The relative importance of two different mathematical abilities to mathematical achievement

Rossana Barros

British Journal of Educational Psychology (2011)

The British Psychological Society

www.wileyonlinelibrary.com

The relative importance of two different mathematical abilities to mathematical achievement

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

Background. Two distinct abilities, mathematical reasoning and arithmetic skill, might make separate and specific contributions to mathematical achievement. However, there is little evidence to inform theory and educational practice on this matter.

Aims. The aims of this study were (1) to assess whether mathematical reasoning and arithmetic make independent contributions to the longitudinal prediction of mathematical achievement over 5 years and (2) to test the specificity of this prediction.

Sample. Data from Avon Longitudinal Study of Parents and Children (ALSPAC) were available on 2,579 participants for analyses of KS2 achievement and on 1,680 for the analyses of KS3 achievement.

Method. Hierarchical regression analyses were used to assess the independence and specificity of the contribution of mathematical reasoning and arithmetic skill to the prediction of achievement in KS2 and KS3 mathematics, science, and English. Age, intelligence, and working memory (WM) were controls in these analyses.

Results. Mathematical reasoning and arithmetic did make independent contributions to the prediction of mathematical achievement; mathematical reasoning was by far the stronger predictor of the two. These predictions were specific in so far as these measures were more strongly related to mathematics than to science or English. Intelligence and WM were non-specific predictors; intelligence contributed more to the prediction of science than of maths, and WM predicted maths and English equally well.

Conclusions. There is clear justification for making a distinction between mathematical reasoning and arithmetic skills. The implication is that schools must plan explicitly to improve mathematical reasoning as well as arithmetic skills.

The relative importance of two different mathematical abilities to mathematical achievement

The mathematical problems that we give to children at school make two kinds of demand on them. The most obvious of the two is that children have to be able make the necessary calculation correctly. The other demand, though less obvious, is also important. It is to

**Correspondence should be addressed to Terezinha Nunes, 15 Norham Gardens, Oxford OX2 6PY, UK (e-mail: terezinha.nunes@education.ox.ac.uk).*

DOI:10.1111/j.2044-8279.2011.02033.x

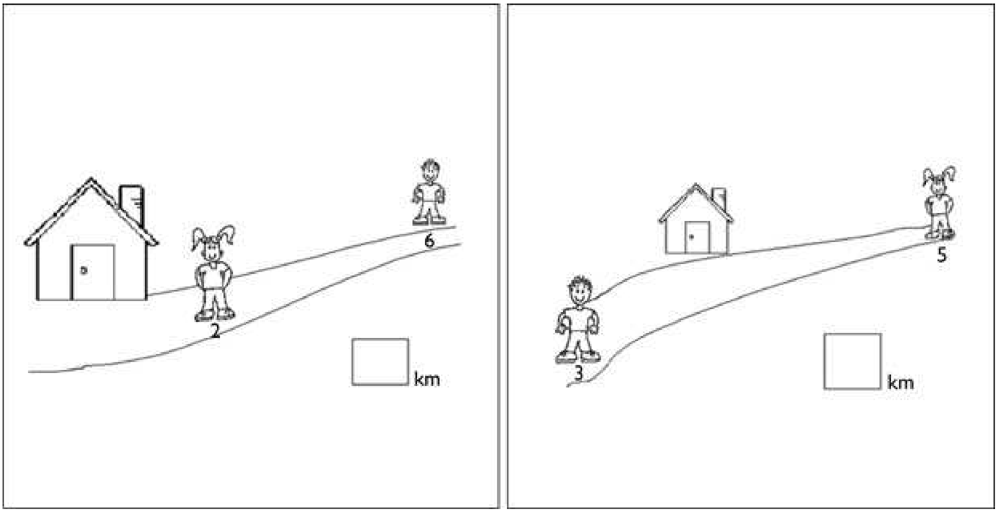

Figure 1. An example of two additive reasoning problems that exemplify how the relations between the quantities determine which calculation should be carried out.

analyse the quantitative relations involved in the problem in order to work out how to manipulate the numbers (e.g., what calculation they have to make or how to count).

For example, the children's task in both the problems in Figure 1 is to work out the distance between the boy and the girl when they both stop walking. Children certainly have to be able to calculate that $6–2 = 4$ to solve the first of the two problems and that $5 + 3 = 8$ to solve the second. However, they also need to reason that, because the two children travelled in the same direction in the first problem but in opposite directions in the second, the final gap between them is the shorter distance walked by the girl subtracted from the longer distance walked by the boy in the first problem, and in the second problem, the gap is the distance walked by the girl added to that walked by the boy. This additive reasoning involves an understanding of spatial relations and also of additive composition (e.g., that a 6-km length consists of a 2-km length plus a 4-km length).

The distinction between reasoning about the relevant mathematical relations and doing the calculation applies in the same way to other kinds of arithmetical problems, and the actual quantitative relations that the child has to reason about vary with the type of problem. In many problems, the main decision that has to be made is about the nature of the relations, whether these are additive or multiplicative relations (Vergnaud 1982; 1983). The quantitative relations that the child must consider in additive reasoning problems are part-whole relations, and in multiplicative reasoning, problems are one-to-many correspondences (Becker 1993; Nunes, Bryant, Evans, & Bell 2010) and proportional relations (Vergnaud 1983). Figure 2 presents two problems, one additive and one multiplicative, in which there are no verbal expressions that could give a clue to the type of relation, such as 'got x from', 'gave y to', 'each has z'. In such problems, which include a reference to a practical situation, the oral presentation does not contribute to the choice of operations.

Children who succeed in these problems often do not carry out computations but count, and what they count and how they count depends on the relations they establish between the different quantities in the problems (for a discussion of the use of counting

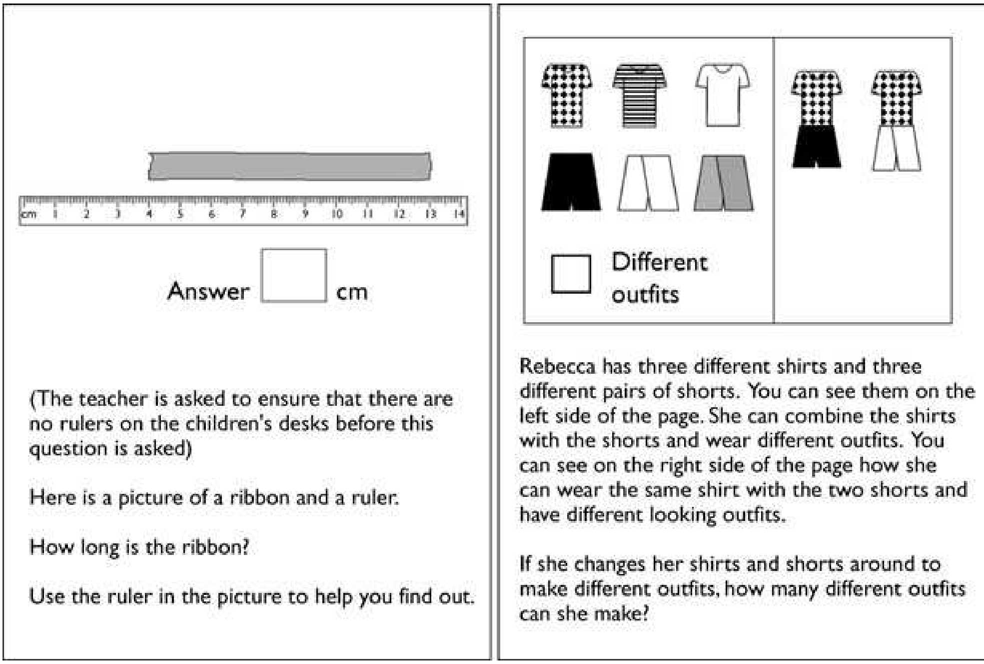

Figure 2. An additive reasoning (left) and a multiplicative reasoning (right) problem that children often solve by reasoning and counting rather than calculating.

in both problems, see Nunes & Bryant 1996). Previous research (e.g., Brown 1981; Cramer, Post, & Currier 1993; De Bock, Van Dooren, Janssens, & Verschaffel 2002; Van Dooren, Bock, Hessels, Janssens, & Verschaffel 2004; Verschaffel, Greer, & de Corte 2007) shows that children sometimes use additive reasoning when multiplicative reasoning is required and that the opposite is also true. Even after apparently correctly deciding that a problem is multiplicative, there are still decisions to be made, which involve whether to multiply or divide and, in the latter case, which number is the dividend and which is the divisor. These are is not always easy decisions for children, who sometimes consider the irrelevant feature number size when making these decisions rather than the relation between the quantities (e.g., De Corte, Verschaffel, & Van Coillie 1988; Greer 1988).

These examples do not exhaust the types of reasoning that children need to carry out in mathematics in school: they clearly also need to reason about space relations in order to learn geometry, about relations between numbers within an operation (e.g., for the same dividend, the bigger the divisor, the smaller the quotient; the order of the addends in an addition does not affect the total), and about relations between operations (addition and subtraction are inverse operations; multiplication and division are inverse operations). Much interesting research has been carried out about children's understanding of these relations (e.g., see Siegler & Stern 1998, for the inverse relation between addition and subtraction; see Baroody & Gannon 1984, for commutativity). Some of this research shows that children understand such relations first in the context of quantities and only later as number relations (e.g., Bryant, Christie & Rendu 1999) and that the type of situation affects whether the number relations are understood or not (e.g., Nunes & Bryant 1995; Squire & Bryant 2003).

The distinction between calculating and reasoning raises an interesting question about how well children learn mathematics at school. It is whether two quite separable abilities may play a part in children's mathematical learning: these are the ability to reason about the underlying quantitative relations in arithmetical problems and the ability to calculate. These two abilities may make separate and independent contributions to individual children's progress in mathematics at school. If this were so, it would be interesting and important to know about the relative strengths of these two determinants of schoolchildren's success in mathematical learning.

There is little direct evidence, at the moment, either on the question of the two separate abilities or on the relative importance of the two abilities, in children's success in mathematics at school. There is also little evidence on whether these abilities, which are easy to distinguish conceptually, can also be distinguished empirically, using quantitative methods. There are longitudinal studies of how well children's number knowledge and computational skills predict their success in learning about mathematics (De Smedt, Verschaffel, & Ghesquière 2009; Durand, Hulme, Larkin, & Snowling 2005; Jordan, Levine & Huttenlocher 1994; Jordan & Montani 1997; Krajewski & Schneider 2009), and there is one study on how children's mathematical reasoning at kindergarten predicts their mathematics achievement in the first year in school (van de Rijt, van Luit, & Pennings 1999). The results of this research, on the whole, have been positive. Individual children's success in tests of arithmetic and number skills and mathematical reasoning do predict how well they learn mathematics at school later on. However, these are separate studies and, thus, cannot provide information on the independence and the relative contribution of mathematical reasoning and of arithmetic to the prediction of mathematics achievement. Stern (1999) carried out an interesting and comprehensive longitudinal study of children's mathematical abilities between the ages of 4 and 12 years, which did involve both calculation and reasoning problems. However, her report of the results of this study did not include measures of the children's achievement in mathematics at school.

To our knowledge, only one longitudinal study has included measures of both arithmetic and mathematical reasoning and analysed their predictive value for mathematics achievement at school (Nunes *et al.*, 2007). Its results were quite positive and indicated that mathematical reasoning and arithmetic make independent contributions to the prediction of mathematics achievement. The prediction was specific, in the sense that it could not be explained by general intelligence or working memory (WM), because these were controlled for in the regression analyses.

It is often, and very plausibly, argued that children's general ability to handle information has an important effect on how well they learn mathematics. The aspect of information processing that has received the most attention as a possible determinant of children's mathematical learning is WM. When children use a procedure, such as addition, to solve a problem, they need to keep in mind the information in the problem and the steps they must take to implement the solution, while monitoring what they have done and what still needs to be done. WM should affect how well they can keep this information in mind and, thus, their success with the procedure. Different researchers have shown a connection between WM and arithmetic computations (e.g., Adams & Hitch 1997; Bull & Johnston 1997; D'Amico & Guarnera 2005; De Smedt, Janssen, Bouwens, Verschaffel, Boets, & Ghesquière 2009; Hitch & McAuley 1991; McLean & Hitch 1999; Siegel & Linder 1984; Siegel & Ryan 1989; Towse & Hitch 1995). The connection between WM and mathematical achievement is also supported by experimental studies that show that disrupting WM interferes with arithmetic performance and by evidence that children who have mathematical difficulties tend to produce low scores in WM tasks (Barrouillet & Lépine 2005; Passolunghi & Siegel 2004). Because it is quite plausible that WM not only predicts children's arithmetic skills but also explains why arithmetic skills and mathematical reasoning predict mathematics achievement (Geary & Brown 1991; Swanson & Beebe-Frankenberger 2004), WM should be used as a control in predictive studies whose aim is to show a specific connection of mathematical reasoning and arithmetic with mathematical achievement.

General intelligence is an ability defined more broadly than WM because it includes crystallized intelligence as well, which might affect how well children do in mathematics. So, predictive studies should control for WM and general intelligence in order to test whether predictors such as mathematical reasoning and arithmetic have a specific connection with mathematical achievement.

It is most unlikely that general intelligence and WM affect children's mathematical learning only or even that they affect their progress in mathematics more than in non-mathematical subjects, since it is hard to think of any intellectual activity that does not involve intelligence and WM. This raises a second issue related to specificity. WM and intelligence measures predict children's success in mathematics but would probably be just as strongly related to non-mathematical subjects as well, whereas, according to our hypothesis, mathematical reasoning and arithmetic should be specific predictors of mathematical achievement and not very good predictors of non-mathematical school subjects such as English.

Thus, research on the relationship between children's abilities and their success in a particular subject, such as mathematics, should deal with the question of specificity in two senses. Mathematical reasoning and arithmetic seem on the face of it to be abilities that will specifically affect children's progress in mathematics but not in other, non-mathematical subjects, such as English. This, however, needs to be checked empirically, which can be done in a longitudinal study by putting in control outcome measures. The issue of specificity is relevant to the choice of predictive, as well as of outcome measures. One simply has to see how well the abilities in question predict children's success in non-mathematical subjects as well as in mathematics. If the role that the two abilities play in children's progress in school is specific to mathematics, the children's scores on measures of these abilities should predict their mathematical success much better than their success in English or in some other non-mathematical outcome measure. This kind of design is all too rare in developmental and educational research. It has occasionally been adopted in longitudinal studies of children's reading (Bradley & Bryant 1983), but we know of no study of children's mathematics that has included such controls, though in our view, they should be regarded as an important part of longitudinal research on any aspect of children's learning.

To summarize, our knowledge of the abilities that influence children's mathematical learning at school is fragmentary. Measures of children's mathematical reasoning, of their calculation skills and of their intelligence and WM are related to their mathematical knowledge, but there is little research to show to what extent these represent separate abilities or whether each of them makes independent predictions of children's mathematical progress or what their relative importance is. Nor do we know how specific the links are between these predictive measures and mathematics. Yet, the answers to these unsolved questions are important. They will affect what children are taught in mathematics and how this teaching is organized. To put this in the form of concrete questions, should teachers emphasize mathematical reasoning more than they do now? Would it be better to concentrate on strengthening the children's calculation skills? Is there a case for trying to improve children's WM, and would such an improvement influence their progress in non-mathematical as well as in mathematical skills?

The present study

We shall describe the results of longitudinal research over a period of 5 years with a large number of children that provides some answers to these questions. We employed four predictive measures, which were of children's quantitative reasoning, of their calculation skills, of their general intelligence, and their WM. Working with a sample from the Avon Longitudinal Study of Parents and Children (ALSPAC), we investigated the links between these measures and the children's performance in three school subjects, Mathematics, Science, and English over the subsequent 5 years.

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 psychological tests such as the Wechsler Intelligence Scale for Children (WISC-III; Wechsler 1992) and a Test of Mathematical Reasoning; the latter two assessments were given to the children when they were in their fourth year in school.

The WISC includes (1) a sub-test, backward digit recall, which is a measure of WM that Gersten, Jordan and Flojo (2005) identified as a reliable predictor of mathematical difficulties and (2) a sub-test, Arithmetic, designed to assess numerical skills.

The Test of Mathematical Reasoning was designed by Nunes and Bryant (see Nunes, Campos, Magina, & Bryant 2001), drawing on the work of van den Heuvel-Panhuizen (1990). The items in these tests require very simple arithmetic computations but make clear demands on relational reasoning.

The children's mean age when they were given the WISC was 8 years 6 months, and their mean age when they took the mathematical reasoning test was 8 years 9 months.

The project also contains information about the participants' results in two standardized tests of mathematical achievement, designed by the UK government and administered by teachers, referred to as Key Stage Assessments. One assessment, Key Stage 2 (KS2), was given to the children when they were in sixth grade; their mean age at the time was 11 years and 2 months. The second assessment, Key Stage 3 (KS3), was given to the children when they were in ninth grade; their mean age at the time was 14 years and 1 month. Both KS tests measure a variety of aspects of mathematics and are seen as ecologically valid measures of mathematical achievement because of the role that they play in the British educational system.

Method

Participants

ALSPAC is a longitudinal study of children born in Avon in the West of England in 1991–92. Golding, Pembey, Jones, and the ALSPAC team (2001) described the variety of methods used to engage the interest of pregnant women in participating, which were wide reaching in the community and engaged the cooperation of health professionals working with pregnant women. At the time of recruitment, the ALSPAC team compared the sample with that described in the British national sample of the Child Health and Education study and found the ALSPAC sample to be comparable to the national sample in many ways, including rural versus urban living, ethnic background, and prevalence of different health indicators (for details, see Golding *et al*. 2001). The characteristics of the sample originally recruited for ALSPAC are reflected in the samples analysed here. Thus, there was not much selective attrition in the sample in terms of ethnic background or maternal occupation at KS2 or KS3 (for details on the sample, see Nunes, Bryant, Sylva, & Barros 2009). In the sample analysed for KS2 results, 48% of the children are boys and 52% are girls; in the sample analysed for KS3 results, 47% of the children are boys and 53% are girls.

The children were recruited for participation in the measures considered here in two ways. For the individually administered measure (the WISC), they were recruited by a letter sent to the mother inviting her to attend a clinic. If no response was obtained, a second letter was sent. The ALSPAC data base contains information on the IQ of 7,354 children.

The mathematical reasoning assessment was administered to the children by the teachers. Teachers were invited to participate if there was a child in the class who was included in the ALSPAC sample; 5,234 children were given the mathematical reasoning measure.

KS2 and KS3 mathematics are also administered by teachers but these are not optional; they are part of the information required by the government from the schools for monitoring school performance. The ALSPAC database contains presently information on 12,472 for KS2 and 8,519 children for KS3 results. The decrease in sample size from KS2 to KS3 results from the fact that, when the data were analysed, KS3 data were not available for the children born later, in 1992.

In total, full information regarding all the measures used in this study was available on 2,579 participants for the analyses of KS2 mathematical achievement and for 1,680 for the analyses of KS3 achievement.

Measures

Predictors

Mathematical reasoning. The mathematical reasoning task included three types of items, additive reasoning about quantities, additive reasoning about relations, and multiplicative reasoning items. Additive reasoning items involve part-whole relations (Carpenter, Ansell, Franke, Fennema, & Weisbeck 1993; Carpenter & Moser 1982; Vergnaud 1982); the parts can be static (i.e., two parts form a whole) or involve change (i.e., a quantity is added to or subtracted from an original one). Within the domain of part-whole relations, one can ask questions about quantities or about relations. The latter can involve, for example, comparisons (e.g., how much more does x have than y?) or distance (how far is x from y?), as illustrated by the items in Figure 1 (developed from an item used by Brown 1981). Examples of additive reasoning items are presented in Figures 1 and 2, left.

Multiplicative reasoning items involve proportional relations between two quantities (Vergnaud 1983). Figures 2, right, and 3 illustrate multiplicative reasoning items used in this study; the example on the left, Figure 3, was adapted from van den Heuvel-Panhuizen (1990).

All items are presented orally with the support of pictures, which reduces memory demands and affords the use of a variety of strategies in finding the numerical answers: for example, counting, addition, or multiplication might be used to solve the item presented in Figure 3, left. The children's booklets, where they are asked to write their answers, contain no text, only drawings; the story is read by the teacher to the class.

The assessment contains a total of 17 items and it is not timed; administration usually takes approximately 25–30 min. The child's score is the number of correct answers.

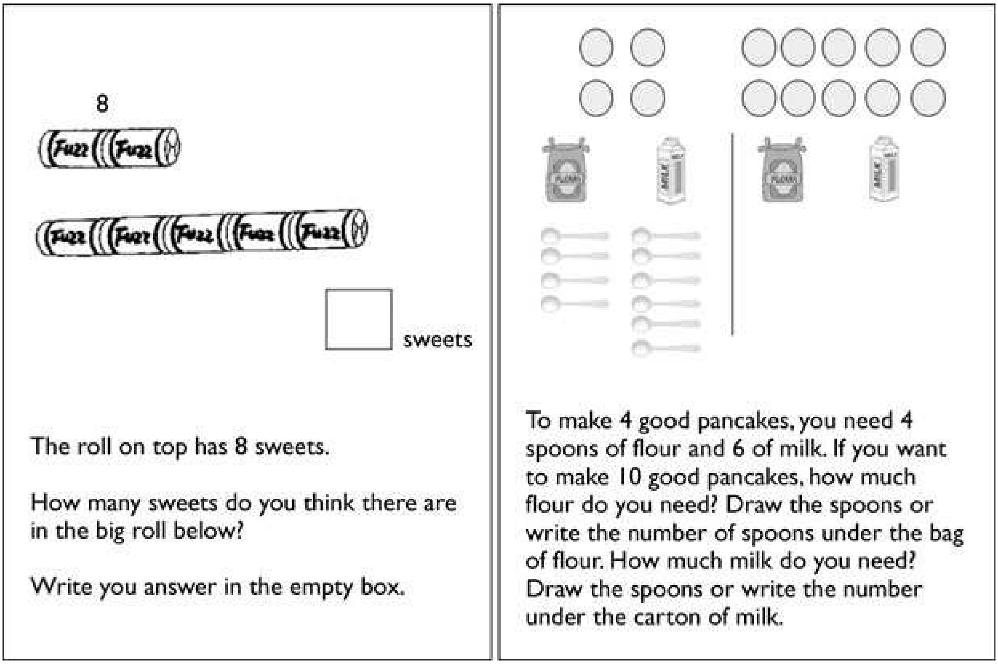

Figure 3. Two multiplicative reasoning problems in which the arithmetic is quite simple, once the student knows how to think about the relations between quantities.

Cronbach's alpha inter-item reliability for this assessment was 0.74; thus, the assessment had a good level of internal consistency (Kline 1999).

Arithmetic. The WISC Arithmetic sub-test was used to assess children's computational ability. It is a standardized measure of arithmetic knowledge, in which the questions are presented as word problems that place little demand on relational reasoning. The arithmetic required to solve the problems becomes progressively more difficult, although, the relational reasoning demanded of the child does not increase. For example, item 8 is: 'Joseph has 5 cakes. He gives 1 to Sam and 1 to Alice. How many does he have left?' and item 18 is: 'A shop had 25 cartons of milk and sold 14 of them. How many cartons were left?' When small numbers are used, pre-school children show high rates of success in such problems (Becker 1993; Carpenter, Hiebert, & Moser, 1981; Carpenter & Moser 1982; Carpenter *et al*. 1993); a result interpreted in the literature as demonstrating that these problems require little relational reasoning (Vergnaud 1979; 1982). The child is required to solve all the problems without the use of paper and pencil and the test is interrupted after three consecutive failures. The child's score is the number of correctly answered questions. The split-half reliability for 8-year-olds is 0.78 (Wechsler 1992, p. 60); and the average correlation with the Wide Range Achievement Test Arithmetic Score is .62 (Wechsler 1992, p. 76), which makes this a valid and economical assessment of children's arithmetic knowledge, thus, suited for large-scale studies such as this.

Controls

Working memory (WM). The most common measure of WM used in longitudinal predictive research of mathematical achievement and difficulties is the Backward Digit

Recall, which is one of the sub-tests in the WISC-III (Wechsler 1992) and also a sub-test in the Working Memory Battery for Children (Pickering & Gathercole 2001). In this sub-test, children hear a series of digits and are asked to repeat them in the reverse order. The number of digits to be recalled is increased by one over the trials until the children can no longer recall the digits in the correct order. The child's score is the highest number of digits correctly recalled in the reverse order in four (of six) trials of the same span. According to Pickering and Gathercole (2001), backward digit recall has a high loading on the central executive factor of WM, which is a strong predictor of mathematical achievement (Gathercole & Pickering 2000). The split-half reliability of the WISC-III Digit Recall for age 8 is 0.84 (Wechsler 1992, p. 60).

As argued earlier on, it is important to control for WM to assess whether the relation of mathematical reasoning and arithmetic to mathematical achievement is specific. This measure will be used in regression analyses, each with one of the outcome measures, where this subtest will be the only measure of information processing used as a control.*General intelligence.* It was expected that children's performance in a measure of general intelligence correlates significantly with all three of the above predictors and also with the children's mathematical achievement. It was, therefore, desirable to include a measure of general intelligence as a control in the regression analyses. The measure used in this study was the WISC-III (Wechsler 1992). For the regression analyses, the general IQ was estimated on the basis of 10 sub-tests; the sub-tests Arithmetic and Digit Span (forward and backward recall) were excluded from the estimation, as these were entered separately in the analyses.

Outcome measures

The outcome measures of mathematical achievement were standardized assessments designed by the British government to measure school standards. They are administered and scored by the teachers, and often used to make decisions about students' placement in mathematics attainment streams. Therefore, these are not only ecologically valid measures but also high-stake tests. The tests are redesigned each year; the participants in this study took the tests in different years as they are from different birth cohorts; the majority took KS2 tests in 2003 and KS3 tests in 2006. The descriptions presented here are taken from KS2 in 2003 and KS3 in 2006.

KS2 had three papers (see QCA 2003: http://www.emaths.co.uk/KS2SAT.htm); students were allowed to use a calculator only in one of these. All papers were timed; the mental arithmetic paper was timed by question and the other two were timed as a whole. The papers assess a variety of aspects of mathematical knowledge that children are taught about by the time they reach their sixth year in primary school: for example, knowledge of decimals, arithmetic (calculation and problem solving), geometric reasoning, measurement of space and time, identification of number patterns of sequences of figures, graph reading (line and bar graphs).

The mental arithmetic paper includes mostly questions with no references to quantities but simply to numbers (e.g., 'divide ninety by three'; 'subtract one point nine from two point seven'; 'when h has the value twelve, calculate five h minus two'), but there are also questions that asses knowledge of scales of measurement and involve calculation (e.g., 'how many grams are in 12 kilograms'; 'how much must I add to four point ninety to make six pounds') and questions related to geometry (e.g., 'look at the figures on the paper; put a ring around the figure which has only one line of symmetry' and 'look at the clock; what angle is made by the hands of the clock when at four o'clock'; 'calculate the perimeter of a rectangle which is eleven meters long and four meters wide'.) Only two of the 20 questions in this paper made reference to a practical situation (e.g., 'A yogurt costs forty-five pence. How many yogurts can be bought for five pounds?'). This paper is more similar to the WISC arithmetic than to the mathematical reasoning assessment.

The second paper to be answered without a calculator included 26 items, distributed in different categories such as calculation either as a direct command (e.g., 'calculate 309 – 198'; 'calculate 2307 × 8') or in numerical expressions with one missing number represented by a box (e.g., '600 × 4 = box'; '50 ÷ box = 2.5'). Item also referred to measurement and money (e.g., 'how many coins of only 1 p, only 10 p or only 20 p do you need to have £1.60'), reading tables and graphs, rounding numbers, identifying patterns in series of numbers and figures, naming geometrical figures, calculating percentages, and perimeter. Six items referred to practical situations; five of these do not require much relational reasoning (e.g., 'Tom and Nadia have 16 cards each, Tom gives Nadia 12 of his cards, how many cards do Tom and Nadia each have now?') and one does require relational reasoning because it is about unequal division ('30 children are going on a trip. It costs £5 including lunch. Some children take their own packed lunch and pay only £3. The 30 children pay a total of £110. How many children are taking their packed lunch?')

The paper in which the children are allowed to use a calculator contains similar types of questions but with larger numbers and fractions. There are seven (of 24) items that use numbers without a reference to quantities; in general, the calculations are more difficult in this paper than in the preceding one. For example, the missing number problems with a box include questions such as '37 × box = 111'; '225 – box = 115' and 'box × box = 378'. Calculations also involve larger numbers such as 'what is 3/8 of 980'. Number relations are explored in questions such as 'here are five digit cards [the digits are from 1 to 5]; fill in the boxes to make this sum correct [the children have to fill in three addends, one with a single digit and two with two digits]; the result is 60' and 'Karen makes a fraction using two number cards. She says, my fraction is equivalent to1/2. One of the number cards is 6. What could Karen's fraction be? Give both possible answers'; ' k + m + n = 1500; m is three times as big as n; k is twice as big as n; calculate the numbers m, k and n'. Although these problems do not involve reference to practical situations, they involve thinking about relations between numbers. Problems that refer to practical situations used in this paper sometimes require the students to obtain information from graphs or tables and sometimes present the information in words; seven of 24 items can be described as word problem in this sense. An example of a simple word problem is 'there are 5 balloons in a packet. There are 18 packets in a box. How many balloons are there altogether in a box?' An example of a more difficult word problem is '250 000 people visited a theme park in one year. 15% of the people visited the park in April and 40% visited the park in August. How many people visited the park in the rest of the year?'). The remaining 10 items are about geometry (e.g., 'draw two straight lines from point A to divide the shaded shape into a square and a triangle'; 'which of the diagrams below shows is a reflection of the mirror line for this figure'), measurement [e.g., 'Here is a clock (a digital display shows 14:53). What time will the clock show in 20 minutes'; 'write these lengths in order, starting with the shortest: 1/2 m; 3.5 cm; 25 mm; 20 cm'] and probability (a square pinner is divided into unequal sections; the students are asked to verify which statements about the probability of the spinner stopping at one number are correct).

This detailed description shows that there is a wide range of items in the papers. Although the mental calculation paper is mostly about numbers without reference to quantities, and could be seen as giving greater weight to arithmetic than reasoning in the KS tests, calculation and reasoning are for success in the three papers.

KS3 tests are administered 3 years after the KS2 tests. Similar to KS2, there are three papers in KS3 tests, one of which is mental arithmetic, and calculators are allowed in only one of the papers. KS3 papers are designed for four different levels of difficulty to avoid giving the most difficult papers to students who would find them too frustrating. Thus, there are 12 different papers for KS3; a detailed description of these tests is beyond the scope of this paper. Suffice it to say here that they include questions designed to assess the same topics included in KS2 at a higher level of difficulty. Three new topics appear, proof, probability, and algebra; questions about fractions, calculations with decimals, and missing numbers in expressions involve larger values. Some calculation questions in the paper in which the students are not allowed to use a calculator explore the students' understanding of properties of operations. For example, students are asked in one question: 'part a: show that 9×28 is 252; part b: What is 27×28? You can use part a to help you'.

The KS tests provide two types of score, one in attainment level, which varies between 1 and 9, and a points measure, which is a finer numerical scale. Our analyses showed no difference in the pattern of results produced by the two scales. We report here the analyses carried out with the finer score. The two scales do not vary in range across the years in which the tests were given.

The children in the sample included two cohorts, as they were born on different years, and thus, they took different tests. In each case, we ran the overall analyses with both cohorts and also separately by cohort. The results replicated the patterns across cohorts; so, we report here only the overall results, in which both samples are combined.

Results

Correlations

The correlations in Table 1 provide some preliminary information about the relationships between predictors and outcome measures. The control measures, WM, and general intelligence, and the two predictor variables, arithmetic and mathematical reasoning, were correlated with each other as well as with each of the outcome measures.

The correlations between mathematical reasoning and the mathematics tests at 11 (.66) and 14 years (.68) were stronger than those between arithmetic and each of these outcome measures. It is noteworthy that mathematical reasoning and general intelligence show almost the same correlations with each of these two measures of mathematics achievement. WM shows weaker correlations with the outcome measures of mathematics than mathematical reasoning. Thus, the children's ability to reason about quantitative relations was a particularly good predictor of their progress in mathematics over the next 5 years.

The children's arithmetic scores also predicted their performance in the national tests of mathematics well (0.57 with the 11-year and 0.58 with the 14-year national tests of mathematics), though not as strongly as their mathematics reasoning scores had done. The children's general intelligence was more strongly related with their mathematics achievement than arithmetic, but arithmetic scores had higher correlations with the outcome measures than WM.

Table 1 also provides some preliminary evidence on the question of specificity. We had argued that the children's scores in the mathematical reasoning and arithmetic tasks would predict their progress in mathematics more strongly than in science or English.

Table 1. Correlations between each of the predictors (mathematical reasoning and arithmetic), control measures (IQ estimate and backward digit span) and outcome measures (KS2 and 3 Mathematics)

Predictors	IQ estimate without WM and arithmetic	Backwards digit span	Mathematical reasoning	Arithmetic
IQ estimate without WM and Arithmetic	1			
Backwards digit span	.29**	1		
Mathematical reasoning	.51**	.28**	1	
Arithmetic	.50**	.32**	.49**	1
KS2 Math	.63**	.33**	.66**	.57**
KS3 Math	.68**	.34**	.68**	.58**
KS2 Science	.65**	.26**	.55**	.45**
KS3 Science	.69**	.28**	.58**	.49**
KS2 English	.60**	.31**	.48**	.44**
KS3 English	.57**	.29**	.50**	.42**

**Correlation is significant at the .01 level (two-tailed).
*Correlation is significant at the .05 level (two-tailed).
$N = 2{,}413$ for KS2 outcome analyses; $N = 1{,}588$ for KS3 outcome analyses

We expected that the two measures of general processing ability, WM and general intelligence, would show similar correlations with mathematics, science, and English. This hypothesis receives some support from the correlational analysis. Mathematical reasoning and arithmetic both show stronger correlations with the mathematics tests than with the science and English tests; the differences between the correlations between the predictors with mathematics and with science or English are significant statistically both at KS2 and KS3. This result is not surprising because the sample size was rather large in both analyses, and N-3 is one of the terms in the numerator of the formula for calculating the t value when comparing correlations in correlated samples (Ferguson 1971). Therefore, this information becomes more important when one considers whether the correlations between the cognitive measures of intelligence and WM with KS2 and KS3 mathematics also differed significantly from those observed for English and science KS tests.

First, one should note that the correlation between general intelligence and achievement in science is actually higher both at KS2 and KS3 than the correlation between general intelligence and mathematical achievement. Although the difference in the coefficients is small, one must conclude that general intelligence predicts science achievement at least as well as it predicts mathematical achievement. It is noteworthy that the opposite is true when it comes to English achievement: general intelligence predicts mathematics (and science) better, and significantly so, than English achievement. At KS2, the difference between the correlations of general intelligence with mathematical achievement ($r = .63$) and with English achievement ($r = .60$), although small, still reaches significance at .01 level ($t = 3.37$; $p < .01$).

In contrast, the children's WM scores predicted the students' performance in mathematics and English better than it predicted their science achievement at both KS tests. The correlation between WM and mathematical achievement did not differ significantly from the correlation between WM and English achievement at KS2 ($t = 1.17$; ns), but this difference was significant at KS3 ($t = 3.933$; $p < .01$).

This close look at the correlations, therefore, suggests that both general intelligence and WM cannot be seen as specific predictors of mathematics, as the first predicts science achievement better than mathematics, whereas, the second predicted KS2 mathematics and English achievement equally well, in spite of the effect of the very large sample size on the statistical comparison between correlations.

Finally, we turn to the relations among the predictors and the controls. The correlation between mathematical reasoning and arithmetic was quite high ($r = .49$) but far from perfect. The reason for the strength of this correlation may be that the children could use an arithmetical calculation in each of the mathematical reasoning items (although, as pointed out, they could also solve many problems by reasoning and counting). Although we did our best to make sure that the calculations did not tax the arithmetical skills, they may still have done so for some of the children in the project. This would have led to the correlation between mathematical reasoning and the arithmetic task, which only measures ability to calculate correctly. We, therefore, had to control for the link between the two predictors in any further examination of the relationship between mathematical reasoning, and we shall report how we did this in multiple regressions in the next section.

The correlations between these two tasks and the controls, WM and general intelligence, were positive and significant; they were higher for general intelligence and lower for WM. The existence of these correlations also emphasized the importance for us, when considering the relationship between each of the predictors and the various outcome measures, to control for the impact of the other two predictors. This is an important test of the specificity of the relationship between the predictors, mathematical reasoning and arithmetic, and the outcome measures, KS2 and KS3 mathematics attainment.

Multiple regressions: Prediction of mathematical achievement

The next question was whether the two predictors – mathematical reasoning and arithmetic – made independent contributions to the prediction of mathematics achievement. In order to test whether these contributions are specific to the measures, rather than explained by more general abilities, we will in these analyses control for WM and general intelligence. We used four hierarchical, fixed-order multiple regressions to answer this question. The outcome measure in two analyses was the children's performance in mathematics tests at 11 years (KS2, see Table 2), and in the other two analyses, it was their performance in mathematics tests at 14 years (KS3, see Table 3). The difference between each pair of analyses was the predictor that was entered as the last step in the equation. Mathematical reasoning was the last step in one analysis and arithmetic in the other. This allowed us to see if each of the two variables accounted for a significant amount of variance in the outcome measure after the effect of the other predictor had been controlled.

The first point to note about the two pairs of regressions is that they accounted for a highly satisfactory amount of the variance in the children's performance in the two mathematics assessments. The multiple regressions in which the children's mathematical achievement at 11 years (KS2, Table 2) was the outcome measure accounted for 58% of the variance in that measure. The analyses accounted for 62% of the total variance in the children's achievement in mathematics at 14 years (KS3; Table3), which is a very high level for a longitudinal predictive analysis over a period of 5 years.

Both tables show that each of the two predictors accounted for a significant amount of additional variance in KS mathematics 2 and 3, after controlling for all the other independent variables in the analysis. Thus, each predictor made an independent contribution to predicting children's mathematical achievement over the next 5 years.

Table 2. Prediction of achievement in KS2 mathematics. Two multiple regressions in which the first three variables entered were the controls: (1) age at key stage assessment; (2) IQ; and (3) WM. The fourth and fifth steps are changed across analyses to test whether the main predictors make independent contributions to the prediction of KS2 attainment. The *B* and β coefficients are those for the regression when all the predictors have been entered ($N = 2{,}413$)

Step in regression		R^2 change	β coefficient	*B*	Standard error of *B*
All regressions first step	Age at outcome	.033**	.129**	0.658	0.070
Second step	WISC IQ estimate without arithmetic and WM	.369**	.316**	0.389	0.020
Third step	WISC WM	.031**	.086**	2.027	0.335
First regression					
Fourth step	WISC arithmetic	.073**	.211**	1.202	0.092
Fifth step	Maths reasoning task	.076**	.344**	2.272	0.109
Second regression					
Foutth step	Maths reasoning task	.119**			
Fifth step	WISC arithmetic	.030**			

**Significant at $p < .001$

This prediction is specific in the sense that it cannot be explained by general factors, such as age, intelligence, and WM.

The tables also show that the β value for mathematical reasoning was higher than for all the other measures, including general intelligence, in the prediction of children's mathematical attainment at 11 years. The β co-efficient was higher for general intelligence than for mathematical reasoning in predicting the children's mathematical achievement at 14 years but was still far greater for mathematical reasoning than for

Table 3. Prediction of achievement in KS3 mathematics. Two multiple regressions in which the first three variables entered were the controls: (1) age at key stage assessment; (2) IQ; and (3) WM. The fourth and fifth steps are changed across analyses to test whether the main predictors make independent contributions to the prediction of KS3 attainment. The *B* and β coefficients are those for the regression when all the predictors have been entered ($N = 1{,}595$)

Step in regression		R^2 change	β coefficient	*B*	Standard error of *B*
All regressions first step	Age at outcome	.011**	.085**	0.028	0.005
Second step	WISC IQ estimate without arithmetic and WM	.463**	.404**	0.030	0.001
Third step	WISC WM	.019**	.063**	0.092	0.024
First regression					
Fourth step	WISC arithmetic	.059**	.184**	0.065	0.007
Fifth step	Math reasoning task	.075**	.340**	0.138	0.008
Second regression					
Fourth step	Math reasoning task	.111**			
Fifth step	WISC arithmetic	.023**			

**Significant at $p < .001$

the other variables. There is a remarkable consistency in the results of the two analyses: the order of importance of the two predictors is the same – mathematical reasoning is a stronger predictor than arithmetic – and the β values are also quite similar. This is an impressive replication, considering that the two outcome measures of mathematics achievement are different and were given to the children 3 years apart.

Multiple regressions: Specificity of prediction

The children were also given national assessments of science and of English at the same time as they took the mathematics assessments. This allowed us to investigate how specific to mathematics was the pattern of relationships that we have just described.

The children's mathematical reasoning scores, and to a lesser extent their arithmetic scores, predicted their mathematical achievement over the following 5 years relatively well. Was this relative success in the predictive power of these two variables specific to their mathematical achievement, or did they predict other outcome measures just as well? We expected the first of these two alternative answers would be the right one, because our hypothesis was that the relative success of these two predictors was specifically due to mathematical factors. For example, we claimed that the outstanding success of the mathematical reasoning task as a predictor was due to its success in testing children's ability to reason about quantities.

The second question was whether the children's WM and intelligence scores would predict their achievement in non-mathematical subjects as well as in mathematics. Our hypothesis was that variation in children's WM and intelligence would play as large a part in the children's achievement in subjects such as English as in their mathematics because of the general information processing skills measured by these tests.

Our aim was to compare the strength of the relations among each of the two predictors and each of the controls and each of these three kinds of outcome measures. We had already done the appropriate analyses with achievement in mathematics as the outcome measure (see Tables 2 and 3). We now carried out additional multiple regressions with exactly the same five predictor variables as in the regressions that we described in Tables 2 and 3, but with different outcome measures. There were four new outcome measures: the national test scores in science at 11 and at 14 years and the English scores at 11 and 14 years.

Table 4 presents the β values for the relationship between the three main predictors and the three kinds of outcome measure (achievement in mathematics, science and English at 11 and 14 years). The figures in this table support our hypothesis about the predictive power of the arithmetic and also about WM and general intelligence. We had predicted that the arithmetic scores would be much more strongly related to achievement in the national tests in mathematics than in science and English and that WM and general intelligence scores, in contrast, would predict achievement in non-mathematical as well as in mathematical subjects. Table 4 shows that the children's arithmetic scores had higher β coefficients in the regressions with mathematics as the outcome measure than in regressions with science and English as the outcomes. The table also shows that the WM and intelligence scores were as strongly related to the children's achievement in English as in mathematics and science both in the 11- and in the 14-year national achievement tests. As we expected, there was no evidence that these measures of general processing are more important in learning about mathematics than about either of the other two subjects.

Table 4 did provide one surprise. We had expected that the mathematical reasoning scores would be far more strongly related to achievement in mathematics than either

Table 4. The relations between the specific predictors and the controls and the children's achievement in mathematics, science, and English

Outcome measures: Achievement tests at 11 years			
	Mathematics	Science	English
Percent of total variance explained by all five steps in the regression	58.0	47.7	40.8
β coefficients			
Mathematical reasoning	0.34	0.19	0.14
Arithmetic	0.21	0.11	0.15
WM	0.09	0.04	0.12
General Intelligence	0.32	0.49	0.41
Outcome measures: Achievement tests at 14 –years			
	Mathematics	Science	English
Percent of total variance explained by all five steps in the regression	62.7	54.8	38.3
β coefficients			
Mathematical reasoning	0.34	0.21	0.14
Arithmetic	0.18	0.11	0.13
WM	0.06	0.03	0.07
General Intelligence	0.40	0.54	0.42

in science or in English, and the results showed that this was strikingly the case. The children's mathematical reasoning scores were very strongly related to their achievement in mathematics, as we have already noted, while the relations between this predictor and science and English were much lower. So, these figures do demonstrate the specificity in the relationship between mathematical reasoning and children's mathematical achievement that we had expected.

However, the children's mathematical reasoning scores also predicted their achievement in science a great deal better than their achievement in English, as well as predicting science achievement much better than arithmetic and WM did. What is the reason for this relatively strong relationship between the children's mathematical reasoning and their achievement in science? One possibility is that it is due to the strong mathematical element that undoubtedly is a part of the scientific curriculum. Scientific exercises involve a great deal of measurement and calculation, of course, and many scientific concepts, such as density and temperature, are intensive quantities (Howe, Nunes, & Bryant 2010; Howe, Nunes, Bryant, Bell, & Desli 2010; Nunes & Bryant 2008; Nunes, Desli, & Bell 2003) and their measurement is based on ratios; children, therefore, have to be able to reason about proportions to understand several aspects of science. None of this is true of English lessons. Therefore, the reason for the quite high connection between children's mathematical reasoning and their eventual achievement in science may exist for specifically mathematical reasons. So, the pattern of predictions that we have just described actually supports the idea of a highly specific connection between mathematical reasoning and children's understanding and use of mathematics.

It is also interesting that the children's scores in arithmetic did not predict their achievement in science any better than in English. This suggests that, in learning about science, it is more important for children to be able to reason about quantities than to manage to do the actual calculations correctly.

Discussion

As far as we know, ours is the first large-scale longitudinal study to have measured the contribution of mathematical reasoning separately from mathematical calculation, to children's school mathematical achievement. Therefore, the finding of a particularly strong link between children's reasoning and their mathematical achievement in school is a result of considerable importance. This result makes a significant contribution to the debate on how much emphasis teachers should give to mathematical reasoning and to knowledge of arithmetic in the classroom. Time is a precious resource in the classroom, and our results suggest that greater investment in developing students' mathematical reasoning should produce a higher pay-off in terms of students' mathematical achievements.

Arithmetic made a smaller, but nevertheless significant and independent contribution, to the children's achievement in mathematics. WM and general intelligence also made independent contributions to the prediction of mathematical achievement. The contribution of WM was relatively modest in size, but still consistently significant even after controlling for general intelligence. The contribution of general intelligence was comparable to that of mathematical reasoning at age 11 years but slightly higher at age 14 years.

Although mathematical reasoning and knowledge of arithmetic were significantly and moderately correlated, their separate and independent contributions to the prediction of mathematics achievement signify that they should be treated as distinct constructs. For some time, researchers in psychology and mathematics education have made a distinction between procedural and conceptual knowledge, most often using qualitative descriptions of students' problem solving activities, in which the students reveal that they have some procedural knowledge without understanding or, alternatively, some conceptual knowledge in the absence of related procedural skills (see, e.g., Hiebert & Lefevre 1986; Rittle-Johnson & Siegler 1998). However, procedural and conceptual knowledge are highly correlated (Hallett, Nunes, & Bryant 2010; Rittle-Johnson, Siegler, & Alibali 2001), and this makes it difficult to justify their independence with quantitative methods such as factor analyses, which only consider the relationship between measures of procedural and conceptual knowledge. This study provides a clear empirical basis for distinguishing mathematical reasoning as a form of conceptual knowledge and knowledge of arithmetic as separate constructs.

An important next step will be to explore the boundary between these two forms of knowledge. We are certain that all the items in our reasoning task were genuine tests of quantitative reasoning: this reasoning was about part-whole relations and additive composition, and one-to-many correspondence and proportions, and had an adequate level of difficulty for this age level. Working with a younger cohort, Nunes *et al.* (2007) used items that assessed reasoning about one-to-one and one-to-many correspondence, additive composition and the inverse relation between addition and subtraction as predictors of mathematics achievement for younger children. Future research on the impact of children's reasoning on their mathematical achievement at school could explore other aspects of reasoning as well as different aspects of children's knowledge of arithmetic.

We are also confident that the task that we used to test children's ability to calculate was also valid. All the questions in the arithmetic task are quite explicit about the calculation that is needed, and thus, the constraint on the children's performance is not how well they reason but how well they do the calculation that they are asked to do. However, there are many other tasks that could be used in predictive studies, and which could shed light on the relationship between arithmetic and mathematical reasoning. Recently, for example, there has been a great deal of interest in the possibility of children using an internal number line to judge which of two numbers is the larger. The evidence for the existence of this form of representation is the 'distance effect' that takes the form of discriminations between numbers that are further apart from each other being easier than discriminations between numbers that are close to each other (Butterworth 2005; Dehaene 1997; Durand *et al*. 2005; De Smedt *et al*. 2009). Because the judgements in such comparison tasks are about the number system, they could be relevant to children's ability to use numbers to make calculations. However, it could be argued that they are also about simple quantitative relations (larger, smaller) and would, therefore, count as reasoning as well. Longitudinal research that includes such tasks as well as the reasoning and arithmetic tasks that were included in our study could contribute to a better understanding of the role that the abilities measured by such tasks play in mathematical achievement.

The multiple regressions, in which the outcome measures were science and English, demonstrated a strong specificity in the relations between mathematical reasoning and arithmetic and the children's progress in mathematics at school. Both variables predicted children's mathematical achievement much better than their achievement in English, which is an entirely non-mathematical subject. This strongly suggests that these two variables predict mathematics because they are measures of specific mathematical abilities.

The relatively strong relation between the children's mathematical reasoning scores and their achievement in science suggests that mathematical reasoning also plays an important part in children's learning about science at school. We need to know the reason for this undoubtedly important connection. Our suggestion is that it is at least partly due to the mathematical nature of some basic scientific concepts. Temperature and density, for example, are intensive quantities: this means that both variables are based on ratios and, thus, make demands on children's proportional reasoning. Thus, density is the ratio of mass to volume, and children will only understand how to vary density and what the effects of such variations will be if they grasp density's proportional nature. This idea, and other possible alternative ideas, about the reason for the strong relationship between mathematical reasoning and children's scientific achievement need to be investigated. In general, the link between children's mathematical knowledge and their progress in learning about science is a neglected topic, but our results suggest that it is an important one.

Our study provides strong evidence that mathematical reasoning should receive a greater emphasis than calculation skills from the early years in primary school and arguably to the end of secondary school. This innovation should produce gains in students' mathematical and scientific achievement in the future.

References

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

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

Barrouillet, P., & Lépine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology*, *91*, 183–204.

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

Bradley, L., & Bryant, P. (1983). Categorising sounds and learning to read – A causal connection. *Nature*, *301*, 419–421.

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., 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.

Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology*, *65*, 1–24.

Butterworth, B. (2005). Developmental dyscalculia. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 455–467). Hove: Psychology Press.

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., Hiebert, J., & Moser, J. M. (1981). Problem structure and first grade children's initial solution processes for simple addition and subtraction problems. *Journal for Research in Mathematics Education*, *12*, 27–39.

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem solving. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 10–24). Hillsdale, NJ: Lawrence Erlbaum.

Cramer, K., Post, T., & Currier, S. (1993). Learning and teaching ratio and proportion: Research implications. In D. T. Owens (Ed.), *Research ideas for the classroom: Middle grades mathematics* (pp. 159–178). New York: Macmillan.

D'Amico, A., & Guarnera, M. (2005). Exploring working memory in children with low arithmetical achievement. *Learning and Individual Differences*, *15*, 189–202.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology*, *103,* 186–201.

De Bock, D., Van Dooren, W., Janssens, D., & Verschaffel, L. (2002). Improper use of linear reasoning: An in-depth study of the nature and the irresistibility of secondary school students' errors. *Educational Studies in Mathematics*, *50*, 311–334.

De Corte, E., Verschaffel, L., & Van Coillie, V. (1988). Influence of number size, problem structure and response mode on children's solutions of multiplication word problems. *Journal of Mathematical Behavior*, *7*, 197–216.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 469–479.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. London: The Penguin Press.

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.

Ferguson, G. A. (1971). *Statistical analysis in psychology and education* (3rd ed.). New York: McGraw-Hill.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at seven years of age. *British Journal of Educational Psychology*, *70*, 177–194.

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.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*,*38*, 293–304.

Golding, J., Prembey, M., Jones, R., & and the ALSPAC team (2001). ALSPAC – The Avon Longitudinal Study of parents and children I. Study methodology. *Paediatric and Perinatal Epidemiology*, *15*, 74–87.

Greer, B. (1988). Nonconservation of multiplication and division: Analysis of a symptom. *Journal of Mathematical Behavior*, *7*, 281–298.

Hallett, D., Nunes, T., & Bryant, P. (2010). Individual differences in conceptual and procedural knowledge when learning fractions. *Journal of Educational Psychology*, *102*, 395–406.

Hiebert, J. & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Lawrence Erlbaum and Associates.

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

Howe, C., Nunes, T., & Bryant, P. (2010). Intensive quantities: Why they matter to developmental research. *British Journal of Developmental Psychology*, *28*, 307–329.

Howe, C., Nunes, T., Bryant, P., Bell, D., & Desli, D. (2010). Intensive quantities: Towards their recognition at primary school level *British Journal of Educational Psychology – Monograph Series II, Number 7 – Understanding number development and difficulties*, 101–118.

Jordan, N., Levine, S., & Huttenlocher, J. (1994). Development of calculation abilities in middle and low income children after formal instruction in school. *Journal of Applied Developmental Psychology*, *15*, 223–240.

Jordan, N., & Montani, T. (1997). Cognitive arithmetic and problem solving: A comparison of children with specific and general mathematics difficulties. *Journal of Learning Disabilities Quarterly*, *30*, 624–634.

Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London: Routledge.

Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual-spatial working memory, and preschool quantity-number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology*, *103*, 516–531.

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.

Nunes, T., & Bryant, P. E. (1995). Do problem situations influence children's understanding of the commutativity of multiplication? *Mathematical Cognition*, *1*, 245–260.

Nunes, T., & Bryant, P. (2008). Rational numbers and intensive quantities: Challenges and insights to pupils' implicit knowledge. *Anales de Psicología*, *24*(2), 262–270.

Nunes, T., Bryant, P., Evans, D., & Bell, D. (2010). The scheme of correspondence and its role in children's mathematics. *British Journal of Educational Psychology Monograph Series II, Number 7 – Understanding number development and difficulties*, *1*(1), 83–99.

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., Sylva, K., & Barros, R. (2009). *Development of maths capabilities and confidence in primary school* (No. Research Report DCSF-RR118). London: Department for Children, Schools and Families. Retrieved from: http://publications.dcsf.gov.uk/eOrderingDownload/DCSF-RB118.pdf (last visited 20 January 2011).

Nunes, T., Desli, D., & Bell, D. (2003). The development of children's understanding of intensive quantities. *International Journal of Educational Research*, *39*, 652–675.

Nunes, T., Campos, T. M. M., Magina, S., & Bryant, P. (2001). *Introdução à Educação Matemática. Os números e as operações numéricas* (An introduction to mathematics education: Number and operations). São Paulo: Proem Editora. (republished: 2005, São Paulo, Editora Cortez.)

Passolunghi, M. c., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology*, *88*, 348–367.

Pickering, S., & Gathercole, S. (2001). *Working memory test battery for children (WMTB-C) Manual*. London: The Psychological Corporation.

QCA. (2003). Key stage tests. Department for education and skills, UK. Retrieved from: http://www.emaths.co.uk/KS2SAT.htm

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 46–36.

Rittle-Johnson. B. & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills* (pp. 75–110). East Sussex, UK: Psychology Press.

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. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology-General*, *127*, 377–397.

Squire, S., & Bryant, P. (2003). Children's understanding and misunderstanding of the inverse relation in division. *British Journal of Developmental Psychology*, *21*, 507–526.

Stern, E. (1992). Spontaneous used of conceptual mathematical knowledge in elementary school children. *Contemporary Educational Psychology*, *17*, 266–277.

Stern, E. (1999) Development of mathematical competencies. In F.E. Weinert & W. Schneider (Eds.) *Individual development from 3 to 12: Findings from the Munich Longitudinal Study* (pp. 154–170). Cambridge: Cambridge University Press.

Swanson, H. L., & Beebe-Frankenberger, M. E. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, *96*, 471–491.

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

van de Rijt, B. A. M., van Luit, J. E. H., & Pennings, A. H. (1999). The construction of the Utrecht early mathematical competence scales. *Educational and Psychological Measurement*, *59*, 289–309.

Van Den Heuvel-Panhuizen, M. (1990). Realistic arithmetic/mathematics instruction and tests. In K. Gravemeijer, M. van den Heuvel, & L. Streefland (Eds.), *Contexts free productions tests and geometry in realistic mathematics education* (pp. 53–78). Utrecht, The Netherlands: Research group for Mathematical Education and Educational Computer Centre, State University of Utrecht.

Van Dooren, W., Bock, D. D., Hessels, A., Janssens, D., & Verschaffel, L. (2004). Remedying secondary school students' illusion of linearity: A teaching experiment aiming at conceptual change. *Learning and Instruction*, *14*, 485–501.

Vergnaud, G. (1979). The acquisition of arithmetical concepts. *Educational Studies in Mathematics*, *10*, 263–274.

Vergnaud, G. (1982). A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. In T. P. Carpenter, J. M. Moser, & R. T. A (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 60–67). Hillsdale (NJ): Lawrence Erlbaum.

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

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 557–628). Greenwich, CT: Information Age Publishing.

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

Received 7 September 2010; revised version received 22 March 2011