Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale

It Is a “Small World”: Relations Between Performance on Five Spatial Tasks and Five Mathematical Tasks in Undergraduate Students

Véronic Delage, Richard J. Daker, Geneviève Trudel, Ian M. Lyons, and Erin A. Maloney

Online First Publication, April 11, 2024. https://dx.doi.org/10.1037/cep0000329

CITATION

Delage, V., Daker, R. J., Trudel, G., Lyons, I. M., & Maloney, E. A. (2024). It is a “small world”: Relations between performance on five spatial tasks and five mathematical tasks in undergraduate students.. *Canadian Journal of Experimental Psychology / Revue canadienne de psychologie expérimentale*. Advance online publication. https://dx.doi.org/10.1037/cep0000329

Canadian Journal of Experimental Psychology /
Revue canadienne de psychologie expérimentale

ISSN: 1196-1961

https://doi.org/10.1037/cep0000329

It Is a "Small World": Relations Between Performance on Five Spatial Tasks and Five Mathematical Tasks in Undergraduate Students

Véronic Delage[1], Richard J. Daker[2], Geneviève Trudel[1], Ian M. Lyons[2], and Erin A. Maloney[1]
[1] Faculty of Social Sciences, School of Psychology, University of Ottawa
[2] Department of Psychology, Georgetown University

One of the most robust relations in cognition is that between spatial and mathematical reasoning. One important question is whether this relation is domain general or if specific relations exist between performance on different types of spatial tasks and performance on different types of mathematical tasks. In this study, we explore unique relations between performance on five spatial tasks and five mathematical tasks. An exploratory factor analysis conducted on Data Set 1 ($N = 391$) yielded a two-factor model, one spatial factor and one mathematical factor with significant cross-domain factor loadings. The general two-factor model structure was replicated in a confirmatory factor analysis conducted in a separate data set ($N = 364$) but the strength of the factor loadings differed by task. Multidimensional scaling and network-based analyses conducted on the combined data sets reveal one spatial cluster, with a central node and one more tightly interconnected mathematical cluster. Both clusters were interconnected via the math task assessing *geometry and spatial sense*. The unique links identified with the network-based analysis are representative of a "small-world network." These results have theoretical implications for our understanding of the spatial–mathematical relation and practical implications for our understanding of the limitations of transfer between spatial training paradigms and mathematical tasks.

Public Significance Statement
We investigated unique relations between five different mathematical tasks and five different spatial tasks in two different data sets of undergraduate students. We found that the mathematical tasks were strongly interconnected forming a tight group and the spatial tasks were also interconnected, forming another group. Both groups were connected via a geometry task that had unique relations with both the mathematical and spatial tasks. The findings indicate specific relations between the mathematical and spatial domains.

Keywords: mathematical reasoning, spatial reasoning, small-world networks

It has been argued that one of the most robust relations in cognition and learning is that between spatial reasoning and mathematical reasoning (Cheng & Mix, 2014; Hawes & Ansari, 2020; Mix & Cheng, 2012; Mix et al., 2016; Tosto et al., 2014). Spatial reasoning refers to mentally visualizing, rotating, and transforming spatial and visual information (Gardner, 1993). It is generally found that poor performance on tasks thought to require spatial reasoning is associated with weaker performance on tasks thought to require mathematical reasoning (Rotzer et al., 2009). This relation has been observed in many stages of development, from children to adolescents and adults (Geary et al., 2021; Geer et al., 2019; Gilligan et al., 2019; Haciomeroglu, 2016; Rohde, 2008; Sewell, 2008; Webb et al., 2007). Further, evidence of a potential causal relation between spatial reasoning and mathematical reasoning can be seen in a meta-analysis by Hawes et al. (2022), in which it was found that the average effect size of spatial training on mathematics was 0.28 (Hedges's g) relative to control conditions, with age serving as a significant moderator. Important to note, however, is the fact that there are inconsistent findings within the literature wherein, in some cases, spatial training transfers to mathematics reasoning, and in some cases, it does not (for a review, see Hawes et al., 2022).

Although the general relation between performance on spatial tasks and performance on mathematical tasks is well established, there are many questions regarding the nature of this relation.

Véronic Delage https://orcid.org/0000-0002-1610-3894
This research was funded by a Natural Sciences and Engineering Research Council of Canada's Discovery Grant (RGPIN-2018-04672) to Erin A. Maloney. The authors have no conflicts of interest to disclose.

The results of this article were presented in a poster presentation of the same name at the 2022 Canadian Society for Brain, Behaviour and Cognitive Science conference in Halifax, Nova Scotia. In addition, Data Set 1 of this study has been previously used in other published articles (Daker et al., 2022; Delage et al., 2022), but the research questions and analyses included in this article are all new and unique to this article. All materials and data are available at https://osf.io/a78qj/?view_only=0f4b407a1a6c4708b402ce5d011b0bbe.

Correspondence concerning this article should be addressed to Véronic Delage or Erin A. Maloney, Faculty of Social Sciences, School of Psychology, University of Ottawa, 136 Jean Jacques Lussier, Vanier Hall, Ottawa, ON K1N 6N5, Canada. Email: vdela018@uottawa.ca or erin.maloney@uottawa.ca

Notably, we have not yet determined why and how spatial reasoning and mathematical reasoning are related, and whether these relations are domain general or domain specific (Hawes & Ansari, 2020). In other words, it is unknown whether spatial reasoning in general relates to mathematical reasoning in general due to some type of shared underlying representation or processing, or whether specific types of spatial reasoning are related to specific types of mathematical reasoning and not to other types of mathematical reasoning. Within the more domain-general argument, one proposed explanation for why these two general constructs are related revolves around the theory that numbers are spatially represented (Hawes & Ansari, 2020; Mix & Cheng, 2012). Indeed, many researchers believe that numbers are represented in the mind along a mental number line (akin to a ruler) and that this representation is inherently visuospatial in nature. Thus, numerical and mathematical reasoning is believed to be related to spatial reasoning via recourse to the mental number line (Restle, 1970). Other researchers within the domain-general camp theorize that spatial processes are recruited for solving mathematical tasks (e.g., Presmeg, 2006; Rasmussen & Bisanz, 2005). Consistent with this claim, performing mathematical tasks and spatial tasks activate similar neural circuits, suggesting the recruitment of some of the same underlying cognitive processes (Hubbard et al., 2005; Walsh, 2003).

In contrast, it may be the case that relations between spatial reasoning and mathematical reasoning are not domain general but rather that unique relations exist between specific spatial skills and specific mathematical skills (e.g., Bailey, 2017; Caviola et al., 2012; Robert & Lefevre, 2013; Trbovich & LeFevre, 2003). For example, visual–spatial working memory (WM) demands vary depending on the layout of an addition problem (horizontal vs. vertical; Caviola et al., 2012; Trbovich & LeFevre, 2003) and the difficulty of substructions (small-operand vs. large-operand; Robert & Lefevre, 2013). This domain-specific hypothesis may explain the inconsistent findings within the spatial training literature. If the relations between space and math are domain general, then one would expect to see the transfer as a function of training. However, if there are unique relations between certain types of spatial reasoning and certain types of mathematical reasoning, and researchers opt to train one type of spatial reasoning and test for transfer to one type of mathematical reasoning that happens to be unrelated, then one would not expect to see transfer.

The domain-specific hypothesis hinges on the fact that within the domains of both space and math, there are different tasks that are believed to be indexing different types of spatial and mathematical reasoning. This heterogeneity within each domain makes understanding the link between math and space complex. Adding even more complexity, there remains debate within the field regarding how to divide, define, and name the subcategories of spatial reasoning (Carroll, 1993; Höffler, 2010; Linn & Petersen, 1985; Mix & Cheng, 2012; Mix et al., 2018; Uttal et al., 2013). Similarly, there are multiple strands of mathematics that likely index different types of mathematical reasoning (Gilmore, 2023; Hjelte et al., 2020). Despite this, in the literature examining the relations between spatial reasoning and mathematical reasoning, different categories of spatial reasoning and different categories of mathematical reasoning are often used interchangeably and simply discussed under the umbrella of "spatial reasoning" and "mathematics," respectively.

In this article, we take the stance that spatial reasoning and mathematical reasoning both represent heterogeneous categories, and we test the hypothesis that the relations between these categories are domain specific. Below we discuss the existing studies that have led us to this hypothesis.

Correlational Studies

In this overview, we synthesize correlational studies that investigate the relations between specific spatial skills and mathematical skills (see Mix & Cheng, 2012, for a thorough review). These studies often involve children and adolescents as well as adults, and findings across age groups are not sufficiently isolated to permit age-specific conclusions. We categorize these relations based on the type of spatial skill assessed for ease of presentation.

Spatial Visualization

Performance on spatial visualization tasks, which require mental transformation of 2D and 3D objects, is related to performance on several types of mathematics. For example, performance on the Block Design task is positively related to performance on arithmetic word problems in 12-year-olds (Hegarty & Kozhevnikov, 1999) and on a counting task in 6-year-olds (Kyttälä et al., 2003). In Harris et al. (2021), performance on the Spatial Visualization subtest of the Spatial Reasoning Instrument was a positive predictor of fifth graders' (10- to 12-year-olds) performance on geometry and measurement questions and number sense questions. Harris et al. (2021) also found the same pattern in eighth graders (13- to 14-year-olds) with performance measured with a paper-folding task. The positive relation between performance on spatial visualization tasks and performance on a variety of mathematical tasks is observed in kindergarteners up to high school students (Haciomeroglu, 2016; Harris et al., 2021; Lachance & Mazzocco, 2006; Mazzocco & Myers, 2003; Sherman, 1980).

Mental Rotation

Performance on mental rotation tasks, which require individuals to mentally rotate objects, a substrand of spatial visualization, has also been shown to correlate positively with performance in geometry, word problems, and arithmetic tasks, as well as general standardized math tests in adults and adolescents (Casey et al., 1995; Delgado & Prieto, 2004; Gilligan et al., 2019; Harris et al., 2021; Kyttälä & Lehto, 2008; Lombardi et al., 2019; Moè, 2018; Reuhkala, 2001; van Tetering et al., 2019; L. Wang & Carr, 2020; Weckbacher & Okamoto, 2014). More recently, the relationship between performance on mental rotation tasks and arithmetic as well as overall mathematics tasks has also been found in children (Gilligan et al., 2019; Lombardi et al., 2019; Moè, 2018; van Tetering et al., 2019), although this relation is not always found for every age group (Gilligan et al., 2019; Moè, 2018).

Disembedding

Research exploring the relation between disembedding performance, which requires finding figures in a distracting background, and mathematics performance has found that there were significant but weak relations between disembedding performance and performance on math tests, number line estimation tasks, and approximate number sense tasks in children (Gilligan et al., 2019; Mazzocco & Myers, 2003). The relation between performance on disembedding tasks and mathematical tasks does not seem to have been studied in adolescents or adults.

Perspective-Taking

Perspective-taking performance has been previously related to geometry and measurement tasks, number sense tasks, and number line estimation tasks, but again only in children (Cameron et al., 2019; Frick, 2019; Gilligan et al., 2019; Harris et al., 2021; Kulp, 1999; Kurdek & Sinclair, 2001; Nesbitt et al., 2019; Sortor & Kulp, 2003). For example, Harris et al. (2021) found that performance on the Spatial Orientation subtest of the Spatial Reasoning Instrument was positively correlated with performance on geometry and measurement questions and number sense questions in fifth and eighth graders. As with disembedding, the relation between performance on perspective-taking tasks and mathematical tasks does not seem to have been studied in adolescents or adults.

Navigation

Navigation, which refers to the ability to navigate through environments, has recently been identified as an important subcomponent of spatial reasoning (Ferguson et al., 2015; Lyons et al., 2018; Sokolowski et al., 2019). Navigation was one of the three subcomponents of spatial reasoning identified by the factor analyses used to create a scale measuring anxiety about tasks involving spatial reasoning (Lyons et al., 2018). Navigation tasks are also related to the concept of sense of direction and large-scale spatial skills, as found in the Santa Barbara Sense-of-Direction Scale (Hegarty et al., 2002), which measures environmental spatial abilities. These studies have explored the relation between navigation performance and math anxiety but no studies to our knowledge have yet to explore the relation between navigation performance and mathematics performance.

Factor Analyses

Another method of exploring relations between different spatial and mathematics tasks is to use methods that incorporate many tasks within the same analysis, such as factor analysis. Mix et al. (2016) had over 800 children from kindergarten ($M_{\text{age}} = 6.04$), third grade ($M_{\text{age}} = 9.04$), and sixth grade ($M_{\text{age}} = 11.74$) complete a series of mathematical and spatial tasks. Using cross-domain factor analysis and cross-domain multiple regression analyses, they found that the tasks converged into two factors, one mostly spatial and one mostly mathematical. However, the cross-domain loadings obtained in the factor analyses were not consistent across age groups. A possible explanation the authors suggest for their results is the "novel versus familiar hypothesis" (Mix et al., 2016). In this hypothesis, spatial reasoning is more important in learning and completing novel math problems than familiar ones. Indeed, the math tasks that cross-loaded onto the spatial factor, calculation, fractions, and algebra are new and challenging concepts for kindergarteners, third graders, and sixth graders, respectively. Interestingly, the authors replicated this study (Mix et al., 2017), and although the general two-factor model (one spatial, one mathematical) and all within-domain factor loadings replicated, not all cross-domain loadings were the same in this second wave of participants. The authors determined that "cross-domain loadings, even those that were replicated, appear much more fragile and context-specific than the within-domain loadings" (Mix et al., 2017, p. 478).

These results may be because relations between different types of spatial and mathematical tasks in elementary school children are variable. It is unclear, however, if these relations stabilize throughout development and are more stable in adults. To our knowledge, no research has taken a parallel approach in adults wherein the relations between performance on multiple spatial tasks and multiple math tasks have been assessed within participants. Against this background, we sought to identify specific relations between spatial and mathematical reasoning tasks in undergraduate students and determine if these relations are consistent across two different data sets.

The Present Study

The objective of the present study is to extend our understanding of the relationship between mathematical and spatial reasoning. More specifically, we are interested in identifying specific relations between performance on five different spatial and mathematical tasks in undergraduate students. First, we will identify these relations by conducting cross-domain factor analyses on two separate data sets. Given some of the inconsistent findings in the literature on the latent structure between mathematical and spatial reasoning (e.g., Mix et al., 2017), we felt it prudent to obtain two independent data sets and to test the replicability of our findings. In the second data set, we included a task that indexes WM capacity (i.e., the backward letter span task) as a control variable. Given the characterization of WM as a "dissociable cognitive skill with unique links to learning outcomes" (Alloway & Alloway, 2008, p. 2) and the findings that WM is a significant predictor of academic achievement (Alloway & Alloway, 2010), we included it as a domain-general measure to help control for the possibility that people who are strong in one academic area (e.g., math) are simply likely to be strong in all academic areas. In both data sets, undergraduate participants completed five spatial tasks and five sets of math questions (each set from a different strand of mathematics). Second, we then performed a second set of analyses, multidimensional scaling (MDS), and network-based analyses on both data sets combined to determine whether there was support for the factor analyses and help visualize the findings.

We have taken three relatively new approaches to exploring this question. First, we opted to use the math strands outlined in the mathematics public curriculum in the geographic region in which the study herein was conducted because we felt that this would have the greatest chance of representing the types of mathematics most commonly encountered (at least during schooling years) by the participants in this study. The Ontario mathematics curriculum Grades 1–8 (Ontario Ministry of Education, 2005) divides math knowledge into five strands: (a) *data management and probability*, (b) *measurement*, (c) *number sense and numeration*, (d) *geometry and spatial sense*, and (e) *algebra*. This approach is different from previous studies in that most studies exploring the relation between mathematical and spatial reasoning use basic arithmetic tasks or simply use one or two general math tasks that do not distinguish between the different types of math taught in schools (e.g., Cameron et al., 2019; Nesbitt et al., 2019; L. Wang & Carr, 2020). Second, selecting strands from the education system is an ecologically valid approach and will make our findings more directly applicable to real-world situations and education as the math strands selected are the ones commonly used in schools. The third novel approach consists of using a network-based analysis approach (in addition to classic factor analysis) to better visualize and draw more detailed inferences about the relations between specific math and spatial tasks. Behavioural sciences are replete with theoretical frameworks in which the structure of associative networks is of central interest, and mapping these associative networks in graphical form is likely to be of significant value (J. Wang et al., 2010; Watts & Strogatz, 1998). We believe that to be the case here to map the relations between tasks in the spatial and mathematical domains.

Taken together, the results of this study have both theoretical implications for our understanding of the space–math relation, whether it is domain general or domain specific, and practical implications for our understanding of the limitations of transfer between spatial training paradigms and mathematical tasks.

Because the relation between mathematics and spatial reasoning has been found throughout development (e.g., Geary et al., 2021; Haciomeroglu, 2016; Rohde, 2008; Sewell, 2008; Webb et al., 2007), including adulthood, we anticipate the general findings from Mix et al. (2016, 2017) to be replicated in the present data set. More specifically, we expect to find two factors, one spatial and one mathematical, when performing the cross-domain factor analyses, with all tasks loading onto their respective domains. Based on the studies discussed earlier, we predict that (a) the tasks measuring *spatial visualization* and *mental rotation* performance will both be related to the math domain and will be uniquely related to the *geometry and spatial sense* and the *number sense and numeration* strands, (b) *disembedding* will not be related to the math domain and will not be uniquely related to any strand of math, (c) perspective-taking will be related to the math domain and uniquely related to the *geometry and spatial sense*, the *measurements*, and the *number sense and numeration* strands. We do not have any predictions for the *navigation* spatial task because its correlation to mathematics performance has not yet been studied.

Method

Participants Data Set 1

Four hundred twenty-five participants were recruited from an undergraduate student research pool. The pool is a research participation programme designed to allow students in introductory psychology courses to participate in research to receive credits towards the course. The students received 1 credit for their participation in this 1-hr long study, as per institutional guidelines. After reading the consent form, six students did not consent to participate, three participants withdrew from the study, and 25 participants were excluded from the analysis after the data cleaning procedure (see the Scoring section, for details). Thus, in total, 391 students (121 identified as males; 265 identified as females; two identified “other” as their gender; and three did not provide their gender) were included in the analyses. Students who did not consent still received course credit as per local Research Ethics Board regulations. Of the participants included in the analyses, 57% identified English and 15.1% identified French as their first language, 0.8% identified both French and English, and 23.3% identified another language, while 3.8% did not provide this information. The mean age of the participants is 19.3 years (seven participants did not provide their age). Note that we use the term “gender” here rather than “sex” as students were asked to identify their gender by selecting “female,” “male,” or “other” in the demographic portion of the study. Data Set 1 has been previously used in other articles (Daker et al., 2022; Delage et al., 2022), but the research question and analyses included in this article are novel.

Participants Data Set 2

Undergraduate students were recruited through the same student research pool from Study 1, at the University of Ottawa. To be eligible for this study, participants needed to have not been participants in Data Set 1. Five hundred forty-seven responses were recorded. Of those responses, seven did not consent to participate, 79 withdrew from the study (did not complete the study or did not provide a response for an entire task), and 36 were identified as a double response (i.e., participated in the study twice). Double responses were identified by the student pool identification code of each student collected before the consent form. Only the first completed response for each identification code was retained, and all subsequent responses were identified as double responses and were discarded before analyses were conducted. Finally, 56 participants were excluded from the analysis after the data cleaning procedure (see the Scoring section, for details). After these exclusions, 364 participants remained (117 identified as males; 241 identified as females; four identified as other; two did not respond to the gender question) and were included in the analyses. Of those included in the analysis, 54.2% identified English and 17.1% identified French as their first language and 28.7% identified another language. The mean age of the participants was 19.9 years (note that seven individuals did not provide age information).

Procedure Data Set 1

The research presented herein received Research Ethics Board approval from the University of Ottawa’s Office of Research Ethics. All research was conducted following ethical research guidelines. Data collection took place during the 2019 fall semester, and the study was approximately 1 hr long. All tasks and questionnaires were completed via Qualtrics, in English. While an online platform was used to administer the tasks, data collection was nonetheless completed in person. When participants arrived at the lab, they were brought to a testing station that included a computer, a pencil, a paper, and a calculator. Participants were first asked to complete the consent form on the Qualtrics survey, and once informed consent was obtained, the study commenced. The math questions were completed first, followed by the spatial tasks and the demographic questionnaire.[1]

Procedure Data Set 2

The procedure for Data Set 2 was identical to Data Set 1 with the following exceptions. Data collection took place during the 2020 summer and fall semesters. Participants completed the study remotely due to the COVID-19 pandemic. Participants were instructed that they needed a computer, a pencil, a paper, a calculator, and a quiet room to participate in the study. Participants were first asked to

[1] Participants also completed questionnaires measuring their math anxiety, general anxiety, and spatial anxiety. These questionnaires were not included in these analyses because they did not correspond to our objectives for this article.

complete the consent form on the Qualtrics survey, and once informed consent was obtained, the study commenced. Participants completed the same math questions and the same five spatial tasks used in Data Set 1 and a measure of WM. Unlike the first study, the presentation order of these three sections was randomized across participants. The presentation order of the five math strands and the five spatial tasks was also randomized. Finally, participants completed the demographic questionnaire last. Although this study was not preregistered, all materials and data for both data sets are available at https://osf.io/a78qj/?view_only=0f4b407a1a6c4708b402ce5d011b0bbe.

Measures

Spatial Tasks. Our five measures of spatial reasoning were selected based on the current literature on spatial reasoning (Mix & Cheng, 2012; Uttal et al., 2013) and tasks easily available to us that were appropriate for adults. We attempted to select a variety of different spatial tasks found in the literature, but we would like to point out that our tasks do not cover all possible types of spatial tasks. In addition, we would like to remind the reader that there are inconsistencies in the literature on how to classify and name spatial tasks.

Mental Rotation. Manipulation performance was measured with a computerized mental rotation task (Shepard & Metzler, 1971). Participants were shown two 3D objects made of 10 adjoining cubes, which were oriented in different directions (Figure 1a). The participants were asked to identify whether they thought the two objects shown were the same objects oriented differently or if they were two different objects by pressing "Same" or "Different" on the screen. There were 15 trials in this task presented one at a time without a time limit. Once a participant selected a response, their response time (RT) for the trial was recorded and they were automatically directed to the next trial.

Dot Localization. A modified computerized version of the dot localization task was used to evaluate participants' performance in spatial visualization (Manna et al., 2010). For this task, the participant was first presented with a rectangle containing two dots for 125 ms. Once this rectangle had disappeared, the participant was presented with another rectangle containing a grid and was asked to press locations within the grid to identify where the two dots, previously shown, would have been located if both rectangles had been superimposed (Figure 1b). There were 15 trials in this task presented one at a time without a time limit for answering.

Navigation. To measure participants' navigation performance, a modified and computerized version of the Road Map Test of Directional Sense was used (Ferguson et al., 2015; Money et al., 1965). For this task, participants were presented with a map that contained a dotted path. On each "street" corner, participants were shown the letter R (right turn) or the letter L (left turn) to demonstrate the direction they would be turning if they were walking along the dotted path. However, not every turn was labelled correctly. Therefore, participants were required to press "Y" (yes) or "N" (no) to identify whether they agreed or disagreed with the direction provided (see Figure 1c). Three maps were presented one at a time with three, 17 and 33 turns without a time limit. Only the third map with 33 turn was used to score performance, the other two were used as practice trials as per Ferguson et al. (2015).

Disembedding. Participants' *disembedding* performance was measured with a computerized and modified version of the embedded figures task (Ekstrom et al., 1976; Lyons et al., 2018). For this task, a complex two-dimensional line drawing is shown to the participant, and they are asked to identify which figure out of five simple line figures is present in the complex drawing (see Figure 1d). There were nine trials in this task presented one at a time without a time limit. The five simple figures presented on each trial were always the same. Once a participant selected a response, their RT for the trial was recorded and they were automatically directed to the next trial.

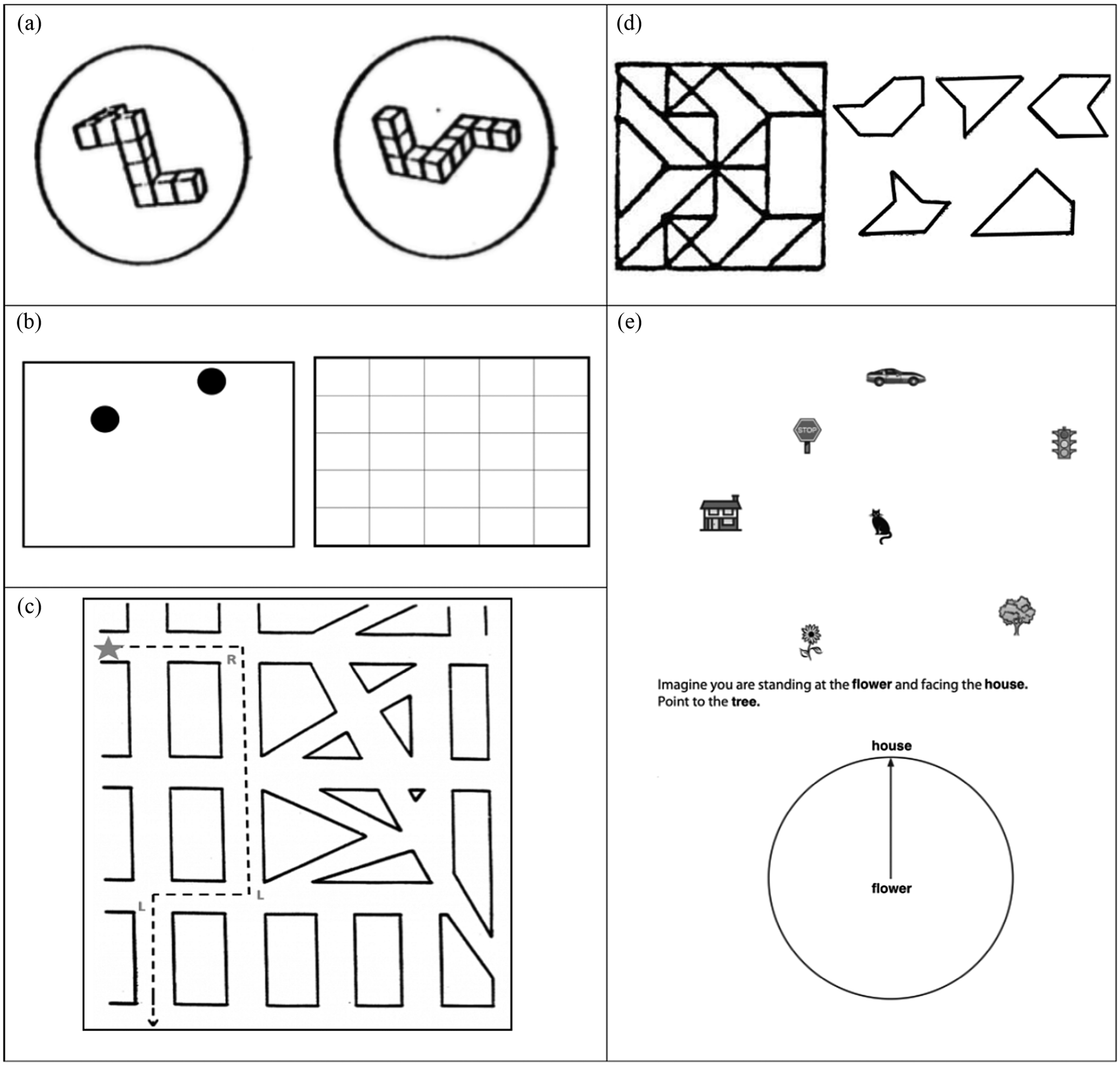
Perspective-Taking. To measure participants' performance in perspective-taking, individuals were asked to complete the Hegarty test (Hegarty & Waller, 2004). For this task, participants were shown a screen with a variety of common objects (cat, car, house, etc.) and an "arrow circle." The participants were asked to imagine that they were standing in the location of the object in question (e.g., Object A in the middle of the circle) and facing a particular point (e.g., Object B at the top of the circle). They were then asked to determine in which direction they would find a third object (e.g., Object C) by using the mouse to click the appropriate area (see Figure 1e). There were 15 trials in this task presented one at a time without a time limit. Once a participant selected a response, their RT for the trial was recorded and they were automatically directed to the next trial.

Math Test. Participants were asked to complete a short mathematical test composed of a variety of fifth (average ages 10–11) to seventh (average ages 12–13) grade questions. These questions were designed by a math curriculum specialist, and they were designed to evaluate concepts taught in the five strands of mathematics evaluated in fifth to seventh grade by the Ontario mathematics curriculum (Ontario Ministry of Education, 2005). More specifically, individuals were asked to answer 10 questions in each of the following strands: data management and probability, geometry and spatial sense, number sense and numeration, measurement, and algebra. These questions were designed to be similar to questions that could be present on a provincially standardized math test. Thus, participants were presented with a math problem and were given four multiple-choice answers to choose from. Once the participant had selected their choice, they were automatically directed to the next question. Questions were presented one at a time without a time limit. These questions were used as the mathematical concepts addressed in these questions are presented early enough in Canadian education that participants should be familiar with these types of questions (i.e., increasing the likelihood of having participants be able to solve the math problems). See Figure 2 for a sample of the math questions.

Data Management and Probability. This strand deals with different ways to gather, organize, display, and analyze data, as well as probability models and situations while applying this knowledge to real-world situations. In this strand, we asked questions to evaluate an individual's ability to interpret graphs, calculate the probability that something may or may not occur, calculate a missing value based on an average, and calculate means and medians.

Measurement. In this strand, students learn about units and processes involved in measurements and apply them to real-life scenarios. Participants were asked to calculate areas, perimeters, and volumes; complete conversions of grams to kilograms; and calculate distances in kilometres per hour.

Figure 1
Stimuli Used for the Five Spatial Tasks



Note. (a) Mental rotation task, (b) dot localization task, (c) navigation task, (d) disembedding task, and (e) perspective-taking task.

Number Sense and Numeration. This strand involves understanding basic numbers, operations, and strategies to solve problems. In this strand, we asked questions about fractions, time (e.g., number of minutes in 1 year), adding and subtracting decimals, and calculation of monetary costs.

Geometry and Spatial Sense. In this strand, students learn basic shapes and figures, their attributes, and geometric properties, as well as skills related to location and movements and the use of the Cartesian plane. To evaluate an individual's proficiency in this strand, we asked questions related to translation, rotation, and reflection of shapes; naming geometric shapes and properties; and asked individuals to locate objects on the Cartesian plane.

Algebra. Algebra is the study of patterns and relationships and deals with solving equations. To evaluate an individual's proficiency in this strand of math, we asked participants to use addition, subtractions, multiplications, and divisions to find missing numbers in equations and to demonstrate equality in equations with unknown quantities on both sides.

Working Memory. Participants in Data Set 2 also completed a measure of WM capacity, a computerized version of the backward letter span task, which was adapted from the backward digit span task on the Wechsler Intelligence Scale for Children, Third Edition (Maloney et al., 2010; Wechsler, 1997). The participants were visually shown a series of letters at a rate of 1 letter per second.

Figure 2

Sample of the Mathematics Questions

Data management and probability

Kyla is a member of the starting lineup of the school's basketball team. The heights of the other starting players are shown below.

160 cm, 156 cm, 148 cm, 147 cm

The mean height of the starting lineup is 152.4 cm. What is Kyla's height?

a. 153 cm
b. 151 cm
c. 191 cm
d. 160 cm

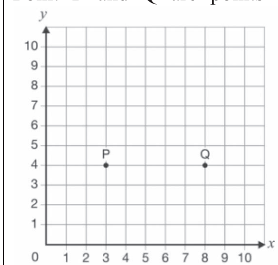
Number sense and numeration

A store sells 1 orange for 3.00$, what is the cost of 8 oranges?

a. 48.00$
b. 16.00$
c. 96.00$
d. 24.00$

Geometry and spatial sense

Point P and Q are points of a square.



Which set of ordered pairs could represent the other two points needed to form this square?

a. (3, 8) and (8, 8)
b. (3, 9) and (8, 9)
c. (8, 3) and (9, 8)
d. (7, 4) and (4, 9)

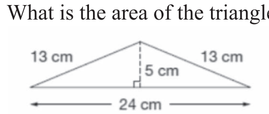
Algebra

Solve this equation: $1 = 10 - W$

a. W = -9
b. W = -11
c. W = 9
d. W = 11

Measurements

What is the area of the triangle shown below?



a. 60 cm^2
b. 65 cm^2
c. 120 cm^2
d. 156 cm^2

Participants had to recall the letters shown in backward order. The series of letters ranged from two letters to eight, starting with two letters and increasing by one letter every two trials (i.e., Level 1 = two letters, Level 2 = three letters, … Level 7 = eight letters). Scoring was based on the highest number of letters in a series correctly recalled by the participant, thus scoring ranged from 0 to 8. Scoring was discontinued when the participants failed both trials at any given level.

Scoring

Accuracy (error rate [ER]; proportion incorrect) and RTs were recorded for each math question and spatial task trials. Several of the math tasks displayed evidence of ceiling effects (see Table 1). In such cases, RTs can add important variation, especially when one's analytic approach emphasizes individual differences, as it does here. Conversely, not all tasks showed ceiling effects, indicating ERs likely captured meaningful variance for those tasks. Hence, to use a single, consistent measure for all tasks in the study—across the mathematical and spatial domains—behavioural performance was operationalized using z-scores that averaged standardized (z-scored) ERs and RTs.

When calculating the average RT for a given task, it is important to note that the high ecological validity of the math tasks meant that there was a very high degree of variability in the expected completion time (RT) for different trials. For instance, on a given math test, one might expect one type of problem to take students only a few seconds on average, while another problem might be expected to take several minutes or more. Hence, when calculating RTs, it was important to treat different trials within a task separately during the data triage process.

When calculating RTs, it is common practice to remove trials on which no response was made (RT = undefined), or when RTs were implausibly low (e.g., RT < 250 ms for a trial expected to take 30 s or more), or when RTs were implausibly high (due to participant distraction, etc.). Here, we arbitrarily determined unfeasibly low RTs as <250 ms and unfeasibly high as >3 *SD*s higher than the average RT for that trial. Across the entire data sets, a total of 1.8% of trials for Data Set 1 and 2.6% of trials for Data Set 2 were removed because of these "outlier" RTs. Note that this outlier removal procedure was determined using Data Set 1 and then applied to both Data Sets 1 and 2.

Again, due to the high variability in expected mean RTs across trials even within a task, removed trials created the potential for biased estimates of an individual's mean RT for a given task. For instance, imagine Person A omitted a response for a trial with an expected RT of 1.5 s on average, and Person B omitted a response for a trial with an expected RT of 45 s on average. The calculated

Table 1
Descriptive Statistics of All Measures, Data Set 1

Measure	*M* ER	ER *SD*	*M* RT	RT *SD*
Mental rotation	.28	0.20	10.46	5.29
Dot localization	.32	0.23	3.69	0.93
Disembedding	.71	0.17	22.72	17.02
Perspective-taking	.38	0.28	17.72	5.85
Navigation	.22	0.28	3.86	2.25
Data management	.16	0.18	53.18	17.74
Measurement	.19	0.15	36.85	15.64
Number sense and numeration	.06	0.09	20.02	12.09
Geometry and spatial sense	.19	0.19	29.47	10.19
Algebra	.08	0.13	16.77	7.47

Note. For the navigation task, the mean RT per turn for Trial 3 is shown. ER = error rates; RT = response time in seconds.

average will appear to be much lower for Person B simply because their data set does not include a valid RT data point for the very long (~45 s) trial. To accommodate this issue, when a trial was dropped from a given task, we computed a weighted RT score that takes into account the average RT of that trial across the whole sample. Each trial of a task was assigned a weight that represented the proportion of overall task RT accounted for by that trial. For example, if a trial was found, on average, to make up 12% of the total RT for that task, that trial received a weight of .12. If a participant's RT was an outlier for that trial, when computing their weighted task RT, we summed the RT on each other trial of that task and divided by .88.

Participants who were missing 30% or more of the total RT time for a given task were excluded from the analysis altogether (Data Set 1: 20 participants, Data Set 2: 48 participants[2]). For the navigation task, only one trial is used to calculate performance, thus participants with an RT equal to or greater than 3 *SD* away from the mean for this trial were excluded from the analysis (Data Set 1: five participants, Data Set 2: eight participants). ERs were computed based only on using the trials that were included in the RT analysis. Note that the weighted RT procedure was constructed using the data in Data Set 1 and applied to both Data Sets 1 and 2.

To create our composite ER and RT scores, we *z*-scored both their ERs and RTs for each task and took the average. Shapiro–Wilk tests of normality showed that this method of creating composite ER and RT scores produced more normal distributions than the untransformed ER scores (which suffered from ceiling effects) and other methods of creating composite ER and RT scores (i.e., inverse efficiency, Townsend & Ashby, 1978, and combined performance, as used in Lyons et al., 2014).

Analysis

All factor analyses were performed with a maximum likelihood estimation model in R Version 4.0.2 using participants' *z*-scores as described in the Scoring section. We used maximum likelihood estimation over orthogonal estimation because we wanted to allow the factors to correlate as mathematics and spatial reasoning are two domains known to be correlated (Mix & Cheng, 2012). For the same reason, a promax rotation was used as it is an oblique rotation and therefore allows factors to be correlated. For the exploratory factor analysis of the first data set, we determined the number of informative factors using eigenvalues above 1.00. Then we evaluated the model fit using the root-mean-square error of approximation (RMSEA), the root-mean-square of the residual (RMSR), and the Tucker–Lewis index of factoring reliability (TLI). All three of these indices are not strongly affected by sample size (Schermelleh-Engel et al., 2003). RMSEA and RMSR values below .05 are viewed as a good fit and below .08 as an acceptable fit, as lower values indicate a better fit (Schermelleh-Engel et al., 2003). TLI value larger than .97 is a good cutoff for a good model fit and .95 is acceptable (Schermelleh-Engel et al., 2003). In addition to estimating the factor loadings, we also used a bootstrapping method to obtain 95% confidence intervals around those estimates to evaluate the significance of each factor loading.

Results

Descriptive Statistics

We first ran descriptive statistics for all the variables for both data sets. See Table 1 for descriptive statistics for each measure of Data Set 1 and Table 2 for Data Set 2.

Factor Analyses

Next, we performed a cross-domain factor analysis on each data set including performance on all the spatial and mathematical tasks. These analyses will give insight into specific relations between performance on mathematical and spatial tasks. We expected to find two factors, one spatial and one math, with all tasks loading onto their respective domains. We also expected to have the dot localization, mental rotation, and perspective-taking tasks to cross-load onto the math domain.

For the first data set, we found convergence onto two factors; the eigenvalue of the first factor (3.59) and the second factor (1.22) were above 1. Two of the model fit measures were good (RMSR = 0.04, RMSEA = 0.04 [0.02, 0.06]) and the other was acceptable (TLI = 0.97). As shown in Table 3, for the first factor, all tasks have significant factor loadings except for the disembedding task, but the spatial tasks (factor loadings ranging from 0.28 to 0.66, except for disembedding) are more strongly loading onto this factor than most of the math tasks (factor loadings ranging from 0.11 to 0.15, except for geometry and spatial sense). The one math task that is highly loading onto this factor is geometry and spatial sense (factor loading of 0.53). For the second factor, all math tasks have significant factor loadings with geometry having the lowest factor loading of 0.22 and with the measurement strand having the highest factor loading of 0.72. In addition, two of the spatial tasks have a significant factor loading for this factor, dot localization (factor loading of 0.18) and interestingly perspective-taking has a negative factor loading of

[2] Breakdown of excluded participants per task (participants were excluded altogether from analyses for having more than 30% of trials removed for a given task because of implausibly short or long RT). Data Set 1: *data management and probability*, one participant; *measurements*, two participants; *algebra*, three participants; *mental rotation*, three participants; *dot localization*, three participants; *disembedding*, five participants; and *perspective-taking*, three participants. Data Set 2: *data management and probability*, two participants; *measurements*, six participants; *number sense and numeration*, two participants; *geometry and spatial sense*, 13 participants; *algebra*, two participants; *dot localization*, two participants; *disembedding*, 20 participants; and *perspective-taking*, one participant.

Table 2
Descriptive Statistics of All Measures, Data Set 2

Measure	*M* ER	ER *SD*	*M* RT	RT *SD*
Mental rotation	.31	0.21	11.23	6.69
Dot localization	.37	0.18	3.78	1.09
Disembedding	.71	0.17	27.83	24.25
Perspective-taking	.42	0.31	19.56	9.13
Navigation	.25	0.30	3.62	2.41
Data management	.26	0.25	53.18	20.92
Measurement	.26	0.22	37.34	18.15
Number sense and numeration	.10	0.16	19.05	5.30
Geometry and spatial sense	.24	0.24	31.42	14.62
Algebra	.11	0.18	18.49	9.67
Backward letter span	5.85[a]	2.06		

Note. For the navigation task, the mean RT per turn for trial three is shown. ER = error rates; RT = response time in seconds.
[a] The mean score/span reached by participants for the backward letter span task.

−0.21. The other spatial tasks do not have significant factor loadings on the second factor. Overall, these results suggest one mainly spatial factor and one mainly mathematical factor with some cross-loading, with geometry and spatial sense and dot localization positively loading onto both factors. Consistent with the literature that these two domains are positively related, the correlation between the two factors is 0.56.

Next, we performed two cross-domain confirmatory factor analyses with the second data set. The first is the same factor analysis we performed with Data Set 1 (not controlling for WM). The second factor analysis also consisted of a cross-domain factor analysis but we residualized each measure with respect to performance on the backward letter span task. This allowed us to test whether our factor analytic results depended on individual differences in WM capacity. For these factor analyses, we hoped to replicate the findings of the factor analysis for the first data set, thus we set the number of factors for these analyses to two factors. The fit of the two-factor model without controlling for WM was good for one measure (RMSR = 0.04), acceptable for another measure (RMSEA = 0.05 [0.04, 0.08]), and slightly below the acceptable benchmark for the last measure (TLI = 0.94). As shown in Table 4a, for the first factor, all math tasks have significant factor loadings with geometry and spatial sense having the lowest factor loading of 0.48 and measurement the highest of 0.81. This is consistent with the two-factor model of Data Set 1. In contrast, only one of the spatial tasks has a significant factor loading for the mathematical factor, dot localization (factor loading of 0.20), this task was also loading onto the mathematical factor for Data Set 1. The other spatial tasks do not have significant factor loadings for the mathematical factor. For the second factor, three of the five spatial tasks have significant factor loadings (ranging from .31 to .69), with perspective-taking having the highest factor loading. In addition, as for Data Set 1, geometry and spatial sense (factor loading of 0.35) is also significantly loading onto this mostly spatial factor. The disembedding task did not have a significant factor loading for either of the factors. Overall, these results replicate the ones found with Data Set 1, suggesting one mainly spatial factor and one mainly mathematical factor with some cross-loading with geometry and spatial sense and dot localization loading onto both factors. However, not all factor loadings were replicated. Consistent with the literature that these two domains are positively related and Data Set 1, the correlation between the two factors is 0.62.

Next, we performed the same cross-domain factor analysis while controlling for WM. As seen in Table 4b, we obtained similar results with factor loadings only varying by 0.03 or less, but there were differences in the significance of the factor loadings. For the first factor, all math tasks still have a significant factor loading, but the spatial task dot localization is no longer significantly loading onto this math factor. In addition, the second factor now only has three significant factor loadings, mental rotation, perspective-taking, and geometry and spatial sense, as the navigation task's factor loading is no longer significant.

MDS and Network-Based Analyses

Next, we performed a second set of analyses on both data sets combined to determine whether there was support for the factor analyses and help visualize the findings. MDS and network-based analyses are novel analyses that may bring insight into understanding the specific relations between spatial and mathematical tasks.

Consistency Across Data Sets. We first sought to establish consistency of results across data sets. To visualize this, one can see the zero-order correlation matrices from each data set in Figure 3. Visually, these appear quite similar; however, it would be useful to supplement this qualitative evaluation with a more quantitative approach. Because the data sets were collected with different sets

Table 3
Cross-Domain Standardized Pattern Matrix Loadings for Factor Analysis of Data Set 1

Predictor	Factor 1 Coefficient	Factor 1 Confidence interval	Factor 2 Coefficient	Factor 2 Confidence interval
Mental rotation	**0.37***	[0.18, 0.52]	−0.04	[−0.21, 0.13]
Spatial visualization	**0.28***	[0.13, 0.42]	0.18*	[0.02, 0.37]
Disembedding	0.06	[−0.13, 0.19]	−0.12	[−0.27, 0.03]
Perspective-taking	**0.97***	[0.66, 1.17]	−0.21*	[−0.37, −0.01]
Navigation	**0.43***	[0.27, 0.54]	0.02	[−0.11, 0.19]
Data management	0.14*	[0.05, 0.26]	**0.69***	[0.57, 0.91]
Measurement	0.15*	[0.07, 0.25]	**0.72***	[0.61, 0.94]
Number sense and numeration	0.11*	[0.02, 0.22]	**0.47***	[0.36, 0.64]
Geometry and spatial sense	**0.53***	[0.35, 0.68]	0.22*	[0.08, 0.42]
Algebra	0.15*	[0.05, 0.27]	**0.56***	[0.45, 0.76]

Note. Values in bold indicate the (significant) factor loading that was greater, considering Factor 1 versus Factor 2.
* $p < .05$.

Table 4
Cross-Domain Standardized Pattern Matrix Loadings for Factor Analysis of Data Set 2

Predictor	Factor 1		Factor 2	
	Coefficient	Confidence interval	Coefficient	Confidence interval
(a)				
Mental rotation	0.15	[−0.07, 0.42]	**0.49***	[0.15, 1.02]
Dot localization	**0.20***	[0.04, 0.47]	0.21	[−0.03, 0.42]
Disembedding	−0.09	[−0.27, 0.14]	0.29	[0.00, 0.56]
Perspective-taking	0.06	[−0.15, 0.53]	**0.69***	[0.18, 1.09]
Navigation	0.03	[−0.16, 0.36]	**0.31***	[0.01, 0.51]
Data management	**0.72***	[0.58, 0.88]	0.01	[−0.10, 0.28]
Measurement	**0.81***	[0.63, 0.97]	−0.05	[−0.23, 0.33]
Number sense and numeration	**0.72***	[0.56, 0.83]	−0.16	[−0.34, 0.19]
Geometry and spatial sense	**0.48***	[0.35, 0.78]	0.35*	[0.17, 0.54]
Algebra	**0.56***	[0.40, 0.76]	0.03	[−0.12, 0.25]
(b)				
Mental rotation	0.17	[−0.12, 0.54]	**0.48***	[0.11, 1.06]
Dot localization	0.21	[0.00, 0.54]	0.19	[−0.12, 0.49]
Disembedding	−0.10	[−0.29, 0.15]	0.28	[−0.03, 0.60]
Perspective-taking	0.09	[−0.17, 0.63]	**0.66***	[0.13, 1.09]
Navigation	0.01	[−0.17, 0.35]	0.30	[−0.02, 0.51]
Data management	**0.71***	[0.42, 1.02]	0.00	[−0.25, 0.43]
Measurement	**0.80***	[0.47, 1.11]	−0.06	[−0.36, 0.47]
Number sense and numeration	**0.70***	[0.47, 0.88]	−0.17	[−0.42, 0.27]
Geometry and spatial sense	**0.50***	[0.24, 0.92]	0.34*	[0.04, 0.67]
Algebra	**0.54***	[0.31, 0.82]	0.02	[−0.24, 0.37]

Note. (a) Shows the cross-domain standardized pattern matrix loadings without controlling for working memory performance. (b) Shows the cross-domain standardized pattern matrix loadings controlling for working memory performance on the backward letter span task. Values in bold indicate the (significant) factor loading that was greater, considering Factor 1 versus Factor 2.
* $p < .05$.

of participants, one cannot compute test–retest reliability in the usual fashion. However, we can instead quantify the similarity between the patterns of results by correlating the correlation matrices themselves. One can think of the correlation matrices in Figure 3 as sets of summary statistics (r-values) that quantify the relationship between each pair of variables in each data set. The question is whether the relative magnitudes of these pairwise relations are preserved across data sets. Because we are dealing with summary statistics, we can directly relate results between data sets without referring to the underlying data. Operationally, we achieve this by vectorizing each correlation matrix and, to convert bounded r-values into normally distributed Fisher's z-values, by taking the inverse hyperbolic tangent of each r-value, $z_{ij} = \text{atanh}\,(r_{ij})$, where r_{ij} is the correlation between a given pair of variables i and j. Finally, we simply correlate the resulting vectors of z-values. Doing so here showed strong agreement between data sets: $r = 0.854$ ($p < .001$). In other words, the relative magnitudes of the pairwise relations characterizing the data sets were well preserved, despite being collected with entirely separate samples of individual participants.

A related approach, especially relevant here, is to correlate the factor loadings from the factor analyses across data sets. In essence, we can ask whether a given task loaded onto each factor to a similar extent in each data set. Here, we took the rotated loading coefficients from the two-factor solutions (Tables 3 and 4a). Note that we took values from Table 4a instead of 4b as the former did not control for WM, which is more comparable to Data Set 1 (though results were highly similar whether one used factor loadings from Table 4a or 4b). In addition, because what constituted Factors "1" and "2" were arbitrarily reversed, we aligned factors according to the domain each primarily captured: spatial (Data Set 1, Factor 1; Data Set 2, Factor 2) and mathematical (Data Set 1, Factor 2; Data Set 2, Factor 1). In other words, for each task, we aligned the "mathematical factor" loadings across studies and the "spatial factor" loading across studies. Table 5 shows these aligned factor loadings (columns labelled Data Set 1 and Data Set 2). As with the analysis above, we Fisher's z-transformed the factor loadings to be normally distributed. The above sequence of steps revealed that agreement between factor loadings across studies was high: $r = 0.734$ ($p < .001$).

In summary, whether examining agreement in terms of pairwise relations between variables ($r = 0.854$) or relative factor loadings ($r = 0.734$), results across data sets demonstrated a high degree of consistency.

Data Visualization. In this section, we sought to further characterize the interrelations between task domains (mathematical and spatial) using a network-based approach. A simple starting point is to visualize the results from the factor analyses, as doing so can potentially reveal larger patterns that may not be immediately evident from examining statistical tables alone.

To this end, Figure 4a visually summarizes the factor loadings for all 10 tasks. For a given task, we averaged its loading on a given factor (mathematical or spatial) across data sets by computing row-wise averages from Table 5 ("Average" column).[3] Next, we

[3] Given (a) the high degree of consistency between data sets established in the previous section, (b) a desire to avoid overfitting conclusions to any one data set, (c) the fact that the analyses in this section use summary statistics as their starting point, the following analyses take their inputs as summary data averaged over the two data sets. Doing so should also yield conclusions most likely to replicate beyond this article.

Figure 3

Zero-Order Correlation Matrices for Data Set 1, Data Set 2, and the Average Thereof

Dataset 1

	MRT	DL	DEM	PERS	NAV	DAT	MEAS	NUM	GEO	ALG
MRT		.12	.13	.29	.23	.14	.29	-.02	.35	.23
DL	.12		-.09	.30	.09	.18	.29	-.06	.36	.28
DEM	.13	-.09		.00	-.08	.09	.24	-.10	.24	.20
PERS	.29	.30	.00		.38	.24	.34	.00	.53	.23
NAV	.23	.09	-.08	.38		.17	.22	-.08	.33	.20
DAT	.14	.18	.09	.24	.17		.63	.44	.46	.51
MEAS	.29	.29	.24	.34	.22	.63		.41	.47	.56
NUM	-.02	-.06	-.10	.00	-.08	.44	.41		.32	.34
GEO	.35	.36	.24	.53	.33	.46	.47	.32		.35
ALG	.23	.28	.20	.23	.20	.51	.56	.34	.35	

Dataset 2

	MRT	DL	DEM	PERS	NAV	DAT	MEAS	NUM	GEO	ALG
MRT		.20	.29	.41	.16	.34	.22	.02	.37	.14
DL	.20		-.04	.27	.17	.39	.25	.07	.33	.18
DEM	.29	-.04		.14	.01	.21	.16	.07	.26	.08
PERS	.41	.27	.14		.26	.42	.24	.08	.50	.21
NAV	.16	.17	.01	.26		.21	.28	.08	.31	.15
DAT	.34	.39	.21	.42	.21		.57	.44	.53	.41
MEAS	.22	.25	.16	.24	.28	.57		.50	.54	.43
NUM	.02	.07	.07	.08	.08	.44	.50		.37	.40
GEO	.37	.33	.26	.50	.31	.53	.54	.37		.42
ALG	.14	.18	.08	.21	.15	.41	.43	.40	.42	

Average

	MRT	DL	DEM	PERS	NAV	DAT	MEAS	NUM	GEO	ALG
MRT		.16	.21	.35	.19	.24	.25	.00	.36	.18
DL	.16		-.06	.29	.13	.29	.27	.01	.34	.23
DEM	.21	-.06		.07	-.03	.15	.20	-.01	.25	.14
PERS	.35	.29	.07		.32	.33	.29	.04	.51	.22
NAV	.19	.13	-.03	.32		.19	.25	.00	.32	.18
DAT	.24	.29	.15	.33	.19		.60	.44	.50	.46
MEAS	.25	.27	.20	.29	.25	.60		.45	.51	.49
NUM	.00	.01	-.01	.04	.00	.44	.45		.35	.37
GEO	.36	.34	.25	.51	.32	.50	.51	.35		.38
ALG	.18	.23	.14	.22	.18	.46	.49	.37	.38	

Note. MRT = mental rotation; DL = dot localization; DEM = disembedding; PERS = perspective-taking; NAV = navigation; DAT = data management; MEAS = measurement; NUM = number sense and numeration; GEO = geometry and spatial sense; ALG = algebra. See the online article for the color version of this figure.

plotted these factor loadings in a two-dimensional space, with "spatial" arbitrarily assigned to the *x*-axis and "mathematical" arbitrarily assigned to the *y*-axis. (Note this is very similar to a MDS approach.)[4] Figure 4a shows four of the five math tasks, coloured red, tightly clustered together along the *y*-axis (the "mathematical" dimension). The five spatial tasks (blue) are less tightly clustered but still largely occupy a distinct portion of the graph. Thus, the graph visually captures several key features of the factor analyses from previous sections: The math tasks load more strongly onto the mathematical factor than did the spatial tasks onto the spatial factor, as indicated by tighter clustering for the former. Moreover, a given task tended to load more highly on the factor representing its "native" domain, which can be seen in Figure 4a as two sets of tasks visually separated into two subareas of the graph.

Turning to what Figure 4a reveals about specific tasks, we see a key exception to the above pattern in the geometry and spatial sense

[4] In a traditional MDS approach, one projects factor loadings into a desired set of dimensions, represented by the same number of factors (the loading on Factor-1 is the *x*-coordinate, that on Factor-2 is the *y*-coordinate, and so on). One typically ensures orthogonality of coordinates by extracting factors that are a priori forced to be orthogonal. Here, we allowed factors to be correlated given strong theoretical assumptions in the literature that these domains are related (and indeed the two factors were correlated in both data sets, per the results discussed in previous sections). To achieve orthogonal coordinates here, we thus relied on the pattern matrix portion of the factor analysis output. The pattern matrix represents the loadings of a given input variable (in our case a task) onto each factor, controlling for all other factors. Rotated factor loadings from the pattern matrix are thus by definition orthogonal. Indeed, the rotated factor loadings in Tables 3 and 4a are taken from their respective pattern matrices. Hence, it is reasonable in the current case to express these loadings as relative weights on orthogonal dimensions, for instance as *x*- and *y*-axes in a scatterplot. Figure 3a is thus effectively a scatterplot depicting each task's relative position on "spatial" (*x*) and "mathematical" (*y*) dimensions.

Table 5
Aligned Factor Loadings for Data Set 1, Data Set 2, and the Average Thereof

Task	Data Set 1	Data Set 2	Average
Spatial factor			
Mental rotation	0.37	0.49	0.43
Dot localization	0.28	0.21	0.25
Disembedding	0.06	0.29	0.18
Perspective-taking	0.97	0.69	0.83
Navigation	0.43	0.31	0.37
Data management	0.14	0.01	0.08
Measurement	0.15	−0.05	0.05
Number sense and numeration	0.11	−0.16	−0.03
Geometry and spatial sense	0.53	0.35	0.44
Algebra	0.15	0.03	0.09
Mathematical factor			
Mental rotation	−0.04	0.15	0.06
Dot localization	0.18	0.20	0.19
Disembedding	−0.12	−0.09	−0.11
Perspective-taking	−0.21	0.06	−0.08
Navigation	0.02	0.03	0.03
Data management	0.69	0.72	0.71
Measurement	0.72	0.81	0.77
Number sense and numeration	0.47	0.72	0.60
Geometry and spatial sense	0.22	0.48	0.35
Algebra	0.56	0.56	0.56

task, which showed significant loadings on both domain factors (the only task to do so consistently—i.e., in both data sets). More specifically, the geometry and spatial sense task loaded more highly onto the spatial factor than any of the other math tasks. This result is visually captured in Figure 4a by the fact that the geometry and spatial sense task is "positioned" roughly halfway between the mathematical and spatial clusters. However, this qualitative assessment does not capture whether geometry bridges the two domains such that it shares specific features with unique aspects of both mathematical and spatial reasoning.

Network-Based Analysis. In both data sets, we observed significant factor loadings for the geometry and spatial sense task on both the mathematical and the spatial factor, and in Figure 4a, geometry and spatial sense appears to occupy a position roughly in between the two domains. However, on further reflection, it is not entirely clear what this means. Most of the other math tasks loaded onto the spatial factor to some extent, as did several of the spatial tasks on the mathematical factor. Hence, one can ask whether the cross-loading seen for the geometry and spatial sense task is in fact specific to that task. Further, one can ask whether geometry is *uniquely* linked to *specific* tasks within each domain. In this way, we can more directly probe whether geometry comprises a unique combination of processes that share features with both mathematical and spatial processes. More broadly, this approach can help us better uncover the specific cognitive mechanisms that link the mathematical and spatial domains.

To test whether the geometry and spatial sense task is uniquely related to specific tasks with each of the two domains, we computed the partial correlation matrix characterizing unique relations between all 10 tasks. As above, we combined the two data sets by averaging summary statistics—in this case, the zero-order correlation matrices in Figure 3. To quantify unique associations between tasks, the (averaged) zero-order matrix was adjusted to reflect only estimates of shared variance that cannot be attributed to any of the other variables in the matrix. In other words, the matrix of zero-order r-values was reduced[5] to a matrix of partial r-values. Figure 4b shows the same scatterplot as in Figure 4a, with a statistically significant unique relation (a significant partial correlation) indicated by a link between those two tasks. Red lines indicate significant unique links between two math tasks; blue lines indicate significant unique links between two spatial tasks; pink lines indicate significant unique links between a math and a spatial task. Finally, it is important to note that this approach represents a fairly stringent test: A link that is present indicates a significant ($p < .05$) relation between the two tasks, *after controlling for the influence of all eight other tasks*. In other words, a link in Figure 4b represents a unique relation *specific* to that pair of tasks.

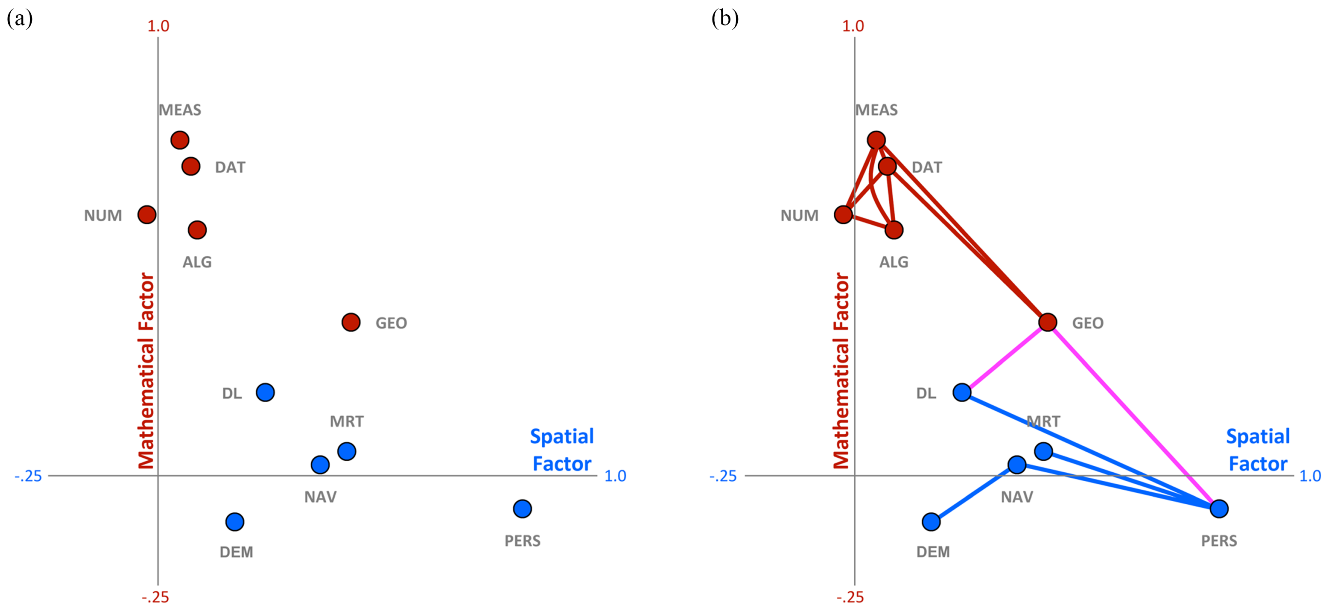
Figure 4b shows that the five math tasks were highly interconnected, with each math task connected to a minimum of two other math tasks. Conversely, the five spatial tasks were largely connected through a central node: the perspective-taking task. In addition, the two domains were connected via links to the geometry and spatial sense task. This sort of configuration—two distinct "neighbourhoods" of dense connections that in turn connect with one another via a central (or "hub") node—is indicative of what is often referred to as a "small-world network." Indeed, one can quantify the "small-worldness" of a network by estimating the extent to which path length is kept relatively short via a small number of connections between densely connected neighbourhoods (Watts & Strogatz, 1998). We quantified "small-worldness" via the Small World Index (SWI; Neal, 2017), where SWI values >1 are conventionally taken to indicate a network exhibiting small-world properties. The SWI of the network depicted in Figure 4b was 1.89, indicating the network indeed exhibited small-world-like properties.

In summary, if we distill the relations between this set of five math and five spatial tasks to only those characteristics shared by each pair of tasks, we see that they organize themselves into distinct domains—a result that corroborates the factor analyses above. Further, we see that these domains are linked via the geometry and spatial sense task. In particular, geometry and spatial sense shares unique characteristics with both specific math tasks (measurement and data management) and specific spatial tasks (perspective-taking and dot localization). In this way, we can answer in the affirmative that geometry indeed bridges the math and spatial domains. Moreover, this bridge consists of a specific constellation of shared cognitive processes. That is, geometry consists of a set of underlying cognitive processes, some of which are unique to specific types of spatial processing and others are unique to specific types of mathematical processing. It is thus not the case that all math tasks are inherently spatial, and vice versa. Instead, the links between domains arise from a specific subset of skills that are common to both domains, and geometry happens to comprise a particularly large number of these cross-domain skills.

[5] We did so using a pseudoinverse procedure pioneered by Strimmer (Opgen-Rhein & Strimmer, 2007; Schäfer & Strimmer, 2005); see also the R package implementing this technique at https://www.strimmerlab.org/software/corpcor/index.html. In essence, inverting a covariance matrix by definition orthogonalizes the off-diagonal elements with respect to one another. One must then return these elements to standardized units—that is, partial correlations (hence "pseudo"-inverse). These can be computed in the usual manner by dividing each element by the square root of the product of its constituent variances, which can be taken from the main diagonal of the inverted matrix.

Figure 4

Visual Representations of the MDS and Network-Based Analyses of Both Data Sets Combined



Note. (a) Scatterplot visualizing the factor loadings of each task. Factor loadings were averaged across both data sets. Spatial tasks are shown in blue; math tasks are shown in red. (b) Lines between nodes indicate a significant correlation between those two tasks, controlling for the influence of the eight other tasks. Blue lines indicate relations between two spatial tasks; red lines indicate relations between two math tasks; pink lines indicate relations between a spatial and a math task. MDS = multidimensional scaling; MEAS = measurement; DAT = data management; NUM = number sense and numeration; ALG = algebra; GEO = geometry and spatial sense; DL = dot localization; MRT = mental rotation; NAV = navigation; DEM = disembedding; PERS = perspective-taking. See the online article for the color version of this figure.

Discussion

On average, people who perform better on measures of spatial skill tend to perform better on measures of mathematical skill. While this relation is often discussed in general terms, it is imperative to remember that "mathematical reasoning" and "spatial reasoning" are umbrella terms used to represent a constellation of different skills. Further, research in this area is leaning towards domain-specific relations between these two constructs and not only a domain-general relation (e.g., Bailey, 2017; Caviola et al., 2012; Robert & Lefevre, 2013; Trbovich & LeFevre, 2003). Indeed, within any given mathematical task or spatial task, there are a host of cognitive processes at play. As such, understanding if specific links exist between performance on different mathematical tasks and spatial tasks is highly important to begin uncovering what these cognitive processes might be. In this work, we sought to explore specific relations, using factor analysis, MDS, and network-based analyses, between performance on five different spatial tasks and five different math tasks in undergraduate students. The factor analyses represented a more classic method of exploring relations between two constructs while the MDS analysis allowed us to visualize the results of the factor analyses and gain more information on the unique relations between performance on the spatial and math tasks. The network-based analysis, where we tested for significant correlations between each pair of tasks after controlling for all eight other tasks (significant partial correlations), then helped us to identify which pairs of tasks share unique variance.

Space and Math: Related but Distinct Constructs

Results from the factor analyses revealed a two-factor model with a spatial factor and a mathematical factor. In Data Set 1, all of the math tasks loaded onto the mathematical factor, and all but one spatial task loaded onto the spatial factor (the disembedding task did not significantly load onto either factor). Further, the two factors were significantly and positively related to one another. The two-factor model was replicated in the confirmatory factor analyses conducted using Data Set 2, the only difference being that only three spatial tasks loaded significantly onto the spatial factor. These results replicated those of Mix et al. (2016, 2017) who also found a two-factor model in factor analyses of spatial and mathematical tasks in children. This evidence is consistent with the idea that spatial reasoning and mathematical reasoning are two related but distinct constructs.

The results generated via the MDS and network-based analyses were also consistent with the related but distinct constructs relation between spatial and mathematical reasoning. Indeed, the tasks were visually organized into two separate but related clusters of tasks. The math tasks formed a tight, highly interconnected cluster that was strongly associated with the mathematical factor, except for the geometry and spatial sense task, which was farther away from the math cluster and seemed as strongly associated with the mathematical factor as the spatial factor. In contrast, the spatial tasks were not as tightly clustered and interconnected but were all more strongly associated with the spatial factor than the mathematical factor.

Overlap of the Spatial Tasks

Interestingly, there was less interconnectedness within the spatial factor compared to the math factor. The disembedding task did not significantly load onto the spatial factor in either data set, the dot localization task did not load onto the spatial factor in the second factor analysis, and there were fewer unique relations between the spatial tasks than between the math tasks in the network-based analysis. These data may indicate that the math tasks share more overlapping variance with one another than do the spatial tasks.

The data presented herein are inconsistent with those reported by Mix et al. (2016, 2017) who explored relations between math and spatial tasks in children and found that spatial tasks always correlated with the spatial factor in children. While it is not clear why we see a different pattern of data than Mix et al. (2016, 2017) with respect to the degree of overlap of the spatial tasks, there are multiple possible explanations for this discrepancy. First, this difference may be because, although similar, different spatial tasks were used between the Mix et al. (2016, 2017) studies and the present study and correlations between tasks and the spatial factor may be task specific. A second possible explanation for the difference between our findings and those of Mix et al. (2016, 2017) may have to do with the age of participants. Indeed, it is possible that as we age, spatial skills become more differentiated, and/or a wider array of strategies can be used to complete spatial tasks. We cannot determine, from the current data set, why this small discrepancy exists between our results and those of Mix et al. (2016, 2017). We nonetheless believe that uncovering the explanation for this discrepancy represents an interesting avenue for future research, one that will require studies with a longitudinal design that follows a cohort of people over a number of years.

Cross-Domain Factor Loadings and Relations

Our results also revealed two cross-domain factor loadings in the factor analyses and interesting patterns of specific relations in the network-based analysis, suggesting domain-specific relations between the spatial and mathematical tasks. First, the strongest cross-domain factor loading was the geometry and spatial sense task. When we combined data sets and averaged factor loadings from both factor analyses, the relation between the geometry and spatial sense task and each factor were very similar, $r = 0.44$ for the spatial factor and $r = 0.35$ for the mathematical factor. In addition, the network-based analyses revealed that the mathematical domain and the spatial domain are interconnected via the geometry and spatial sense task with this task acting as a bridge between both clusters of tasks. More specifically, the geometry and spatial sense task shared unique variance with two other math tasks (measurement and data management) and two spatial tasks (perspective-taking and dot localization). From this, we can predict that geometry and spatial sense, measurement, and data management rely on common cognitive processes that are not, or to a much lesser degree, used in the other math tasks. In the same way, we can predict that the geometry and spatial sense task shares common cognitive processes with the perspective-taking task and the dot localization task. This may not be surprising given that "spatial sense" is incorporated into this strand and many of the questions used to measure this strand require students to hold in mind and manipulate spatial representations. For example, in one question in the Geometry and Spatial Sense section of the math task, participants are asked to perform a series of translations and reflections on a shape in a Cartesian plane and to identify its new position.

These data highlight an interesting point. That is, while we have been arguing that the concepts of "mathematical reasoning" and "spatial reasoning" represent concepts that differ from one another, even within the substrands of math tested here, there exists variability. In other words, each of the math strands also requires a diverse set of cognitive processes and these may vary between question types within the same strand. Indeed, the individual questions within the geometry and spatial sense strand may vary with respect to their degree of overlap with any given spatial skill. For example, transformations on a Cartesian plane may overlap more heavily with mental rotation, whereas determining how many lines of symmetry a given shape has may overlap more heavily with perspective-taking. Thus, while we have taken an important step towards understanding the intricate relations between math and space, there is certainly more work to be done.

Second, the dot localization task is the only spatial task that consistently cross-loaded onto the mathematical factor. In the exploratory factor analysis of Data Set 1, performance on the dot localization task cross-loaded onto the mathematical factor, suggesting that this spatial task may be more strongly related to math performance compared to other spatial tasks. In the confirmatory factor analysis, the dot localization task cross-loading replicated, with this task only loading significantly onto the mathematical factor and not the spatial factor. However, when considering both data sets and visualizing the factor loadings in the MDS analysis, it became clear that this task was part of the "spatial" cluster and is slightly more strongly associated with the spatial factor than the mathematical factor. Further, the network-based analysis revealed that this task is uniquely related to one math task, geometry and spatial sense, and one spatial task, the perspective-taking task.

The cross-domain factor loadings here can add to evidence regarding the novel versus familiar hypotheses proposed by Mix et al. (2016; discussed in the introduction). In our study, the mathematical content should consist of familiar math skills that participants would have learned in grade school because our questions were at a Grade 5–7 level. In a way, our results tend to support the novel versus familiar theory because only one out of five of each math task and spatial task cross-loaded onto the other domain in the factor analyses, and only one mathematical task was uniquely related to two spatial tasks in the network-based analysis. Perhaps as math content becomes more familiar, some relations that were once unique become more general (as our two factors were highly correlated, $r = 0.56–0.62$). This may be because as math becomes more familiar, we can begin to rely on multiple solution strategies to solve the same math. Take, for example, two-digit additions. Children may first reply on stacking the two numbers, adding the ones, and then adding the 10s (e.g., $17 + 30$). But, as the person becomes more familiar with the math, they may be able to use a decomposition strategy (e.g., $20 + 30 - 3$) or even use direct retrieval. Nonetheless, some cross-domain factor loadings were observed within our adult population, suggesting that some specific relations between spatial and mathematical tasks remain even when math operations become familiar and more automatic.

Controlling for Working Memory

Our second data set included a measure of WM as a proxy for general ability. We selected this task because it is a good predictor of academic achievement (Alloway & Alloway, 2010). The results of the confirmatory factor analysis were similar when we controlled for WM performance and when we did not. This result suggests that the relations found between the mathematical and spatial tasks in the factor analysis cannot simply be explained by a general tendency for individuals who perform well in one academic area to perform well in other areas. This finding is consistent with the hypothesis that the relation between spatial reasoning and math reasoning is not only domain general but also that specific relations exist between spatial and mathematical tasks.

Perspective-Taking, the Central Node

Another interesting finding revealed by the network-based analyses is that the perspective-taking task is a central node of the spatial task cluster and is uniquely associated with three of the other four spatial tasks, as well as the geometry and spatial sense task. This suggests that the perspective-taking task shares cognitive processes with most other spatial tasks and the geometry and spatial sense task. Perhaps the perspective-taking task indexes a cognitive process that is required to complete most other spatial tasks or rather the perspective-taking task requires multiple cognitive processes indexed by each of the other spatial tasks. For example, to complete the perspective-taking task, one may need to use two subcomponents from Kosslyn and colleague's Mental Imagery framework (Kosslyn, 1980; Kosslyn et al., 2006), image maintenance (to maintain the array of objects in mind) and image transformation (to change the perspective), while the mental rotation task may only require mental transformation and the dot localization only image maintenance. Future studies should seek to test this hypothesis and further explore these cognitive processes.

Small-World Network Properties

Lastly, we found that the unique links identified with the network-based analysis are representative of a "small-world network." A small-world network is characterized by dense neighbourhoods or clusters of nodes that are interconnected by central nodes and thus the path lengths between each node are "small." As such, the two characteristics of a small-world network is a high clustering coefficient and a small mean path length between each node (Watts & Strogatz, 1998). One example of small-world networks is social networks. The high clustering coefficient or neighbourhoods are represented by social groups, and the small mean path length is represented by the "6 degrees of separation" between all individuals in the world. This small-world network pattern has been found in many different contexts including social networks (Guimerà et al., 2003; Uzzi & Spiro, 2005), transportation networks (Neal, 2014), biological networks (Wagner & Fell, 2001), and neural networks (Guye et al., 2010; Telesford et al., 2011). Research suggests that some brain networks are organized in small-world networks, and this can help explain how information is transferred in the brain (Gallos et al., 2012). Thus, the small-world network we found when exploring unique relations between performance on mathematical and spatial tasks could inform future brain imaging and neurocognitive studies on how these two domains are related in the brain. To our knowledge, this is the first time that the small-world network has been investigated in the context of the relation between the spatial and mathematical domains. This representation may provide insight into how math and space are related and how cognitive processes between the two domains interact in the brain. More specifically, some math and spatial tasks may be central nodes that are closely connected to multiple other tasks in each domain, while other tasks that do not have direct links to the other domain may still be related but not as strongly as the "central node" tasks. In this study, we identified geometry and spatial sense as a central node of the small-world network wherein the spatial domain is one neighbourhood or cluster and the math domain is another.

The small-world network may also be a good way to reconcile the domain-general relation and domain-specific relation theories about the relation between spatial and mathematical reasoning. Whereby the central node tasks create specific unique links between some spatial and mathematical tasks and domain-general relations exist via the "6 degrees of freedom" between the other tasks that are not central nodes.

Limitations and Future Directions

Although our study employed many innovative approaches, there remains room for extension and there are important limitations to consider. First, our second data set was collected online, which means participants completed the study in their homes. This means that we were not able to control the environment and thus confirm the participant's engagement in the tasks and limit distractions. Second, one of our spatial tasks, the disembedding task, nearly had an "at chance" ER (ER = .71, at chance ER for this task = .80), which indicates it may be too difficult and not capturing participants' true disembedding skills. This task also did not load onto either factor during the factor analysis. Future studies investigating specific relations between different spatial and mathematical tasks should include a task that better captures disembedding performance.

Another limitation of the present study is the ease with which participants completed the math tasks. We worked to mediate the ceiling effects by using an accuracy and RT composite score to obtain a more normal distribution of scores. However, the task difficulty may still have created a bias in the results. While the ceiling effect allows us to confidently conclude that these tasks were familiar to the participants, it does mean that we cannot generalize our findings to math tasks that are more novel. Future studies using a mix of novel and familiar math questions are needed to determine the generalizability of these findings.

Lastly, our study was conducted using undergraduate adults. This sample has both advantages and disadvantages with respect to the interpretation of our data in light of those of Mix et al. (2016, 2017), Specifically, an advantage of our study is that it is the first, to our knowledge, to test the cross-modal relations between math and space in adults. However, the fact that we did not also include children as a comparison group within our design, means that we cannot be sure whether the discrepancies observed between our results and those of Mix et al. (2016, 2017) are due to developmental differences in the relation between mathematical and spatial reasoning, to something specific to the exact mathematical and spatial tasks employed, or to something else entirely. Nonetheless, the present study has yielded data with important implications for our understanding of the relations between math and space and has illuminated interesting questions for future research.

Conclusion

In two separate data sets, we explored unique links between different strands of mathematics and different measures of spatial skill. Our results suggest that, in North American undergraduate students, spatial and mathematical processing are two distinct but related domains that have "small-word network" properties, with geometry, a math task that seems to require more spatial strategies, as a central node that interconnects the two domains. These findings, taken together with those of previous research (e.g., Mix et al., 2016, 2017), have important theoretical implications for our understanding of the complex relations between mathematical and spatial reasoning, highlighting important domain-specific relations that may vary across developmental periods. The findings of these studies can also help guide future training studies aiming to improve different areas of mathematics using spatial training. Finally, future studies may want to explore the relation between other spatial tasks of the same type as the ones used in our study to determine if the small-world network properties found between the spatial and familiar mathematical domains are specific to these tasks at these levels of difficulty or whether small-world network properties generalize to other types of spatial and mathematical tasks.

Résumé

L'un des liens les plus robustes dans le domaine de la cognition est celui entre le raisonnement spatial et le raisonnement mathématique. Or, une question importante se pose : cette relation est-elle généralisée au domaine, ou existe-t-il des liens précis entre l'exécution de certaines tâches spatiales et l'exécution de certaines tâches mathématiques? Dans cette étude, nous évaluons les liens uniques entre l'exécution de cinq tâches spatiales et cinq tâches mathématiques. Une analyse exploratoire des facteurs effectuée sur l'ensemble de données 1 ($N = 391$) a généré un modèle bifactoriel, soit un facteur spatial et une mathématique, assorti de saturations factorielles inter domaines significatives. La structure générale du modèle bifactoriel a été reproduite lors d'une analyse de confirmation des facteurs menée sur un ensemble de données distinct ($N = 364$). Cependant, l'importance des saturations factorielles différait selon la tâche. Le positionnement multidimensionnel et les analyses en réseau menées sur les ensembles de données combinés mettent en lumière un groupe « spatial » avec un noyau central, ainsi qu'un groupe « mathématique » plus étroitement interconnecté. Les deux groupes étaient interconnectés par la tâche mathématique *géométrie et orientation spatiale*. Les liens uniques relevés par l'analyse en réseau sont représentatifs d'un réseau du type « petit monde ». Ces résultats ont des conséquences théoriques sur notre compréhension de la relation entre le spatial et la mathématique, ainsi que des conséquences pratiques sur notre compréhension des limites du transfert entre les paradigmes de la formation spatiale et les tâches mathématiques.

Mots-clés : raisonnement mathématique, raisonnement spatial, réseaux « petit monde »

References

Alloway, T. P., & Alloway, R. G. (2008). Working memory: Is it the new IQ? *Nature Precedings*, *3*, 1–17. https://doi.org/10.1038/npre.2008.2343.1

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

Bailey, D. H. (2017). Causal inference and the spatial–math link in early childhood. *Monographs of the Society for Research in Child Development*, *82*(1), 127–136. https://doi.org/10.1111/mono.12288

Cameron, C. E., Kim, H., Duncan, R. J., Becker, D. R., & Mcclelland, M. M. (2019). Bidirectional and co-developing associations of cognitive, mathematics, and literacy skills during kindergarten. *Journal of Applied Developmental Psychology*, *62*, 135–144. https://doi.org/10.1016/j.appdev.2019.02.004

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press. https://doi.org/10.1017/CBO9780511571312

Casey, M. B., Nuttall, R., Pezaris, E., & Persson Benbow, C. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, *31*(4), 697–705. https://doi.org/10.1037/0012-1649.31.4.697

Caviola, S., Mammarella, I. C., Cornoldi, C., & Lucangeli, D. (2012). The involvement of working memory in children's exact and approximate mental addition. *Journal of Experimental Child Psychology*, *112*(2), 141–160. https://doi.org/10.1016/j.jecp.2012.02.005

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

Daker, R. J., Delage, V., Maloney, E. A., & Lyons, I. M. (2022). Testing the specificity of links between anxiety and performance within mathematics and spatial reasoning. *Annals of the New York Academy of Sciences*, *1512*(1), 174–191. https://doi.org/10.1111/nyas.14761

Delage, V. (2024, January 9). *It's a "small world": Relations between performance on five spatial tasks and five mathematical tasks in undergraduate students*. https://osf.io/a78qj/?view_only=0f4b407a1a6c4708b402ce5d011b0bbe

Delage, V., Trudel, G., Retanal, F., & Maloney, E. A. (2022). Spatial anxiety and spatial ability: Mediators of gender differences in math anxiety. *Journal of Experimental Psychology: General*, *151*(4), 921–933. https://doi.org/10.1037/xge0000884

Delgado, A. R., & Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence*, *32*(1), 25–32. https://doi.org/10.1016/S0160-2896(03)00061-8

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*. Educational.

Ferguson, A. M., Maloney, E. A., Fugelsang, J., & Risko, E. F. (2015). On the relation between math and spatial ability: The case of math anxiety. *Learning and Individual Differences*, *39*, 1–12. https://doi.org/10.1016/j.lindif.2015.02.007

Frick, A. (2019). Spatial transformation abilities and their relation to later mathematics performance. *Psychological Research*, *83*(7), 1465–1484. https://doi.org/10.1007/s00426-018-1008-5

Gallos, L. K., Makse, H. A., & Sigman, M. (2012). A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proceedings of the National Academy of*

Sciences of the United States of America, *109*(8), 2825–2830. https://doi.org/10.1073/pnas.1106612109

Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences*. Basic Books.

Geary, D. C., Hoard, M. K., Nugent, L., & Scofield, J. E. (2021). In-class attention, spatial ability, and mathematics anxiety predict across-grade gains in adolescents' mathematics achievement. *Journal of Educational Psychology*, *113*(4), 754–769. https://doi.org/10.1037/edu0000487

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), 637–652. https://doi.org/10.1037/dev0000649

Gilligan, K. A., Hodgkiss, A., Thomas, M. S. C., & Farran, E. K. (2019). The developmental relations between spatial cognition and mathematics in primary school children. *Developmental Science*, *22*(4), Article e12786. https://doi.org/10.1111/desc.12786

Gilmore, C. (2023). Understanding the complexities of mathematical cognition: A multi-level framework. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *76*(9), 1953–1972. https://doi.org/10.1177/17470218231175325

Guimerà, R., Danon, L., Díaz-Guilera, A., Giralt, F., & Arenas, A. (2003). Self-similar community structure in a network of human interactions. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *68*(6 Pt 2), Article 065103. https://doi.org/10.1103/PhysRevE.68.065103

Guye, M., Bettus, G., Bartolomei, F., & Cozzone, P. J. (2010). Graph theoretical analysis of structural and functional connectivity MRI in normal and pathological brain networks. *Magnetic Resonance Materials in Biology Physics and Medicine*, *23*(5–6), 409–421. https://doi.org/10.1007/s10334-010-0205-z

Haciomeroglu, E. S. (2016). Object-spatial visualization and verbal cognitive styles, and their relation to cognitive abilities and mathematical performance. *Educational Sciences: Theory & Practice*, *16*(3), 987–1003.

Harris, D., Lowrie, T., Logan, T., & Hegarty, M. (2021). Spatial reasoning, mathematics, and gender: Do spatial constructs differ in their contribution to performance? *The British Journal of Educational Psychology*, *91*(1), 409–441. https://doi.org/10.1111/bjep.12371

Hawes, Z. C. K., & Ansari, D. (2020). What explains the relationship between spatial and mathematical skills? A review of evidence from brain and behavior. *Psychonomic Bulletin & Review*, *27*(3), 465–482. https://doi.org/10.3758/s13423-019-01694-7

Hawes, Z. C. K., Gilligan-Lee, K. A., & Mix, K. S. (2022). Effects of spatial training on mathematics performance: A meta-analysis. *Developmental Psychology*, *58*(1), 112–137. https://doi.org/10.1037/dev0001281

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual–spatial representations and mathematical problem solving. *Journal of Educational Psychology*, *91*(4), 684–689. https://doi.org/10.1037/0022-0663.91.4.684

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*(5), 425–447. https://doi.org/10.1016/S0160-2896(02)00116-2

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175–191. https://doi.org/10.1016/j.intell.2003.12.001

Hjelte, A., Schindler, M., & Nilsson, P. (2020). Kinds of mathematical reasoning addressed in empirical research in mathematics education: A systematic review. *Education Sciences*, *10*(10), Article 289. https://doi.org/10.3390/educsci10100289

Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—A meta-analytic review. *Educational Psychology Review*, *22*(3), 245–269. https://doi.org/10.1007/s10648-010-9126-7

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

Kosslyn, S. M. (1980). *Image and mind*. Harvard University Press.

Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. *Recherche* (Issue August 2012). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195179088.001.0001

Kulp, T. M. (1999). Relationship between visual motor integration skill and academic performance in kindergarten through third grade. *Optometry and Vision Science*, *76*(3), 159–163. https://doi.org/10.1097/00006324-199903000-00015

Kurdek, L. A., & Sinclair, R. J. (2001). Predicting reading and mathematics achievement in fourth-grade children from kindergarten readiness scores. *Journal of Educational Psychology*, *93*(3), 451–455. https://doi.org/10.1037/0022-0663.93.3.451

Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., & Hautamäki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology*, *20*(3), 65–76. https://doi.org/10.53841/bpsecp.2003.20.3.65

Kyttälä, M., & Lehto, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence. *European Journal of Psychology of Education*, *23*(1), 77–94. https://doi.org/10.1007/BF03173141

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

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*(6), 1479–1498. https://doi.org/10.2307/1130467

Lombardi, C. M., Casey, B. M., Pezaris, E., Shadmehr, M., & Jong, M. (2019). Longitudinal analysis of associations between 3-D mental rotation and mathematics reasoning skills during middle school: Across and within genders. *Journal of Cognition and Development*, *20*(4), 487–509. https://doi.org/10.1080/15248372.2019.1614592

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science*, *17*(5), 714–726. https://doi.org/10.1111/desc.12152

Lyons, I. M., Ramirez, G., Maloney, E. A., Rendina, D. N., Levine, S. C., & Beilock, S. L. (2018). Spatial anxiety: A novel questionnaire with subscales for measuring three aspects of spatial anxiety. *Journal of Numerical Cognition*, *4*(3), 526–553. https://doi.org/10.5964/jnc.v4i3.154

Maloney, E. A., Risko, E. F., Ansari, D., & Fugelsang, J. (2010). Mathematics anxiety affects counting but not subitizing during visual enumeration. *Cognition*, *114*(2), 293–297. https://doi.org/10.1016/j.cognition.2009.09.013

Manna, C. B. G., Tenke, C. E., Gates, N. A., Kayser, J., Borod, J. C., Stewart, J. W., McGrath, P. J., & Bruder, G. E. (2010). EEG hemispheric asymmetries during cognitive tasks in depressed patients with high versus low trait anxiety. *Clinical EEG and Neuroscience*, *41*(4), 196–202. https://doi.org/10.1177/155005941004100406

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

Mix, K. S., & Cheng, Y. L. (2012). The relation between space and math. Developmental and educational implications. In J. B. Benson (Ed.), *Advances in child development and behavior* (pp. 197–243). Elsevier Science & Technology. https://doi.org/10.1016/B978-0-12-394388-0.00006-X

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. https://doi.org/10.1016/j.cognition.2018.07.012

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., Levine, S. C., Cheng, Y. L., Young, C. J., Hambrick, D. Z., & Konstantopoulos, S. (2017). The latent structure of spatial skills and mathematics: A replication of the two-factor model. *Journal of Cognition and Development*, *18*(4), 465–492. https://doi.org/10.1080/15248372.2017.1346658

Moè, A. (2018). Mental rotation and mathematics: Gender-stereotyped beliefs and relationships in primary school children. *Learning and Individual Differences*, *61*, 172–180. https://doi.org/10.1016/j.lindif.2017.12.002

Money, J., Alexander, D., & Walker, H. T. (1965). *A standardized road-map test of direction sense: Manual*. Johns Hopkins Press.

Neal, Z. (2014). The devil is in the details: Differences in air traffic networks by scale, species, and season. *Social Networks*, *38*(1), 63–73. https://doi.org/10.1016/j.socnet.2014.03.003

Neal, Z. (2017). How small is it? Comparing indices of small worldliness. *Network Science*, *5*(1), 30–44. https://doi.org/10.1017/nws.2017.5

Nesbitt, K. T., Fuhs, M. W., & Farran, D. C. (2019). Stability and instability in the co-development of mathematics, executive function skills, and visual-motor integration from prekindergarten to first grade. *Early Childhood Research Quarterly*, *46*, 262–274. https://doi.org/10.1016/j.ecresq.2018.02.003

Ontario Ministry of Education. (2005). *The Ontario curriculum grades 1–8: Mathematics*. https://www.edu.gov.on.ca

Opgen-Rhein, R., & Strimmer, K. (2007). Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. *Statistical Applications in Genetics and Molecular Biology*, *6*(1), Article e9. https://doi.org/10.2202/1544-6115.1252

Presmeg, N. (2006). Research on visualization in learning and teaching mathematics. In A. Gutiérrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education* (pp. 205–235). Brill. https://doi.org/10.1163/9789087901127_009

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology*, *91*(2), 137–157. https://doi.org/10.1016/j.jecp.2005.01.004

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*(2), 274–278. https://doi.org/10.1037/h0028573

Reuhkala, M. (2001). Mathematical skills in ninth-graders: Relationship with visuo-spatial abilities and working memory. *Educational Psychology*, *21*(4), 387–399. https://doi.org/10.1080/01443410120090786

Robert, N. D., & Lefevre, J.-A. (2013). Ending up with less: The role of working memory in solving simple subtraction problems with positive and negative answers British Society for Research into Learning Mathematics. *Research in Mathematics Education*, *15*(2), 165–176. https://doi.org/10.1080/14794802.2013.797748

Rohde, T. E. (2008). *An examination of how visual perception abilities influence mathematics achievement*. Case Western Reserve University. https://rave.ohiolink.edu/etdc/view?acc_num=case1196193538

Rotzer, S., Loenneker, T., Kucian, K., Martin, E., Klaver, P., & von Aster, M. (2009). Dysfunctional neural network of spatial working memory contributes to developmental dyscalculia. *Neuropsychologia*, *47*(13), 2859–2865. https://doi.org/10.1016/j.neuropsychologia.2009.06.009

Schäfer, J., & Strimmer, K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, *4*(1), Article e32. https://doi.org/10.2202/1544-6115.1175

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

Sewell, R. D. (2008). *Cognitive and behavioral correlates in mathematical ability subtypes*. https://ocul-uo.primo.exlibrisgroup.com/discovery/openurl?institution=01OCUL_UO&vid=01OCUL_UO:UO_DEFAULT&%3Fsid=OVID:psycdb&id=pmid:&id=doi:&issn=0419-4209&isbn=&volume=69&issue=6-A&spage=2193&pages=2193&date=2008&title=DissertationAbstractsInternationa

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Sherman, J. A. (1980). Predicting mathematics grades of high school girls and boys: A further study. *Contemporary Educational Psychology*, *5*(3), 249–255. https://doi.org/10.1016/0361-476X(80)90048-X

Sokolowski, H. M., Hawes, Z., & Lyons, I. M. (2019). What explains sex differences in math anxiety? A closer look at the role of spatial processing. *Cognition*, *182*, 193–212. https://doi.org/10.1016/j.cognition.2018.10.005

Sortor, J. M., & Kulp, M. T. (2003). Are the results of the Beery-Buktenica Developmental Test of Visual-Motor Integration and its subtests related to achievement test scores? *Optometry and Vision Science*, *80*(11), 758–763. https://doi.org/10.1097/00006324-200311000-00013

Telesford, Q. K., Joyce, K. E., Hayasaka, S., Burdette, J. H., & Laurienti, P. J. (2011). The ubiquity of small-world networks. *Brain Connectivity*, *1*(5), 367–375. https://doi.org/10.1089/brain.2011.0038

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

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 200–239). Erlbaum.

Trbovich, P. L., & LeFevre, J. A. (2003). Phonological and visual working memory in mental addition. *Memory & Cognition*, *31*(5), 738–745. https://doi.org/10.3758/BF03196112

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

Uzzi, B., & Spiro, J. (2005). Collaboration and creativity: The small world problem. *American Journal of Sociology*, *111*(2), 447–504. https://doi.org/10.1086/432782

van Tetering, M., van der Donk, M., de Groot, R. H. M., & Jolles, J. (2019). Sex differences in the performance of 7–12 year olds on a mental rotation task and the relation with arithmetic performance. *Frontiers in Psychology*, *10*, Article 107. https://doi.org/10.3389/fpsyg.2019.00107

Wagner, A., & Fell, D. A. (2001). The small world inside large metabolic networks. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1478), 1803–1810. https://doi.org/10.1098/rspb.2001.1711

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

Wang, J., Zuo, X., & He, Y. (2010). Graph-based network analysis of resting-state functional MRI. *Frontiers in Systems Neuroscience*, *4*, Article 16. https://doi.org/10.3389/fnsys.2010.00016

Wang, L., & Carr, M. (2020). Understanding the relationship between spatial ability and mathematics achievement through the lens of information processing theory. *North American Journal of Psychology*, *22*(2), 173–192. https://researchgate.net/publication/339471900_Understanding_the_

Relationship_between_Spatial_Ability_and_Mathematics_Achievement_through_the_Lens_of_Information_Processing_Theory

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology*, *99*(2), 397–420. https://doi.org/10.1037/0022-0663.99.2.397

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. The Psychological Corporation.

Weckbacher, L. M., & Okamoto, Y. (2014). Mental rotation ability in relation to self-perceptions of high school geometry. *Learning and Individual Differences*, *30*, 58–63. https://doi.org/10.1016/j.lindif.2013.10.007

Received May 30, 2023
Revision received January 12, 2024
Accepted February 5, 2024 ■