The contribution of spatial ability to mathematics achievement in middle childhood

Katie A. Gilligan [a], Eirini Flouri [a] & Emily K. Farran [a]

[a] Department of Psychology and Human Development, UCL Institute of Education, University College London, 25, Woburn Square, London WC1H 0AA, UK

Corresponding author contact details:

Katie Gilligan, Department of Psychology and Human Development, UCL Institute of Education, University College London, London WC1H 0AA, UK. Tel: 07349134106; E-mail: Katie.gilligan.15@ucl.ac.uk

Additional contact details:

Eirini Flouri, e.flouri@ucl.ac.uk

Emily Farran, e.farran@ucl.ac.uk

Word Count (Text and references): 11,013 words

Abstract

Strong spatial skills are associated with success in Science, Technology, Engineering and Mathematics (STEM) domains. Although there is convincing evidence that spatial skills are a reliable predictor of mathematical achievement in pre-school children and in University students, there is a lack of research exploring associations between spatial and mathematics achievement in the primary school years. To address this question, this study explored associations between mathematics and spatial skills in children aged 5 and 7 years. The study sample included 12,099 children who participated in both Wave 3 (mean age: 5; 02, years; months) and Wave 4 (mean age: 7; 03, years; months) of the Millennium Cohort Study. Measures included a standardised assessment of mathematics and the Pattern Construction subscale of the BAS-II to assess intrinsic-dynamic spatial skills. Spatial skills at 5 and 7 years explained a significant 8.8 % of the variation in mathematics achievement at age 7, above that explained by other predictors of mathematics including gender, socio-economic status, ethnicity and language skills. This percentage increased to 22.6 % without adjustment for language skills. This study expands previous findings by using a large scale, longitudinal sample of primary school children, a population that have been largely omitted from previous research exploring associations between spatial ability and mathematics achievement. The findings that early and concurrent spatial skills contribute to mathematics achievement at 7 years highlight the potential of spatial skills as a novel target in the design of mathematics interventions for children of this age range.

Key words

Development, Mathematics, Millennium Cohort Study, Spatial Cognition, STEM

The contribution of spatial ability to mathematics achievement in middle childhood

Recent studies have proposed the use of spatial ability training as a means to improve both spatial and mathematics achievement in children. The prospective benefits of such an intervention would be significant for many countries including the UK. From an educational perspective students from the UK typically perform at or below the average level of their international counterparts in assessments of mathematics and science (Mullis, Martin, Foy, & Arora, 2012; Organisation for Economic Co-operation and Development [OECD], 2013). Improving Science, Technology, Engineering and Mathematics (STEM) success is also a particularly pertinent economic issue. STEM-related industries contribute over 400 billion pounds to the UK economy per year (Centre for Economics and Business Research [CEBR], 2015; also see Berressem, 2011) yet over 39% of firms in need of STEM employees have reported difficulties recruiting suitably qualified candidates (Confederation of British Industry [CBI], 2013). Spatial ability has been identified as a reliable predictor of adult achievement in STEM domains in many large-scale (N > 500) longitudinal studies following both normative and intellectually gifted populations through adolescence and adulthood (Shea, Lubinski, & Benbow, 2001; Wai, Lubinski, & Benbow, 2009). For example, it has been reported that students who pursue STEM careers and complete STEM degrees at both undergraduate and masters level have higher spatial ability scores at 13 years (Wai et al., 2009). If effective, spatial training interventions could offer a promising alternative to traditional attempts at improving STEM achievement, and could in turn confer both educational and economic benefits.

The theory that spatial training interventions could improve mathematical ability is supported by findings that spatial ability is malleable and changeable, that changes and improvements in spatial ability are durable over time and that improvements in certain spatial skills transfer to other non-trained spatial skills (Ehrlich, Levine, & Goldin-Meadow, 2006; Uttal et al., 2013). Spatial training interventions have demonstrated improvements in spatial skills as early as in primary school children (Bruce & Hawes, 2015; Taylor & Hutton, 2013; Uttal et al., 2013). However, few studies with children have investigated the transfer of spatial training gains to mathematical domains. Cheng and Mix, (2014) were among the first to report significant gains in a mathematical calculation task following spatial skill training in children aged 6-8 years. However, these gains were specific to missing term problems and were not observed on number fact or multi-digit calculation problems. Furthermore, subsequent results from Hawes, Moss, Caswell and Poliszczuk (2015) failed to demonstrate improvements in non-verbal arithmetic or missing number problems following a similar spatial training intervention. Taken together, these findings suggest a need for further investigation of the associations between spatial cognition and mathematics achievement in children, in order to best design effective training interventions.

Understanding spatial cognition

Spatial cognition as described by Hart and Moore, (1973) is 'the knowledge and internal or cognitive representation of the structure, entities, and relations of space; in other words, the internalised reflection and reconstruction of space in thought" (p. 248). Despite almost half a century of research, attempts to define sub-divisions within spatial cognition have led to the emergence of many contrasting typologies (Linn & Petersen, 1985). This study will adopt Uttal et al.'s. (2013) top-down model of spatial typology which has gained significant endorsement and popularity owing to the convincing neurological, behavioral and linguistic evidence supporting it (Chatterjee, 2008; Palmer, 1978; Talmy, 2000). This model is built on two fundamental, theoretical distinctions. The first distinction is between intrinsic and extrinsic representations and the second is between static and dynamic representations. Intrinsic representations are those that relate to the size and orientation of an object, its parts and their relationships. In contrast, extrinsic representations relate to the location of an object, the relationship between objects as well as the relationship between objects and their reference frames. Dynamic representations require movement such as bending, moving, folding, scaling or rotation, whilst static representations do not. By combining these two fundamental distinctions, Uttal et al. (2013) propose a two by two classification of spatial skills thus rendering four distinct sub-domains; intrinsic-static, intrinsic-dynamic, extrinsic-static and extrinsic-dynamic. Uttal et al.'s (2013) model offers a useful framework within which to investigate spatial cognition. However, it is possible that Uttal's proposed distinctions are over refined and spatial thinking can also be described using similar, broader categories (Mix et al., 2016; Newcombe, in press). Furthermore, the use of this model is sometimes complicated by the fact that some spatial tasks recruit a number of Uttal et al.'s (2013) spatial sub-domains in combination, and cannot be easily mapped onto one sub-domain within this framework (Okamoto, Kotsopoulos, McGarvey, & Hallowell, 2015).

Spatial ability and success in mathematics in child populations

Findings from cross-sectional studies on spatial thinking and mathematics in childhood populations render mixed results. Significant correlations have been reported between mental rotation (an intrinsic-dynamic spatial skill) and both calculation and arithmetic in children aged 6-8 years (Cheng & Mix, 2014; Hawes et al., 2015). In contrast, Carr, Steiner, Kyser and Biddlecomb (2008) found no significant association between mental rotation and standardised maths performance in children aged 7 years. For other intrinsic spatial tasks including dis-embedding and spatial visualisation, task performance has been associated with a range of mathematics achievement measures at approximately 10 and 11.5 years respectively (Markey, 2010; Tosto et al., 2014). Mix and colleagues (2016) completed an in-depth investigation of cross-domain connections between spatial and mathematical cognition in children, using factor analysis. The results indicated that, when entered into a single factor analysis, spatial and mathematics scores emerged as distinct factors with significant cross-domain loadings for some tasks. They also highlighted that some spatial sub-domains were more highly related to mathematics at certain ages. Mental rotation (an intrinsic-dynamic spatial skill) was singled out as an important predictor of mathematics[1] for kindergarten children (mean age 6 years).

Longitudinal studies in younger child populations have also found associations between spatial skills and mathematics. For example, Verdine et al. (2014) report that spatial skills at age 3, as assessed using the Test Of Spatial Assembly (TOSA), a measure of intrinsic-dynamic spatial ability, predict a significant proportion of the variation in mathematical problem solving measured using the Wechsler Individual Achievement Test (WIAT) at age 4. Similarly, a preliminary report from Farmer et al. (2013) indicates that spatial performance on the TOSA at 3 years is significantly correlated with a combined mathematics measure at 5 years. Wolfgang, Stannard, & Jones, (2001) also demonstrated that early spatial play, in particular adaptiveness and integration in block play, is associated with later mathematics achievement. However, these results should be interpreted cautiously as

[1] Mathematics as described in this study refers to a mathematical factor derived using factor analysis. The variables included in the factor analysis included place value, word problems, missing terms, calculation and number line estimation (Mix et al., 2016).

interpretation of free block play is subject to errors and cannot easily be mapped onto spatial typology frameworks (Wolfgang et al., 2001).

The longitudinal studies reviewed above used intrinsic-dynamic tasks, namely a single measure of spatial ability. The use of an intrinsic-dynamic task is theoretically useful, given the strong association of this spatial subdomain with mathematics observed in studies with older children and adults. Furthermore, findings from correlational studies suggest that intrinsic-dynamic spatial skills may have particular associations with mathematics in the early primary school years (Mix et al., 2016). Nonetheless, studies that explored associations beyond the intrinsic-dynamic subdomain also report similar findings, which suggests that the association between spatial ability and mathematics competence is wide-ranging. For example, early spatial visualisation (a task which requires input from a number of spatial sub-domains) at age three was found to be a significant predictor of arithmetic at age 10 (Zhang et al., 2014). Furthermore, a composite measure of spatial performance (assessing performance on a range of spatial sub-domains) at age 7 significantly predicted mathematics achievement levels at approximately 10 years (Carr et al., 2017). Similarly, Casey and colleagues (2015) reported, in a study with girls, that spatial skills (a composite measure generated from block design and mental transformation tasks) at age 7 were significant predictors of mathematics reasoning at age 11. Longitudinal studies of primary school students have also reported correlations between visuospatial skills (including visual perception and motor integration) at age 6 and mathematics achievement at age 9. However, these findings were not specific to a single spatial domain and the results were confounded by the visual and motor demands of the tasks used (Lachance & Mazzocco, 2006; Mazzocco & Myers, 2003).

Overall, there is evidence that spatial skills, particularly intrinsic-dynamic spatial skills, are associated with mathematics in the pre-school and early school years. However, the majority of studies to date are based on relatively small populations of children at specific ages. There is a need to replicate and extend these findings with large scale, longitudinal studies. Additionally, many of these studies focus on specific sub-components of mathematics such as arithmetic or calculation (Cheng & Mix, 2014; Hawes et al., 2015). Given that some content areas of mathematics, such as geometry and trigonometry, are inherently spatial (Hawes et al., 2015; Newcombe, 2013) there is a need to investigate the influence of spatial skills on mathematics more generally, using more holistic measures of mathematics. The use of more comprehensive measures of mathematics including algebra, geometry, problem solving and number processing among other skills, may be more reflective of the range and diversity of mathematical content areas that children are required to master in the mathematics classroom, beyond calculation skills.

Furthermore, no known studies investigate the longitudinal associations between intrinsic-dynamic spatial skills and mathematics in the early primary school years. Identifying the nature of this relationship is particularly relevant given evidence that early interventions have significantly higher rates of success, compared to those implemented at later stages of development (Heckman, 2006).

Explaining associations between spatial cognition and mathematics

As described, there is evidence of associations between spatial skills and mathematics in both child and adult populations. However, the nature of these associations is largely unknown. Findings from brain imaging studies suggest similar patterns of brain activation in the completion of spatial and mathematics tasks (Hubbard, Piazza, Pinel, & Dehaene, 2005;

Umiltà, Priftis, & Zorzi, 2009; Walsh, 2003). This may reflect shared processing requirements for the completion of both mathematics and spatial tasks. The notion that mathematics and spatial cognition may share processing requirements is not surprising given that; many mathematics tasks such as geometry and trigonometry are spatial (Hawes et al., 2015; Newcombe, 2013); and many mathematical constructs are believed to be represented spatially in the brain using a mental number-line (Barsalou, 2008; Lakoff & Núñez, 2000). This is further supported by findings that differences in the spatial presentation of equations and numbers influence performance in mathematics tasks (Fisher, Borchert, & Bassok, 2011; Landy & Goldstone, 2007; McNeil & Alibali, 2004). For these reasons one might expect individuals with better spatial abilities to have better mathematical performance. Spatial scaling, spatial visualisations and form perception have all been proposed as shared processes that may be required in the completion of both mathematics and spatial tasks (Mix et al., 2016). However, it is unclear whether all mathematics and spatial tasks require similar processes, and whether these shared processes are stable over development.

Other predictors of mathematics achievement

Beyond the spatial domain, success in mathematics has been associated with cognitive abilities including general cognitive skills (von Aster & Shalev, 2007), working memory (Alloway & Alloway, 2010), executive function (Verdine et al., 2014) and attention (Merrell & Tymms, 2001). Furthermore, early language skills have also been proposed to influence mathematical success. LeFevre et al. (2010) propose that linguistic measures are a reliable early predictor of achievement in mathematics, while Moll, Snowling, Göbel and Hulme (2015) propose that individuals with language difficulties or reading problems also demonstrate poor performance in mathematical achievement tests. However, while language might play a distinct role in mathematical development, correlations between numeracy and literacy achievement may also reflect the presence of an underlying general intelligence or "g" factor (Alloway & Alloway, 2010; Mayes, Calhoun, Bixler, & Zimmerman, 2009).

Differences in mathematical performance have also been associated with social and demographic factors including socio-economic status (SES) (Byrnes & Wasik, 2009) and gender (Halpern et al., 2007). Children from low SES backgrounds typically perform less favorably on mathematical measures when compared to their higher SES counterparts (Byrnes & Wasik, 2009; Oakes, 2005). In contrast, evidence for gender differences in mathematics achievement is less well supported and many studies argue against gender differences in this domain (Lindberg, Hyde, Petersen, & Linn, 2010). Some studies that report better male performance in mathematics and science have attributed gender differences to differences in interests, neurological, or cognitive outcomes, in turn shaped by biological, genetic and environmental influences (Halpern et al., 2007; Penner & Paret, 2008). However, beyond these findings, few studies have explored associations between spatial and mathematical skills, controlling for these predictors of mathematics achievement (Lachance & Mazzocco, 2006; Mazzocco & Myers, 2003). It is also unclear whether associations between mathematics and spatial skills differ with gender or SES. In one recent study, Carr and colleagues (2017) found no significant gender differences in longitudinal associations between spatial and mathematics skills.

The current study

As described, spatial skills in adolescence have been associated with both educational and occupational success in STEM domains in adulthood. Furthermore, in childhood populations, there is evidence that spatial skills in the pre-school years reliably predict success in mathematics later in development. More specifically, there is evidence from correlational studies showing associations between intrinsic-dynamic spatial skills and mathematics in the early primary school years, above those seen for other spatial sub-domains (Mix et al., 2016). This is the first study to investigate both concurrent and longitudinal relationships between intrinsic-dynamic spatial skills and mathematics in the early primary school years (age 5-7 years). Additionally, while the majority of the studies to date focus on specific sub-components of mathematics such as arithmetic or calculation, this study explores associations between spatial skills and mathematics achievement more generally. In particular, it explores the value of intrinsic-dynamic spatial skills as a longitudinal predictor of mathematics achievement. Using data from the Millennium Cohort Study (MCS) it explores associations between spatial skills and mathematics in middle childhood using a large scale, general population longitudinal sample. It investigates changes in intrinsic-dynamic spatial skills over time and identifies the contribution of spatial skills at 5 and 7 years to achievement in mathematics at age 7. It also extends previous research by exploring spatial skills and mathematics while accounting for the roles of other known predictors of mathematics performance, i.e. gender, SES and language skills. Thus, this study will identify reliable associations between a specific spatial skill and mathematics achievement at primary school ages which, if significant, could enable the effective design of targeted age-based mathematics interventions, the outcomes of which may have both educational and economic implications.

Materials and Methods

The Millennium Cohort Study

The Millennium Cohort Study (MCS) is a longitudinal population-based study of children born in the United Kingdom between 2000 and 2002. Participants of the MCS were sampled using a stratified, clustered design, ensuring adequate representation of disadvantaged and ethnic minority groups and over-representation of children living in the smaller UK countries including Scotland, Northern-Ireland and Wales. To date, the MCS has collected 5 waves of data during which the children in the study were approximately aged 9 months, and 3, 5, 7 and 11 years respectively. The MCS uses questionnaires, interviews and a range of cognitive assessments with cohort members, their families and teachers to collect information on a wide range of variables including; cognitive development; child and parental physical and mental health; income and poverty; parenting; ethnicity and schooling among others.

The current study focuses on the Millennium cohort during Waves 3 and 4, for which suitable measures of spatial ability are available. Wave 3 was completed between February 2006 and January 2007 when the study participants (N = 15,460) were around 5 years old. Wave 4 was completed between January 2008 and February 2009 when the participants (N = 14,043) were aged around 7 years. The Centre for Longitudinal Studies, who manage the MCS, attained ethical approval for Wave 3 of the MCS from the London Multi-Centre Research Ethics Committee of the National Health Service (NHS) while ethical approval for Wave 4 of the study was obtained from the Northern and Yorkshire Research Ethics Committee of the NHS. No additional ethical approval was required for this study. Further information on sampling, response rates and participation is published elsewhere (Hansen, 2012).

Participants

The initial study sample included the eldest cohort child from each MCS family (N = 19,244). The inclusion of a single participant from each family ensured that clustering effects did not occur. Participants with missing data on any of our cognitive and educational measures (see below) were subsequently excluded from the sample rendering a sample size of 12,537 participants. Furthermore, participants who did not indicate that they spoke English only or mostly English at home were excluded from this study in order to remove variance created by differences in language comprehension (438 participants excluded). Therefore, the final sample size for this study was 12,099 participants. The Organisation for Economic Co-Operation and Development (OECD) equivalised income scores[2] (Hansen & Joshi, 2008) at Wave 4 were used as a measure of SES in this study. As shown in table 1, the final results of this study should be viewed in light of; the slight under-representation of participants in low income families and; the slight over-representation of white participants relative to all other ethnic groups. Further information on the demographics of the final sample compared to those of the excluded sample can be found in Appendix 1 of the supplementary material.

[2] OECD equivalised income scores convert reported household income into a modified scale based on the number and age of all members of the family (Hansen & Joshi, 2008).

		Final Sample	
		N	*% total*
Gender			
	Male	6079	50.2
	Female	6020	49.8
Ethnic group			
	White	10463	86.5
	Mixed	324	2.7
	Indian	259	2.1
	Pakistani & Bangladeshi	534	4.4
	Black or Black British	340	2.8
	Other Ethnic group	122	1.0
	Missing	57	.5
OECD Equivalised Income Quintiles			
	Lowest	2267	18.7
	Second quintile	2394	19.8
	Third quintile	2502	20.7
	Fourth quintile	2475	20.5
	Highest quintile	2450	20.2
	Missing	11	0.1
		Mean	*SD*
Age Wave 3 (years)			
	Male	5.215	0.247
	Female	5.210	0.244
Age Wave 4 (years)			
	Male	7.227	0.247
	Female	7.219	0.245

Table 1: Demographic characteristics of the study sample (unweighted data)

Measures

As shown in Table 2, all participants completed a series of cognitive measures across Wave 3 and Wave 4 of the MCS. This included a subset of items from a standardised test of mathematics for seven-year-olds (National Foundation for Educational Research [NFER], 2004) in addition to a selection of measures taken from the British Ability Scales II (BAS II), a standardised test battery that measures cognitive ability (Elliott, Smith, & Mc Cullock, 1996). For all test measures, age-based standardised test scores, converted to z-scores are reported.

Table 2. Cognitive measures included in the MCS waves 3 and 4

Test Measure	**Wave 3**	**Wave 4**
BAS II Pattern Construction	✓	✓
BAS II Naming Vocabulary	✓	
BAS II Word Reading		✓
NFER Progress in Maths		✓

BAS II British Ability Scales II, *NFER* National Foundation for Educational Research

Mathematics skills

A shortened version of the National Foundation for Educational Research Progress in Maths (NFER PiM) test for seven-year-olds was administered at Wave 4 as a measure of mathematics (NFER, 2004). The NFER PiM is an assessment of mathematics ability and includes a wide assortment of items on all aspects of the National Mathematics Curricula including questions on numbers, shapes, measurement and data handling. Age-based standardised scores were based on 6 month age intervals and were calculated based on the full length NFER PiM test normed in 2004.

Spatial skills

This study used the Pattern Construction subscale of the BAS-II as a measure of spatial ability (BASII; Elliott et al., 1996; Hill, 2005). This non-verbal reasoning task is modelled on Kohs' traditional Block Design Test (Kohs, 1919). The task requires participants to copy a stimulus pattern using a set of blocks. The block faces are either all yellow, all black, or half-yellow, half-black. Participants must re-create a stimulus pattern by rotating, re-arranging and joining the blocks. As such the task falls within the intrinsic-dynamic sub-domain of spatial cognition as described by (Uttal et al., 2013). In easier trials the stimulus pattern is presented using 3-D blocks. Harder trials use 2-D picture representations of the stimulus pattern. Task success is measured as accuracy in block orientation and positioning, and response time. Age-based standardised scores were calculated based on three month age intervals (BASII; Elliott et al., 1996; Hill, 2005).

Control variables

Additional sub-tests of the BAS-II included as covariates in analyses were the Naming Vocabulary subscale (Wave 3) which measures expressive verbal ability and the Word Reading subscale (Wave 4) which measures educational knowledge of reading. In the Naming Vocabulary scale children are shown a series of pictures and are asked to name each of them. In the Word Reading scale children are shown words on cards and are asked to read them aloud. Age-based standardised scores for these measures were based on three-month age intervals. Due to the age difference of participants at different waves of the MCS, different language measures were included at Wave 3 and Wave 4. No single language measure was available for both waves.

Analysis strategy

Statistical analyses were completed using IBM SPSS Statistics for windows (version 22). In analyses of variance, where equal variances could not be assumed, the results for unequal variance were reported. Post-hoc Games-Howell or Hochberg's GT2 tests were used appropriately in cases where the assumption of homogeneity of variance was violated or met, respectively. Missing OECD equivalised income values, which accounted for 0.1% of cases, and missing ethnicity values, which accounted for 0.5% of cases, were calculated using the multiple imputation function in SPSS. MCS weights to account for the original stratified, clustered design of the MCS sample and sample attrition and non-response were applied to all analyses unless otherwise stated. All N's reported are based on unweighted data.

Initial descriptive statistics were completed to provide an overview of overall performance patterns across tasks. T-tests and ANOVA were used to investigate the main effects of gender and SES (income groups) on task performance for all test measures including both language and spatial based cognitive tasks, and mathematics achievement. A correlation matrix was completed to investigate the relative associations between performance measures and to inform subsequent general linear models. To explore the role of spatial skills as a predictor of mathematics achievement, we used general linear models in SPSS. General linear models allowed us to use the MCS weights to account for sample design, attrition and non-response. Furthermore, the use of age adjusted z-scores for all cognitive task measures and age allowed for meaningful comparison of unstandardised b values within models. (Although age-based standardised scores were used throughout, these scores were based on three month (BAS II) or six month (NFER PiM) age intervals and did not account for age-based variability within these age brackets. Hence, exact age at Wave 4 was included as a predictor in all models. While this extra adjustment for age is a more conservative approach, comparable results were found in models where age was not included as a predictor.)

Regression models

Below we explain in detail the models fitted. The first model investigated the influence of spatial skills on mathematics in light of other confounds including gender, SES (income groups), ethnicity, age and language skills. We present the additional variation explained by spatial skills above that explained by demographic and language measures (Naming Vocabulary and Word Reading at Wave 3 and 4 respectively). In this model, spatial task performance (Pattern Construction task performance) at Wave 3 and 4 was considered simultaneously. In addition, following identification of gender and SES (income group) differences in spatial task performance in the bivariate analysis (described below), it was postulated that spatial skills may be differentially associated with mathematics in males and females, and in individuals with different SES (income group status). Hence within this model we explored the role of gender and SES as moderators in the relationship between spatial and mathematics skills by adding interaction terms for gender*spatial skills and SES*spatial skills at Wave 4. Model 2 investigated the presence of shared variation between language and spatial skills. This model explored the role of spatial skills as a predictor of mathematics when controlling for demographic factors only. Language skills were included after spatial skills in this model. Model 3 investigated the value of spatial and language skills at Wave 3 as longitudinal predictors of mathematics achievement at Wave 4. In previous models it is likely that the individual value of Wave 3 measures in predicting mathematics achievement was underestimated due to shared variance between Wave 3 and Wave 4 spatial and language measures respectively. Hence no Wave 4 measures were included as predictors in this model in order to ascertain the value of Wave 3 measures as longitudinal predictors of later mathematics performance. To allow for the comparison of concurrent and longitudinal predictors of mathematics, model 4 investigated the role of spatial and language measures at Wave 4 as concurrent predictors of mathematics achievement. To allow meaningful comparison of model 3 and model 4 (the contributions of wave three and wave four predictors respectively), the order of inclusion of variables in model 4 was identical to model 3. Finally, in order to explore whether the proposed relationship between spatial thinking and later academic achievement is unique to mathematical domains, similar regression models to those described above were also completed with word reading (as opposed to mathematics achievement) at age 7 as the outcome measure. Full details of these analyses can be found in Appendix 2 of the supplementary material.

Results

Overall task performance

Descriptive statistics for each of the cognitive and academic measures used in this study are shown in Table 3. While these results are specifically based on the sample included in this study, they are comparable to those describing the performance of the total MCS sample at Waves 3 and Wave 4 (Hansen, Jones, & Budge, 2010; Hansen & Joshi, 2008).

Table 3. Descriptive statistics for child task performance across Waves 3 and 4 (z-scores, unweighted data)

Wave	Test Measure	N	Maximum	Minimum	Mean	SD
Three						
	BAS II- Pattern Construction	12,099	2.950	-3.123	.000	1.000
	BAS II- Naming Vocabulary	12,099	2.34	-3.249	.000	1.000
Four						
	BAS II- Pattern Construction	12,099	2.436	-3.053	.000	1.000
	BAS II- Word Reading	12,099	1.859	-3.162	.000	1.000
	NFER Progress in Maths	12,099	2.418	-1.883	.000	1.000

BAS II British Ability Scales II, *NFER* National Foundation for Educational Research

Performance differences based on gender and SES

Gender differences

Independent T-tests were carried out to identify differences in task performance based on gender. Due to the large sample size normality was assumed. As shown in Table 4, the results indicate that there was a significant difference in performance between males and females for all tasks. The mean score for females exceeded that for males on all tasks with the exception of mathematics performance where male scores were above those of females. These results should be viewed in light of the relatively small effect sizes reported for all t-tests. Cohen described values of d below .2 as small effects (Cohen, 1988). Hence, the magnitude of

Cohen's d observed in Table 4, ranging from .053 to .177, suggests that the reported differences in performance of males and females on academic and cognitive measures are relatively small but not insignificant.

Table 4. Gender differences in cognitive and mathematics task performance (z-scores, weighted data)

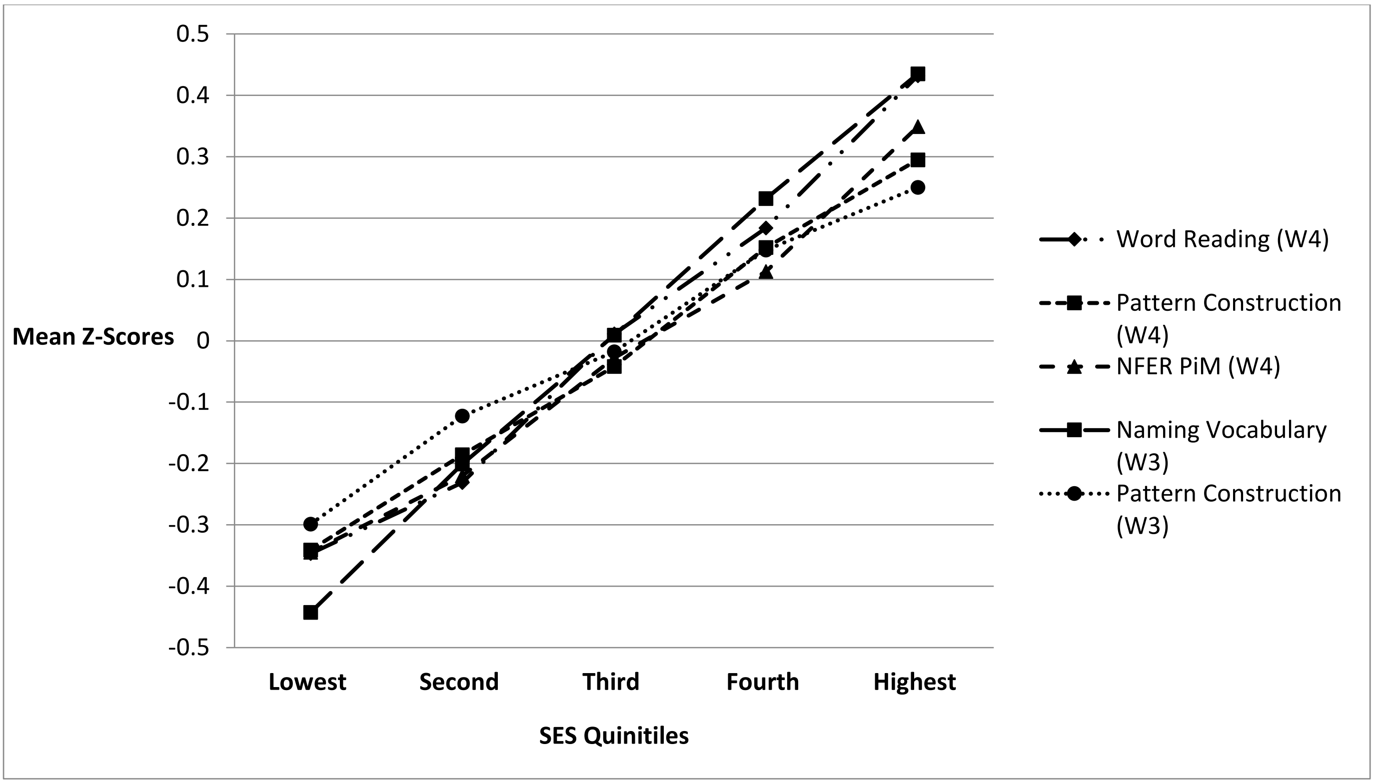
Test Measure	**Gender**				**Statistics**	
	Male (n=6079)		*Female (n= 6020)*		*Test statistic*	*Effect size*
	Mean	*SD*	*Mean*	*SD*	T value	*Cohen's D*
Wave 3						
BAS II- Pattern Construction	-.085	1.043	.09	0.937	-9.810 **	.177
BAS II- Naming Vocabulary	-.005	1.014	.051	0.947	-3.175 **	.057
Wave 4						
BAS II- Pattern Construction	-.04	1.035	.015	0.959	-3.091 **	.055
BAS II- Word Reading	-.051	1.052	.101	0.915	-8.599 **	.154
NFER Progress in Maths	.012	1.036	-.041	0.948	2.973 **	.053

** indicates p < .01, * indicates p < .05, all n's are based on unweighted data, *BAS II,* British Ability Scales II, *NFER*, National Foundation for Educational Research.

SES differences

One-way ANOVA tests with SES as a between participant factor (5 levels) demonstrated significant differences in cognitive and mathematics performance across income levels. As shown in Figure 1, significant differences in performance across income groups were reported for all tasks as follows. Word Reading (Wave 4); $F (4, 12320) = 268.182$, $p < .001$, $\eta_p^2 = .075$; Pattern Construction (Wave 4); $F (4, 12320) = 146.05$, $p < .001$, $\eta_p^2 = .05$; NFER PiM; $F (4, 12320) = 197.929$, $p < .001$, $\eta_p^2 = .058$; Naming Vocabulary (Wave 3); $F (4, 12320) = 291.961$, $p <.001$, $\eta_p^2 = .096$ and; Pattern Construction (Wave 3); $F (4, 12320) = 120.275$, $p <.001$, $\eta_p^2 = .036$). Post-hoc tests revealed significant differences between all SES groups (p <.01 for all). All effect sizes reported are less than .10 and can be classified as small (Cohen, 1988).

Figure 1. Cognitive and mathematics task performance across SES groups (income quintiles) (z-scores, weighted data)



Associations between mathematics and cognitive measures

Bivariate correlations between mathematics performance scores at 7 years (NFER PiM) and all cognitive measures included in this study are shown in Table 5. As expected, there were strong correlations between mathematics and all cognitive measures. Word Reading at Wave 4 had the largest correlation with NFER mathematics scores (r = .529, p < .001), followed by Pattern Construction scores at both Wave 4 (r = .479, p < .001) and Wave 3 (r = .430, p < .001).

Table 5. Correlations between mathematics and cognitive measures (z-scores, unweighted data)

	Measure	Wave 4		Wave 3	
		BAS-II Pattern Construction	BAS-II Word Reading	BAS-II Naming Vocabulary	BAS-II Pattern Construction
Wave 4	NFER Progress in Maths	.479	.529	.386	.430
	BAS-II Pattern Construction		.334	.318	.556
	BAS-II Word Reading			.372	.348
Wave 3	BAS-II Naming Vocabulary				.332

All correlations were significant at the $p < .001$ level, unweighted N = 12,099, *BAS II*, British Ability Scales II, *NFER*, National Foundation for Educational Research

Regression analyses

Model 1

The results of all models are summarised in Table 6. Model 1 sought to determine the contribution of spatial ability to the variation in mathematics achievement while controlling for other known or possible predictors of mathematics ability including language skills, gender, age, ethnicity and SES. This model is the most conservative. Word Reading at Wave 4 and Naming Vocabulary at Wave 3 were both included as language measures, accounting for language skills across two time points. Spatial measures included Pattern Construction scores at both Wave 3 and Wave 4. As the correlations between language and mathematics scores at Wave 4 were greater than those between spatial skills and mathematics performance (shown in Table 5), language measures were added to the model before spatial measures.

Overall the model accounted for 42.4% of the variation in mathematics scores at 7 years. The demographic measures entered in step 1 including gender, age at Wave 4, ethnicity and SES accounted for 7.3% of the variation, while the language measures added in step 2 accounted for 26.3% of the variation. The spatial measures entered in step 3 accounted for an additional 8.8% of the variation, even after accounting for all other predictors. No significant interactions between gender and spatial skills, or SES and spatial skills were reported in step 4 ($p>.05$ for both). All other variables, with the exception of ethnic group, were significant predictors in the final model. The b values, t-statistics and effect sizes indicated that Word Reading and Pattern Construction at Wave 4 make the most significant impact on predicting mathematics achievement.

Model 2

Model 2 explored the role of spatial skills as a predictor of mathematics when controlling for demographic factors only. As seen in model 1, the demographic measures accounted for 7.3% of the variation in mathematics. Spatial scores at Waves 3 and 4 were entered simultaneously in step 2 and accounted for 22.6% of the variation. The language measures entered in step 3 explained an additional 19.8% of the variation. Both spatial and language measures were significant predictors in this model ($p < .001$ for all).

Model 3

Model 3 explored the variation in mathematics achievement predicted by cognitive measures at Wave 3 only. Overall the model accounted for 27.7% of the variation in mathematics scores at 7 years, with demographic measures accounting for 7.3% of this variation. Based on the magnitude of correlations between Wave 3 measures and mathematics achievement (shown in Table 5), spatial scores were added to the model before language scores. The spatial measure accounted for 15.4% of the variation in mathematics, while the language measure accounted for an additional 5.0% of the variation. The b values, t- statistics and effect sizes suggest that Pattern Construction makes the most significant impact on predicting mathematics achievement, followed by Naming Vocabulary.

Model 4

Finally, model 4 explored the variation in mathematics achievement predicted by cognitive measures at Wave 4 only. The final model accounted for 40.1% of the variation in mathematics scores at 7 years, with demographic measures accounting for 7.3% of this variation. Spatial skills were entered in step 2 accounting for 19.2% of the variation in mathematics, while the language measure added in step 3 accounted for an additional 13.5% of the variation. Word Reading made the most significant impact on predicting mathematics achievement, followed by Pattern Construction scores.

Additional Information

For all models, the data analysed obeyed the assumptions of normality. Outliers were defined as any individuals falling outside three standard deviations of the mean for at least one of the continuous variables in a given model. In models 1 and 2, 396 cases were identified as outliers (3.27% of the sample). In model 3, 289 cases (2.39% of the sample) and in model 4, 141 cases (1.17% of the sample) were identified as outliers. All outliers were included in analyses as they account for very small proportions of the sample population and do not significantly influence the findings reported. In addition, there was no justifiable reason to exclude these cases as it is likely that they reflect natural variation in the population.

Table 6. General linear models predicting mathematics achievement at 7 years (weighted data)

Model 1		**B**	***SE***	***t***	***p***	**Partial η^2**	***F***	***df***	***p***	**Partial η^2**	***Adj. R^2***	**Δ *Adj. R^2***
Step 1												
SES (income quintiles)[a]	Lowest	-0.107	0.024	-4.403	< .001	0.002	85.134	11, 11667	<.001	0.074	0.073	
	Second	-0.110	0.023	-4.469	< .001	0.002						
	Third	-0.065	0.023	-2.859	0.004	0.001						
	Fourth	-0.096	0.023	-4.248	< .001	0.002						
Ethnicity[b]	White	0.017	0.066	0.261	0.794	0.000						
	Mixed	0.041	0.075	0.549	0.583	0.000						
	Indian	-0.009	0.085	-0.103	0.918	0.000						
	Pakistani, Bangladeshi	-0.049	0.079	-0.619	0.536	0.000						
	Black, Black British	-0.097	0.076	-1.270	0.204	0.000						
Gender	Male	0.143	0.014	10.116	< .001	0.009						
Age		-0.048	0.007	-6.812	< .001	0.004						
Step 2												
Word Reading (W4)		0.352	0.008	42.674	< .001	0.135	456.543	13, 11665	<.001	0.337	0.336	0.263
Naming Vocabulary (W3)		0.124	0.008	14.767	< .001	0.18						
Step 3												
Pattern Construction (W4)		0.251	0.019	13.504	< .001	0.015	575.005	15, 11663	<.001	0.425	0.424	0.088
Pattern Construction (W3)		0.128	0.009	14.329	< .001	0.017						
Step 4												
Gender* Pattern Construction (W4)		-0.005	0.014	-0.349	0.727	.000	431.627	20,11658	<.001	0.425	0.424	0
SES* Pattern Construction[a] (W4)	Lowest	-0.009	0.022	-0.383	0.701	.000						
	Second	0.012	0.023	0.525	0.600	.000						
	Third	-0.032	0.023	-1.375	0.169	.000						
	Fourth	0.021	0.023	0.919	0.358	.000						

Model 2	b	*SE*	*t*	*P*	Partial η ²	*F*	*Df*	*p*	Partial η ²	*Adj. R^2*	Δ *Adj. R^2*
Step 1											
As seen for model 1[c]						85.134	11, 11667	<.001	0.074	0.073	
Step 2											
Pattern Construction (W4)	0.247	0.007	28.119	< .001	0.063	384.304	13, 11665	<.001	0.300	0.299	0.226
Pattern Construction (W3)	0.127	0.009	14.318	< .001	0.017						
Step 3											
Word Reading (W4)	0.124	0.008	14.787	< .001	0.018	575.005	15,11663	<.001	0.425	0.424	0.198
Naming Vocabulary (W3)	0.352	0.008	42.657	< .001	0.135						
Model 3											
Step 1											
As seen for model 1[c]						85.134	11, 11667	<.001	0.074	0.073	
Step 2											
Pattern Construction (W3)	0.339	0.008	39.891	< .001	0.120	287.004	12, 11666	<.001	0.228	0.227	0.154
Step 3											
Naming Vocabulary (W3)	0.255	0.009	28.260	< .001	0.064	344.472	13, 11665	<.001	0.277	0.277	0.050
Model 4											
Step 1											
As seen for model 1[c]	0.408	0.008	51.287	< .001	0.184	85.134	11, 11667	<.001	0.074	0.073	
Step 2											
Pattern Construction (W4)	0.329	0.008	28.260	< .001	0.133	352.825	12, 11666	<.001	0.266	0.266	0.192
Step 3											
Word Reading (W4)	0.408	0.008	51.287	< .001	0.184	601.428	13, 11665	<.001	0.401	0.401	0.135

[a] The reference category is highest SES quintile [b] The reference category is other ethnic group. *W3*, Wave 3, *W4*, Wave 4 [c] The parameter estimates for the demographic measures entered in step 1 varied very subtly for each of models 1-4, due to differences in the predictors included in each of the models. The exact parameter estimates for step 1 of each model are available on request.

Discussion

In this study, intrinsic-dynamic spatial skills were shown to explain a significant proportion of the variance in mathematics achievement in the early primary school years above that explained by other demographic factors, or language skills alone. Based on a sample of over 12,000 participants, these findings add substantial support to the results of Mix et al. (2016). The current findings highlight both the concurrent and longitudinal roles of spatial skills for general mathematics achievement, assessed in this study by a more comprehensive measure of mathematics than calculation skills alone. The results of this study also extend previous longitudinal findings in pre-school populations and older children (Casey et al., 2015; Verdine et al., 2014; Zhang et al., 2014) to children in the early primary school years. More specifically, this study demonstrates that spatial skills at 5 years explain a unique proportion of the variance in mathematics achievement at 7 years, in middle childhood. Owing to the design of this study it was also possible to investigate shared variation between spatial and language measures. By comparing models that include and exclude language skills, we could estimate the true proportion of variation in mathematics explained by spatial skills. This value is predicted to fall between the more conservative 8.8% result and the more liberal 22.6% result, generated by models that either include or exclude shared variance with language skills respectively.

Further analyses highlighted the individual and unique contributions of Wave 3 measures at 5 years and Wave 4 measures at 7 years of age to the variation in mathematics outcomes at age 7 (Wave 4). In both models, spatial skills explained a substantial proportion (over 15%) of the variation in mathematics performance at age 7. It is interesting to note that the profile of associations between spatial versus language predictors and mathematics achievement at Wave 4 contrasts with that seen for Wave 3. Based on the observed b values, t-statistics and effect sizes, language at age 7 is a stronger predictor of mathematics when compared to spatial skills. In contrast, at age 5, spatial skills are a stronger predictor of subsequent mathematics achievement at age 7, when compared to language skills. Although this pattern of findings may be due to the different languages measures used in the two waves, it may also suggest that while language skills are stronger concurrent predictors of mathematics, spatial skills are stronger longitudinal predictors of mathematics achievement. This is not to say that spatial skills do not have an important concurrent role in mathematics performance, but to highlight the particular longitudinal connections between spatial skills and mathematics performance between 5 and 7 years, in the context of language measures. Previous findings show that spatial skills may be more important for novel mathematics tasks compared to practiced, automatic mathematics skills (Ackerman, 1988; Uttal & Cohen, 2012). At 5 years of age, children in the UK begin formal schooling and thus are faced with large amounts of new mathematics material. The findings of this study may support the notion that children with strong spatial skills at 5 years are better able to learn novel mathematical concepts, which in turn impacts their later mathematics performance. This finding is particularly interesting as it may indicate a particular, positive role for early spatial skills in later mathematics achievement.

Another notable finding was the difference in performance on the Pattern Construction task between Wave 3 and Wave 4. While this may reflect the test-retest reliability of the Pattern Construction task, previous test-retest correlations of 0.88 have been reported for this measure (Elliott, Smith, & McCulloch, 1997). This suggests that performance differences seen in Pattern construction scores across waves may reflect the malleability of spatial skills in middle childhood. As the spatial scores calculated account for age, the findings suggest that other environmental factors or experiences, aside from age-dependent developmental change alone, may influence spatial development between age 5 and 7. Factors influencing this change in spatial ability may include developmental strategy change or environmental factors such as early schooling experiences, exposure to technology or gaming (OFcom, 2015; Spence & Feng, 2010). Identifying these factors could improve understanding of individual differences in spatial skills.

This study also demonstrated that both gender and income were significantly associated with differences in task performance across all measures investigated here. In line with other studies such as Byrnes and Wasik, (2009), these findings show that children from higher socio-economic backgrounds consistently outperformed their lower SES counterparts. This finding was consistent across all tasks. Gender differences were also reported such that females outperformed males in all test measures except for mathematics achievement where male performance was above that of females. However, it is important to recognise that the effect sizes of these findings were very small, suggesting that although gender differences in performance may exist, the size of these differences may be negligible. Importantly however, the findings of this study highlight a slight female advantage in spatial task completion. This contrasts with previous studies in which males (in the pre-school and primary school years) have been reported to outperform females on a range of spatial measures (e.g. Carr et al., 2008; Casey et al., 2008; Casey, Pezaris, & Nuttall, 1992; Johnson & Meade, 1987; Levine, Huttenlocher, Taylor, & Langrock, 1999; Levine, Vasilyeva, Lourenco, Newcombe, & Huttenlocher, 2005). Thus, our findings add to a growing body of literature challenging the existence of significant gender differences, in particular a male advantage, in spatial cognition in young children (Alyman & Peters, 1993; Halpern et al., 2007; Lachance & Mazzocco, 2006; LeFevre et al., 2010; Manger & Eikeland, 1998; Neuburger, Jansen, Heil, & Quaiser-Pohl, 2011).

Future directions and limitations

An important strength of our study was the use of large-scale, general population data, which ensured the generalisability of our findings. The nature of the sampling protocol employed in the MCS enhances the generalisability of the results reported, due to the inclusion of participants from a range of socio-economic backgrounds, in turn associated with both mathematics achievement and performance on language and spatial skills tests, as we showed. However, the use of secondary data to answer novel research questions is dependent on the availability of suitable test measures. While the MCS dataset provides an excellent resource for the examination of the relationship between intrinsic-dynamic spatial skills and mathematics achievement in children aged 7 years, these findings cannot be generalised to other spatial sub-domains. Previous findings indicate that intrinsic-dynamic spatial domains, including the accurate completion of mental transformations, may be particularly useful to mathematics. It has been proposed that intrinsic-dynamic skills can be applied in the completion of measurement tasks, lines of symmetry tasks and equations that are presented in atypical formats (Bruce & Hawes, 2015; Mix & Cheng, 2012). Strong intrinsic-dynamic skills may be useful for tasks of this type as they may allow children to cognitively manipulate aspects of a given task, for example, by folding shapes or re-arranging the order of equations. While associations between other domains of spatial thinking and mathematics are less well understood, there is some indication that different spatial sub-domains may be particularly important for different mathematics tasks at different developmental ages (Mix et al., 2016). For example, extrinsic tasks such as spatial scaling may be particularly important for the ordinal comparison of numbers (Mix, Prather, Smith, & Stockton, 2014) and the use of a mental number line (Dehaene, Bossini, & Giraux, 1993). Future research could establish whether the findings reported here are also applicable to intrinsic-static, extrinsic-static and extrinsic-dynamic skills. Similarly, this study focused on associations between spatial and mathematics skills at age 5 and 7 years only. Using a similar design, it would be important to assess older primary school children, for whom there is limited research in this domain. Using the Pattern Construction scores, used in this study, and the Spatial Working Memory task, the only spatial measure included in Wave 5 (age 11) of the MCS, future studies might link spatial skills in the primary school years with mathematics achievement at secondary school and beyond.

It will also be important for future research to test our findings experimentally. In combination with previous evidence that spatial skills are malleable (Uttal et al., 2013) and that training effects in spatial tasks are transferable to calculation skills (Cheng & Mix, 2014), the findings of this study suggest that, due to the strong associations between intrinsic-dynamic spatial skills and mathematics achievement over time, the implementation of an intervention targeting intrinsic-dynamic spatial skills at age 5 or age 7 may render effective improvements in both spatial skills and later mathematics performance. It has been suggested that "improving children's spatial thinking can have a "two-for-one" effect" (p. 341, Bruce & Hawes, 2015) benefitting both spatial and mathematics outcomes, beyond the gains seen for training interventions targeting mathematics alone (Bruce & Hawes, 2015). That is, improving spatial skills may increase children's capacity to acquire novel mathematics skills, in line with suggestions that spatial thinking is particularly important in the understanding of novel mathematics concepts (Ackerman, 1988; Uttal & Cohen, 2012). In addition, improving spatial skills may directly improve performance on tasks with high spatial demands such as geometry, algebra and symmetry tasks (Bruce & Hawes, 2015; Mix & Cheng, 2012). Training protocols should investigate both of these benefits, i.e., the use of spatial training that is independent to the spatial content of mathematics instruction, as well as the use of spatial thinking within specific subdomains of mathematics. Future intervention paradigms could therefore test if success in mathematics in the classroom requires spatial processing in addition to number-skills. They could also investigate whether spatial training interventions may have far-reaching implications for other aspects of education including language. There is limited research on the influence of spatial thinking on other domains of learning. As we showed in this study, however, it may affect word reading, too.

Conclusion

This study reports significant associations between intrinsic-dynamic spatial skills and mathematics achievement such that spatial task performance at both 5 and 7 years can explain a significant proportion of variation in 7-year-olds' mathematics scores above that described by socio-demographic or language measures. This highlights the potential of training early intrinsic-dynamic spatial skills as a novel method of improving mathematics achievement.

Acknowledgements

Funding for this research was provided by The Bloomsbury Colleges Ph.D. Scholarship Programme and was supported by the National Centre for Curriculum and Assessment (NCCA) Ireland.

Supplementary Material

Appendix 1

Demographics of the final sample compared to those of the excluded sample[3] are shown in Table 7. These demographics are based on unweighted data[4]. As shown, the selection criteria used to generate the final study sample led to small but significant differences in the ages of the samples at Wave 3 and Wave 4. Across both waves, the mean age for the excluded sample was higher in comparison to the included sample. Furthermore, although there is a significant difference in the gender ratio between the samples, the table indicates that the final sample has a more balanced gender distribution, compared to the excluded sample. As expected, the percentage of participants in all non-white ethnic groups was reduced in the final sample leading to a 13.4% increase in the percentage of white participants in the study compared to the relative percentage of white participants in the excluded sample. This can likely be explained by language exclusions. The excluded sample also has significantly higher proportions of participants in the lowest and second quintiles. This may be explained by higher rates of non-response and attrition in the lower income groups. In comparison, the final sample includes approximately even percentages of participants in each income-based quintile, with a slight under-representation of the lower income groups.

[3] The excluded sample comprises all participants present in the original MCS sample who were excluded from this analysis.

[4] As some of the excluded sample were not present at Wave 4, application of Wave 4 weights accounting for sampling design, non-response and attrition was not suitable for this group. Hence Wave 4 weights were applied to neither the excluded nor the final samples.

Table 7: Demographic characteristics of the final study sample compared to participants excluded from analysis (unweighted data)

		Excluded Sample		Final Sample		Test	
		N	*% total*	*N*	*% total*	*Pearsons χ^2*	
Gender							
	Male	3818	53.4	6079	50.2	18.325 ***	
	Female	3327	46.6	6020	49.8		
Ethnic group							
	White	5220	73.1	10463	86.5	578.570***	
	Mixed	265	3.7	324	2.7		
	Indian	237	3.3	259	2.1		
	Pakistani & Bangladeshi	800	11.2	534	4.4		
	Black or Black British	384	5.4	340	2.8		
	Other Ethnic group	177	2.5	122	1.0		
	Missing	62	.9	57	.5		
OECD Equivalised Income Quintiles							
	Lowest	589	8.2	2267	18.7	344.353***	
	Second quintile	468	6.6	2394	19.8		
	Third quintile	295	4.1	2502	20.7		
	Fourth quintile	224	3.1	2475	20.5		
	Highest quintile	173	2.4	2450	20.2		
	Missing	5396	75.4	11	0.1		
		Mean	*SD*	*Mean*	*SD*	*T Value*	*Cohens D*
Age Wave 3 (years)							
	Male	5.232[a]	.252	5.215	0.247	3.554***	.068
	Female	5.227[a]	.258	5.210	0.244		
Age Wave 4 (years)							
	Male	7.304[a]	.297	7.227	0.247	10.901***	.296
	Female	7.296[a]	.283	7.219	0.245		

*** indicates p <.001, ** indicates p < .01, * indicates p < .05, [a] For the excluded sample, ages at Waves 3 and Wave 4 are based on a sample size of 3146 and 1745 participants respectively. This reduction in sample size is due to the large number of participants in the initial sample who did not participate in Wave 3 and/or Wave 4.

Appendix 2

As shown in Table 8, regression models investigating the role of spatial skills in predicting language measures were fitted. Model 1, the most conservative estimate, demonstrates that spatial skills explain a small but significant 2% of the variation in word reading skills even after accounting for demographic and mathematics measures. Model 2, which accounts for demographic measures only (i.e. shared variance between mathematics and spatial skills is not accounted for), demonstrates that spatial skills explain up to 12% of the variation in word reading. This suggests that spatial training interventions may have far reaching implications to other aspects of education beyond mathematics, likely improving both mathematics and language skills (also see Franceschini, Gori, Ruffino, Pedrolli, & Facoetti, 2012; Gabrieli & Norton, 2012).

Table 8. General linear models predicting word reading achievement at 7 years (weighted data)

Model 1		**b**	***SE***	***t***	***p***	**Partial η^2**	***F***	***df***	***p***	**Partial η^2**	***Adj. R^2***	**Δ *Adj. R^2***
Step 1												
SES (income quintiles)[a]	Lowest	-0.404	0.025	-16.283	< .001	0.022	111.736	11, 11667	<.001	0.095	0.094	
	Second	-0.369	0.024	-15.232	< .001	0.02						
	Third	-0.208	0.024	-8.780	< .001	0.007						
	Fourth	-0.132	0.023	-5.606	< .001	0.003						
Ethnicity[b]	White	-0.199	0.07	-2.856	0.004	0.001						
	Mixed	-0.096	0.079	-1.207	0.227	0.000						
	Indian	0.169	0.091	1.868	0.062	0.000						
	Pakistani, Bangladeshi	0.233	0.084	2.789	0.005	0.001						
	Black, Black British	0.280	0.081	3.462	0.001	0.001						
Gender	Male	-0.153	0.015	-10.187	< .001	0.009						
Age		-0.031	0.008	-4.099	< .001	0.001						
Step 2												
NFER Progress in Maths (W4)		0.427	0.009	47.752	< .001	0.164	467.751	12,11666	<.001	0.325	0.324	0.23
Step 3												
Pattern Construction (W3)		0.118	0.009	12.537	< .001	0.013	430.811	14, 11664	<.001	0.341	0.340	0.016
Pattern Construction (W4)		0.044	0.01	4.555	< .001	0.002						
Model 2		**b**	***SE***	***t***	***p***	**Partial η^2**	***F***	***df***	***p***	**Partial η^2**	***Adj. R^2***	**Δ *Adj. R^2***
Step 1												
As seen for model 1[c]							111.736	11, 11667	<.001	0.095	0.094	
Step 2												
Pattern Construction (W3)		0.215	0.01	21.303	< .001	0.037	241.378	13, 11665	<.001	0.212	0.211	0.117
Pattern Construction (W4)		0.184	0.01	18.374	< .001	0.028						

[a] The reference category is highest SES quintile [b] The reference category is other ethnic group. *W3*, Wave 3, *W4*, Wave 4 [c]. The parameter estimates for the demographic measures entered in step 1 varied very subtly for each of models 1-4, due to differences in the predictors included in each of the models. The exact parameter estimates for step 1 of each model are available on request.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, *117*(3), 288–318. https://doi.org/10.1037/0096-3445.117.3.288

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Alyman, C., & Peters, M. (1993). Performance of male and female children, adolescents and adults on spatial tasks that involve everyday objects and settings. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale*, *47*(4), 730–747.

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, *59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Berressem, J. (2011). STEM - Getting to the root of the STEM problem. *Times Education Supplement*. UK.

Bruce, C. D., & Hawes, Z. (2015). The role of 2D and 3D mental rotation in mathematics for young children: what is it? Why does it matter? And what can we do about it? *ZDM*, *47*(3), 331–343. https://doi.org/10.1007/s11858-014-0637-4

Byrnes, J. P., & Wasik, B. A. (2009). Factors predictive of mathematics achievement in kindergarten, first and third grades: An opportunity–propensity analysis. *Contemporary Educational Psychology*, *34*(2), 167–183. https://doi.org/10.1016/j.cedpsych.2009.01.002

Carr, M., Alexeev, N., Wang, L., Barned, N., Horan, E., & Reed, A. (2017). The Development of Spatial Skills in Elementary School Students. *Child Development*. https://doi.org/10.1111/cdev.12753

Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences*, *18*(1), 61–75. https://doi.org/10.1016/j.lindif.2007.04.005

Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The Development of Spatial Skills Through Interventions Involving Block Building Activities. *Cognition and Instruction*, *26*(3), 269–309. https://doi.org/10.1080/07370000802177177

Casey, B. M., Pezaris, E., Fineman, B., Pollock, A., Demers, L., & Dearing, E. (2015). A longitudinal analysis of early spatial skills compared to arithmetic and verbal skills as predictors of fifth-grade girls' math reasoning. *Learning and Individual Differences*, *40*, 90–100. https://doi.org/10.1016/j.lindif.2015.03.028

Casey, B., Pezaris, E., & Nuttall, R. (1992). Spatial ability as a predictor of math achievement: The importance of sex and handedness patterns. *Neuropsychologia*, *30*(1), 35–45. https://doi.org/10.1016/0028-3932(92)90012-B

Centre for Economics and Business Research (CEBR). (2015). *The contribution of engineering to the UK economy- the multiplier impacts*. London.

Chatterjee, A. (2008). The Neural Organization of Spatial Thought and Language. *Seminars in Speech and Language*, *29*(03), 226–238. https://doi.org/10.1055/s-0028-1082886

Cheng, Y. L., & Mix, K. S. (2014). Spatial Training Improves Children's Mathematics Ability. *Journal of Cognition and Development*, *15*(1), 2–11. https://doi.org/10.1080/15248372.2012.725186

Cohen, D. (1988). *Statistical power analysis for the behavioural sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Confederation of British Industry (CBI). (2013). *Changing the pace: CBI/Pearson education and skills survey 2013*. London: CBI.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371–396. https://doi.org/10.1037/0096-3445.122.3.371

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology*, *42*(6), 1259–1268. https://doi.org/10.1037/0012-1649.42.6.1259

Elliott, C., Smith, P., & Mc Cullock, K. (1996). *British Ability Scales Second Edition (BAS II). Administration and Scoring Manual*. London: Nelson.

Elliott, C., Smith, P., & McCulloch, K. (1997). *British Ability Scales Second Edition (BAS II). Technical Manual.* London: Nelson.

Farmer, G., Verdine, B., Lucca, K., Davies, T., Dempsey, R., Hirsh-Pasek, K., & Golinkoff, R. (2013). Putting the pieces together: Spatial skills at age 3 predict to spatial and math performance at age 5. Presented at the Society for Research in Child Development. Seattle, WA.

Fisher, K. J., Borchert, K., & Bassok, M. (2011). Following the standard form: effects of equation format on algebraic modeling. *Memory & Cognition*, *39*(3), 502–515. https://doi.org/10.3758/s13421-010-0031-6

Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A Causal Link between Visual Spatial Attention and Reading Acquisition. *Current Biology*, *22*(9), 814–819. https://doi.org/10.1016/j.cub.2012.03.013

Gabrieli, J. D. E., & Norton, E. S. (2012). Reading Abilities: Importance of Visual-Spatial Attention. *Current Biology*, *22*(9), R298–R299. https://doi.org/10.1016/j.cub.2012.03.041

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The Science of Sex Differences in Science and Mathematics. *Psychological*

Science in the Public Interest, *8*(1), 1–51. https://doi.org/10.1111/j.1529-1006.2007.00032.x

Hansen, K. (2012). *Millennium cohort study: First, second, third and fourth Surveys. A guide to the datasets. (7th ed.)*. Centre for Longitudinal Studies, Institute of Education, London.

Hansen, K., Jones, E., & Budge, D. (2010). *Millennium Cohort Study Fourth Survey: A user's guide to Initial Findings (2nd ed.)*. Centre for Longitudinal Studies, Institute of Education, University of London, London.

Hansen, K., & Joshi, H. (2008). *Millennium Cohort Study Third Survey: A user's guide to Initial Findings.* Centre for Longitudinal Studies, Institute of Education, London.

Hart, R. A., & Moore, G. T. (1973). The development of spatial cognition: a review. In *Image and Environment: Cognitive Mapping and Spatial Behavior*. Chicago: Aldine Publishing Company.

Hawes, Z., Moss, J., Caswell, B., & Poliszczuk, D. (2015). Effects of mental rotation training on children's spatial and mathematics performance: A randomized controlled study. *Trends in Neuroscience and Education*, *4*(3), 60–68. https://doi.org/10.1016/j.tine.2015.05.001

Heckman, J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312(5782),* 1900–1902.

Hill, V. (2005). Through the past darkly: A review of the British Ability Scales Second Edition. *Child and Adolescent Mental Health*, *10*, 87–98. https://doi.org/10.1111/j.1475-3588.2004.00123.x

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, *6*(6), 435–448. https://doi.org/10.1038/nrn1684

Johnson, E. S., & Meade, A. C. (1987). Developmental Patterns of Spatial Ability: An Early Sex Difference. *Child Development*, *58*(3), 725. https://doi.org/10.2307/1130210

Kohs, S. C. (1919). *Block Designs Test*. Wood Dale, IL: Stoelting.

Lachance, J. A., & Mazzocco, M. M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and Individual Differences*, *16*(3), 195–216. https://doi.org/10.1016/j.lindif.2005.12.001

Lakoff, G., & Núñez, R. (2000). *Where mathematics comes from: How the embodies brings mathematics into being.* New York: NY: Basic Books.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 720–733. https://doi.org/10.1037/0278-7393.33.4.720

LeFevre, J., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: longitudinal predictors of performance. *Child Development*, *81*(6), 1753–1767. https://doi.org/10.1111/j.1467-8624.2010.01508.x

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, *35*(4), 940–949.

Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic Status Modifies the Sex Difference in Spatial Skill. *Psychological Science*, *16*(11), 841–845. https://doi.org/10.1111/j.1467-9280.2005.01623.x

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New Trends in Gender and Mathematics Performance: A Meta-Analysis. *Psychological Bulletin*, *136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Linn, M. C., & Petersen, A. C. (1985). Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis. *Child Development*, *56*(6), 1479. https://doi.org/10.2307/1130467

Manger, T., & Eikeland, O.-J. (1998). The effects of spatial visualization and students' sex on mathematical achievement. *British Journal of Psychology*, *89*(1), 17–25. https://doi.org/10.1111/j.2044-8295.1998.tb02670.x

Markey, S. (2010). The relationship between visual-spatial reasoning ability and math and geometry problem *solving. Dissertation Abstracts International: Section B. The Sciences and Engineering*, 70, 7874.

Mayes, S. D., Calhoun, S. L., Bixler, E. O., & Zimmerman, D. N. (2009). IQ and neuropsychological predictors of academic achievement. *Learning and Individual Differences*, *19*(2), 238–241. https://doi.org/10.1016/j.lindif.2008.09.001

Mazzocco, M. M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia*, *53*(1), 218–253. https://doi.org/10.1007/s11881-003-0011-7

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science*, *28*(3), 451–466. https://doi.org/10.1016/j.cogsci.2003.11.002

Merrell, C., & Tymms, P. B. (2001). Inattention, hyperactivity and impulsiveness: their impact on academic achievement and progress. *The British Journal of Educational Psychology*, *71*(Pt 1), 43–56. https://doi.org/10.1348/000709901158389

Mix, K. S., & Cheng, Y.-L. (2012). The relation between space and math: developmental and educational implications. *Advances in Child Development and Behavior*, *42*, 197–243.

Mix, K. S., Levine, S. C., Cheng, Y.-L., Young, C., Hambrick, D. Z., Ping, R., & Konstantopoulos, S. (2016). Separate but correlated: The latent structure of space and mathematics across development. *Journal of Experimental Psychology. General*, *145*(9), 1206–1227. https://doi.org/10.1037/xge0000182

Mix, K. S., Prather, R. W., Smith, L. B., & Stockton, J. D. (2014). Young Children's Interpretation of Multidigit Number Names: From Emerging Competence to Mastery. *Child Development*, *85*(3), 1306–1319. https://doi.org/10.1111/cdev.12197

Moll, K., Snowling, M. J., Göbel, S. M., & Hulme, C. (2015). Early language and executive skills predict variations in number and arithmetic skills in children at family-risk of dyslexia and typically developing controls. *Learning and Instruction, 38*, 53–62. https://doi.org/10.1016/j.learninstruc.2015.03.004

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

National Foundation for Educational Research (NFER). (2004). *NFER Progress in Maths 4-14 Test Series*. London: NFER-NELSON.

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents' mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences, 50*(8), 1238–1242. https://doi.org/10.1016/j.paid.2011.02.017

Newcombe, N. (in press). Three kinds of spatial cognition. In *J. Wixted (Ed.), Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* (4th edition).

Newcombe, N. S. (2013). Seeing Relationships: Using Spatial Thinking to Teach Science, Mathematics, and Social Studies. *American Educator, 37*(1), 26.

Oakes, J. (2005). *Keeping Track: How Schools Structure Inequality, Second Edition* (2nd Revised ed. edition). New Haven, Conn.; London: Yale University Press.

Office of Communication (OFcom). (2015). *Children and Parents: Media Use and Attitudes Report*. London: OFcom.

Okamoto, Y., Kotsopoulos, D., McGarvey, L., & Hallowell, D. (2015). The development of spatial reasoning in young children. In *Spatial Reasoning in the Early Years: Principles, Assertions, and Speculations* (pp. 15–29). New York: Routledge.

Organisation for Economic Cooperation and Development (OECD). (2013). *Pisa 2012 Results in Focus. What 15- year-olds know and what they can do with what they know*. Paris: OECD Publishing.

Palmer, S. . (1978). Fundamental aspects of cognitive representation. In *E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization* (pp. 259–303). Hillsdale, NJ: Erlbaum.

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research*, *37*(1), 239–253. https://doi.org/10.1016/j.ssresearch.2007.06.012

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, *93*, 604–614. https://doi.org/10.1037/0022-0663.93.3.604

Spence, I., & Feng, J. (2010). Video games and spatial cognition. *Review of General Psychology*, *14*(2), 92–104. https://doi.org/10.1037/a0019491

Talmy, L. (2000). How Language Structures Space. In *Towards a Cognitive Semantics* (Vol. 1, pp. 177–254). MIT Press.

Taylor, H. A., & Hutton, A. (2013). Think3d!: Training Spatial Thinking Fundamental to STEM Education. *Cognition and Instruction*, *31*(4), 434–455. https://doi.org/10.1080/07370008.2013.828727

Tosto, M. G., Hanscombe, K. B., Haworth, C. M. A., Davis, O. S. P., Petrill, S. A., Dale, P. S., Kovas, Y. (2014). Why do spatial abilities predict mathematical performance? *Developmental Science*, *17*(3), 462–470. https://doi.org/10.1111/desc.12138

Umiltà, C., Priftis, K., & Zorzi, M. (2009). The spatial representation of numbers: evidence from neglect and pseudoneglect. *Experimental Brain Research*, *192*(3), 561–569. https://doi.org/10.1007/s00221-008-1623-2

Uttal, D., & Cohen, C. (2012). Spatial Thinking and STEM Education. In *Psychology of Learning and Motivation* (Vol. 57, pp. 147–181). Elsevier. https://doi.org/10.1016/B978-0-12-394293-7.00004-2

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352–402. https://doi.org/10.1037/a0028446

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing Building Blocks: Preschoolers' Spatial Assembly Performance Relates to Early Mathematical Skills. *Child Development*, *85*(3), 1062–1076. https://doi.org/10.1111/cdev.12165

von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology*, *49*(11), 868–873. https://doi.org/10.1111/j.1469-8749.2007.00868.x

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817–835. https://doi.org/10.1037/a0016127

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Wolfgang, C., Stannard, L., & Jones, I. (2001). Block play performance among preschoolers as a predictor of later school achievement in mathematics. *Journal of Research in Childhood Education*, *15*(2), 173-181. https://doi.org/10.1080/02568540109594958

Zhang, X., Koponen, T., Räsänen, P., Aunola, K., Lerkkanen, M.-K., & Nurmi, J.-E. (2014). Linguistic and Spatial Skills Predict Early Arithmetic Development via Counting Sequence Knowledge. *Child Development*, *85*(3), 1091–1107. https://doi.org/10.1111/cdev.12173