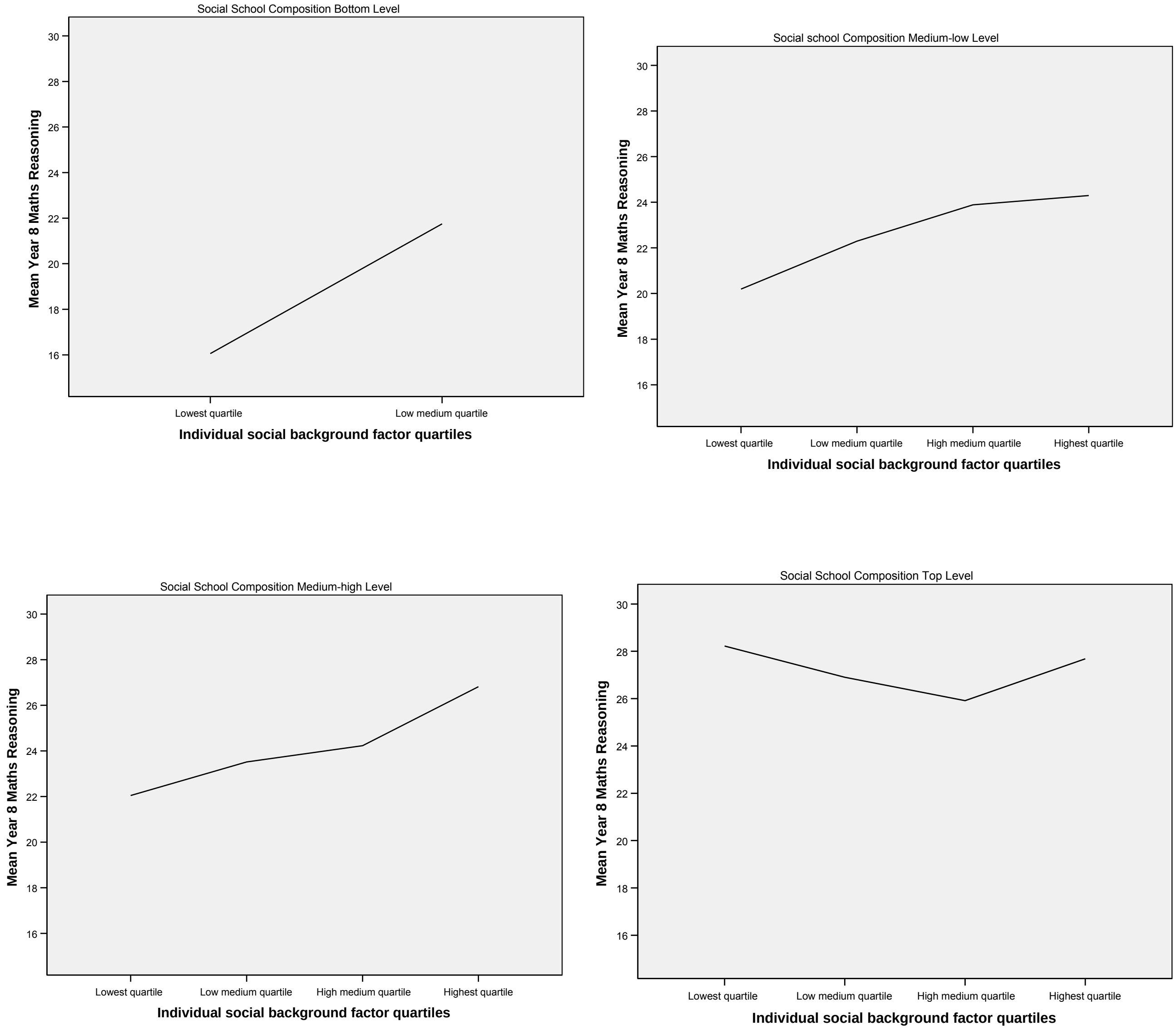
To simplify an already quite complicated manoeuvre, we also divided the total sample of children in this analysis into four social background quartile groups and then we calculated the mean mathematical reasoning score for each of the quartiles in each school group. Thus, each child belonged to the Highest quartile of the SES range or to the High Medium quartile or to the Low Medium quartile or to the Lowest quartile. We then calculated how the children in each quartile did in the Year 8 Mathematics Reasoning task separately within each group of schools with the same social composition.

The four graphs in Figure 22 present the results of these calculations. Each graph shows the relationship between the children's social background and their mathematical reasoning scores within each school group. The line in each graph represents the overall relationship between the children's individual social background and their reasoning scores.

Before we discuss the nature of the interaction, we should like to point out how well the graphs demonstrate the strong simple school effect. The figure shows how children from the same social background do much better if they go to a school in which the social composition of the pupils is relatively high than to one in which is relatively low.

Notice, for example, that the mean mathematical reasoning score for children in the lowest social background quartile was close to 16 in the Bottom social composition group of schools, close to 20 in the Medium Low group, close to 22 in the Medium High group and close to 24 in the Top social composition group. Much the same pattern emerged for children in all other three SES quartiles. Thus, children at a school with a high level social composition are better at reasoning mathematically than children whose social backgrounds are exactly the same but who go to schools with a low level social composition.

Figure 24 - The relation between children's SES and their mathematical reasoning scores



Turning to the interaction, the four graphs do indeed show that the relationship between children's own social background and their mathematical reasoning differed between the four social composition school groups. The main difference was between the highest social composition group and the other three groups. There was no progressive increase in mathematical reasoning as a function of social background among the children in the Top group of schools, but this increase in results was strong and near to linear in the three other social composition groups. One, but not the only, reason why the highest group was different from the other three may have been that there were very few children in the lowest social quartile in these Top schools, and they may therefore have been unusual children. At any rate, the average mathematical reasoning score of these particular children was remarkably high. However, there must have been other reasons as well for the different pattern in the Top group of schools, since, even if the analysis omitted the children who came from the lowest SES backgrounds in the Top group, the relation between SES and mathematical reasoning would still have been different in this group than in the others. One possibility is that the education given to the children in the Top group of schools was more effective than in the schools in the other three groups and thus obliterated the usual relationship between social background and mathematical ability. However, another possibility is that the result may have been due to selection procedures in the top group schools.

In contrast, the relationship between SES and mathematical reasoning was an important part of the pattern of mathematical reasoning scores in the other three social composition school groups. Even in the lowest social composition group, which contained only children from the bottom two quartiles, there was a striking difference in the mathematical reasoning scores of children from the different quartiles. The quality of the education in these schools did not remove, and may not even have attenuated, the disadvantages that are clearly part of coming from a less privileged social background.

Ref: DCSF-RR118

ISBN: 978 1 84775 457 8

© University of Oxford 2009

www.dcsf.gov.uk/research

Published by the Department for
Children, Schools and Families

department for
children, schools and families