BIROn - Birkbeck Institutional Research Online

Aged-based differences in spatial language skills from 6 to 10 years: Relations with spatial and mathematics skills

Authors

Katie A. Gilligan-Lee[ab*], Alex Hodgkiss[bd], Michael S. C. Thomas[bc], Pari K. Patel[a] & Emily K. Farran[ab]

[a] School of Psychology, University of Surrey, Guilford, GU2 7XH, UK

[b] Centre for Educational Neuroscience, University of London, London, UK

[c] Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London WC1E 7HX, UK

[d] Department of Education, University of Oxford, 15 Norham Gardens, Oxford, OX2 6PY

*Corresponding author: k.gilligan@surrey.ac.uk; +44 (0)1483 684888; address above

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Data Sharing

The data from this study is available from the corresponding author Katie Gilligan-Lee (k.gilligan@surrey.ac.uk) upon reasonable request.

Author Statement

- *Katie A. Gilligan-Lee:* Conceptualization, Methodology, Investigation, Writing - Original Draft, Formal Analysis

- *Alex Hodgkiss:* Conceptualization, Methodology, Investigation, Writing - Review & Editing

- *Michael S. C. Thomas:* Conceptualization, Methodology, Writing - Review & Editing, Supervision

- *Pari K. Patel:* Validation, Formal Analysis, Data Curation

- *Emily K. Farran:* Conceptualization, Methodology, Writing - Review & Editing, Supervision

Highlights

- Spatial language skills show progression between 6 and 10 years

- Spatial language comprehension is stronger than production in middle childhood

- Spatial language is associated with performance on some spatial measures

- Spatial language is a significant predictor of a range of mathematics skills

- These findings lay a foundation for future spatial language interventions

Abstract

Spatial language is the language of spatial concepts and spatial relationships. Prior research has demonstrated an association between spatial language and spatial thinking in pre-school children. However, there is limited evidence exploring age-based differences in spatial language in older childhood. This cross-sectional study has three main aims. First, we present a novel spatial language measure and show differences in spatial language performance across age groups from 6-10 years (N = 155). Second, having demonstrated that our measure is sensitive to age-based progression, we use regression analyses to determine relations between spatial language and performance on a range of spatial tasks (r2:1.2% -9.0%). Third, we investigate the relations between spatial language and different mathematics skills (r2:0.2% -15.4%) and propose mechanisms that may explain these associations. We discuss how these findings lay a foundation for future spatial language interventions as a novel tool which may lead to educational improvements in mathematics.

Aged-based differences in spatial language skills from 6 to 10 years: Relations with spatial and mathematics skills

1. Introduction

Spatial thinking involves perceiving the location and dimension of objects, and their relationships with other objects. There is convincing evidence that improving spatial thinking in children can positively influence mathematical performance (Cheng & Mix, 2012; Gilligan, Thomas & Farran, 2019; Hawes, Moss, Caswell, Naqvi & MacKinnon, 2017). This has led to a movement to encourage "spatialisation" of the primary school classroom, whereby spatial thinking is embedded into existing school curricula and lessons. It is proposed that this spatialisation will give children greater opportunities to engage in spatial thinking which may consequently improve both their spatial and mathematical outcomes (Newcombe, 2013).

Spatial language is the language of spatial concepts and spatial relationships. In this study spatial language specifically refers to relational terms, i.e., location and direction terms. Spatial language has been proposed as one tool for supporting and enhancing spatial thinking (Newcombe, 2010). However, while there is evidence that spatial language is positively associated with spatial ability during the pre-school years (Pruden, Levine & Huttenlocher, 2011), there are fewer studies showing associations between spatial language and either spatial thinking or mathematics in later childhood. For the first time, this study explores age-based differences in spatial language in children aged 6 to 10 years. We have used spatial relational terms to create a novel spatial language measure and provide evidence that spatial language skills are positively correlated with spatial and mathematical outcomes.

1.1 Spatial language and spatial performance

Most studies on the relations between spatial language and spatial thinking focus on pre-school populations (6 years and younger). There is evidence that knowledge of, or exposure to, specific spatial language terms is associated with higher performance on tasks that require these spatial concepts. For example, Miller, Patterson and Simmering, (2016) found that spontaneous verbal coding of spatial relations (*by, next to, between*) was associated with higher performance on a spatial reference frame selection task at 4 years. Similarly, knowledge of the spatial relation terms *middle* and *between* correlated with performance on a midpoint search task at 2.5-5.5 years (Simms & Gentner, 2019), and production of the spatial relation terms *left* and *right* was associated with higher performance on a spatial disorientation task requiring these concepts at 6 years (Hermer-Vazquez, Moffet & Munkholm, 2001). In an experimental study, Loewenstein and Gentner (2005) reported that exposing 3-year-old children to relevant spatial relational language cues (*on, in, under, at the top, at the bottom, in the middle*) before completing a spatial relations task improved their performance. Similar findings from Farran and O'Leary (2015) demonstrated higher feature matching task performance, when 4-year-olds were exposed to appropriate spatial relational language cues (*top, bottom, left, right*). In short, in early childhood familiarity with, and exposure to, spatial relational language terms is associated with improved spatial task performance, especially for tasks requiring these terms.

However, spatial language may have wider benefits to spatial performance beyond an understanding of specific spatial concepts. Szechter and Liben (2004) demonstrated that parental spatial-graphic behaviours (conveying information about spatial relationships) during picture book reading were associated with spatial ordering skills at 5 years. The spatial-graphic behaviours were not directly related to

the spatial ordering task, which suggests a broad relationship rather than a specific mapping between uses of individual terms. Beyond parental language, Pruden and colleagues (2011) found that children who spontaneously produced more spatial language regarding spatial properties and features during infancy (14-46 months) had significantly higher spatial task performance (spatial transformation, block design and spatial analogies tasks) at 4 years. Furthermore, intervention studies suggest that exposure to spatial language may have a causal effect on spatial skills. Casasola et al. (2020) reported that play-based spatial language training led to increased mental rotation scores at 4 years. Taken together, both receptive knowledge and productive use of spatial language is associated with spatial task performance in pre-schoolers. This is not solely attributable to an understanding of specific spatial relations and their verbal labels.

In older children, associations between spatial language ability and spatial skills are less well understood. Hawes et al. (2017) reported a significant association between spatial language (comprehension of spatial location terms, shape and figure names) and spatial performance (mental rotation and visuo-spatial processing in geometry tasks) at 6 years. Similarly, Farran and Atkinson (2016) reported significant associations between spatial relation language production (location and direction terms) and performance on a spatial categorisation task (assessing a range of spatial relationships) at 4 to 7 years. However, there were no significant associations between spatial language comprehension and spatial categorisation, perhaps emphasising the role of verbalisation during spatial task completion. In contrast, Mizzi et al. (2017) reported no significant difference in the spatial language discourse strategies used during a collaborative block construction task between 11-year-olds with high and

low spatial ability. The lack of consistency across these studies calls for further research.

Findings on the relations between spatial language and spatial performance are limited in two main ways. First, there is a lack of suitable tools to assess spatial language in older children (above 7 years). By extension, there is limited research on age-based differences in spatial language and the associations between spatial language and spatial thinking in older children. Second, previous studies investigating spatial language do not account for the multi-dimensionality of spatial ability in their research design (Farran & Atkinson, 2016; Hawes et al., 2017). Spatial thinking is not a unitary construct. The Uttal et al. (2013) classification of spatial thinking is adopted in the current study (see also Newcombe and Shipley, 2015). It is based on distinctions between *intrinsic* and *extrinsic*, and *static* and *dynamic* representations giving rise to four spatial sub-domains: intrinsic-static, intrinsic-dynamic, extrinsic-static and extrinsic-dynamic skills. *Intrinsic* representations pertain to the size and structure of objects, their parts and the relationship between them. *Extrinsic* representations pertain to object locations, and the location of objects relative to other objects, and their reference frames. The *intrinsic* vs. *extrinsic* distinction is supported by confirmatory factor-analysis (Mix et al., 2018), and neurological evidence that highlights processing differences between "what" (intrinsic) and "where" (extrinsic) information in the brain (Chatterjee, 2008). For the second distinction in the model, *dynamic* representations require movement or transformation e.g. scaling, rotating, while *static* representations do not. While there is less evidence supporting the *static* vs. *dynamic* distinction (Mix et al., 2018), this may be because static spatial skills are a necessary pre-requisite to dynamic ones.

As highlighted by Pruden et al. (2011), one explanation for associations between spatial language and spatial skills is that verbal coding of spatial relations may be an effective strategy for spatial task completion. A verbal coding strategy may be particularly useful when determining the relations between objects and between objects and an observer (i.e., extrinsic spatial tasks), or when completing manipulations such as rotations or transformations (dynamic spatial tasks). In this study, we compare the role of spatial language for different spatial sub-domains in primary school children, for the first time. We do this by measuring spatial thinking across each of Uttal et al.'s (2013) spatial categories using a carefully-selected task to target each sub-domain: intrinsic-static (disembedding task), intrinsic-dynamic (mental rotation task), extrinsic-static (spatial scaling task) and extrinsic-dynamic (perspective taking task).

1.2   Spatial language and mathematics performance

The relative lack of literature exploring the role of spatial language for mathematics is somewhat surprising given findings that both spatial skills and general language ability are correlated with mathematics outcomes. For spatial skills, both intrinsic and extrinsic spatial sub-domains are associated with mathematics outcomes in children (Geer, Quinn & Ganley, 2019; Gilligan, Flouri & Farran, 2017; Mix et al., 2016). Furthermore, intervention studies suggest that training spatial skills improves mathematics performance, i.e., a causal effect of spatial skills on mathematics (Cheng & Mix, 2014; Hawes et al., 2017; Gilligan, Thomas & Farran, 2019). For language ability, verbal and print-based language (literacy) are associated with both general mathematics outcomes and early numeracy skills. Expressive language predicts numeracy skills at 4-6 years (Purpura & Ganley, 2014), and vocabulary skills, print knowledge, and phonological awareness are correlated with numeracy outcomes at 3-

5 years (Purpura, Hume, Sims, & Lonigan, 2011). Language skills (receptive vocabulary and phonological awareness) at 5-6 years are also a significant predictor of mathematical outcomes 2 years later (LeFevre et al., 2010).

To explain these associations, there are several mechanisms through which language may facilitate mathematics learning. As outlined by Robertson and Graven (2020), student discourse is essential to mathematical learning, enabling students to explain, reason, argue and defend their mathematical thinking. The process of verbally justifying and defending a method may support future problem solving when similar reasoned decisions are required. This is supported by evidence that children who are trained to use talk (pupil-pupil talk) more effectively as a tool for reasoning show improved mathematical reasoning, understanding and problem-solving (Mercer & Sam, 2006). Furthermore, in other cognitive domains, it has been demonstrated that 'external' talk supports individual's 'inner' verbal problem-solving processes (Alderson-Day & Fernyhough, 2015). Second, language is used by teachers to describe new mathematical ideas and explain new concepts, Findings show that students who are taught and assessed in a non-native language have lower mathematical performance (Prediger et al. 2019).

Third, understanding mathematics specific-vocabulary is vital for mathematics learning as it gives students access to difficult concepts (Riccomini et al., 2015). There are several sub-components of mathematics vocabulary including spatial language, quantitative language (vocabulary that enables comparisons between groups or numbers, e.g., less than, many, fewer [Purpura et al., 2017]), and number and shape labels, among others. Both Purpura and Reid (2016), and Purpura and Logan (2015) reported that mathematical language, including both quantitative (e.g. more, less, fewer, some) and spatial language (e.g. below, under, middle), explained additional

variation in numeracy outcomes at 3-5 years, above general language skills. Toll and Van Luit (2014) also found that mathematics language (both spatial and quantitative) was associated with numeracy development between 4 and 6 years, after controlling for verbal and visuospatial working memory, and symbolic and non-symbolic number comparison. In older children of approximately 8 years, Ufer and Bochnik (2020) found that mathematical vocabulary and mathematical text comprehension both predicted arithmetic skills, even after controlling for prior arithmetic skills, general language skills, general cognitive skills and socio-economic status. However, none of these studies disentangle the differential relationships between spatial and quantitative language, and mathematics. Therefore, it is unknown whether the findings reported are driven by the role of quantitative or spatial language on mathematics outcomes.

To date, only one study has explored the role of spatial language specifically for mathematics outcomes in primary school children. Hawes et al. (2017) reported significant correlations between spatial language comprehension and numeracy outcomes (including number knowledge, and symbolic and non-symbolic number comparison) in 6-year-olds. However, as the spatial language measure included both shape labels and relational language, the findings cannot be used to determine the extent to which an understanding of spatial relations is uniquely associated with mathematics. Another limitation of research into spatial language is that previous studies typically explore the role of spatial language for numeracy skills only. This does not encompass the range of skills required in the mathematics classroom. In contrast, the current study uses von Aster and Shalev's (2007) model of numerical cognition. This model proposes that individuals are born with an innate system for representing number, an approximate number sense (ANS). Through development, a second cognitive number system is acquired for the representation of symbolic

numerals, the symbolic number system. Von Aster and Shalev (2007) propose that these two systems provide a platform for the development of complex mathematical skills such as multi-digit calculation, word problem solving, algebra, measurement and data handling skills. The current study extends previous findings by including a diverse set of mathematics tasks, assessing each component of von Aster and Shalev's (2007) model; the ANS (dot comparison task), symbolic number skills (number line estimation) and more complex mathematics skills (standardised classroom-based mathematics task).

1.3 Current Study

This study has three main research question. First, how does performance on spatial language comprehension and production differ between age-groups from 6 to 10 years? We will assess this using a novel spatial language measure that has been designed for this study. This measure distinguishes between comprehension and production skills. From a theoretical perspective, we propose that these skills are distinct and will be differentially associated with outcome measures. For example, spatial language production may be more related to verbal coding of spatial relations while comprehension may be more related to spatial concept development. This is supported by evidence from previous studies that have reported differential findings for spatial language comprehension and production (Farran & Atkinson, 2016). We predict age-based performance differences for both spatial language production and comprehension. We also anticipate higher comprehension compared to production scores at all ages.

Second, how does spatial language performance relate to performance on different spatial sub-domains, defined using the Uttal et al. (2013) classification of spatial skills? We predict that spatial language, specifically spatial language

11

production, will be associated with performance on extrinsic spatial sub-domains (both static and dynamic) where verbal coding of spatial relations may be used to determine the relations between objects or between objects and an observer. For intrinsic spatial sub-domains, we predict that spatial language production will be associated with tasks where manipulations, and by extension verbal coding of transformations, are required, i.e., dynamic tasks. To account for the possible influence of language ability (non-spatial language) on spatial thinking, receptive vocabulary performance is included as a control variable.

Third, how does spatial language performance relate to performance on different mathematical outcomes? As outlined for the spatial measures, we predict that spatial language, particularly spatial language production, will account for more variance for mathematics tasks where it is useful to strategically manipulate information in a mathematical context, i.e., tasks that may require verbal coding of spatial relations. Language ability, age and spatial skill scores are included as control variables.

2　Materials and Methods

2.1　Participants and Procedure

The data reported here is part of a larger study investigating the role of spatial thinking for mathematics and science learning (anonymised for review, 2018; 2018). Participants were 155 children aged 6 to 10 years, from a culturally diverse, London-based school (see Table 1). The overall ethnicity of the school population at the time of the study was 44% Asian, 29% White, 13% Black and 14% mixed/other. In this school, 19% of children were eligible for free school meals. This provides an indicator of socioeconomic disadvantage, as the national average of free school meals in the UK is 14% (Department for Education, 2017).

　　　Across five school-based sessions, participants completed spatial, mathematics, science and language measures. During Session 1, a one-hour classroom-based session, a standardised measure of mathematics, the National Foundation for Education Research (NFER) Progress in Mathematics (PiM)Test and (for children aged 8 years and older) the Number-Line Task, were completed. Session 2 (35-minutes) was completed in the school's computer room in groups of 8 children. Children completed mathematics tasks (the ANS Task, the Child Maths Anxiety Questionnaire [not discussed here] and the Number-Line Task [children aged 7 and younger]) followed by spatial measures (the Mental Rotation Task and a folding task [not discussed here]). For session 3 (45 minutes), participants were tested individually beginning with a spontaneous focus on number task (not discussed here), followed by spatial tasks (the Perspective Taking Task, the Children's Embedded Figures Test and the Scaling Task), the spatial language measure and the vocabulary measure (the BPVS). Two science assessments were also completed (not discussed here).

Table 1.

Demographic features of the study sample

Age group	Sample size	% Male	Age years (mean ± SD)
6 years	30	53	5.99 ± 0.34
7 years	31	42	6.98 ± 0.29
8 years	32	56	8.03 ± 0.28
9 years	31	45	8.99 ± 0.33
10 years	31	52	9.95 ± 0.33

2.2 Spatial Language Comprehension and Production Task

2.2.1 Task Design

Modelled on spatial language tasks for younger participants (Farran & Atkinson, 2016), the novel task used in this study had two blocks, a production block followed by a comprehension block. The fixed order of block presentation ensured that participants were not exposed to the correct spatial language terms prior to the production task. As this task required script reading, it was not possible to randomise experimental items. Instead, two task forms with different item orders were generated (Form A and Form B). Equal numbers of participants completed each form.

2.2.2 Selection of spatial language terms

For each block, participants completed two practice items using the spatial language terms *On* and *Under*. Findings from Farran and Atkinson (2016) show that by 6 years children can successfully complete both production and comprehension trials using these terms (Percentage accuracy- Comprehension *On*: 100%; *Under*: 100%; Production *On*: 100%; *Under*: 83% [Farran & Atkinson, 2016]). Feedback was provided for practice items. All participants achieved 100% on the practice items.

Each block included 12 experimental items. The items assessed the spatial relation terms *Above, Below, Right, Left, Between, Around, Through, Higher, Lower, Closer, Further, Parallel.* The initial selection of terms was based on Farran and Atkinson's (2016) findings for 6-year-olds. The terms *In, In front* and *Behind* were excluded, the terms *On* and *Under* were included as practice items, and the terms *Above, Below, Left and Right* were included as experimental items in the current study. Based on the high scores reported for the Farran and Atkinson (2016) spatial language task at 7 years, additional terms were added to increase task difficulty for older children. As all the Farran and Atkinson (2016) terms were relational, and for consistency with most previous studies of spatial language, additional terms were chosen from the same spatial language category within the Cannon, Levine and Huttenlocher (2007) Spatial Language Coding Guide, i.e., Category C- Location and Direction. The additional terms were: *Between, Around, Through, Higher, Lower, Closer, Further, Parallel.*

### 2.2.3 *Spatial Language Production*

For spatial language production items, participants were shown an image (using an iPad mini). Based on the image, participants were asked to verbally complete the missing word in a statement, read aloud by the experimenter (see Figure 1). All responses were manually recorded. Responses were later coded using an adapted version of the Spatial Language Coding Guide, a tool for coding spontaneous, naturalistic language, for spatial language content (Cannon, Levine & Huttenlocher, 2007). The Spatial Language Coding Guide divides spatial language terms into 8 spatial categories based on the 8 domains of spatial phenomena that are present in parent-child spoken interactions, e.g., spatial dimensions; locations and directions (Cannon, Levine & Huttenlocher, 2007). Each category is further divided into spatial

concepts, e.g., the category "Location and Direction" includes: along a vertical axis, along a horizontal axis, proximal/distal to another point. Based on these spatial categories and concepts, a decision-making tree for coding spatial language was designed for use in this study of non-spontaneous language production (see Figure 2). Responses were classified using one of six codes (Table 2). To determine spatial language production scores, only responses classified as Code 1- Expected Correct Response or Code 2[1]- Unexpected Correct Response, were accepted as correct answers (awarded a score of 1). All other responses (Code 3- Between Concept Error, Code 4- Within Concept Error, Code 5- Between Category Error, and Code 6-Non-Spatial Response) were coded as errors (awarded a score of 0). Participants were awarded an overall spatial language production score out of 12. The inter-rater reliability for 20% of responses was 100%.

---

[1] Only 4% of responses were categorised as Code 2.
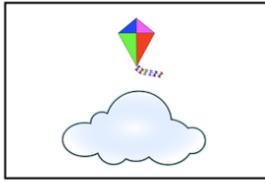
The kite is (above) the cloud.
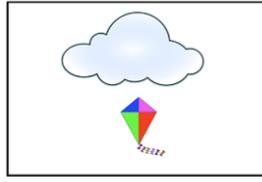
The kite is (below) the cloud.
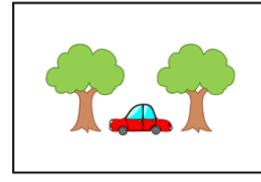
The car is (between) the trees

The ball is on the (left) side of the present
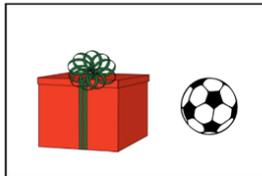
The ball is on the (right) side of the present.
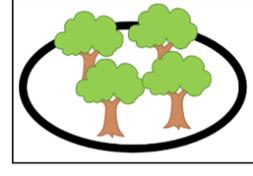
The black line is a road. The road goes (around) the trees.
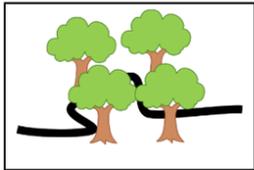
The black line is a road. The road goes (through) the trees

The present is on a (higher) shelf then the ball.

The present is on a (lower) shelf then the ball.
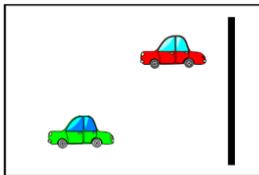
The black line shows the finish line of a race. The red car is (closer) to the finish line
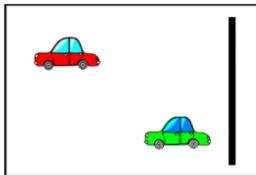
The black line shows the finish line of a race. The red car is (further) from the finish line
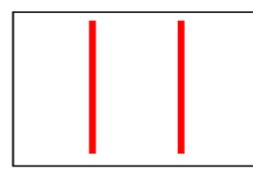
The ribbons are always the same distance apart. They are (parallel).

Figure 1. Spatial language production items
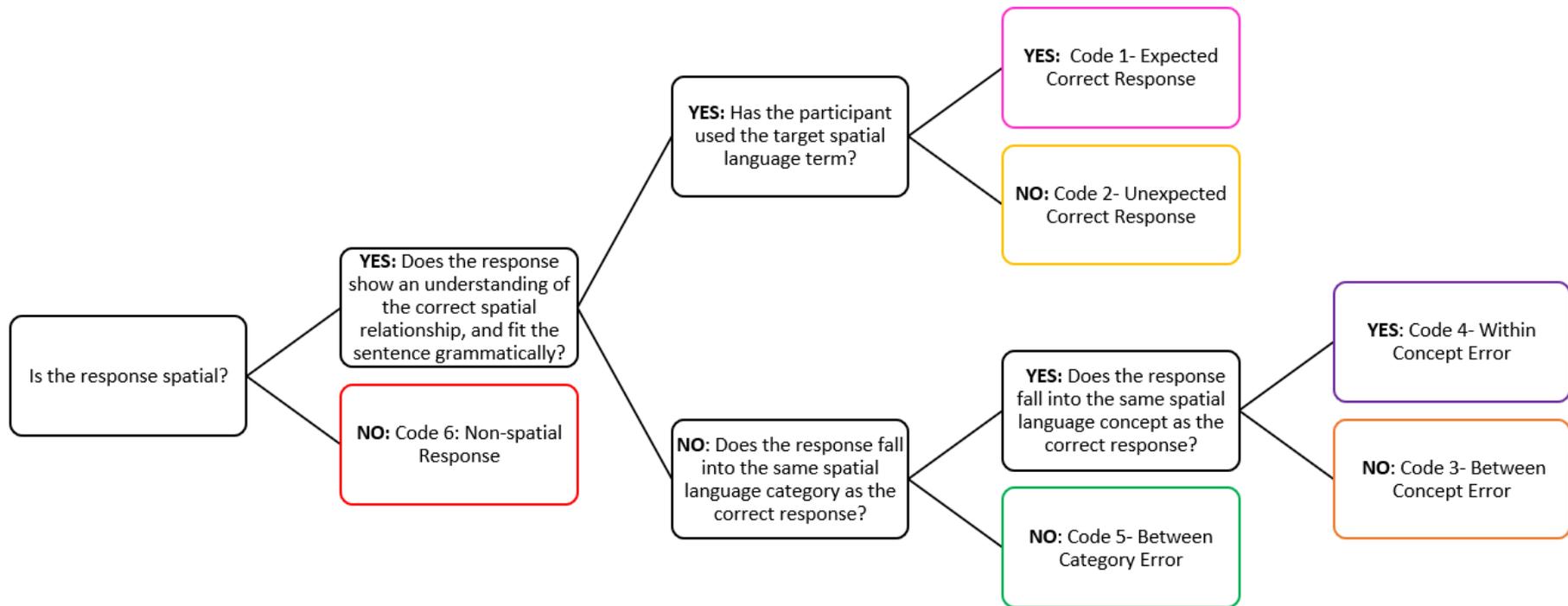
Figure 2. Decision-making tree for coding spatial language production items**.**

Table 2.

Table for coding spatial language production responses including sample responses

for each code. Note, except for Unexpected Correct Responses, the examples are not

an exhaustive list.

| Code	Response type	Explanation	Example	
			Target Word	**Response**
1	Expected Correct Response	The participant uses the target spatial language term	*On*	On
2	Unexpected Correct Response*	The participant doesn't use the target spatial language term but uses a term that shows an understanding of the correct spatial relationship. The word also fits the sentence grammatically	*Above*	Over
			Further	Far
			Below	Under, Beneath
			Between	In the middle of
			Closer	Nearest, Nearer
3	Between-concept error	The participant responds with an incorrect spatial language term that falls under a different spatial concept (but same spatial category) than the correct answer.	*Above*	In front
4	Within-concept error	The participant responds with an incorrect spatial language term that falls under the same concept as the target word.	*Above*	Below
			Left	Right
5	Between-category error	The participant responds with an incorrect spatial language term that falls into a different spatial category than the correct answer.	*Above*	Triangle
			In	Enough
6	Non-Spatial Response	The participant gives any other incorrect answer and the response is not spatial. This includes Homonyms, Metaphors, Spatially ambiguous answer, Nominatives and Ambiguous/vague answers.	*Left*	There is no chocolate left
			Around	I will be home around five pm
			Closer	He is closer to me in age

*To determine how to allocate a code 2 (Unexpected Correct Response), twenty adults

(Mean Age: 32.8 years; Males 30%) were presented with the stimuli used in the

production block of the task (Figure 1). They were asked to choose which terms, from

a list generated from our child responses, correctly completed the statement. Only terms

selected by ≥ 80% of adults were awarded a code 2.

2.2.4   *Spatial Language Comprehension*

Participants completed 12 items using the same terms as the spatial language

production block. For each item, the experimenter read a spatial language statement

aloud and asked the participant to select which of four onscreen images best matched

the statement (see Figure 3). The images were presented in a two-by-two array using a

13-inch Hewlett Packard touchscreen laptop. The position of the correct image was

counterbalanced. In line with Phillips et al. (2004) distractor images contained the

same object(s) as the target image. For each item, one distractor was designed to be

the opposite of the target image. The two additional distractor images displayed

neither the correct answer nor the inverse answer to the correct answer, i.e., they

could not be described by the target statement or the opposite statement to the target

statement. Participants were awarded a score of 1 for every correct answer.

Performance accuracy was recorded.

The ball is on the <u>left</u> side of the present   **The present is on a <u>lower</u> shelf than the ball**
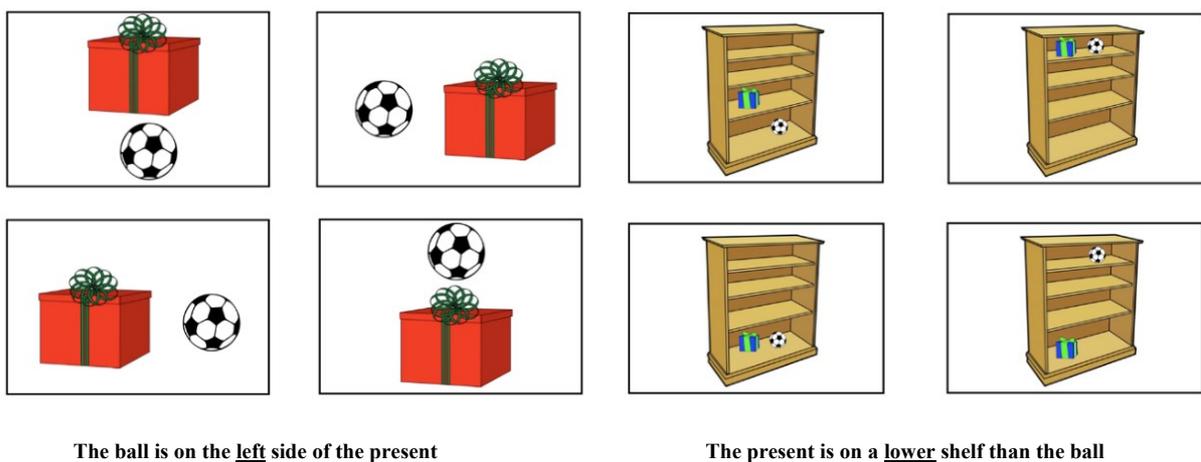
Figure 3: Sample spatial language comprehension items

2.3   Spatial Measures

Participants completed a spatial task assessing each of Uttal et al.'s (2013) spatial sub-domains. For more detailed information on the tasks included in this study see anonymised for review (2018).

2.3.1   *Disembedding (intrinsic-static skills)*

The Children's Embedded Figures Task (CEFT), a disembedding task, was included to measure *intrinsic-static* spatial ability (Witkin, Otman, Raskin, & Karp, 1971). The task was presented as two blocks in a fixed order. For each block, participants first completed 4 discrimination trials where they were required to identify a target shape (house or tent shape) from a selection of other similar shapes. Following this, participants completed practice trial(s) (Block A: 2 practice items; Block B: 1 practice item) in which they were required to locate the target shape within a more complex picture. Feedback was given. This was followed by experimental trials where participants were again required to locate the target shape within more complex pictures (11 items for Block A; 14 items for Block B). No feedback was given. Percentage accuracy was recorded.

2.3.2   *Mental Rotation (intrinsic-dynamic skills)*

A Mental Rotation Task was included as a measure of *intrinsic-dynamic* spatial ability. Participants were asked to identify which of two monkey images located above a horizontal line matched a third target image below the line (Broadbent, Farran, & Tolmie, 2014). The choice images included a mirror image of the target, and a version of the target image rotated by a fixed degree. Participants completed four practice trials at 0° followed by 36 experimental trials. Equal numbers of trials

were included at 0° 45°, 90°, 135° and 180°. Participants responded using a keyboard response. Percentage accuracy was recorded.

2.3.3 *Spatial Scaling (extrinsic-static skills)*

A Spatial Scaling Task was used to assess *extrinsic-static* skills (Gilligan, Hodgkiss, Thomas, Tolmie, & Farran, 2018). Participants used a model map with a target, to identify a corresponding referent map from four options. The task included 2 practice trials (scaling factor of 1) followed by three blocks of 6 experimental trials. Scaling factor varied across blocks and was set at 1, 0.5 and 0.25. Feedback was given for practice trials only. Participants responded manually using a touchscreen laptop. Percentage accuracy was recorded.

2.3.4 *Perspective Taking (extrinsic-dynamic skills)*

A Perspective Taking Task was included as a measure of *extrinsic-dynamic* spatial thinking (Frick, Mohring, & Newcombe, 2014). Participants were asked to identify which of four photographs had been taken from the perspective of a toy photographer. Participants completed four practice trials with real, 3-D objects and Playmobil characters holding cameras. Feedback was given. The task included 18 computer-based experimental trials. Participants completed equal numbers of trials in which they were positioned at 0°, 90° and 180° from the photographer respectively. Complexity was also added by increasing the number of objects in the stimulus picture (one, two or four objects). No feedback was given. Percentage accuracy was recorded.

2.4   Mathematics Measures

2.4.1   *Mathematics Achievement- NFER Progress in Mathematics Test Series*

The National Foundation for Education Research (NFER), Progress in Mathematics (PiM) test series were administered as a standardised measure of mathematics achievement (NFER, 2004). This is a series of classroom measures of mathematics that has been designed to address the National Curriculum in the UK. It assesses: number; algebra; shape, space and measures; and data handling. In this study, the children in each age group were all in the same grade at school, and all completed the same test, i.e., all children in the subgroup 8-year-olds were all in Year 3, and all completed the same Year 3 test. As such, specific tests were administered to each age group of participants, as per the test guidelines (NFER, 2004). Each specific test included content appropriate for that school grade group. Standardised scores (mean of 100; standard deviation of 15) were calculated based on national calibration samples for this test. These national standardisations were at grade level, i.e., a child's results were standardised against the national standardisation scores for other children in their grade. Standardised scores were used in all analyses.

2.4.2   *Dot Comparison (ANS task)*

A dot comparison task was used to assess approximate number sense (Gilmore, Attridge, De Smedt, and Inglis, 2014). In each of 64 experimental trials, participants compared two dot arrays and selected (using a key board response) the more numerous array. The quantity of dots in each comparison array ranged from 5 to 22 and the ratio between the dots varied between 0.5, 0.6, 0.7 and 0.8. The color of the more numerous array (red or blue) in addition to the size and the density of dot presentation were counterbalanced. Performance accuracy was recorded.

2.4.3   Symbolic Number Skills (Number Line Estimation)

A Number-Line Estimation Task was administered to assess symbolic number skills (adapted from Siegler and Opfer, 2003). Participants completed trials using number lines ranging from 0-10 and 0-100. Two trial types were included, number estimation (NP) and position estimation (PN) trials. For NP trials, participants were asked to position a target number on a number line by drawing a straight line at their selected location. For PN trials participants were shown a vertical hatch mark on a number line and were asked to estimate what number was represented by the mark. Within each block participants completed two practice trials (one NP and one PN) followed by eight experimental trials. Performance was measured as Percentage Absolute Error (PAE) and linear response patterns (R^2_{LIN}). PAE is the distance from a participant's answer to the correct answer, divided by the length of the line. R^2_{LIN} is the correlation between participants' estimates and the target numbers.

2.5   Vocabulary Measure

The British Picture Vocabulary Scale (BPVS) (III), was included to assess receptive vocabulary (Dunn, Dunn, Styles, & Sewell, 2009). This standardised measure is equivalent to the Peabody Picture Vocabulary Test in the United States (Dunn &Dunn, 1997). It was delivered as per the administration guidelines. Participants were required to select which of four coloured pictures best illustrated the meaning of a given word.

2.6   Analysis Strategy

Parametric analyses were used as all variables were broadly normal and the sample size was greater than 30 in each age group, i.e., sufficiently large for the central limit

theorem to apply (Field, 2013). There were few outliers (scoring ± 2sd from the mean): NFER PiM test (two participants), BPVS (one participant), 0-10 Number Line Estimation block (three participants) and 0-100 Number Line Estimation block (2 participants). All outliers were retained as they were deemed to reflect normal variation in the population.

Participants who were missing data for the Spatial Language Task (5 participants), were excluded from task-based analysis. However, to optimise statistical power for between task analyses, missing scores for all tasks were replaced by mean scores on that task for a participant's age group. The data replaced amounted to less than 1% (0.82%) of all data[2]. An exception to this was the Number Line Task where missing data was not replaced. Where a participant's mean percentage absolute error (PAE) scores for the practice trials in a block were greater than 15%, or where participants failed to answer 80% of items in a block, they were excluded from analysis for this block. The sample sizes for the 0-10 and 0-100 blocks of the Number Line Task were 131 and 136 respectively.

For regressions, all predictors were converted to z-scores. All Tolerance and VIF scores were within acceptable levels. Where significant interactions were reported, follow up analyses were completed using younger (6 and 7 years) and older (8, 9 and 10 years) participants respectively. These age groups reflect the organisation of the UK Primary Education system which is divided into two blocks: Key Stage 1 (age 5-7 years) and Key Stage 2 (age 8-11 years).

Post-hoc sensitivity analysis revealed that the t-test (d=0.4), and main regression analyses were powered to detect small effect sizes(f=0.06). However, the ANOVA

---

[2] Missing data: Spatial Language Task (N=5); CEFT (N=1); Perspective Taking (N=2); NFER PiM (N=2); ANS (N=2); BPVS (N=2)

analysis (f=0.24), and the regression analyses that were completed with the older (f=0.11) and younger (f=0.17) age groups separately, were only powered to detect medium sized effects. The age-based analyses should therefore be interpreted cautiously.

3   Results

3.1   The Spatial Language Production and Comprehension Task

T-tests indicated no significant effect of item order on spatial language comprehension, t(148)=1.18, p=.238, d=.21, or production, t(148)=0.50, p=.617, d=.08. Hence, task form was not considered in subsequent analysis. Spatial language production and comprehension were treated as separate constructs for all analysis. Beyond the theoretical rationale for this (outlined in the introduction), a paired samples t-test found a significant difference in performance between spatial language blocks, t(149)=26.12, p<.001, d=1.94, and the correlation reported between spatial language production and comprehension, r(150)=.60, p<.001, suggests that there may be differences in performance on the two measures i.e., they may assess different abilities. The reliability of each block was measured using Cronbach's α. Medium to high values of Cronbach's α (comprehension=.67; production=.72) were reported (Field, 2013). The omission of no single item caused a significant change in the reliability scores.

A mixed ANOVA was completed with gender (between participant factor) and spatial language block (within participant factor). There was no main effect of gender, F(1,148)=2.13, p=.146, η_p^2=.014, or interaction between spatial language block and gender, F(1,148)=0.05, p=.817, η_p^2=.001. Gender was therefore not included in

subsequent analysis. No significant gender differences were reported for any other measures (see anonymised for review 2018).

3.2   Age-based differences in spatial language performance

One-sample t-tests against chance (25%) demonstrated that all age groups performed significantly above chance ($p>.05$ for all) on the spatial language comprehension block. It was not possible to calculate above chance performance for the open-ended responses on the spatial language production task. A mixed ANOVA with a between participant factor of age group (6, 7, 8, 9, 10 years) and within participant factor of spatial language block (comprehension, production) indicated a main effect of block. Language comprehension scores ($M= 83.94$, $SD= 16.86$) were significantly higher than production scores ($M= 46.44$, $SD= 21.27$), $F(1,145)=739.78$, $p<.001$, $\eta_p^2=.84$. There was also a main effect of age group, $F(4, 145)=17.66$, $p<.001$, $\eta_p^2=.33$, which interacted with spatial language block, $F(4,145)=4.18$, $p=.003$, $\eta_p^2=.10$. As such, the main effect of age is best explained within the context of the interaction.

As shown in Figure 4, two follow up one-way ANOVAs both demonstrated a significant effect of age group ($p's<.001$, $\eta_p^2's <.23$). For comprehension, Tukey post-hoc tests revealed significantly lower performance at 6 compared to 8, 9 and 10 years, and 7 compared to 9 and 10 years. For production, there was significantly lower performance at 6,7 and 8 years compared to at 9 and 10 years. These post-hoc comparisons reveal subtle differences in the age-based performance patterns for the spatial language production and comprehension tasks.
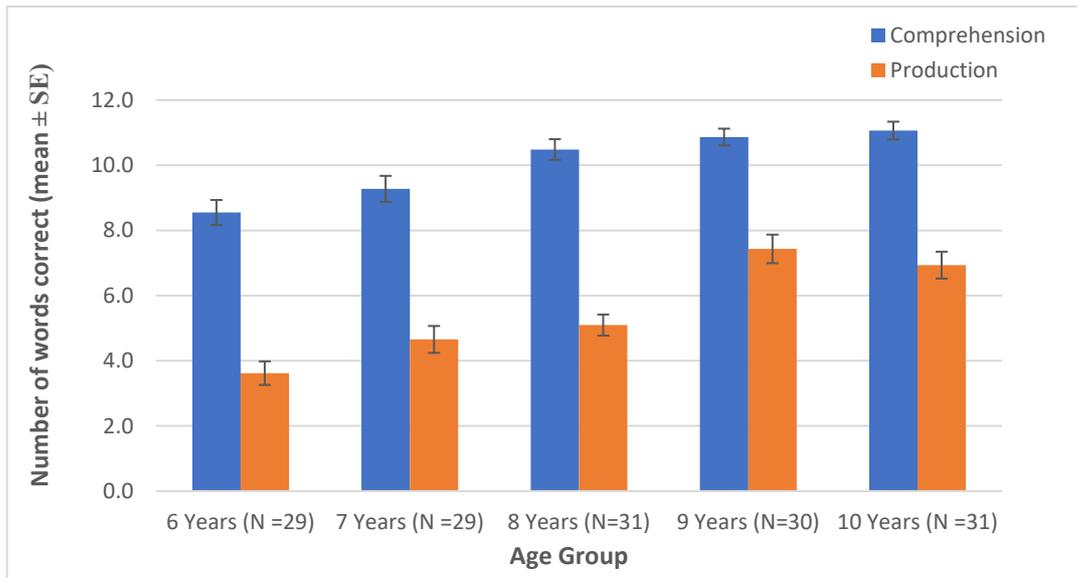
Figure 4. Spatial language performance across age groups[3]

3.3   Associations between spatial language and other measures

Pearson's correlations indicated large, significant associations between both spatial language measures, receptive vocabulary and age (Table 3). Hence, receptive vocabulary skills and age were included as covariates in subsequent correlations. Partial correlations corrected for multiple comparisons (alpha level: 0.05/8=.006) found significant associations between spatial language production and standardised mathematics performance, and between spatial language comprehension and 0-100 number line estimation. The spatial language measures were differentially associated with different spatial skills. Spatial language comprehension was significantly associated with both spatial scaling and mental rotation. Spatial language production was associated with spatial scaling only.

---

[3] Additional information on performance on each spatial language term, by age-group can be found in the supplementary material.

Table 3.

Correlations between spatial language comprehension and production and other

measures, controlling for BPVS and age.

Measure	Spatial Language comprehension	Spatial Language Production
BPVS[1]	**.60 (p < .001)**	**.66 (p < .001)**
Age[1]	**.50 (p < .001)**	**.53 (p < .001)**
Disembedding	.12 (p =.141)	.15 (p =.060)
Mental Rotation	**.23 (p =.004)**	.11 (p =.187)
Spatial Scaling	**.26 (p < .001)**	**.37 (p < .001)**
Perspective Taking	.08 (p =.316)	.21 (p =.009)
Standardised Maths Performance	.20 (p =.016)	**.47 (p < .001)**
ANS Skills	.15 (p =.057)	.20 (p =.015)
0-10 Number Line Estimation R^2_{LIN} (N =131)[2]	.09 (p =.317)	.19 (p =.030)
0-100 Number Line Estimation R^2_{LIN} (N =136)[2]	**.33 (p < .001)**	.20 (p =.022)

[1] This correlation does not control for the BPVS or age. [2] These results are based on R^2_{LIN} values. No significant associations were reported for PAE (see supplementary material).

3.4   The relationship between Spatial Language and Spatial Skills

Hierarchical regressions controlling for receptive vocabulary ability (BPVS scores) and age (months), determined the proportion of variation in each spatial measure explained by spatial language. The control variables (age and BPVS scores) were added in step 1, spatial language comprehension and production scores were added in step 2, and interaction terms between spatial language comprehension and production, and age were added using stepwise entry in step 3. Only significant interactions were

retained. For disembedding, spatial language explained an additional 1.2% of the variation in performance. Age was the only significant predictor in the final model (see Model 1, Table 4). For mental rotation, the spatial language measures explained 2.7% of the variation. The interaction term between age and spatial language comprehension explained an additional 3.1% of the variation. The interaction term, spatial language comprehension and age were significant predictors in the final model (see Model 2, Table 4). To explore this interaction, the analysis was repeated with younger (6 and 7 years) and older (8, 9 and 10 years) groups respectively. Age was the strongest predictor for both groups (Table 4). However, the Beta values suggest that while the role of spatial language production was similar for both age groups, spatial language comprehension was a stronger predictor for the younger (B=0.22) compared to the older group (B=0.12).

Spatial language explained 9.0% of the variation in spatial scaling. In addition to age, both spatial language production and spatial language comprehension were significant predictors in the final model (see Model 3, Table 4). Finally, for perspective taking, spatial language explained 2.3% of the variation. A significant interaction was reported between age and spatial language comprehension, explaining an additional 2.4% of the variation. Spatial language production, spatial language comprehension and the interaction term were significant predictors in the final model (see Model 4, Table 4). As before, to explore this interaction, the analysis was repeated with younger and older groups. Spatial language production was a significant predictor for the older group only (Table 4).

Table 4.

Regression models outlining the variance in spatial skills accounted for by spatial language

	b	SE	ß	t	p	F	df	p	Adj. R^2	Δ Adj.R^2
Model 1: Disembedding (intrinsic-static)										
Age	0.46	0.09	0.46	5.00	**< .001**	30.54	152	**< .001**	.277	
BPVS	-0.07	0.11	-0.07	-0.68	.495					
Spatial Language Comprehension	0.09	0.09	0.09	0.96	.339	16.65	150	**< .001**	.289	.012
Spatial Language Production	0.15	0.10	0.15	1.52	.130					
Model 2: Mental Rotation (intrinsic- dynamic)										
Age	1.23	0.37	1.23	3.35	**< .001**	28.19	152	**<.001**	.261	
BPVS	0.15	0.11	0.15	1.40	.165					
Spatial Language Comprehension	1.24	0.38	1.24	3.22	**< .001**	16.55	150	**<.001**	.288	.027
Spatial Language Production	0.09	0.10	0.09	0.98	.330					
Comprehension * Age	-1.78	0.64	-1.78	-2.80	**.006**	15.42	149	**<.001**	.319	.031
Model 2: Follow up Younger Age Group										
Age	0.55	0.27	0.24	2.00	.051	9.09	58	**<.001**	.212	
BPVS	0.09	0.17	0.08	0.56	.578					
Spatial Language Comprehension	0.22	0.14	0.25	1.55	.127	6.83	56	**<.001**	.280	.068
Spatial Language Production	0.18	0.17	0.16	1.02	.313					
Model 2: Follow up Older Age Group										
Age	0.27	0.11	0.31	2.53	**.013**	26.98	91	**<.001**	.363	
BPVS	0.12	0.14	0.13	0.86	.391					
Spatial Language Comprehension	0.12	0.12	0.11	0.97	.337	14.81	89	**<.001**	.378	.015
Spatial Language Production	0.19	0.12	0.20	1.56	.122					

Model 3: Spatial Scaling (extrinsic-static)										
Age	0.25	0.08	0.25	3.19	**< .001**	50.24	152	**<.001**	.390	
BPVS	0.08	0.09	0.08	0.90	.368					
Spatial Language Comprehension	0.16	0.08	0.16	2.11	**.036**	36.53	150	**<.001**	.480	.090
Spatial Language Production	0.34	0.08	0.34	4.02	**< .001**					
Model 4: Perspective taking (extrinsic-dynamic)										
Age	0.03	0.11	0.03	0.28	.784	30.50	152	**<.001**	.277	
BPVS	-0.54	0.37	-0.54	-1.48	.141					
Spatial Language Comprehension	-0.91	0.38	-0.91	-2.39	**.018**	17.50	150	**<.001**	.300	.023
Spatial Language Production	0.25	0.10	0.25	2.60	**.010**					
Comprehension*Age	1.59	0.63	1.59	2.51	**.013**	15.75	149	**<.001**	.324	.024
Model 4: Follow up Younger Age Group										
Age	0.26	0.26	0.14	0.99	.329	0.77	58	.469	.008	
BPVS	-0.07	0.16	-0.08	-0.46	.647					
Spatial Language Comprehension	-0.04	0.14	-0.06	-0.30	.765	0.61	56	.655	.026	.018
Spatial Language Production	0.16	0.17	0.17	0.94	.352					
Model 4: Follow up Older Age Group										
Age	0.33	0.17	0.20	1.96	.053	9.10	91	**<.001**	.148	
BPVS	0.06	0.15	0.05	0.40	.688					
Spatial Language Comprehension	0.18	0.13	0.14	1.36	.176	6.82	89	**<.001**	.200	.052
Spatial Language Production	0.28	0.13	0.27	2.24	**.028**					

3.5 The relationship between Spatial Language and Mathematics

3.5.1 Models excluding spatial skills

Hierarchical regressions determined the variance in mathematics outcomes explained by spatial language controlling for age and BPVS scores. The process of entering variables into each model was identical to that described in section 3.4. However, for standardised mathematics performance, school grade was also entered as a control variable to account for any differences in grade level sample selection. For standardised mathematics performance, the spatial language measures accounted for an additional 15.4% of the variation in performance. Spatial language production, grade and vocabulary skills were significant predictors in the final model (see Model 5, Table 5). The spatial language variables accounted for an additional 2.4% of the variation in ANS performance with age and spatial language production as significant predictors in the final model (see Model 6, Table 5). For the 0-10 number line estimation task spatial language explained an additional 3.2% of the variation. Spatial language production (*B*=0.24) was the strongest predictor in the final model (see Model 7, Table 5). Finally, for the 0-100 number line estimation task, spatial language explained 7.4% of the variation. Spatial language comprehension and age were significant in the final model (see Model 8, Table 5).

3.5.2 Models including spatial skills

To investigate whether there was shared variation between spatial skills and spatial language in explaining mathematics outcomes, the hierarchical regressions described in 3.5.1 were repeated with age, BPVS scores, and all the spatial skill scores as control variables. For standardised mathematics, the spatial language measures accounted for 8% of the variation. Spatial language production was a significant

predictor (see Model 9, Table 6). After controlling for spatial skills, spatial language

measures did not account for any additional variation in ANS and 0-10 number line

performance (see Model 10&11, Table 6). Spatial language explained 2.9% of the

variation in 0-100 number line estimation performance. Spatial language

comprehension was a significant predictor (see Model 12, Table 6).

Table 5.

Regression models outlining the variance in mathematics accounted for by spatial language controlling for age, BPVS and grade (for standardised maths performance only)

	b	SE	β	t	p	F	df	p	Adj. R²	Δ Adj.R²
Model 5: Standardised Maths Performance										
Age[1]	0.42	0.28	0.42	1.50	.137	23.36	151	**< .001**	.303	
Grade[2]	-0.51	0.20	-0.71	-2.56	.011					
BPVS	0.32	0.10	0.32	3.35	<.001					
Spatial Language Comprehension	0.03	0.08	0.03	0.38	.707	28.89	149	**< .001**	.457	.154
Spatial Language Production	0.54	0.09	0.54	6.24	<.001					
Model 6: ANS Skills										
Age	0.00	0.10	0.00	0.04	.969	37.16	152	**< .001**	.320	
BPVS	0.40	0.09	0.40	4.56	**< .001**					
Spatial Language Comprehension	0.11	0.09	0.11	1.25	.214	21.19	150	**< .001**	.344	.024
Spatial Language Production	0.19	0.10	0.19	1.98	**.050**					
Model 7: 0-10 Number Line Estimation[3]										
Age	0.08	0.13	0.08	0.64	.525	10.35	128	**< .001**	.139	
BPVS	0.12	0.11	0.12	1.11	.270					
Spatial Language Comprehension	0.04	0.11	0.04	0.38	.707	6.52	126	**< .001**	.171	.032
Spatial Language Production	0.23	0.12	0.24	1.97	.051					
Model 8: 0-100 Number Line Estimation[3]										
Age	0.31	0.09	0.29	3.26	**< .001**	29.14	133	**< .001**	.294	
BPVS	0.05	0.11	0.05	0.48	.633					
Spatial Language Comprehension	0.31	0.09	0.30	3.41	**< .001**	20.61	131	**< .001**	.368	.074
Spatial Language Production	0.13	0.10	0.12	1.28	.204					

Table 6.

Regression models outlining the variance in mathematics outcomes accounted for by spatial language controlling for age, BPVS, all spatial performance scores and grade (for standardised maths performance only),

	b	SE	ß	t	p	F	df	p	Adj. R²	Δ Adj.R²
Model 9: Standardised Maths Performance										
Control Measures						17.28	147	**< .001**	.425	
Spatial Language Comprehension	-0.02	0.08	-0.02	-.31	.759					
Spatial Language Production	0.44	0.09	0.44	5.01	**< .001**	18.49	145	**< .001**	.505	.080
Model 10: ANS Skills										
Control Measures						18.38	148	**< .001**	.404	
Spatial Language Comprehension	0.04	0.09	0.04	0.48	.632					
Spatial Language Production	0.06	0.10	0.06	0.59	.553	13.76	146	**< .001**	.399	-.005
Model 11: 0-10 Number Line Estimation										
Control Measures						6.78	124	**< .001**	.211	
Spatial Language Comprehension	-0.01	0.11	-0.01	-0.06	.956					
Spatial Language Production	0.16	0.12	0.16	1.30	.195	5.30	122	**< .001**	.209	-.002
Model 12: 0-100 Number Line Estimation										
Control Measures						14.31	129	**< .001**	.372	
Spatial Language Comprehension	0.25	0.09	0.24	2.71	**.008**					
Spatial Language Production	0.04	0.10	0.04	0.44	.662	12.32	127	**< .001**	.401	.029

4 Discussion

This study presents findings on age-based differences in spatial language comprehension and production in children aged 6-10 years, assessed through a novel/newly developed spatial language task. In line with studies of younger children (Farran & Atkinson, 2016), we found that spatial language comprehension skills are stronger than spatial language production skills at all ages, from 6 to 10 years. Although there were slight differences in age to age progression on spatial language production compared to comprehension tasks, for both skills there were larger differences in accuracy between the younger age groups than the older age groups. These findings provide the first benchmarks of spatial language performance in children older than 7 years and highlight age-based differences in spatial language skills in middle childhood. Although it was not a primary aim, no gender differences in spatial language were found. This is consistent with some (Pruden et al., 2011) but not all (Pruden & Levine, 2017) studies in younger children (2-4 years). A possible explanation for differences reported is that previous studies measured *spatial properties and features* terms, while the current study tested *spatial relational* terms.

Addressing the second research question, the associations between spatial language and spatial skills varied by spatial sub-domains. Some spatial sub-domains had no significant associations with spatial language (disembedding), while others had associations with both spatial language production and comprehension (spatial scaling and perspective taking). Although a causal direction cannot be determined through this correlational work, one theoretical explanation for these differential findings is that spatial language is more important for spatial tasks in which verbal coding of spatial relations is useful (Pruden, et al. 2011). Elaborating on this, we argue, that spatial language *production* specifically is important for tasks requiring verbal coding of spatial relations. In line with this hypothesis,

the current study reported that spatial language production accounted for more variance for extrinsic spatial tasks. In these tasks, verbal coding of spatial relations is hypothesised to be a useful strategy as participants must elucidate the relationship between objects or between an observer and a scene, e.g., spatial scaling: *the target is in the middle, the target is at the top*; perspective taking: *the man is beside the cube, the pyramid is on the left, the block is at the back*. In contrast, the associations reported between intrinsic spatial tasks and spatial language production were weaker.

Our findings also highlighted associations between spatial language comprehension and performance on several spatial tasks (mental rotation, scaling and perspective taking). We propose that these relations may be underpinned by individual differences in spatial concept formation (Mandler, 2012) that influence both spatial task performance and spatial language comprehension skills. This is elaborated on in later paragraphs, with the caveat that causal explanations must be treated cautiously due to the associational design of this study.

Our findings also show age-specific relations between spatial language and some spatial skills. Spatial language comprehension explained a greater proportion of the variation in mental rotation in younger children. In contrast, for perspective taking, the variation explained by spatial language production increased with age. Taken together, these findings may reflect strategy differences for spatial task completion in younger (more reliant on basic strategies/spatial concept formation) and older children (more reliant on verbal coding of spatial relations). It may be the case that as children's spatial language production skills improve, they rely on them more heavily in spatial task completion. If this were the case, it would suggest that spatial language interventions targeting spatial language production at younger ages may confer benefits to a large range of extrinsic spatial tasks. Previous studies suggest that spatial language training may improve spatial task performance in pre-schoolers (Casasola et al., 2020). The associational evidence presented here supports the design of

future studies investigating spatial language production interventions in middle childhood, particularly those investigating transfer of gains to spatial tasks where spatial language strategies are beneficial.

Addressing the third research question, we are the first to demonstrate positive relations between spatial language and general mathematics ability in children aged 6 to 10 years. For most mathematics outcomes, spatial language skills continued to explain unique variation in mathematics performance even after controlling for spatial thinking skills. Although for some of the mathematics tasks the effects are moderate to small, the results highlight spatial language as an important contributor to the variance in standardised mathematics performance in 6-10-year-olds. In particular, spatial language production accounted for a significant proportion of the variance in classroom-based mathematics outcomes. By comparing models that include and exclude spatial skills as control variables, we have shown that: 1) spatial language explains unique variation in mathematics outcomes that cannot be attributed to spatial thinking; 2) for all mathematics outcomes, there is shared variation between spatial language and spatial skills. This may reflect an indirect association between spatial language and mathematics that is mediated by spatial skill. Although a causal direction cannot be determined form our findings, from a theoretical perspective, we propose several mechanistic explanations that may explain the reported associations between spatial language and mathematics.

First, as described for spatial tasks, spatial language production may offer an effective strategy for completing mathematics tasks in which verbal coding of spatial relations is particularly useful, i.e. to strategically manipulate information in a mathematical context. This is supported by the aforementioned evidence that improving mathematical discourse improves student outcomes (Robertson & Graven, 2020). For example, when working with fractions such as ¾, children may benefit from verbally labelling numbers "3 is above the

line, 4 is below the line". Similarly, for place value questions, verbal cues may be used to describe numbers on the left and right of the decimal place, "whole numbers are on the left of the decimal place, all numbers on the right of the decimal place are fractions". For number line estimation tasks, children may use verbal cues to position items, "50 is in the middle, 75 is between 50 and 100, 5 is closer to 0". In contrast, for ANS tasks where participants are asked which is the more numerous of two dot arrays, language cues are not expected to aid performance. The arrays are only briefly displayed, and the more numerous array has no defining features that can be labelled verbally, e.g. colour, size, density. The use of verbal coding of spatial relations in both spatial skill and spatial language tasks may explain the shared variation reported between these skills in predicting mathematics outcomes.

Second, associations may be present between spatial language and mathematics due to shared requirements for grounding symbolic and conceptual representations. Theories of grounded cognition posit that there is a single representational system (schema) for every concept, i.e., perceptual, conceptual and symbolic representations are grounded onto a single schema (Barsalou, 1999; Matheson & Barsalou, 2018). Understanding symbolic number requires grounding of number symbols (arabic symbols and number words), with a conceptual understanding of quantity, e.g. understanding that 5 and five represent the quantity five (De Smedt et al., 2013; Hurst, Anderson & Cordes, 2016). Symbolic number skills are integral to mathematical development and early symbolic number skills predict later mathematics achievement (Friso-van den Bos et al., 2015). Similarly, spatial language tasks require grounding of verbal labels with conceptual representations of spatial concepts, e.g. understanding that the word "above" represents the spatial concept above (Munnich, & Landau, 2003). It is possible, therefore, that some individuals are better at grounding symbolic and conceptual representations across subject domains, i.e., children who have better grounding of symbolic (linguistic) and conceptual representations of spatial concepts,

may also have better grounding of symbolic and conceptual representations of quantity information. There is some support for this in this study, in the slightly larger associations between spatial language and mathematical tasks that require symbolic number skills, i.e., components of the standardised mathematics measure and number line estimation, relative to ANS performance. However, this should be interpreted tentatively as the associations between spatial language and number line estimation association were only slightly larger than those between spatial language and ANS performance. Furthermore, as requirements for grounding symbolic and conceptual representations are low for most spatial skill tasks, this explanation may explain the unique variation explained by spatial language for some mathematics tasks, beyond variation explained by spatial skill performance alone.

Third, from a developmental perspective, spatial concepts may act as the building blocks on which later mathematical concepts are formed (Mandler 2012; Mandler & Pagan-Canovas, 2014). Mandler (2012) proposed that all other concepts are derived from spatial primitives, i.e., there is a spatial basis to all concept formation. Therefore, individuals who form rich spatial concepts may have an advantage in concept development in other domains, e.g., recall, inferences, early language learning, mental problem solving and mathematics (Mandler, 2012). This may be reflected in associations between spatial language (a proxy for spatial concept development) and mathematics (mathematical concept formation). However, this proposal does not align with the differential associations that were found between spatial language and different mathematics tasks in this study. It also cannot explain why associations between spatial language and mathematics persisted even after controlling for spatial skills. Thus, this explanation is difficult to support.

4.1 Limitations and future work

This is the first study to systematically explore the relations between spatial language, spatial skills and mathematics in children aged 6 to 10 years using a cross-sectional approach.

However, the findings would be strengthened by longitudinal research following a single cohort from 6 to 10 years, thus providing more nuanced information on individual variation in spatial language development. To our knowledge, this study is the first to investigate the role of spatial language for mathematics, and to propose theoretical mechanisms that may underpin the reported associations. However, while we assume a direction of causality between spatial language and mathematics based on theoretical arguments, we used an associational design and so future intervention work is required to determine the causal direction of these effects. Furthermore, although several hypotheses have been proposed to explain causal spatial-mathematical relations, including spatial visualisation, form perception and spatial scaling (Mix et al., 2016), it is not yet known, how spatial language fits into this model. This further emphasises the need to establish the causal relationships between spatial skills, spatial language and mathematics.

The generalisability of these findings is limited by the absence of socio-economic status (SES) as a control variable. Previous studies have reported differences in spatial language performance across SES groups (Purpura & Reid, 2016) and this should be explored in future work. Furthermore, although this study controlled for language skills using a vocabulary task, the results would be strengthened by controlling for a broader range of language abilities, e.g., expressive and receptive grammar. Finally, the age-based analyses presented were only powered to detect medium sized effects. While this is the minimum desirable effect for new educational interventions (Hattie, 2008), from a theoretical perspective there is a need to replicate these findings with larger samples.

4.2   Conclusion

Taken together, the results of this study show age-based differences in spatial language skills through middle childhood. Furthermore, significant associations were reported between spatial language and both spatial (particularly extrinsic spatial tasks) and mathematics

(particularly classroom-based mathematics) outcomes. Given that finding novel methods of improving mathematical thinking in children is an educational priority in the UK (National Audit Office UK, 2018), these findings provide initial support for the design of future spatial language intervention studies and highlight spatial language production as a particular intervention target .

5 References

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. Psychological Bulletin, 141(5), 931-965.

Barsalou, L. W. (1999). Perceptual symbol systems. Behavioural and Brain Sciences, 22, 577–660.

Broadbent, H. J., Farran, E. K., & Tolmie, A. (2014). Object-based mental rotation and visual perspective-taking in typical development and Williams syndrome. Developmental Neuropsychology, 39(3), 205–225.

Cannon, J., Levine, S., & Huttenlocher, J. (2007). A system for analyzing children and caregivers' language about space in structured and unstructured contexts. Spatial Intelligence and Learning Center (SILC) technical report.

Casasola, M., Wei, W. S., Suh, D. D., Donskoy, P., & Ransom, A. (2020). Children's exposure to spatial language promotes their spatial thinking. Journal of Experimental Psychology: General.

Chatterjee, A. (2008). The Neural Organization of Spatial Thought and Language. Seminars in Speech and Language, 29(03), 226–238.

Cheng, Y.-L., & Mix, K. S. (2014). Spatial Training Improves Children's Mathematics Ability. Journal of Cognition and Development, 15(1), 2–11.

Department for Education. (2017, January). Schools, pupils and their characteristics: January 2017. Retrieved from https://www.gov.uk/government/statistics/schools-pupils-and-their-characteristics-january-2017

De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in

children's mathematical skills? A review of evidence from brain and behaviour. Trends
in Neuroscience and Education, 2 (2), 48-55.

Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). The British Picture Vocabulary
Scale - 3rd Edition. London: GL Assessment.

Dunn, L. M., & Dunn, L. M. (1997). Peabody picture vocabulary test-III. Circle Pines, MN:
American Guidance Service.

Farran, E. K., & O'Leary, B. (2015). Children's ability to bind and maintain colour-location
conjunctions: The effect of spatial language cues. Journal of Cognitive Psychology,
28(1), 44-51.

Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics. SAGE.

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Picturing perspectives: development of
perspective-taking abilities in 4- to 8-year-olds. Frontiers in Psychology, 5.

Friso-van den Bos, I., Kroesbergen, E. H., Van Luit, J. E. H., Xenidou-Dervou, I., Jonkman,
L. M., Van der Schoot, M., & Van Lieshout, E. C. D. M. (2015). Longitudinal
development of number line estimation and mathematics performance in primary
school children. Journal of Experimental Child Psychology, 134, 12–29.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., ... &
Schatschneider, C. (2010). Do different types of school mathematics development
depend on different constellations of numerical versus general cognitive abilities?.
Developmental psychology, 46(6), 1731.

Geer, E. A., Quinn, J. M., & Ganley, C. M. (2019). Relations between spatial skills and math
performance in elementary school children: A longitudinal investigation.
Developmental psychology, 55(3), 637.

Gilligan, K. A., Flouri, E., & Farran, E. K. (2017). The contribution of spatial ability to mathematics achievement in middle childhood. Journal of Experimental Child Psychology, 163, 107–125.

Gilligan, K. A., Hodgkiss, A., Thomas, M. S. C., & Farran, E. K. (2018b). The use of discrimination scaling tasks: A novel perspective on the development of spatial scaling in children. Cognitive Development, 47, 133–145.

Gilligan, K. A., Thomas, M. S., & Farran, E. K. (2019). First demonstration of effective spatial training for near transfer to spatial performance and far transfer to a range of mathematics skills at 8 years. Developmental science, e12909.

Gilmore, C., Attridge, N., De Smedt, B., & Inglis, M. (2014). Measuring the approximate number system in children: Exploring the relationships among different tasks. Learning and Individual Differences, 29, 50–58.

Hattie, J. (2008). Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement. Routledge Press.

Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S. (2017). Enhancing Children's Spatial and Numerical Skills through a Dynamic Spatial Approach to Early Geometry Instruction: Effects of a 32-Week Intervention. Cognition and Instruction, 35(3), 236–264.

Hermer-Vazquez, L., Moffet, A., & Munkholm, P. (2001). Language, space, and the development of cognitive flexibility: The case of two spatial memory tasks. Cognition, 79, 263-299.

Hurst, M., Anderson, U. S., & Cordes, S. (2016). The acquisition of mappings between number words, numerals, and non-symbolic quantities in preschoolers. Journal of Cognition and Development, 18, 41-62.

LeFevre, J.-A., 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.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. Cognitive Psychology, 50, 315-353.

Mandler, J. M. (2012). On the Spatial Foundations of the Conceptual System and Its Enrichment. Cognitive Science, 36, 421-451.

Mandler, J. M., & Pagan Canovas, C. (2014). On defining image schemas. Language and Cognition, 0, 1-23. https://doi.org/10.1017/langcog.2014.14

Matheson, H. E., & Barsalou, L. W. (2018). Embodiment and grounding in cognitive neuroscience. In J. Wixted (Ed.), Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience (4th ed.).

Mercer, N., & Sams, C. (2006) Teaching Children How to Use Language to Solve Maths Problems. Language and Education, 20 (6), 507-528.

Miller, H. E., Patterson, R., & Simmering, V. R. (2016). Language supports young children's use of spatial relations to remember locations. Cognition, 150, 170-180.

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.

Mix, K. S., Hambrick, D. Z., Satyam, V. R., Burgoyne, A. P., & Levine, S. C. (2018). The latent structure of spatial skill: A test of the 2×2 typology. Cognition, 180, 268-278.

Mizzi, A. (2017). The Relationship between Language and Spatial Ability. An Analysis of Spatial Language for Reconstructing the Solving of Spatial Tasks. Springer.

Munnich, E., & Landau, B. (2003). The effects of spatial language on spatial representation: Setting some boundaries. *Language in mind: Advances in the study of language and thought*, 113-155.

National Audit Office UK. (2018). Delivering STEM (science, technology, engineering and mathematics) skills for the economy. London: Department for Business, Energy & Industrial Strategy Department for Education.

Newcombe, N. S. (2010). Picture This: Increasing Math and Science Learning by Improving Spatial Thinking. American Educator, 34 (2), 29-35.

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

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

Prediger, S., Erath, K., & Moser Opitz, E. (2019). Language challenges for students with mathematics difficulties: An overview on research results and instructional approaches. International handbook of math learning difficulties: From the laboratory to the classroom*, 437–455. Cham: Springer.

Pruden, S. M., & Levine, S. C. (2017). Parents' spatial language mediates a sex difference in preschoolers' spatial-language use. Psychological science, 28(11), 1583-1596.

Pruden, S. M., Levine, S. C., & Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? Developmental Science, 14(6), 1417-1430.

Purpura, D. J., & Ganley, C.M. (2014). Working memory and language: Skill-specific or domain-general relations to mathematics? Journal of Experimental Child Psychology, 122, 104-121.

Purpura, D. J., Hume, L. E., Sims, D. M., & Lonigan, C. J. (2011). Early literacy and early numeracy: The value of including early literacy skills in the prediction of numeracy development. Journal of Experimental Child Psychology, 110, 647–658.

Purpura, D. J., & Logan, J. A. (2015). The nonlinear relations of the approximate number system and mathematical language to early mathematics development. Developmental Psychology, 51(12), 1717.

Purpura, D. J., & Reid, E. E. (2016). Mathematics and language: Individual and group differences in mathematical language skills in young children. Early Childhood Research Quarterly, 36, 259–268.

Purpura, D. J., Napoli, A. R., Wehrspann, E. A., & Gold, Z. S. (2017). Causal connections between mathematical language and mathematical knowledge: A dialogic reading intervention. *Journal of Research on Educational Effectiveness*, *10*(1), 116-137.

Robertson, S., Graven, M. (2020). Language as an including or excluding factor in mathematics teaching and learning. Mathematics Education Research Journal, 32, 77–101.

Riccomini, P.J., Smith, G. W., Hughes, E. M., & Fries, K. M. (2015) The Language of Mathematics: The Importance of Teaching and Learning Mathematical Vocabulary. Reading & Writing Quarterly, 31(3), 235-252.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. Psychological Science, 14(3), 237-243.

Simms, N. K., & Gentner, D. (2019). Finding the middle: Spatial language and spatial reasoning. Cognitive Development, 50, 177-194.

Szechter, L. E., & Liben, L. S. (2004). Parental Guidance in Preschoolers' Understanding of Spatial-Graphic Representations. Child Development, 75 (3), 869 – 885.

Toll, S. W. M., & Van Luit, J. E. H. (2014). Explaining numeracy development in weak performing kindergartners. Journal of Experimental Child Psychology, 124, 97–111.

Ufer, S., & Bochnik, K. (2020). The Role of General and Subject-specific Language Skills when Learning Mathematics in Elementary School. Journal für Mathematik-Didaktik, 41, 81-117.

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.

von Aster, M. G., & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology*, *49*(11), 868–873.

Witkin, H. A., Otman, P. K., Raskin, E., & Karp, S. (1971). A manual for the embedded figures test. California: Consulting Psychologists Press.