Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

INSTRUCTIONS

- Before beginning submission process, download this PDF coversheet if you will need to provide information not on the PDF.
- Fill in all fields—information in this form **must match** the information on the submitted PDF and add missing information.
- Attach completed coversheet to the PDF you will upload to ERIC [use Adobe Acrobat or other program to combine PDF files]—do not upload the coversheet as a separate document.
- Begin completing submission form at https://eric.ed.gov/submit/ and upload the full-text PDF with attached coversheet when indicated. Your full-text PDF will display in ERIC after the 12-month embargo period.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

Scaling up spatial development: A closer look at children's scaling ability and its relation to number knowledge

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
Jirout, J. J.	University of Virginia	
Holmes, C. A.	Trinity College Dublin	
Ramsook, K. A.	Penn State University	
Newcombe, N.	Temple University	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) 2018

Check type of content being submitted and complete one of the following in the box below:

- ☑ If article: Name of journal, volume, and issue number if available
- ☐ If paper: Name of conference, date of conference, and place of conference
- ☐ If book chapter: Title of book, page range, publisher name and location
- ☐ If book: Publisher name and location
- ☐ If dissertation: Name of institution, type of degree, and department granting degree

Mind Brain and Education, 12

DOI or URL to published work (if available)

Acknowledgement of Funding— Grantees should check with their grant officer for the preferred wording to acknowledge funding. If the grant officer does not have a preference, grantees can use this suggested wording (adjust wording if multiple grants are to be acknowledged). Fill in Department of Education funding office, grant number, and name of grant recipient institution or organization.

"This work was supported by U.S. Department of Education **[Office name]** Institute of Education Scie through **[Grant number]** R305B1500 to **Institution]** The Pennsylva.The opinions expressed are those of the authors and do not represent views of the **[Office name]** Institute of Education Scie or the U.S. Department of Education.

Jirout Jamie (Orcid ID: 0000-0003-4832-930X)

Scaling up spatial development: A closer look at children's scaling ability and its relation to number knowledge

Jamie J. Jirout[1]

Corinne A. Holmes[2]

Kizzann Ashana Ramsook[3]

Nora S. Newcombe[4]

[1] University of Virginia, [2] Trinity College Dublin, [3] Penn State University, [4]Temple University

Author Note

Work on this project was funded as part of the Spatial Intelligence and Learning Center grant from the National Science Foundation, SBE-1041707. Special thanks to Paula Yust and Lindsey Hildebrand for their help with data collection, and the staff and students at Miquon

School for their participation. Correspondence concerning this article should be addressed to Jamie Jirout, Curry School of Education, University of Virginia, 405 Emmet Street South, VA, 22903, Jirout@virginia.edu, 434-284-0835.

Abstract

Spatial skills are consistently linked to mathematical reasoning, and are sensitive to intervention. One important spatial skill is spatial scaling. We evaluated whether (1) a playful scaling game might promote learning by providing feedback during play, and (2) spatial scaling is related to number-line estimation based on the mutual reliance on relative magnitude reasoning. Forty-eight children, ages 5.5-8.3, completed a playful scaling game and a number line estimation task. Results show that children improve from the first to second half of the task, especially for more difficult trials and for the lowest-performing children. In addition, scaling and number-line estimation relate when controlling for age and vocabulary. Similar improvement on the task and relations to number-line estimation were observed in a conceptual replication (N=52). These results provide support for further study of improving spatial scaling in children, with the possibility to test whether scaling could support related mathematical skills as well.

Keywords: Spatial Scaling, Spatial Development, Number Knowledge, Space-Math Relation

Spatial thinking is central to many everyday activities, such as using a map, navigating to a destination, or fitting groceries into a car's trunk, as well as to achievement in the STEM fields (e.g., science, technology, engineering, and mathematics; Sorby, 2009; Wai, Lubinski, & Benbow, 2009). Importantly, spatial thinking is not fixed. Practice over time yields significant improvements, with sustained gains that generalize to untrained tasks (Uttal et al., 2013; Wright, Thompson, Ganis, Newcombe, & Kosslyn, 2008). Because spatial thinking and math knowledge appear deeply intertwined (e.g., Cheng & Mix, 2014), early spatial interventions have the potential to boost spatial *and* mathematical cognition (Verdine, Golinkoff, Hirsh-Pasek, & Newcombe, 2014). However, the web of interelations observed across different types of spatial and mathematical skills suggests the need for clearer understanding of how and why specific spatial and mathematical processes are related, both from a theoretical point of view and to determine the most effective means of promoting development.

Relative magnitude understanding is one possible link between spatial and mathematical domains (Newcombe, Levine, & Mix, 2015; Jirout & Newcombe, in press). There is evidence from the earliest points in cognitive development that children associate space and number, and that experience influences this association (McCrink & de Hevia, 2018). Spatial scaling, the focus of the current study, involves reasoning about different-sized, but proportionally equivalent, visuospatial representations. For example, when navigating with a map, one must reason about the distance (magnitude) from the current to goal locations on the map relative to the actual distance using the map scale. Children use spatial scaling during play, for example, when using an image of a puzzle to estimate where individual puzzle pieces belong. Individual pieces can be matched to the smaller image, but then scaling must be used to estimate where they belong in the larger puzzle space. Although this can be done using spatial estimation, more symbolic proportional information likely becomes useful (e.g., moving from thinking "the same distance from each corner" to "halfway between the corners"). Several mathematical skills also include relative magnitude understanding, such as number-line estimation and magnitude comparisons, and processes such as proportional reasoning and fraction understanding rely on relative magnitude information (see Newcombe, Frick, & Möhring, in press; Jirout & Newcombe, in press, for more thorough discussion). Consistent with this overlap in relative magnitude reasoning, spatial scale of stimuli impacts performance on proportional reasoning tasks (Boyer & Levine, 2012), suggesting that improving spatial scaling in children might have beneficial effects on spatial and mathematical development. However, no studies have investigated improving spatial scaling. The current study sought to investigate the relation between spatial scaling and a relative magnitude math task (number line estimation), and to investigate whether spatial scaling might improve with practice and feedback.

Children's Use of Scaled Representations

Scaled representations are ubiquitous in our world – besides using maps to navigate, we use pictorial instructions for assembling things, diagrams and charts for understanding data or processes, scaled images to show things otherwise too small or too large to be visible, etc. Children also use scaled representations, such as instructions in a Lego kit, playing Pokemon

Go!, or using an image to assemble a puzzle. Although children understand the nature of simple representations from the age of three (DeLoache, 2002, 2004), the ability to *use* them in spatially accurate ways develops considerably over the subsequent years. For example, when using maps scaled along a single dimension, 3-year-olds can successfully locate a target (Huttenlocher, Newcombe & Vasilyeva, 1999). However, scaling along two dimensions is more difficult, with performance affected by the magnitude of the scale factor (i.e., the difference between the size of the map and the size of the space it represents) and developing considerably from preschool into kindergarten (Vasilyeva & Huttenlocher, 2004). Experience using scaled representations likely supports this development, perhaps by promoting attention to relational structures (Davies & Uttal, 2007; Gentner, 1988).

Frick and Newcombe (2012) created a measure of spatial scaling to assess development of scaled map use. Children used a map showing where an “egg” was hidden in a field, and were asked to show the same location on a referent field, which was a larger rectangle otherwise identical to the map (except without the egg). Accuracy in placing the marker at the correct location was higher for 1:2 than 1:4 scaled maps, and significantly increased from three to six years, at which point children’s performance did not differ from that of adults (Frick & Newcombe, 2012). Using a similar task, Möhring and colleagues presented both children and adults with trials ranging in scale, including 1:1, 1:1.14, 1:1.3, 1:1.6, 1:2, 1:2.6, and 1:4 scaling factors. They tested the theory that if scaling involved mental transformations, the increasing size difference between map and referent space would require greater transformations, increasing response time and decreasing accuracy, which is what results showed (Möhring, Newcombe, & Frick, 2014). It is unknown if this table-top map uses similar processes as more typical map use. Maps and scaled diagrams are often not visually identical to the space they represent, giving more symbolic relational information without much visual similarity, and the scale factors are typicall much larger than 1:4 (Liben & Downs, 1989). Thus, more research is needed to understand the impact of map-space similarity and the development of chlidren's scaling with a wider range of scale factors.

Though prior research shows children's capability in using scaled representations, less is known about how children learn to to do this. For example, it could be that practice using maps to represent space could help develop this skill in children, but should these maps be simiplistic, like the measures described above, or more naturalistic, which requires more abstract reasoning to use them? Prior research suggests that, while visual similarity simplifies the task and intuitively seems easier, maps that appear less similar to the space they represent facilitate the understanding of the map as a representation rather than a separate object (Gentner, Ratterman, Markman, Kotovsky, 1995).

Naturalistic experience with representations may provide spatial learning, such as using scaled images like a picture to guide jigsaw puzzle completion or building a Lego set from instructions, because children receive feedback when pieces don't fit. Jirout and Newcombe (2014) found that spatial scaling related to children's overall performance on mazes with scaled maps and their frequency of map use (i.e., children with higher scaling ability used maps more), and children's frequency of map use positively related to children's maze performance. More generally, toys such as blocks and puzzles relate to spatial thinking (Jirout & Newcombe, 2015; Verdine et al., 2014; Fisher, Hirsh-Pasek, Newcombe, & Golinkoff, 2013; Casey et al., 2008; Levine, Ratliff, Huttenlocher, & Cannon, 2012) and playful activities show potential for improving spatial skills (e.g., Weisberg, Kittredge, Hirsh-Pasek, Golinkoff, & Klahr, 2015). This latter point is in need of empirical testing to assess whether spatial play does, in fact, lead to improved spatial skills, beginning with developing effective spatial play activities. Our aim was to apply this research to a playful spatial activity, a spatial scaling search game, that might improve spatial scaling. More specifically, we examine whether children improve across a hide-and-seek game that provides implicit feedback during spatial play, i.e., success or failure in finding a hidden object using a small-scacled map.

Number Line Estimation

Number line estimation (NLE) is commonly used in research on children's developing representations of number (Siegler & Opfer, 2003), with research suggesting that children's representations become more linear with age, shifting away from logarithmic representations (e.g., Siegler, Thompson, & Opfer, 2009). The physical number line used in the NLE task is a spatial representation of a numberical scale, used to assess chidren's mental representation of numbers by their placement of numerical locations. A linear representation is the true representation of whole numbers, and requires estimating the location of a number (magnitude) based on its relative position to the overall scale (i.e., 10 is the middle of a 1-20 scale, but 1/10th of the distance of the line for a 1-100 scale).

Recent research suggests that children's NLE performance might be supported by their use of spatial positions of numbers, such as dividing the line into categories of equal size by using mental "benchmarks" (i.e., using the midpoint; Slusser, Santiago, & Barth, 2013; Peeters, Verschaffel, & Luwel, 2017), similar to children's use of spatial categories to aid spatial memory (e.g., halving space; Huttenlocher et al., 1999; Huttenlocher, Newcombe, & Sandberg, 1994). For example, 4- to 6-year-old children use one or both endpoints as landmarks, and 7- to 10-year-old children also use the mid-point (Slusser et al., 2013), and even adults benefit when benchmarks are provided on NLE tasks (Peeters, Verschaffel, & Luwel, 2017). While evidence for the use of proportional reasoning is inconcsistent, several studies have shown that children consider the midpoint of the number line when completing estimations (Schneider et al., 2008; Ashcroft & Moore, 2012), and children's proportional reasoning ability relates to NLE performance, (Möhring, Frick, and Newcombe, 2018). However, there is debate around children's development of linear representations, such as whether children shift from logarithmic to linear representations on increasing scales vs. develop more advanced use of proportional reasoning strategies (see Opfer, Thompson, & Kim, 2016; Opfer, Siegler, & Young, 2011; and Slusser, Santiago, & Barth, 2011 as examples). In this initial examination of a specific space-math relation with scaling, we explore the relation of scaling performance on a spatial scaling game to children's linearity and error in NLE. Establishing this relation could support future research exploring how to support children's use of spatial categorical thinking to promote the use of spatial categories on mathematical tasks like NLE, and perhaps contribute to understanding of children's developing knowledge of number. There is some evidence of NLE relating to spatial estimation. Thompson and colleagues asked children to estimate where different pages fell within a book 100 un-numbered pages, and their error related to NLE error, suggesting a relation between NLE and spatial magnitude reasoning (Thompson, Morris, & Sidney, 2017). Further, NLE relates to scaling performance specifically, even when controlling for proportional reasoning ability (Möhring, Frick, and Newcombe, 2018), and spatial information can influence children's attention to symbolic numbers (Opfer, Thompson, & Furlong, 2010). Building on these results, the current study tests for a similar relation between number line estimation and spatial scaling performance on our novel playful scaling game.

Spatial Scaling and Mathematical Reasoning

The development of a spatial scaling search game that relates to NLE performance would allow future studies to further explore the direction of the space-math link. Spatial skills are important for math learning, and research consistently shows a relation between spatial and mathematical reasoning (Mix & Cheng, 2012). Reasoning about relative spatial locations - a process critical for using scaled representations - is especially important for reasoning about number and magnitude (Jirout & Newcombe, in press). For example, relations are important when considering the difference between two vs. three and two vs. ten, or what makes a number "big" or "small". Thus magnitude understanding, comparison, and relational reasoning plausibly lay the foundation for later math learning – for example, understanding the meaning of fractions and proportions, which are essentially relations between numbers. Because spatial scaling specifically is shown to be important for relative magnitude-related mathematics like proportional reasoning, fraction learning, and understanding algebra (Newcombe et al., 2015; Möhring, Newcombe, Levine, & Frick, 2016a; Möhring, Newcombe, Levine, & Frick, 2016b), improvements in spatial scaling could provide related support for learning these mathematical skills as well. The first step in exploring this is to test for a relation between spatial scaling and a magnitude reasoning task, in the current study a numbr line estimation task.

Aims and Hypotheses

The current study explored two research questions. First, we examined whether scaling accuracy improves with experience on a scaling game by comparing the first half to the second half of the task, with the expectation that performance would improve across trials of the task. Second, we tested whether spatial scaling during the game relates to performance on the number line estimation task.

Methods

Participants

Participants included forty-eight children, ages 5.5-8.3 (*M*=6.90 years, *SD*=0.82; 53% male), from three classrooms at a suburban independent private school serving predominantly middle/upper-middle class families. All children in kindergarten through 2nd grade at the school were recruited; those whose parents provided consent and children gave assent included 19 children from the kindergarten class and 29 children from two combined 1st/2nd grade classes. The school reports 31% students of color and 45% students receiving financial aid. One kindergarten female stopped early and was excluded from analyses. A university IRB and the participating school approved the study, and parental consent and child assent was collected. All children received a small toy for participation.

Materials and Procedure

Children were tested individually in a sectioned-off area of the classroom. One experimenter administered the floor-based scaling game followed by a second experimenter administering the NLE task and the picture vocabulary task, in that order.

Spatial scaling search game. Our goal was to develop a spatial scaling game that was playful and provided experience and feedback on using scaled repesentations. We based the game on measures from prior studies of spatial scaling, using a task in which children would use a map to get information about a referent space. In those studies, children saw a map and referent space that were identical except for size and a marked location on the map, with the task presented on a tabletop. Children would then use the map to mark the same spot on the referent space, which was later measured for the distance between the child's mark and the correct location. Rather than marking a space, our game asked children to use the map to identify a hidden star in a large floor-sized search space. The resulting scaling game was a "hide-and-seek" activity. Children used scaled maps to locate a target within a grid of flaps, with feedback received through search, i.e., finding a star or not.

Twenty-four square pieces of laminated colored paper (white, light blue, and dark blue), were mounted onto posterboard to make a 6x4 unit (72"x48") grid – i.e., the search space (see Figure 1). Flaps could be lifted using yellow ribbon attached to the bottom of each flap. Maps contained a black rectangle outline on white paper representing the search space and a star within each rectangle indicated under which flap a 5" star was hidden. Of the 15 total maps, three practice maps were scaled at 1:8 (9 x 6 in.), six were scaled at 1:12 (6 x 4 in.), and six were scaled at 1:48 (1.5 x 1 in.). To allow us to look at improvement from the first half to second half of the task, we included equal proportions of trials for each scale factor in each half (i.e., three trials of each scale factor in each half), with the order of the items within the first and second sets of trials presented in a random order.

[Figure 1 Near Here]

Children first completed three practice trials with locations found to be easiest in pilot testing (one corner target locations and two target locations along the border). Next, children completed 12 test trials, six at each scale factor (1:12 and 1:48), with mirror-matched target locations for the scale factors.

On each trial, children were asked to turn around and close their eyes while the experimenter hid the star under a flap. The child was then handed a map and instructed to "use the map to find where the star is hidden" by lifting the flap that matched the location shown on the map. The experimenter ensured correct orientation of the map and reminded the child to use it (e.g., "Use your map to find the star," or "Don't forget to use your map") for every search attempt. When children found the star or they were shown it after six failed attempts, they handed it back to the experimenter and the procedure was repeated. After testing, children were asked about their map use, however, it was determined that experimenter error resulted in inconsistent questioning, so responses are not analyzed. The task took approximately 12 minutes to complete.

Although the task was modelled off of prior scaling measures, several design decisions were made to address the goal of creating a potentially educational activity. These include feedback provided through play to support learning, a naturalistic playful context, and more similarity with "real" map use, including larger search space with more abstract maps at more extreme scales. Specific differences from past scaling measures included: 1) increased scale factors (1:12 and 1:48 from 1:4), by increasing the referent space to a 4'x6' rectangle; 2) *retrieving* hidden targets rather than indicating a target's location, which provided immediate feedback, supporting motivation and learning; 3) a visually discrete referent space, rather than being a singular, continuous entity (see Figure 1), and 4) plain black/white outlined maps. A similar table-top version of the scaling game related to an existing measure of spatial scaling in pilot tests ($r_{(41)} = .306$, $p = .046$, controlling for age).

Number line estimation task (NL). Thirteen estimation trials were presented as consecutive pages in a binder. A solid, black 25 cm horizontal number line with anchors labeled as 0 and 100 was centered on each page, with the target number printed above the center of the line (Siegler & Opfer, 2003). The anchors were half-inch verticle lines on the two ends of the horizontal lines, with no other marks or hashes on the horizontal line. Two target numbers were randomly selected from each of the first three decades and one target number randomly selected from each the remaining decades. Children were told that all of the numbers from 0 to 100 go somewhere on the line, with each number having its own place on the line, and every spot on the line having a different number that goes there. They were then reminded of the anchors for each trial (e.g., "If zero goes here, and 100 goes here, where does 20 go?"). Children used a small rubber peg to indicate their response, and the experimenter recorded their responses before flipping the page to begin the next trial (adapted from the scaling task in Frick & Newcombe, 2012). Only one response was recorded on each line, and children could not look back at prior responses. If the child placed the peg more than an inch from the number line, the experimenter asked the child to place the peg on the number line.

Picture vocabulary task. Children's vocabulary was assessed as a control for general intelligence using the Woodcock-Johnson III-R Picture Vocabulary Test (Woodcock, McGrew, & Mather, 2001). Standardized procedures were followed, in which children were asked to verbally identify pictures (e.g., "What is this called? What kind of insect is this?") and completed trials until they gave incorrect responses – i.e., incorrect label or no response – for all items on a page. Standardized scores were calculated using total correct answers. The number-line and vocabulary tasks together took approximately 8 minutes to complete.

Data Coding and Design

A mixed-methods design controled for vocabulary, with age (i.e., in months) and sex as between-subject variables, and map scale (1:12 and 1:48) and test time (first- or second-half) as within-subject variables. For interactions with age, we included continuous age in all statistical analyses, but present the graphs of results using age groupings to support interpretation of age patterns, including six year-olds (n=19; M_{age} = 73.2 months, age range = 67-77 months), seven year-olds (n=15; M_{age} = 83.2 months, age range = 77-89 months) and eight year-olds (n=14, M_{age} = 95.4, age range=90-100 months).

The scaling game was scored by proportion of perfect searches, trials on which the child chose the correct location on the first attempt[1]. To calculate performance on the NL task, each child's responses were fit to a linear model and percent of absolute error was calculated (PAE; average percent of error between responses and correct location; Siegler & Opfer, 2003). Vocabulary scores were calculated using standardized scoring tables (Woodcock et al., 2001).

Results

We present a descriptive summary of children's performance on the scaling search game and then provide tests of scaling game improvement, followed by tests of the relation between scaling and number line estimation. All analyses control for standardized vocabulary.

[1] After the initial search, children often began lifting adjacent flaps without looking back at the map, and would sometimes accidentally see the star under different flaps with adjacent searches, so first search accuracy seemed the best indicator of performance.

Performance on the Scaling Search Game

Children located the target on the first attempt for 62% of trials. A repeated-measures ANCOVA was run with scale factor as the within-subjects factor, sex as a between-subject factor, and controlling for age and vocabulary as covariates. The ANCOVA showed main effects of both age and sex ($F_{age(1,43)}$=26.41, p<.001, · p^2=.38; $F_{sex(1,43)}$=9.31, p=.004, · p^2=.18; p_{scale}=.22, p_{vocab}=.84). Males (M=.69, SE=.04) significantly outperformed females (M=.53, SE=.04), and age significantly correlated with performance (r_{47}=.590, p<.001). Although the 1:12 scaled trials showed better performance than the 1:48 scaled trials, this difference was not significant ($M_{1:12}$ = .63, $SE_{1:12}$=.03; $M_{1:48}$ = .59, $SE_{1:48}$ = .03; $F_{(1,43)}$=1.58; p = .215, · p^2= .04). Vocabulary was not a significant covariate ($F_{(1,43)}$ = 0.04, p = .846, · p^2= .001).

In addition to the age and sex main effects, the age X scale factor interaction was significant ($F_{(1,43)}$=4.59, p=.038, · p^2=.10). To investigate this interaction, we re-ran the ANCOVAs separately for the two scale factors. While age was still significant for both scale factors, difference in performance across age was greater when maps were scaled 1:48 than 1:12, with an effect size of age more than double for the 1:48 scaled items (1 to 12: $F_{(1,43)}$=9.60, p=.003, · p^2=.183; 1 to 48: $F_{(1,43)}$=31.69, p<.001, · p^2=.424). Figure 2 shows pairwise comparisons from the scale factor univariate ANCOVAs broken up by age group to show the direction of the trends.

[Figure 2 Here]

Improvement Across Search Game

Average performance on the first- to second-six trials was compared to assess improvement using a repeated-measures ANCOVA. We again observed main effects of age ($F_{(1,43)}$=26.59, *p*<.001, $\cdot\ p^2$=.38) and sex ($F_{(1,43)}$=9.57, *p*=.003, $\cdot\ p^2$=.18), with main effect of time not significant ($F_{(1,43)}$=0.84, *p*=.364, $\cdot\ p^2$=.02; M_{time1}=.57, *SE* = .03, M_{time2} =.65, *SE*=.04). However, significant interactions with time (first vs. second half) were observed for both age and sex (age: $F_{(1,43)}$=5.40, *p*=.025, $\cdot\ p^2$=.11, sex: $F_{(1,43)}$=6.39, *p*=.015, $\cdot\ p^2$=.13) (see Figures 3 & 4).

To explore these interactions, we re-ran the ANCOVAS by time. For the first six trials, the effects of both age and sex remained significant (age: $F_{(1,43)}$=37.82, *p*<.001, $\cdot\ p^2$=.468; sex: $F_{(1,43)}$=19.66, *p*<.001, $\cdot\ p^2$=.314). For the second six trials, only the effect of age remained significant, and the effect size was much smaller than for the first six trials (age: $F_{(1,43)}$=6.49, *p*=.014, $\cdot\ p^2$=.131; sex: $F_{(1,43)}$=0.834, *p*=.366, $\cdot\ p^2$=.019). To explore whether males and females showed significant change from the first to second set of trials, we ran our ANCOVA separately by sex. For males, there was only a significant time X vocabulary interaction ($F_{(1,23)}$=4.32, *p*=.049, $\cdot\ p^2$=.16; $F_{(1,23)}$=2.60, *p*=.120, $\cdot\ p^2$=.10; M_{time1}= .70, *SE* = .04; M_{time2}=.69, *SE* =.04). For females, the effect of time was significant ($F_{(1,18)}$=4.83, *p*=.048, $\cdot\ p^2$=.20; M_{time1}= .43, *SE* = .04; M_{time2}=.61, *SE*=.07), indicating significant improvement. Additionally, the time X age interaction was significant for females ($F_{(1,18)}$=8.47, *p*=.009, $\cdot\ p^2$=.32). Figure 3 shows mean performance for age by time, and Figure 4 shows mean performance for sex by time.

[Figure 3 Near Here]

[Figure 4 Near Here]

Number Line Estimation and Scaling Search Game

Children's percent absolute error (PAE) on the number-line task was 16.54% (SE=1.18), and linearity of estimations (using R^2) was .72 (SE=.03). ANCOVAs testing for effects of sex and age showed a main effect of age for estimation error ($F_{(1,42)}$=21.09, p<.001, $\cdot\, p^2$=.33) and linearity of estimations ($F_{(1,43)}$=19.63, p<.001, $\cdot\, p^2$=.31).

Partial correlations (controlling for age and standardized vocabulary) between children's number line estimations and performance on the scaling game were conducted. There was a significant relation between scaling performance and PAE ($r_{(43)}$=-.311, p=.037), but not between scaling performance and linearity ($r_{(43)}$=.179, p=.238).

Conceptual Replication

To attempt to replicate the results observed, we analyzed data from a similar spatial scaling game included in a larger, ongoing study[2]. The procedure for the scaling search game was similar, though participation was in a lab setting in two geographic locations, and the search

[2] All data are available on databrary.org, by request.

game was kept out of view until after children completed several other measures, of which we include number line estimation.

Participants

Participants for the conceptual replication included fifty-two children, ages 5.5-8.5 (*M*=6.98 years, *SD*=1.02; 62% female). Participants were sampled from two psychology lab participant databases, one from a suburb outside of a large Northeastern city (N=37), the other from a mid-size city in the Midsouth (N=15).

Materials and Procedure

The total search space was decreased to 36" x 48", but the size of the discrete search spaces was decreased to 6" squares, resulting in the number of squares increasing to 48 (8 squares horizontally by 6 squares vertically). Map sizes were scaled 1:6 (8" x 6") and 1:24 (2" x 1.5"). Number line estimation was again included, but was administered before the scaling game. Vocabulary was only assessed for the Midsouth sample, and so is not included in analyses.

Performance on the Scaling Search Game

Children located the target on the first attempt for 52.7% (SE=3.0) of trials. Compared to performance with the original scaling game, this was significantly lower ($t_{(97)} = 1.955$, $p = .046$, $d = .39$). A repeated-measures ANCOVA was again run with scale factor as the within-subjects factor, sex as a between-subject factor, and controlling for age a covariate. Only the effect of age was significant ($F_{age(1,49)} = 9.525$, $p = .003$, $\cdot\, p^2=.16$; $F_{sex(1,49)}=0.657$, $p = .42$, $\cdot\, p^2=.01$; $p_{scale}=.54$,

· p^2=.01). Age significantly correlated with performance (r_{52}=.412, p=.002). There were no significant interactions.

Improvement Across Search Game

We again compared average performance on the first- to second-six trials using a repeated-measures ANCOVA. A significant effect of age was observed ($F_{(1,49)}$ = 9.525, p=.003, · p^2=.163). The main effect of time was not significant (p=.617, · p^2=.004, M_{time1}=.46, SE = .03, M_{time2} =.61, SE=.04), but was again qualified by a significant time X gender interaction ($F_{(1,49)}$= 5.625, p=.022, · p^2=.103). We re-ran the ANCOVAS by time. For the first six trials, the effects of sex was now significant ($F_{(1,49)}$=5.469, p = .011, · p^2=.125; M_{Female} = .382, SE = .04, M_{Male} =.539, SE=.05). The effect of sex was not significant for the second six trials ($F_{(1,49)}$ = 0.664, p=.419, · p^2=.013; M_{Female} = .637, SE = .05, M_{Male} =.572, SE=.06).

Number Line Estimation and Scaling Search Game

Children's percent absolute error (PAE) on the number-line task was 13.5% (SE=1.0%), and linearity of estimations (using R^2) was .70 (SE=.04). A multivariate ANCOVA testing for effects of sex and age on NLE showed a main effect of age for estimation error ($F_{(1,49)}$= 62.499, p <.001, · p^2=.561) and linearity of estimations ($F_{(1,49)}$ =47.376, p < .001, · p^2=.492).

Partial correlations (controlling for age) between children's number line estimations and performance on the scaling game were conducted. There was a significant relation between scaling performance and both PAE ($r_{(49)}$=-.469, p=.001), and between scaling performance and linearity ($r_{(49)}$=.394, p=.004).

Discussion

To explore the potential of the scaling game as a spatial-relational intervention, we assessed children's improvement on the spatial scaling game with experience. Although children were already quite successful on the task, performance seemed to improve with experience. There is indication that girls, and possibly younger children, showed greatest improvement, though they also had the most *opportunity* for improvement. Although some indication of sex effects were observed, sex difference were only seen in the first half of the game and were no longer significant by the second half of the game. These results are limited by a short duration – it isn't clear whether children improved from learning or if the practice simply helped them adjust to the task. The task was relatively simple to modify to adjust the level of difficulty, which we did successfully in our replication, so it is possible to create versions appropriate for different age and ability levels. Despite the limitations, the improvement observed shows promise for a play-based spatial scaling game that could have potential for developing children's spatial scaling ability and should be studied futher, including with a broader range of populations represented and with a larger sample size.

Consistent with study hypotheses, results showed that children's performance on a scaling game related to NLE ability, perhaps because both tasks may rely on similar processes relating to relative magnitude (Jirout & Newcombe, in press), expanding prior research showing relations between spatial scaling and both proportional reasoning skills and NLE, and between number-line estimation and mental transformation (Möhring et al., 2016a; Gunderson et al., 2012;

Möhring, Frick, and Newcombe, 2018). Recent research shows that spatial scaling relates to NLE even when controlling for proportional reasoning skill (Möhring, Frick, and Newcombe, 2018), suggesting other related processes involved, for example a more general spatial-relational reasoning. However, this association was mixed across tasks; with the easier scaling game, performance was only related to NLE error, but both PAE and linearity related to the more difficult scaling game in the replication test. Error may have shown more consistent relations due to the little variability at the bottom and, especially, top ends of our age distribution – the children on the lower end of our age distribution have very logarithmic representations, while children on the higher end are strongly linear, consistent with prior findings (Siegler & Booth, 2004).

What does this spatial-relational thinking look like? Learning to use spatial categories can include using directional terms like top and bottom, right and left, or big and small. These terms are relational, in that "big" or "top" are relative. Spatial scaling tasks involve encoding of relative distance, in which a target's location is coded in proportion to one representation and then applied to another (Huttenlocher, Newcombe, & Vasilyeva, 1999). This type of relational thinking is also used for mathematical thinking and learning. For example, children often use manipulatives or visual spatial diagrams of proportions and fractions, which help to develop abstract understanding of relative magnitude by directing attention to the physical features (e.g, size) (Mix, 2010). It is possible that this type of relative magnitude thinking relates to the spatial-relational thinking engaged with the spatial task in this study, or similar cognitive proceses that can explain the link between spatial and mathematical tasks. The spatial scaling game could provide playful practice using methods such as focusing on proportional landmarks like the midpoint. Experience with the scaling game may also provide practice using similar reasoning as using mathematical manipulatives found be beneficial, such as practice using maps as spatial-relational models, embodiment (i.e., moving through space) as a way of understanding spatial relations, and directing attention to the physical space and representation. Having a way to improve spatial-relational thinking would allow the directionality of this relation to be explored in future studies.

Prior work shows that non-symbolic interventions can facilitate the development of numeric symbolic knowledge and reasoning (e.g., spatial training on a mental rotation task yielded significant improvements on calculation problems: Cheng & Mix, 2014; and frequent experience with visuospatial toys improved spatial and mathematical knowledge: Grissmer et al., 2013). Applied to the current study, the link between scaling performance and number line estimation suggests that the scaling game has the potential to support a specific mathematical skill – relative magnitude reasoning, or estimating the relation between non-integer representations (Möhring et al., 2016a; also see Barth & Paladino, 2011), or that relative magnitude understanding could support more advanced spatial-relational reasoning, or that the relation is bidirectional. The next step in this work is to test the direction of this relation by assessing the effectiveness of learning from the scaling game on relative magnitude mathematical skills.

Conclusions

This work contributes to the body of literature testing specific relations between spatial and mathematical tasks. Specifically, this study documents associations between scaling and number-line estimation and provides support for the hypothesis that they both draw on underlying processes related to relative magnitude reasoning (Möhring et al., 2016a; Jirout & Newcombe, in press). Practice on the playful spatial scaling game can improve performance, at least for some children, through the use of feedback and an engaging context. Future work is needed to assess whether improved performance indicates learning and test for causal direction of the spatial and mathematicla relation.

Figure Captions

Figure 1. Spatial scaling search game.

Figure 2. Search game performance by scale factor across age groups.

Figure 3. Search game performance by age group at Time 1 and Time 2.

Figure 4. Search game performance by sex at Time 1 and Time 2.

References

Ashcraft, M. H., & Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *Journal of experimental child psychology*, *111*(2), 246-267.

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science*, *14*, 125-135.

Boyer, T. W., & Levine, S. C. (2012). Child proportional scaling: Is 1/3= 2/6= 3/9= 4/12? *Journal of Experimental Child Psychology*, *111*, 516-533.

Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block-building activities. *Cognition and Instruction*, *26*, 269-309.

Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of Cognition and Development*, *15*, 2-11.

Davies, C, & Uttal, D.H. (2007). Map use and the development of spatial cognition. In J. Plumert & J. Spencer (Eds.), *The Emerging Spatial Mind* (pp. 219-247). New York, NY: Oxford University Press.

Deloache, J. S. (2002). The symbol-mindedness of young children. *Child psychology in retrospect and prospect: In celebration of the 75th anniversary of the Institute of Child Development*, *32*, 73-101.

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences, 8,* 66-70.

Fisher, K. R., Hirsh-Pasek, K., Newcombe, N., & Golinkoff, R. M. (2013). Taking shape:

Supporting preschoolers' acquisition of geometric knowledge through guided play. *Child Development*, *84*, 1872-1878.

Frick, A., & Newcombe, N. S. (2012). Getting the big picture: Development of spatial scaling abilities. *Cognitive Development*, *27*, 270-282.

Gentner, D. (1998). Metaphor as structure mapping: The relational shift. *Child Development, 59*, 47-59.

Gentner, D., Rattermann, M. J., Markman, A. B., & Kotovsky, L. (1995). Two forces in the development of relational similarity. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 263-313). Hillsdale, NJ: LEA.

Grissmer, D., Mashburn, A., Cottone, E., Brock, L., Murrah, W., Blodgett, J., & Cameron, C. (2013, September). *The efficacy of minds in motion on children's development of executive function, visuo-spatial and math skills*. Paper presented at the Society for Research in Educational Effectiveness Conference, Washington, DC.

Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Developmental Psychology*, *48*, 1229-1241.

Huttenlocher, J., Newcombe, N., & Sandberg, E. H. (1994). The coding of spatial location in young children. *Cognitive Psychology*, *27*(2), 115-147.

Huttenlocher, J., Newcombe, N., & Vasilyeva, M. (1999). Spatial scaling in young children.

Psychological Science, *10*, 393-398.

Jirout, J. J., & Newcombe, N. S. (2014). Mazes and maps: Can young children find their way?. *Mind, Brain, and Education, 8*(2), 89-96.

Jirout, J. J., & Newcombe, N. S. (2015). Building blocks for developing spatial skills: Evidence from a large representative U.S. sample. *Psychological Science, 26*(3), 303-310.

Jirout, J. J., & Newcombe, N. S. (in press). Relative magnitude as a key idea in mathematics cognition. In K. S. Mix & M. T. Battista (Eds.) *Visualizing Mathematics: The Role of Spatial Reasoning in Mathematical Thought.* New York, NY: Springer.

Levine, S. C., Ratliff, K. R., Huttenlocher, J., & Cannon, J. (2012). Early puzzle play: A predictor of preschoolers' spatial transformation skill. *Developmental Psychology*, *48*, 530.

Liben, L. S., & Downs, R. M. (1989). Understanding maps as symbols: The development of map concepts in children. *Advances in child development and behavior, 22,* 145-201.

McCrink, K., & de Hevia, M. D. (2018). From Innate Spatial Biases to Enculturated Spatial Cognition: The Case of Spatial Associations in Number and Other Sequences. *Frontiers in psychology*, *9*, 415.

Mix, K. S., & Cheng, Y. L. (2012). Space and math: The developmental and educational implications. In J. Benson (Ed.), *Advances in Child Development and Behavior* (pp. 179–243). New York, NY: Elsevier.

Möhring, W., Frick, A. & Newcobe, N. S. (2018). Spatial scaling, proportional thinking, and numerical understanding in 5-to7-year-old children. *Cognitive Development, 45*, 57-67.

Möhring, W., Newcombe, N. S., & Frick, A. (2014). Zooming in on spatial scaling: Preschool children and adults use mental transformations to scale spaces. *Developmental Psychology, 50*(5), 1614-1619.

Möhring, W., Newcombe, N. S., Levine, S. C., & Frick, A. (2016a). The Relation between Spatial Thinking and Proportional Reasoning in Preschoolers. *Journal of Experimental Child Psychology, 132*, 213-220.

Möhring, W., Newcombe, N. S., Levine, S. C., & Frick, A. (2016b). Spatial proportional reasoning is associated with formal knowledge about fractions. *Journal of Cognition and Development*, *17*(1), 67-84.

Newcombe, N. S., Frick, A. & Möhring, W. (in press). How big is many? Development of spatial and numerical magnitude understanding.

Newcombe, N. S., Levine, S. C. & Mix, K. S. (2015). Thinking about quantity: The intertwined development of spatial and numerical cognition. *WIREs in Cognitive Science, 6*(6), 491-505.

Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The powers of noise-fitting: Reply to Barth and Paladino. *Developmental Science, 14*(5), 1194-1204.

Opfer, J. E., Thompson, C. A., & Furlong, E. E. (2010). Early development of spatial-numeric associations: evidence from spatial and quantitative performance of preschoolers. *Developmental Science*, *13*(5), 761-771.

Opfer, J. E., Thompson, C. A., & Kim, D. (2016). Free versus anchored numerical estimation: A unified approach. *Cognition*, *149*, 11-17.

Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., ... & Stern, E. (2008). A validation of eye movements as a measure of elementary school children's developing number sense. *Cognitive Development*, *23*(3), 409-422.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child development*, *75*(2), 428-444.

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

Siegler, R. S., Thompson, C. A. & Opfer, J. (2009). The Logarithmic-To-Linear Shift: One Learning Sequence, Many Tasks, Many Time Scales. *Mind, Brain, and Education*, *3*(3), 143-150.

Singer-Freeman, K. E. & Goswamani, U. (2001). Does half a pizza equal half a box of chocolates? Proportional matching in an analogy task. *Cognitive Development, 16*, 811-829.

Slusser, E., & Barth, H., (N.D.) An Excel worksheet for proportion-judgment analyses of number-line data. Retrieved from https://works.bepress.com/emily_slusser/4/ on 8/10/16.

Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General, 142*, 193-208.

Sorby, S. (2009). Developing spatial cognitive skills among middle school students. *Cognitive Processing*, *10*, 312-315.

Spinillo, A. G., & Bryant, P. (1991). Children's proportional judgments: The importance of "half". *Child Development*, *62*(3), 427-440.

Thompson, C. A., Morris, B. J., & Sidney, P. G. (2017). Are Books Like Number Lines? Children Spontaneously Encode Spatial-Numeric Relationships in a Novel Spatial Estimation Task. *Frontiers in psychology*, *8*, 2242.

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*, 352-402.

Vasilyeva, M., & Huttenlocher, J. (2004). Early Development of scaling ability. *Developmental Psychology*, *40*, 682-690.

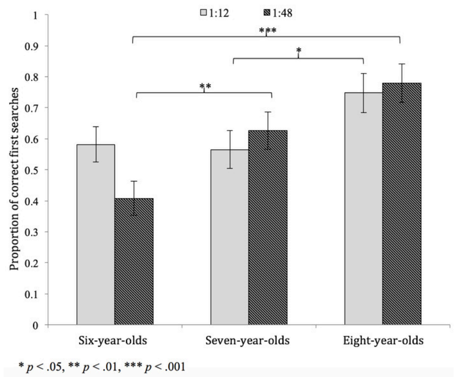
Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2014). Finding the missing piece: Blocks, puzzles, and shapes fuel school readiness. *Trends in Neuroscience and Education*, *3*, 7-13.

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

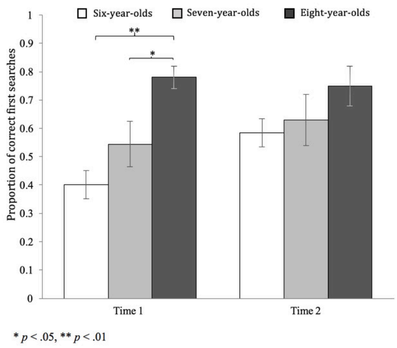
Weisberg, D. S., Kittredge, A. K., Hirsh-Pasek, K., Golinkoff, R. M., & Klahr, D. (2015). Making play work for education. *Phi Delta Kappan, 96*(8), 8-13.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). Examiner's manual. *Woodcock-Johnson III Tests of Achievement.* Itasca, IL: Riverside Publishing.

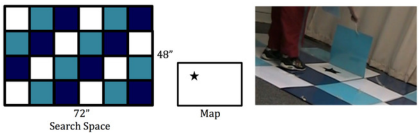
Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychonomic Bulletin & Review*, *15*, 763-771.



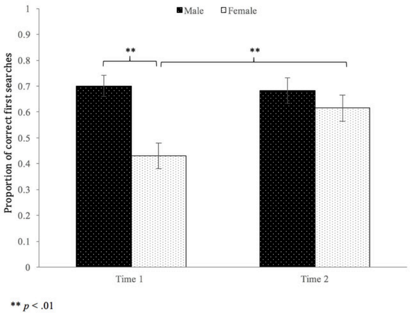
600scaling.fig2.jpg



600scaling.Fig3.jpg



600Scaling.Fig1.jpg



600Scaling.Fig4.jpg

Scaling up spatial development: A closer look at children's scaling ability and its relation to number knowledge

Jamie J. Jirout[1]

Corinne A. Holmes[2]

Kizzann Ashana Ramsook[3]

Nora S. Newcombe[4]

[1] University of Virginia, [2] Trinity College Dublin, [3] Penn State University, [4]Temple University

Author Note

Work on this project was funded as part of the Spatial Intelligence and Learning Center grant from the National Science Foundation, SBE-1041707. Special thanks to Paula Yust and Lindsey Hildebrand for their help with data collection, and the staff and students at Miquon School for their participation. Correspondence concerning this article should be addressed to Jamie Jirout, Curry School of Education, University of Virginia, 405 Emmet Street South, VA, 22903, Jirout@virginia.edu, 434-284-0835.