Cognition and Instruction

ISSN: 0737-0008 (Print) 1532-690X (Online) Journal homepage: https://www.tandfonline.com/loi/hcgi20

Enhancing Children's Spatial and Numerical Skills through a Dynamic Spatial Approach to Early Geometry Instruction: Effects of a 32-Week Intervention

Zachary Hawes, Joan Moss, Beverly Caswell, Sarah Naqvi & Sharla MacKinnon

COGNITION AND INSTRUCTION
2017, VOL. 35, NO. 3, 236–264
https://doi.org/10.1080/07370008.2017.1323902

Enhancing Children's Spatial and Numerical Skills through a Dynamic Spatial Approach to Early Geometry Instruction: Effects of a 32-Week Intervention

Zachary Hawes[a], Joan Moss[a], Beverly Caswell[a], Sarah Naqvi[a], and Sharla MacKinnon[b]

[a]University of Toronto, Applied Psychology & Human Development, Ontario, Canada; [b]Rainy River District School Board, Fort Frances, Ontario, Canada

ABSTRACT
This study describes the implementation and effects of a 32-week teacher-led spatial reasoning intervention in K–2 classrooms. The intervention targeted spatial visualization skills as an integrated feature of regular mathematics instruction. Compared to an active control group, children in the spatial intervention demonstrated gains in spatial language, visual-spatial reasoning, 2D mental rotation, and symbolic number comparison. Overall, the findings highlight the potential significance of attending to and developing young children's spatial thinking as part of early mathematics instruction.

Spatial thinking is fundamental to learning and cognition. Although spatial skills are necessary for day-to-day activities, including moving from one location to another and remembering the location of objects and events, spatial thinking also plays a fundamental role in scientific discovery and innovation (Kell, Lubinksi, Benbow, & Steiger, 2013). For example, the invention of the induction motor, the discovery of the structure of DNA, and Einstein's theory of relativity were all said to have been borne out of spatial thinking (Newcombe, 2010; von Károlyi, 2013). Einstein described his thinking as follows:

> The words or language, as they are written or spoken, do not seem to play any role in my mechanism of thought. The physical entities which seem to serve as elements in thought are certain signs and more or less clear images which can be "voluntarily" reproduced and combined. … Conventional words or other signs have to be sought for laboriously only in a secondary stage, when the mentioned associative play is sufficiently established and can be reproduced at will (Einstein, quoted in Hadamard, 1945, p. 142–143).

This quote eloquently captures the playful and imaginative nature of spatial thinking. It also accurately describes the quality of spatial thinking that is most often defined and measured in the psychological literature. While definitions vary, spatial thinking is broadly defined as the ability to generate, retain, retrieve, and transform well-structured visual images (Lohman, 1996). Although other definitions include reference to physical interactions with the environment (e.g., navigational skills), this article adheres to a definition of spatial thinking that largely deals with forming and manipulating visual-spatial mental images (see Newcombe & Shipley, 2012 for a fine-grained distinction of the various types of spatial thinking). Indeed, it is the ability to form and manipulate mental images of objects that has been found to play a fundamental role in determining which students enjoy, enter, and succeed in STEM disciplines (science, technology, engineering, and mathematics; Wai, Lubinksi, & Benbow, 2009).

CONTACT Zachary Hawes zhawes@gmail.com University of Toronto, Department of Applied Psychology & Human Development, Dr. Eric Jackman Institute of Child Study, 45 Walmer Road, Toronto, Ontario, M5R 2X2
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hcgi.

In light of these and other related findings, large-scale efforts are underway to infuse and integrate spatial thinking into K-12 curricula (National Council of Teachers of Mathematics, 2000; 2006; National Research Council, 2006). The National Research Council (NRC, 2006) has emphasized that spatial thinking should not be

> an add-on to an already crowded school curriculum, but rather a missing link across that curriculum. Integration and infusion of spatial thinking can help achieve existing curricular objectives … and enable students to achieve a deeper and more insightful understanding of subjects. (NRC, 2006, p. 7)

Furthermore, the NRC describes the current state of spatial thinking as a "major blind spot in the American educational system" and that without explicit attention and curricular focus "spatial thinking will remain locked in a curious educational twilight zone: extensively relied on across the K–12 curriculum but not explicitly and systematically instructed in any part of the curriculum" (NRC, 2006, p. 6). Indeed, research has demonstrated a strong link between spatial thinking and performance across a range of academic subjects, including, but not limited to, science (Wai et al., 2009), geography (Orion, Ben-Chaim, & Kail, 1997), physical education (Pietsch & Jansen, 2012), the arts (Goldsmith, Winner, Hetland, Hoyle, & Brooks, 2013), and, perhaps most notably, mathematics (Mix & Cheng, 2012).

Since the inception of the *Curriculum and Evaluation Standards for School Mathematics* (1989), the National Council of Teachers of Mathematics (NCTM) has been advocating that spatial reasoning[1] play an important role throughout pre-K–12 mathematics education. Most recently, NCTM (2006) has recommended that spatial reasoning receive as much emphasis and instructional time as numeracy in kindergarten through 8th-grade mathematics. NCTM (1989, 2000, 2006) also has recommended that young children's initial school experiences with mathematics should include a *spatial reasoning* approach as part of an introduction to the discipline. A review of the NCTM Geometry Standards for Pre-K to Grade 2 (2006) further illustrates the strong emphasis and expectations placed on spatial reasoning. During these school years, children should be able to apply transformations (e.g., mental rotation) and use symmetry to analyze mathematical situations, create mental images of geometric shapes using spatial memory and spatial visualization, recognize and represent shapes from different perspectives, relate ideas in geometry to ideas in number and measurement, and investigate and predict the results of putting together and taking apart two- and three-dimensional shapes (spatial visualization). These expectations align closely with research findings that demonstrate that even children as young as four and five are capable of engaging in spatial reasoning that involves visualization and mental transformations (Frick, Hansen, & Newcombe, 2013; Frick, Möhring, & Newcombe, 2014; Hawes, LeFevre, Xu, & Bruce, 2015; Levine, Huttenlocher, Taylor, & Langrock, 1999). Importantly, other research also suggest that not only do young children enter school with impressive spatial thinking skills, but they show high levels of enjoyment and motivation during spatially oriented lessons and activities (Naqvi, Hawes, Chang, & Moss, 2013; Taylor & Hutton, 2013).

Taken together, the aforementioned NCTM expectations and related research findings suggest that spatial thinking should be an important component of early years mathematics education. This is rarely the case. In fact, the NCTM geometry expectations stand in stark contrast to how geometry and spatial thinking is typically approached. According to Sarama and Clements (2009) geometry and spatial thinking are often ignored and minimized in early education settings (also see Clements & Sarama, 2011). When geometry and spatial thinking are approached, common practice includes labeling and sorting shapes (Clements, 2004), tasks that are primarily verbal in nature and ignore spatial thinking skills that involve visualization and spatial transformations. Compounding this issue further is the finding that early years educators also receive little to no professional development in geometry and spatial thinking (Ginsburg et al., 2006). A study by Lee (2010) found that early years teachers demonstrate significantly lower content knowledge in spatial sense compared to other mathematics topics, including number sense, patterning, ordering, shapes, and comparison problems. Clearly, increased efforts are needed to address these shortcomings.

Coincidentally, the call to improve early years mathematics through increased attention to spatial thinking (see Newcombe, 2010; Ontario Ministry of Education, 2014) comes at a time of unprecedented political and academic focus on the need to raise the quality of early years mathematics instruction

(Clarke, Clarke, & Roche, 2011; Ginsburg, Lee, & Boyd, 2008; MacDonald, Davies, Dockett, & Perry, 2012). Educators, policy-makers, and researchers have jointly identified the need to provide young children with "extensive, high-quality early mathematics instruction that can serve as a sound foundation for later learning in mathematics and contribute to addressing long-term systemic inequities in educational outcomes" (Cross, Woods, & Schweingruber, 2009 p. 2). Economist James Heckman (2006) has shown that early educational interventions have long-lasting effects and can contribute to the narrowing of disparities and inequalities related to socioeconomic status (SES). Early interventions, especially among the most disadvantaged populations, are critical to ensuring equitable opportunities for all children. Interventions aimed at improving children's spatial thinking and mathematics performance might be especially important, given the fundamental importance of spatial thinking and mathematics to later academic and occupational success (Duncan et al., 2007; Wai et al., 2009).

It was with this background in mind that this study was conducted. The primary purpose of this research was to support the development of young children's (4- to 7-year-olds) spatial thinking and mathematics skills. To achieve this goal, we worked closely with early years teachers to plan and carry out a 32-week intervention that involved implementing a *dynamic spatial* approach to early geometry instruction; an approach that strongly emphasizes the dynamic and spatial aspects of geometry (e.g., visualizing and describing geometric transformations). In line with this approach, a central focus of the intervention was working toward developing children's spatial visualization skills over a sustained period of time (i.e., the academic school year) and across many and varied learning opportunities. Furthermore, this study was carried out in classrooms serving primarily First Nation[2] (Anishinaabe) and Métis students, historically underserved by the educational system. The participating schools routinely under perform in mathematics on standardized provincial tests. Thus, one of our aims was to demonstrate the importance and potential effectiveness of providing underserved students with high-quality and engaging mathematics curricula.

In the following sections, we provide a more detailed account of our rationale to focus on developing children's spatial thinking. We begin by reviewing the literature on the relationship between spatial thinking and mathematics and then discuss the extent to which spatial thinking is malleable and can be improved. We conclude by providing relevant background information on the rationale, planning, and design underlying the current intervention.

On the relationship between spatial thinking and mathematics

An extensive body of research indicates a strong positive relationship between spatial thinking and mathematics performance (see Mix & Cheng, 2012). Moreover, the relationship between spatial thinking and mathematics does not appear to be limited to one specific strand or subset of skills within mathematics. For example, spatial thinking has been linked to performance in geometry (Battista, 1990; Delgado & Prieto, 2004), as one might expect, but also to mental arithmetic (Kyttälä & Lehto, 2008), basic magnitude and counting skills (Kyttälä, Aunio, Lehto, Van Luit, & Hautamaki, 2003; Thompson, Nuerk, Moeller, & Cohen Kadosh, 2013), algebra (Tolar, Lederberg, & Fletcher, 2009), calculus (Sorby, Casey, Veurink, & Dulaney, 2013; Zimmerman, 1991), word problems (Hegarty & Kozhevnikov, 1999), and advanced mathematics (e.g., function theory, set theory, mathematical logic; Wei, Yuan, Chen, & Zhou, 2012). It is important to note that spatial thinking, like mathematics, is not a unitary construct, but rather a collection of cognitive and learned skills.

To date, the majority of research that supports a close connection between spatial skills and mathematics have utilized measures of spatial thinking that fall under the related terms of *spatial visualization* and *mental rotation*. Although spatial visualization involves forming, maintaining, and mentally manipulating images of objects, mental rotation more specifically refers to the mental act of being able to imagine the rotation of 2D and 3D objects (Mix & Cheng, 2012). Thus, both spatial visualization and mental rotation involve the shared cognitive skill of forming and manipulating mental images. Given our goal of providing young children with a strong mathematical foundation, most our spatial intervention explicitly aimed to support children's capacity to flexibly form, retain, and manipulate visual-spatial information.

Support for the role of spatial thinking in acquiring mathematics skills comes from a study by Gunderson, Ramirez, Beilock, and Levine (2012). Using two longitudinal data sets, the authors first demonstrated that spatial skills at the beginning of first and second grades predicted growth in linear number line knowledge over the course of the year. Second, the authors found that children's spatial skills at age 5 predicted how well they performed on an approximate symbolic calculation task at the age of 8. Interestingly, this relationship was mediated by children's linear number line performance at age 6, suggesting that spatial skills play an important role in helping children develop a linear spatial representation of numbers. This finding is significant insofar as children's access to an accurate *mental number line*—as measured by mapping numbers to space—has been found to strongly predict concurrent and later mathematics performance (Siegler & Booth, 2004; Booth & Siegler, 2006). This study, along with others (e.g., see Thompson et al., 2013; Viarouge, Hubbard, & McCandliss, 2014) suggests that spatial skills form an important relationship with basic numerical representations, a relationship that might also help in the development of early number knowledge and computational skills.

Some preliminary evidence for this hypothesis comes from Verdine, Irwin, Golinkoff, and Hirsh-Pasek (2014), who found that early spatial skills longitudinally predicted early number knowledge skills, including number concepts such as *more*, *less*, *equal*, and *second*. Spatial skills assessed at 3 years of age, along with executive functioning skills assessed at 4 years of age, predicted over 70% of children's mathematics performance (basic number knowledge skills) at 4 years of age. Even after controlling for executive functioning and vocabulary skills, spatial skills uniquely predicted approximately 15% of the variability in early number knowledge.

For the purpose our study, it is worth noting how the authors assessed spatial thinking. To capture the spatial skills of young children, the researchers created the Test of Spatial Assembly (i.e., TOSA; see Verdine et al., 2014b), in which children were assessed on their ability to accurately replicate 2D designs made with pattern blocks and 3D structures made with Lego. The authors also assessed spatial skills with the Beery Test of Visual-Motor Integration, a test that involves perceiving and copying a series of progressively more complex 2D line drawings. Although these tasks appear relatively simple, they are sensitive to important individual differences in spatial skills. For example, children as young as 3 years old, of lower SES backgrounds, performed significantly worse on these copying measures, as compared to their higher SES peers (Verdine et al., 2014b; Verdine et al., 2014). These findings suggest that early interventions aimed at improving children's skills at perceiving and copying 2D and 3D shapes might be an effective means for enhancing the development of early spatial skills, especially among children lacking sufficient experience in these sorts of tasks.

Given the important role of spatial thinking for early mathematics, researchers have suggested that spatial instruction is expected to have a two-for-one effect, yielding benefits in spatial thinking and mathematics (Verdine, Golinkoff, Hirsh-Pasek, & Newcombe, 2014a; 2014b; also see Bishop, 1980). However, the success of this strategy hinges on the assumption that spatial thinking is a malleable. Thus, it is first necessary to ask whether, and to what extent, spatial thinking can be trained and improved within educational settings.

On the malleability of spatial thinking

Commonly believed to be a fixed intellectual trait—either you have it or you don't—there is an emerging consensus that spatial thinking is malleable and can be improved under certain learning conditions. Indeed, a recent meta-analysis indicates that people of all ages and through a variety of training approaches (i.e., video games, course training, spatial task training) can significantly improve their ability to think spatially (Uttal et al., 2013). To arrive at this conclusion, Uttal and colleagues (2013) analyzed a total of 206 spatial training studies conducted over a 25-year period. Relative to a control, the average effect size of training was approximately one-half standard deviation (0.47). Furthermore, the authors found evidence to suggest that spatial training transferred to other spatial tasks that were not directly trained. Although this study provides substantial evidence for the plasticity of spatial thinking, it should be mentioned that the majority of training interventions were conducted in controlled laboratory settings and largely consisted of spatial task training (i.e., training that uses spatial tasks to improve spatial thinking). Although the authors suggest that a spatially enriched education might pay dividends in improving mathematics and science achievement, more research is needed to test both the external and ecological validity of such training effects. The classroom offers an ideal real world setting to test the educational relevance of spatial training.

Several spatial training programs have been successfully implemented in early elementary school classrooms. For example, Casey and colleagues (2008) investigated the effects of using block-building interventions to enhance the spatial thinking of kindergarten students. More specifically, two intervention conditions and one control condition were included to test whether a block-building intervention within a story context was more impactful than an intervention without a narrative. The researchers found evidence that both interventions resulted in improved spatial visualization skills relative to the control group. However, only those children in the narrative condition demonstrated enhanced block-building skills. In a study with fourth-grade students, Taylor and Hutton (2013) demonstrated the effectiveness of implementing an in-class spatial training program that involved origami and pop-up paper folding exercises. Compared to a control classroom, children who received the intervention demonstrated enhanced visual-spatial skills. Furthermore, children in the spatial training group reported high levels of engagement with the program.

One last early years intervention program deserves mentioning because of its direct relation to our study. Tzuriel and Egozi (2010) examined the extent to which Wheatley's (1996) Quick Draw activities resulted in gains in first-grade students' mental rotation abilities. The intervention consisted of briefly presenting children with 2D geometric line drawings. Children were then asked to reproduce the image from memory using pencil and paper. The experimenter then facilitated small-group discussions around the various ways the drawings were perceived and remembered, encouraging an appreciation for different perspectives. The intervention was intended to develop children's ability to represent and transform spatial information. Indeed, the authors found evidence that this was the case. Compared to an active control group, children who took part in the Quick Draw intervention demonstrated significant gains on two separate measures of mental rotation.

Although these research findings clearly indicate that spatial skills are malleable, many questions remain regarding the extent to which spatial training generalizes to other related aspects of thinking.

On improving mathematics through spatial training

The idea that supporting children's spatial thinking will serve to leverage the learning of mathematics is not new (e.g., Bishop, 1980), but surprisingly remains a rather untested area of research (Mix & Cheng, 2012). To date, only one study has causally examined whether spatial training improves children's mathematics performance. The researchers of this study, Cheng and Mix (2014), randomly assigned children to either a spatial training condition (i.e., mental rotation training) or crossword puzzle condition. Both groups completed the same pre- and post-tests, assessing both spatial and math skills. Children in the spatial training group, but not the crossword condition, demonstrated significant improvements in their calculation skills. Improvements were most evident on missing-term problems (e.g., $5 + \_\_\_ = 7$), whereby the solution can be arrived at through spatially reorganizing the problem (e.g., $5 + \_\_\_\_ = 7$ becomes $\_\_\_\_ = 7 - 5$). Further support that spatial training generalizes to mathematics comes from studies with college students. Spatial training has been shown to result in improvements in students' calculus performance (Sorby et al., 2013) and other mathematics-related disciplines, including engineering (Hsi, Linn, & Bell, 1997), physics (Miller & Halpern, 2013), and chemistry (Small & Morton, 1983).

Taken together, these research findings suggest that spatial training might be an effective means for improving mathematics instruction and learning. However, the scarcity of research in this area gives rise to many questions, but few answers. More research is needed to determine the extent to which spatial training generalizes to improvements in mathematics performance. Furthermore, amidst the call to *spatialize* the early mathematics curricula, is the need to consider the instrumental role of teachers in achieving this objective. The integration and infusion of spatial skills into existing mathematics curricula depends on teachers; teachers who understand the importance of developing young children's spatial thinking and see the instruction of spatial skills as vital in providing a strong foundation in mathematics.

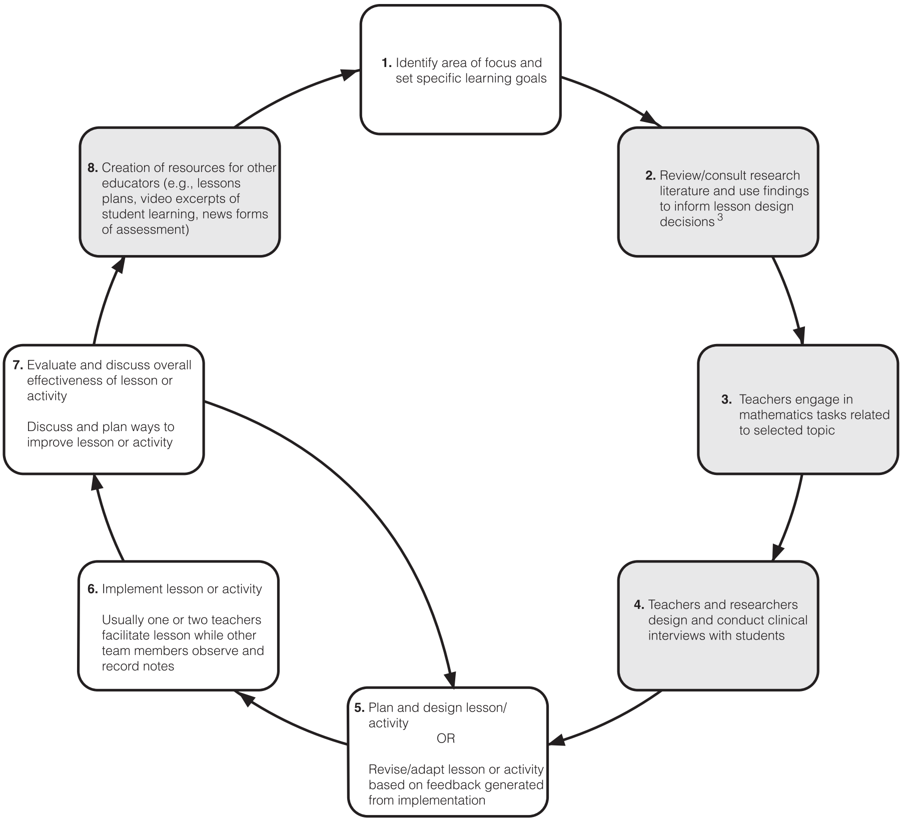


Figure 1. Collaborative lesson and activity design cycle. The process involves teachers and researchers working together to design, implement, evaluate, and revise/adapt lessons and activities based on an identified area of focus or student need. This process is based on a model of teacher professional development known as Japanese Lesson Study (for details, see Lewis et al., 2006). The white boxes (stages 1,5,6,7) refer to the widely recognized four-stage model, while the grey boxes refer to the four additional adaptations.[3]

Building the intervention: Combining research, practice, and theory

Before we describe how we approached the design of the intervention, it is worth summarizing whywe saw the need to build a spatial intervention in the first place. As alluded to, a gap exists between research findings from education, psychology, and cognitive science and classroom practice. On the one hand, findings from developmental cognitive science point to spatial reasoning as a key building block in children's mathematical development. Importantly, research also suggest that spatial thinking is malleable and can be improved through a wide assortment of spatial activities (Uttal et al., 2013). On the other hand, despite the significance and potential utility of such findings, little attention is paid to developing children's spatial thinking in the elementary classroom. It was against this backdrop that we identified the need to begin working toward closing this gap and spatializing the mathematics curriculum (NRC, 2006). Efforts to do so began in earnest 3 years prior to this project.

With several other collaborators, we began a project known as the Math for Young Children (M4YC) project (see Bruce & Hawes, 2015; Moss, Bruce, Caswell, Flynn, & Hawes; 2016; Moss, Hawes, Naqvi, & Caswell, 2015). In short, this project involved working closely with teams of early years teachers (pre-k to grade 3) to design, implement, evaluate, and revise spatial activities or lessons in authentic classrooms settings (see Figure 1 for simplified illustration of the process; for more details, see Moss et al., 2015). Our efforts were primarily focused on spatializing the teaching and learning of geometry. Although geometry is inherently spatial and involves perceiving, manipulating, and reasoning about spatial relationships, our collective experiences working in classrooms (as both teachers and researchers) provided reasons to doubt a widely-adopted approach to instruction that is explicit in its focus on developing children's spatial skills. On the contrary, early geometry instruction is typically limited to sorting and labeling shapes (Clements & Sarama, 2011; for notable exceptions see Clements, Wilson, & Sarama, 2004; Lehrer & Chazen, 2012; Sarama & Clements, 2009). Additionally, in Ontario at least, geometry instruction receives significantly less attention than other strands of mathematics, including number sense, data management, measurement, and algebra and patterning (Bruce, Ross, & Moss, 2012). Taken together, these findings stand in stark contrast to the NCTM pre-k to grade 2 geometry expectations mentioned earlier; expectations that in many ways emphasize the highly dynamic and spatial aspects of geometry.

In line with NCTM's expectations, we worked toward designing lessons and activities that emphasized the explicitly dynamic, spatial, and imaginative aspects of geometry, an approach that focused heavily on developing children's spatial visualization skills and the ability to playfully generate, maintain, and transform visual-spatial images in mind. In this regard, we veered away from teaching geometry solely as the study of static images (e.g., naming shapes), but worked toward also teaching geometry as the study of dynamic images (e.g., understanding how shapes change through various spatial transformations). For this reason, we refer to the present intervention as a *dynamic spatial* approach to early geometry instruction.

The intervention described and evaluated hereafter was the result of 3 years of prior design research and collaborative lesson design with over 100 early years educators and, by extension, over 2,000 of their students. The structure for our design research (described in detail elsewhere, for example, see Moss et al., 2015; Moss, Bruce, et al., 2016) was based on an adapted form of Japanese Lesson Study (Lewis, Perry, & Murata, 2006) and included four additional structures to the typical *lesson study* cycle (see Figure 1). The first adaptation involved sharing and discussing research findings relevant to the group's selected topic of interest (e.g., 2D congruence). The second adaptation involved the lesson study group participants engaging in selected novel geometry tasks, leading to the third adaptation, which involved the design and implementation of clinical interviews (task-based one-to-one dynamic assessments) to learn more about how their students' might reason when working with similar novel geometry tasks. The fourth adaptation involved experimentation with multiple iterations of exploratory lessons and tasks, ultimately leading to the design of a series of novel lessons and activities.

It was through engaging in this process that we gradually accumulated a wide assortment of field-tested spatial activities and lessons. Indeed, it was this bank of activities and lessons that the current intervention was based upon. Teachers involved in this study were offered an assortment of these spatial curricula both at the beginning and throughout the school year (see Methods section for more details). More specifically, teachers were provided with five full-length lesson plans (1 hr each), targeting major geometrical topics and concepts (e.g., area measurement, congruence, symmetry), as well as a wide assortment of short (∼ 10-15 min) and engaging spatial activities. These *quick challenge* activities, as they have become known, were designed with the specific intention of developing students' spatial skills gradually, and adaptively, over the course of the school year. A common feature of both the lessons and quick challenge activities was the strong emphasis on developing children's spatial visualization skills.

Our general approach to the implementation of the intervention, is one that is based on the hypothesized importance of targeting multiple geometric and spatial concepts and skills throughout the school year and across a wide variety of lessons and activities. In other words, we believe that the strength of the intervention lies in both the diversity and combination of lessons and activities.

Purpose of this study

This study sought to determine the effectiveness of a year-long intervention aimed at supporting the development of young children's (4- to 7-year-olds) spatial thinking skills. Our primary objective was to examine the extent to which the in-class spatial activities led to general improvements in young children's spatial and geometry performance. To test for such effects, children were assessed on an assortment of spatial and geometry tasks that were closely aligned with pre-k to grade 2 geometry standards outlined by NCTM (2000). For example, children were assessed on spatial language (including shape recognition and positional language), 2D mental rotation (a task requiring mental translations and rotations), and a comprehensive visual-spatial geometry measure (a task requiring reasoning about 2D shapes, transformations, symmetry, and composition/decomposition of 2D shapes). These skills represent key concepts in the early learning of geometry, and with the possible exception of spatial language, embody geometrical concepts and skills that are closely related—both conceptually and cognitively—to the construct of spatial visualization. Thus, we reasoned to effectively improve children's performance in these areas of geometry would require an instructional approach to geometry that explicitly aims to develop spatial thinking skills, most notably, spatial visualization. Furthermore, unlike other spatial training programs that typically target and train a single spatial skill (e.g., mental rotation), we reasoned that a comprehensive approach to developing spatial thinking would be a more effective (and ecologically valid) method of bringing about general improvements in spatial thinking. For these reasons, we hypothesized that children in spatial intervention classrooms would demonstrate significant gains in spatial thinking compared to the control group.

A secondary purpose of this study was to determine the extent to which supporting children's spatial thinking skills might generalize to improvements in numerical performance. Given the dearth of research in this area, we were less certain about our predictions. However, based on behavioral and neurological research indicating that spatial thinking is both malleable (Uttal et al., 2013) and strongly involved in numerical processing skills (Hubbard, Piazza, Pinel, & Dehaene, 2009), we had reasons to expect that an extensive spatial training program might also yield benefits in children's numerical skills. To test for this possibility, we included measures of both basic numerical processing (magnitude comparison tasks) and more advanced number knowledge and calculation skills.

Methods

Participants

Participation at the school level

Three elementary schools, all part of the same district school board in rural Northwest Ontario, agreed to participate. At the time of this study, these were the lowest performing schools in the district, according to standardized provincial testing results. Utilizing a quasi-experimental design, the two lowest performing schools were assigned to the experimental condition and the third lowest school served as the control group. Thus, group assignment was not random, but was based on a hierarchy of need for intervention.[4] Importantly, however, all three schools were considered well-matched based on standardized provincial test scores in literacy and mathematics, as well as available sociodemographic information. It is important to note both the experimental and the control schools serve a high percentage of Canadian First Nation students who, for the most part, live in First Nation communities but attend public schools regulated by the Government of Ontario.

Participation at the teacher and classroom level

A total of 12 female teachers participated, six as part of the experimental group ($M_{\text{teaching experience}} =$ 9.5 years, $SD = 4.32$) and six as part of the active control group ($M_{\text{teaching experience}} = 11.6$ years, $SD = 3.67$). Whereas teachers in the experimental group carried out the spatial intervention, teachers in the active control group—as opposed to a passive or business as usual control group—carried out an intervention based on an inquiry-based approach to environmental science (more details to follow). In both cases, the interventions were carried out with the entire class of students. Furthermore, in both groups, there were two senior kindergarten teachers, two first-grade teachers, and two second-grade teachers. It was from these participating teachers' classrooms that child participants were randomly selected for pre- and posttesting.

Participation at the student level

During the first week of school, parents and/or guardians were provided with an information letter detailing the project and a request for their child to participate in data collection. In all three schools, over 90% of parents/guardians provided written consent for their child to participate in the study. Due to time constraints, only a small sample of students from each participating classroom were randomly selected to participate in the pre- and posttesting sessions. However, all children, regardless of whether they were selected to participate in testing, took part in the whole-class teacher-led intervention activities. Thus, the interventions were carried out at the classroom level, by the 12 teachers who agreed to have their entire classroom participate in the study

In total, 39 students from the experimental classrooms (M_{age} = 6.25 years, *SD* = .84, range = 4.8–7.6 years, 49% girls) and 28 students (M_{age} = 6.38 years, *SD* = .88, range = 4.8–7.8, 54% girls) from the control classrooms took part in the testing. Two participants in the control classroom were no longer available for testing at posttest, resulting in a total of 26 children in the control group. Information on other missing or incomplete data is further described under the description of each measure.

Demographic information was collected on all child participants regarding their First Nation or Métis status. First Nation people identify themselves by the nation to which they belong. Sixty-nine percent of the total population identified themselves as First Nation or Métis. Of this population, 93% identified as Ojibwe and 7% identified as Métis. In the experimental group, 79% identified as First Nation (93% Ojibwe) or Métis (7%) and in the control group, 54% identified as First Nation (93% Ojibwe) or Métis (7%). All children spoke English as their first language.

Finally, it should be noted that both the University of Toronto and the district school board granted ethical consent prior to the project start date. The Band Councils in the participating First Nation communities also gave permission. Participating teachers and principals provided written consent in agreement with the study's purpose and procedures. Only those children from whom we gathered parental consent were selected for pre- and post-testing.

Design

As mentioned, the study employed a quasi-experimental pre-post research design. Two schools were assigned to the spatial intervention and one school was assigned to the active control condition. Teachers in both the experimental and control group (i.e., six teachers per group; two kindergarten, two first grade, and two second grade classrooms per group), participated in teacher professional development (PD) throughout the academic school year (October 1–May 30). Although the teachers in the experimental group received PD on teaching and learning of spatial reasoning, teachers in the control group received PD on inquiry-based teaching and learning, with a special focus on environmental inquiry. An active control group was employed in an attempt to make the following conditions as equal as possible between the two groups: design, quality, and implementation of PD, time spent engaging in teacher PD, and alignment of pedagogical approaches/beliefs towards practice (i.e., child-centered, inquiry-based approach to education). By holding these factors constant, we hoped to control for any influences or changes in teacher affect (e.g., motivation, self-efficacy, etc.) associated with the general approach to teacher PD. Utilizing such an active control group theoretically allows for a higher degree of certainty that any changes at the student level are less likely due to differences in teacher affect and more likely due to differences in the content and cognitive process targeted in the teacher PD. The two approaches to PD are described in greater detail in the following.

Description of experimental intervention: A dynamic spatial approach to the teaching and learning of early geometry

Process and procedure

Teachers in the spatial reasoning PD were provided with 6 full-days of paid teacher release. The PD was designed and carried out by the first four authors of this article. The PD consisted of both in-person meetings and online meetings via Skype. The in-person meetings occurred at the beginning (October), middle (January), and end of the school year (May). At each time point, meetings occurred over two successive days. The Skype meetings occurred monthly for approximately 2 hr each.

During the first two in-person PD meetings, teachers were introduced to the topic of spatial reasoning. Teachers were presented with (a) a summary of research on spatial reasoning and its link with mathematics, (b) opportunities to experience various types of spatial reasoning activities first hand, and (c) video examples of children's spatial reasoning. Furthermore, teachers were introduced to clinical interviews as a means of more closely examining children's spatial thinking. Prior research suggests the clinical interviews serve the important function of broadening and strengthening teachers' expertise in understanding, assessing, and developing children's mathematical thinking, while also providing important insight into teachers' own mathematical knowledge (Clark et al., 2011; Mast & Ginsburg, 2010; Moss et al., 2015). In our own research, for example, teachers often report being "surprised" by what their children reveal during the clinical interview process, demonstrating competencies when not expected (e.g., on mental rotation tasks) and relatively fragile or weak understandings on tasks assumed to be mastered (e.g., comparing a shapes area and perimeter). As facilitators, we first modeled how to conduct teacher-led clinical interviews (see Ginsburg, 1997) and then invited teachers to carry out interviews with their own students. Observations and findings from the clinical interviews were discussed as a group and helped provide an initial starting place for critically examining various aspects of children's spatial thinking (e.g., children's capacity to visualize the solutions to various composition/decomposition tasks). The aforementioned PD activities served the dual purpose of (a) building relationships among team members and (b) raising excitement, awareness, and better understandings of children's spatial thinking and its potential role in early mathematics teaching and learning.

During the latter half of our first in-person meeting together, teachers were introduced to the intervention lessons and activities. Teachers were provided with paper copies of previously field-tested spatial lessons and activities (i.e., those generated through engaging in the lesson design process outlined in Figure 1). More specifically, teachers were provided with a sample of lessons and activities (three lessons and eight activities) that were selected based on ease of implementation, adaptability, and our own prior experiences of high student engagement and success. Importantly, these were also lessons and activities that with few modifications could be carried out across all three grades (Kindergarten–second grade). Similarly, the activities, referred to hereafter as *quick challenge* activities, were ideally suited to be implemented incrementally throughout the school year. That is, at the beginning of the year, teachers could select from among easier challenges and gradually over the course of the year, and depending on their students' learning needs, select increasingly more difficult challenges. Note that the following section provides further details on the specific lessons and quick challenge activities conducted in the experimental classrooms.

Following our first in-person meeting and throughout the remainder of the year, other lessons and activities were gradually introduced to the team. For example, during the second in-person visit (January), two new lessons and a variety of other quick challenge activities were provided to the team. Our intention of providing lessons and activities gradually, rather than all at once, was to avoid overwhelming the team, especially at the beginning of the year, and to maintain a relatively coherent team focus.

With few exceptions, each lesson and quick challenge activity was first modeled by one of the first authors in one of the participating teachers' classrooms. We did this is an effort to demonstrate how the intervention was intended to be delivered, and to encourage teachers to have fun with the lessons/activities. In each case, we modeled a playful, guided-inquiry approach to pedagogy (e.g., see Fischer et al., 2013), emphasizing ways of fostering student discourse and intended explorations with materials. In particular, we modeled ways of listening and responding to children in an effort to support and draw out students' emerging understandings.[5]

After each time introducing and modeling a lesson or activity, teachers committed to implementing the activity/lesson with their own students. Then, at a later time (e.g., during a Skype meeting), we met as a group and discussed and shared the results of their teaching experiences. Feedback was provided along with, if necessary, recommendations for how to revise or expand the activity or lesson.

The monthly Skype sessions served as an opportunity to hear from teachers on a regular basis and to introduce new activities or ways to expand and/or adapt previously tried activities. As experienced practitioners and experts of their students' learning needs, the teachers played the invaluable role of making recommendations and fine-tuning activities and lessons in an effort to optimize student learning. In some situations, teachers even created their own lesson or activity that they later shared with the rest of the team. In this way, both the in-person and Skype meetings were highly collaborative in nature, as all members brought different strengths to the table and were dedicated to working together to implement the activities and lessons to their fullest potential.

Geometry lessons and quick challenge spatial activities

Geometry lessons

Throughout the school year, each teacher carried out five whole-class lessons designed to target various topics within geometry. Table 1 provides the names of each lesson, as well as the specific geometrical and spatial topics addressed. In brief, the names and focus of these five lessons were as follows: pentomino challenge with a focus on 2D congruence; 3D cube challenge with a focus on mental and physical transformations in three-dimensions; symmetry game where the focus was on reflectional symmetry around vertical and horizontal axes; tile lesson with a spatial approach to a grid structure for area measurement; and garden patio with a focus on composing, decomposing, and transforming area. Each lesson was estimated to last about 1 hr and had previously been field-tested and refined in other pre-k to grade 2 classrooms (see Figure 1). Each lesson was contextualized within a narrative or game-like scenario, and was designed to elicit student inquiry and guided discovery.

Quick challenge activities

In addition to the five lessons, teachers were provided with a series of quick challenge activities. Table 2 provides examples of these activities, as well as the geometric and spatial skills targeted in each. Each quick challenge activity was designed as a brief (10–15 min) and easy-to-implement spatial activity or challenge, requiring short bouts of intense visual-spatial attention. These activities included drawing, building, copying, and visualization exercises (see Table 2 for more details) and formed the bulk of the intervention. The main aim was to develop young children's ability to engage in various features of spatial visualization, including the ability to generate, recall, maintain, and manipulate or transform visual-spatial information in mind and with the aid of manipulatives. The extent to which any one or combination of these features were targeted varied across tasks. For example, some activities placed a heavy emphasis on the recall of visual-spatial information, whereas others focused more on the careful noticing of geometric relationships (e.g., compare Can You Build It? with Shape Transformer in Table 2). In addition to developing children's spatial skills, the activities were also intended to simultaneously target important geometrical concepts, including the composition/decomposition of 2D and 3D shapes, proportional reasoning, and transformational geometry. Through continually adapting the activities gradually over time, it was our intention to maintain student engagement with the tasks, all while building up students' geometric and spatial skills and concepts iteratively and progressively throughout the school year.

Description of active control intervention: Environmental inquiry

Process and procedure

Teachers in the inquiry-based PD were provided with 7 full days of paid teacher release. The PD was facilitated by members of the school board, in collaboration with three experts on inquiry-based learning from the University of Toronto Laboratory School. Two of the authors of this article (Moss and Caswell) also contributed to the facilitation of 2 full days of PD.

The PD consisted of mostly in-person meetings that occurred on a regular basis throughout the school year (October–May). On two occasions, teachers met online and communicated via Skype with those members from the University of Toronto Laboratory School. These meetings were brief, lasting approximately 30 minutes.

The general approach to the in-person PD sessions was comparable to the process and procedure followed by the experimental group. Teachers were first provided with (a) an introduction and background

Table 1. Overview of lessons implemented in the experimental classrooms.

Name of Lesson	Description of Lesson	Geometry and Spatial Skills Targeted
1. Pentomino Challenge 	• Using sets of 5 square tiles, children were challenged to build as many unique configurations as possible (there are 12 possibilities) • Contextualized within a narrative which involved creating "keys" to unlock 12 secret doors • Children were also asked to imagine (and later physically confirm) which pentominoes can be folded into boxes	• Discovering 2D congruence through rotations • Composition/decomposition of 2D shapes • Visualization/mental transformations (mentally folding pentominoes)
2. Cube Challenge 	• Using sets of 3, 4, and eventually 5 multi-link cubes, children were challenged to build as many unique configurations as possible • Eventually the class pooled their structures together and played the role of "shape detectives," comparing their structures and eliminating all structures not unique	• Discovering 3D equivalence through mental and physical rotations of 3D cube structures • Understanding/differentiating mirror images • Composition/decomposition of 3D figures
3. Symmetry Game 	• Playing the game in pairs, children were provided with magnetic shapes and a cookie sheet with a either a vertical or horizontal line of symmetry • Each player was designated his/her own side of the cookie sheet (divided by the line of symmetry) • To begin, one player placed one or two shapes on his/her side of the cookie sheet • The other player then had to create the 'mirror image' on his/her side before placing one or two new shapes for his/her partner (process is repeated until players have filled most of their cookie sheet with a symmetrical design) • As an introduction and follow-up to this lesson, children were presented with a series of 'paper-folding' exercises that involved folding pieces of paper in half and cutting out a design along the folded line of symmetry (e.g., half a heart) • Children then had to reason what the shape would look like when the piece of paper was unfolded (e.g. a heart)	• Understanding reflection symmetry around the vertical and horizontal axes • Paying attention to/understanding the relation of shapes to one another based on location and orientation • Visualization/mental transformations (paper-folding exercises)
4. Tile Lesson 	• Children were provided with a large foam square mat • Teacher showed a smaller square tile (1/4 the size of the square mat) and asked children to imagine how many would be needed to fit perfectly on the mat (no gaps or overlaps) • After making predictions, students were provided with one square tile to confirm/disconfirm their original predictions • Children were then provided with as many square tiles as needed to cover the mat • This process was repeated but with shapes (units) of different relations to both the other units and the larger mat • Throughout the lesson children were required to reason about the proportional relationships (e.g., If the rectangle is half the size of the square, and we know that it takes 4 squares to cover the mat, then how many rectangles will it take to cover the mat?) • Lesson contextualized within a narrative that involved having to purchase tiles of various unit sizes	• Area Measurement • Proportional reasoning (with emphasis on shape/area and number) • Visualization (how many of these squares will it take to cover the mat?) • Mental and physical iteration with one unit (using one unit to reason about the total area) • Tiling

(*Continued on next page*)

Table 1. Continued.

Name of Lesson	Description of Lesson	Geometry and Spatial Skills Targeted
5. Garden Patio Lesson  	• Children were presented with two different shapes, "garden patios" • Children were asked to make predictions about whether the shapes shared the same or different areas • After making predictions, children were provided with one square tile and asked to revise their thinking if necessary • Children were then provided with 3 square tiles and eventually enough tiles to cover the entire area of both shapes (i.e., 9 square tiles)	• Visualization (How many squares will it take to cover one row? How many rows will there be?) • Mental and physical iterations • Tiling • Unitization • Conservation of area

information on inquiry-based teaching and learning, (b) opportunities to experience environmental science inquiry first hand, and (c) time to view and discuss examples of environmental inquiry lessons. As an introduction and guide to environmental science inquiry, teachers were provided with a practice-oriented book that provided examples of lessons told through individual teacher case studies (Natural Curiosity, 2011). This book was written and prepared by members of the University of Toronto Laboratory School (two of the same members who facilitated the environmental PD). Following the introduction to environmental inquiry, the focus of the PD sessions was on designing, implementing, and reviewing co-designed lessons.

Throughout the school year, teachers in the environmental inquiry group carried out lessons and units focused on environmental inquiry. These lessons were either designed during the in-service PD or were based on the lesson/unit designs described in the guide book, *Natural Curiosity*. Topics of inquiry included life cycles (e.g., butterflies, frogs, plants), weather patterns, soil, and gravity.

It is worth noting that in addition to the aforementioned in-service PD, teachers also received 3 full days of mathematics PD with a focus on number sense. This PD was facilitated by the board's numeracy facilitator with the goal of improving the provincial standardized mathematics scores.

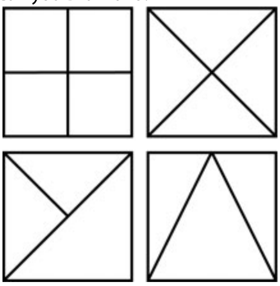
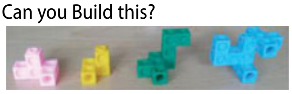
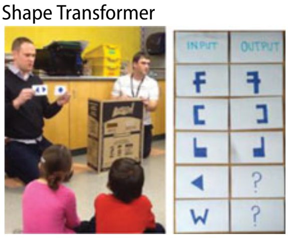
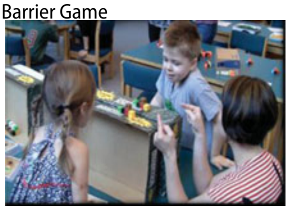
Total time engaged in lessons and activities for the experimental and active control group

Teachers in the experimental group were provided with tracking sheets where they recorded the date, duration, name of activity, and a brief description of and notable observations on the implementation of all activities/lessons conducted. On average, teachers spent approximately 44 hrs of in-class time carrying out the lessons and activities with their students ($M = 44$, $SD = 13.58$, range = 22–55 hr). However, a more accurate estimate of implementation time involves considering an average of time spent per child due to variation in the number of children tested per classroom. Overall, the average child in the experimental group received approximately 47 hr of in-class spatial instruction ($SD = 10.03$). Approximately 90% of the intervention was spent on the quick challenge activities ($M_{\text{quickchallenge activities}} = .88$ vs. $M_{\text{full-lessons}} = .12$, $SD = .05$), which were carried out, on average, about 3 times per week ($M = 2.75$, $SD = .76$).

It is important to note that the experimental intervention was not carried out in addition to the existing mathematics curriculum, but rather as an integrated part of it. Said differently, teachers were not asked to dedicate more time to teaching math. Instead, teachers were asked to carry out the spatial lessons and activities in replacement of time that would have ordinarily been dedicated to business-as-usual mathematics instruction.

Teachers in the active control group recorded the content and number of inquiry-based units tried throughout the year. On average, teachers in the control group carried out approximately four inquiry units based on the planning and lesson design afforded during the PD ($M = 3.8$, $SD = 2.14$, Range =

Table 2. Overview of "quick challenge" activities implemented in the experimental classrooms.

Name Quick Challenge Activity	Description of Activity	Geometry and spatial skills targeted
1. Can you Draw this? 	• Children were provided with pieces of paper with an outline of a square on it • Children were then shown a geometric design composed within the square boundaries • After viewing the design for 10 seconds, children attempted to re-create (using a pencil) the exact design within the boundaries of their own square • Teachers facilitated discussions around strategies and different ways of remembering the designs • Note: this activity was based on Wheatley, 1996; also see Tzuriel & Egozi (2010) for a study on the effectiveness of this activity	• Visual-spatial memory/visualization • Composing/ decomposing/partitioning space • Proportional reasoning
2. Can you Build this? 	• Similar procedures to 'Can you Draw this?' • Children were shown a geometric structure composed of multi-link cubes • After viewing the structure for 10 seconds, children attempted to re-create the structure from memory using their own multi-link cubes • In another version of this activity, children were presented with a structure and asked to re-create it with no memory component	• Visual-spatial memory/visualization • Composing/decomposing 3D figures
3. Building with the Mind's Eye 	• Children were given oral instructions in how to build a 2D or 3D shape/figure (e.g., Take two blue cubes and attach them together, one on top of the other. Stand up the two attached cubes and make them look like a tower. Now take a red cube and attach to...etc.) • Children built images of the shape/figure in mind, based on instructions given • After giving instructions, teacher showed children multiple shapes/figures and had children discuss/reason which one perfectly matched the description	• Visualization • Composition of 2D shapes, 3D figures • Mental transformations (e.g., now take the shape you have in mind and flip it upside down) • Spatial language comprehension • Visual-spatial working memory
4. Shape Transformer 	• Modeled after the 'Function Machine,' an 'input/output' activity typically done with numbers (e.g., input = 2, output = 4; input = 5, output = 10, ...etc. Rule, y = 2x) • In this version, input and output functions deal with spatial relationships (e.g., transformations) • Children were presented with a 'machine' made out of a poster board, with 'input' and an 'output' slots cut out • Teacher (and eventually students) prepared input and output cards to enter and exit into/out of the 'machine' • Children watched and paid attention to relationship between input and output cards and tried to predict the transformation (e.g., each shape that goes into the machine gets rotated 45°)	• Mental transformations/visualization • Visual-spatial reasoning/deductive reasoning • Composition/decomposition of 2D shapes
5. Barrier Game 	• Children worked in pairs with a barrier (folder) in between them and each with their own building materials (e.g., pattern blocks or multi-link cubes) • One partner built a shape/figure and described how to build the shape/figure to his/her partner, who built according to the instructions provided • Children then compared their structures before reversing roles	• Spatial language • Visualization • Composing/decomposing 2D shapes/3D Figures

1 – 7). In terms of mathematics instruction, the teachers implemented the regular Ontario mathematics curriculum (see http://www.edu.gov.on.ca/eng/curriculum/elementary/math18curr.pdf).

Issues of implementation fidelity

This intervention was not intended to be followed and adhered to in a prescriptive fashion. On the contrary, we encouraged teachers to rely on their pedagogical expertise and knowledge of their students to adapt the lessons and activities where they saw fit. With that said, we did support and gauge implementation through the following three ways. First, the teacher's classroom was visited on at least five occasions by two school board employees who were regular members of our team meetings. During these visits, the school board employees either observed a lesson or activity in action or, on some occasions, co-taught the lesson with the classroom teacher. Second, all teachers completed the log sheets in which they provided notes and details on the lessons and activities they implemented. Third, and related to this point, during our regular monthly meetings together teachers shared and discussed the lessons/activities they had carried out between meetings. As part of this sharing process, teachers provided work samples of student work and/or pictures or video footage of the lessons and activities being implemented. Additionally, these classroom artifacts were uploaded to a shared Google Drive folder for all team members to view. Taken together, we had clear evidence that all teachers were actively involved in implementing the lessons and activities throughout the school year.

Procedures for the individual pre- and post-testing

All pre- and posttest measures were administered on a one-to-one basis by trained research assistants in a quiet location of the child's school. The pre- and posttest occurred immediately preceding and following the 32-week intervention. Testing occurred in a fixed order (as presented in the following), over one session, and lasted between 20 and 30 min per child. All but the magnitude comparison tasks were untimed.

Assessment measures

Spatial language test

This test was devised for the purpose of this study to assess both positional language and shape recognition. This test consisted of three parts, with the first part asking children to identify their left and right hand, respectively. The second part (8 questions) required children to identify the location of a ball in relation to a box. For example, children were asked to "point to the picture that shows the ball beside the box." The third part of the test (6 questions) required children to identify various shapes and figures. For example, "Point to the rhombus," "Point to the cube," etc. The measure consisted of 16 items in total and children were allotted one point for each correct answer.

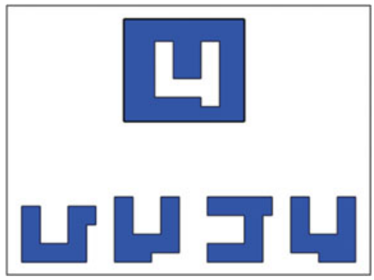
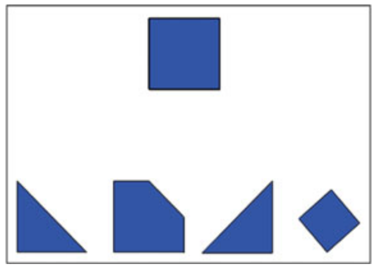
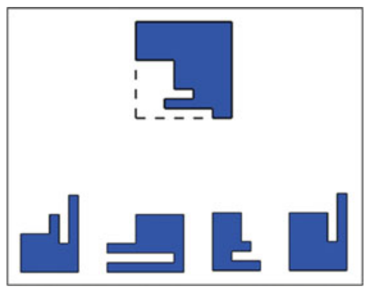
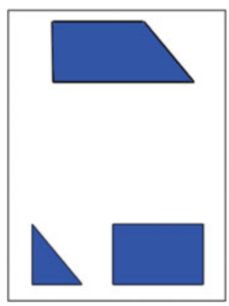
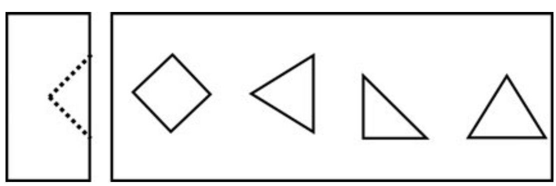
Visual-spatial geometry test

This test was also devised for the purpose of this study and consisted of five different types of questions requiring visual-spatial skills. We created a 20-item test that consisted of five different types of visual-spatial geometry questions, involving missing puzzle pieces (two variations), paper folding, composition/decomposition of 2D shapes, and dissection of 2D shapes. For each question type, we created four items of increasing difficulty. For each question, children were asked to identify the correct answer among four options. Children were allotted one point for each correct response. Table 3 provides an example and brief description of each question type.

2D mental rotation task

The 2D mental rotation task was an adapted version of Levine and colleague's (1999) children's mental transformation task (CMTT; Form D), a widely used measure of intrinsic-dynamic spatial skills designed for young children (4- to 7-year-olds; Ehrlich, Levine, & Goldin-Meadow, 2006; Gunderson et al., 2012;

Table 3. Description and examples of types of test items in the visual-spatial geometry test.

Types of Test Items	Example of Item
Missing Centre Puzzles ● Participants point to the piece that can be made to fit perfectly into the white center space	
Composing/Decomposing 2D Shapes ● Participants point to the two (or three) shapes that can be put together to make the target shape	
Missing Border Puzzles ● Participants point to the piece that can complete the border to make a square	
Dissecting 2D shapes ● Participants use their finger to show where and how to 'cut' the target shape to get the resulting pieces below	
Paper Folding ● Participants presented with a folded piece of paper with the outline of a shape along the fold (i.e., line of symmetry) ● Participants asked to indicate what shape would result after pretending to cut along the outline and unfolding the piece	

Harris, Newcombe, & Hirsh-Pasek, 2013). In the original 32-item test, half the items require mental rotation and the other half involve mental translations. We opted to include only items that required mental rotation. Thus, our test consisted of 16 items. For each item, children were presented with two halves of a shape that has been bisected along either the horizontal or vertical line of symmetry (e.g., a diamond that has been divided into two triangles). The two shapes were separated from one another and rotated 60° away from the vertical axis and presented on either the same plane (direct rotation items) or on the diagonal plane (diagonal rotation items). Directly below the two bisected shapes were four response items (2D shapes) presented in a 2 × 2 array. For each item, children were asked to point to the shape that could be made by putting the two bisected pieces together (e.g., a diamond results when the two triangles are rotated and translated). Half of the items required children to perform direct rotations; the other half required diagonal rotations. For each item there was only one correct response. Children were awarded one point for each correct response.

Symbolic and nonsymbolic magnitude comparison

A paper-and-pencil task developed by Nosworthy, Bugden, Archibald, Evans, and Ansari (2013) was used to assess children's abilities to compare symbolic and nonsymbolic numerical magnitudes. The task is presented in a booklet format and divided into two sections; symbolic comparison items that involve comparing pairs of Arabic numerals and nonsymbolic items that involve comparing pairs of dots. For each section, children were provided with 1 min to complete as many items as possible by crossing off the larger magnitude for each presented pair. Each section has a total of 56 items. In this article, we report on the total number of correct responses achieved within the 1-min time limit. The order of presentation (symbolic vs. nonsymbolic) was random and counterbalanced across individuals. Hereafter, we report on the symbolic and nonsymbolic sections as separate measures of numerical skills. Data are reported only on those children who completed both the pre- and post-test magnitude assessments. For the symbolic comparison tasks, data were not included for seven children in the experimental group and one child in the control group. For the nonsymbolic comparison task, data were incomplete for four children in the experimental group and two children in the control group. In each case, data were incomplete because the child was deemed unable to perform the task (i.e., even after the practice portion, the child still did not grasp the purpose of the activity).

Number knowledge test

The number knowledge test (Okamoto & Case, 1996) was used to assess children's number knowledge and numerical computation skills. We opted to use the most recent version of the test, as it has been used with a large representative sample of Canadian children as part of the National Longitudinal Survey for Children and Youth (Statistics Canada, 2004). The test is standardized and provides normative data for Canadian children. The test consists of 34 items and is divided into five developmental levels. The levels increase in complexity and are based on Case and Okamoto's central conceptual theory for number. At the beginning of the test, children are asked a series of questions that assess basic counting principles (i.e., stable order, one-to-one correspondence, cardinality, order irrelevance, and the abstraction principle). As children progress, they are asked questions that assess the mental counting line (e.g., How many numbers are in between 7 and 9; What number comes 3 numbers before 5?) and computation skills (e.g., How much is 12 plus 54?). With the exception of the first five questions that use counting chips, the test is both administered and responded to orally. The test is discontinued after three consecutive incorrect responses within a developmental level. In this study, children were assigned a raw score based on the total number of items answered correctly. Due to experimenter error, data were incomplete for one child in the control condition.

Peabody picture vocabulary test

Children's vocabulary was measured using the Peabody Picture Vocabulary Test – Revised (Form B; Dunn & Dunn, 1997). This test is widely used as a measure of vocabulary and has been used as a general index of intellectual functioning (Bonny & Lourenco, 2013). The test is standardized with a mean value of 100 and a standard deviation of 15. The test is administered by showing four adjacent pictures and asking the test-taker to indicate the picture that best represents a given word. The test is developmental in structure, allowing for the administrator to establish both basal and ceiling performance. In this study, we report on children's raw scores (ceiling – total number of errors = raw score). Due to time constraints, eight children in the experimental group and five children in the control group were not administered the measure.

Table 4. Correlations among gender, age, and spatial, mathematics, and vocabulary measures at pre-test.

Measures	1	2	3	4	5	6	7	8
1. Gender	—							
2. Age (months)	0.03	—						
3. Spatial Language	0.12	.50**	—					
4. Visual-spatial Geometry	0.15	.34*	.56**	—				
5. 2D Mental Rotation	− 0.01	.49**	.44**	.65**	—			
6. Symbolic Comparison: Numerals	0.17	.54**	.38*	.37*	.44**	—		
7. Nonsymbolic Comparison: Dots	0.11	.65**	.42**	.39*	.57**	.78**	—	
8. Number Knowledge Test	0.08	.54**	.55**	.54**	.46**	.42**	.44**	—
9. Vocabulary (PPVT)	0.08	.59**	.43**	.48**	.62**	.44**	.44**	.64**

Note. $^{**}p \leq .001$, $^{*}p < .01$.

Results

Preliminary data screening and analyses

Prior to analyses, data were screened for potential outliers and violations of assumptions of normality. With one exception, all participants had scores that fell between ± 3 *SD*s of the mean across all seven measures. One child had a score on the number knowledge test that was just above 3 *SD*s of the overall mean. Subsequent analyses were carried out both with and without the inclusion of this participant. In either case, nearly identical results were obtained. Subsequent analyses and results are reported with the inclusion of this participant's data. Histograms of data distributions indicated approximately normal distributions on each measure and across both time points (pre and post). Furthermore, skewness and kurtosis values were all within an acceptable range (i.e., ± 2; George & Mallery, 2010). Overall, the data appeared suitable to analyze using parametric statistical tests, as reported below.

To confirm that no significant differences existed between the experimental and control groups at pretest, a series of independent *t*-tests were conducted. There were no significant differences between groups on any of the following measures: visual-spatial geometry, $t(63) = .948$, $p = .35$; 2D mental rotation, $t(63) = 1.71$, $p = .092$; symbolic comparison, $t(58) = .146$, $p = .885$; nonsymbolic comparison, $t(57) = .034$, $p = .973$; number knowledge test, $t(62) = .482$, $p = .631$: vocabulary (PPVT), $t(50) = 1.16$, $p = .251$. The only difference was found in spatial language, $t(63) = 2.18$, $p = .033$, where the control group outperformed the experimental group. Note that only those children who completed both the pre- and posttest for each respective measure were included in these analyses. A series of independent *t*-tests were also carried out to test for gender differences at pretest. Results revealed no gender differences at the level of the entire sample (for all tests, ps > .22). Analyses were then conducted at the group level (experimental vs. control). There were also no gender differences between the experimental group (*ps* > .23) and the control group (*ps* > .31). Thus, we failed to replicate other findings of early gender differences on measures of spatial thinking (e.g., see Levine et al., 1999); a finding that might be explained in part by the relatively low SES of the sample (e.g., see Levine, Vasilyeva, Lourenco, Newcombe, & Huttenlocher, 2005). Given these findings, the main analyses did not include gender as a factor of interest.

Correlations among age, gender, and cognitive measures

As a first the step in our analyses, we carried out bivariate Pearson correlations to determine the strength of the relations between the various measures. Table 4 shows the simple correlations between gender, age, and the seven measures administered at pre-test across all participants. With the exception of gender, all measures were significantly and positively related to one another (*ps* ≤ .003).

Intervention effects

Training effects were assessed using a series of repeated measures ANOVAS with time (pre, post) as the within-subjects variable and group (experimental, control) as the between-subjects variable.[6] Figure 2

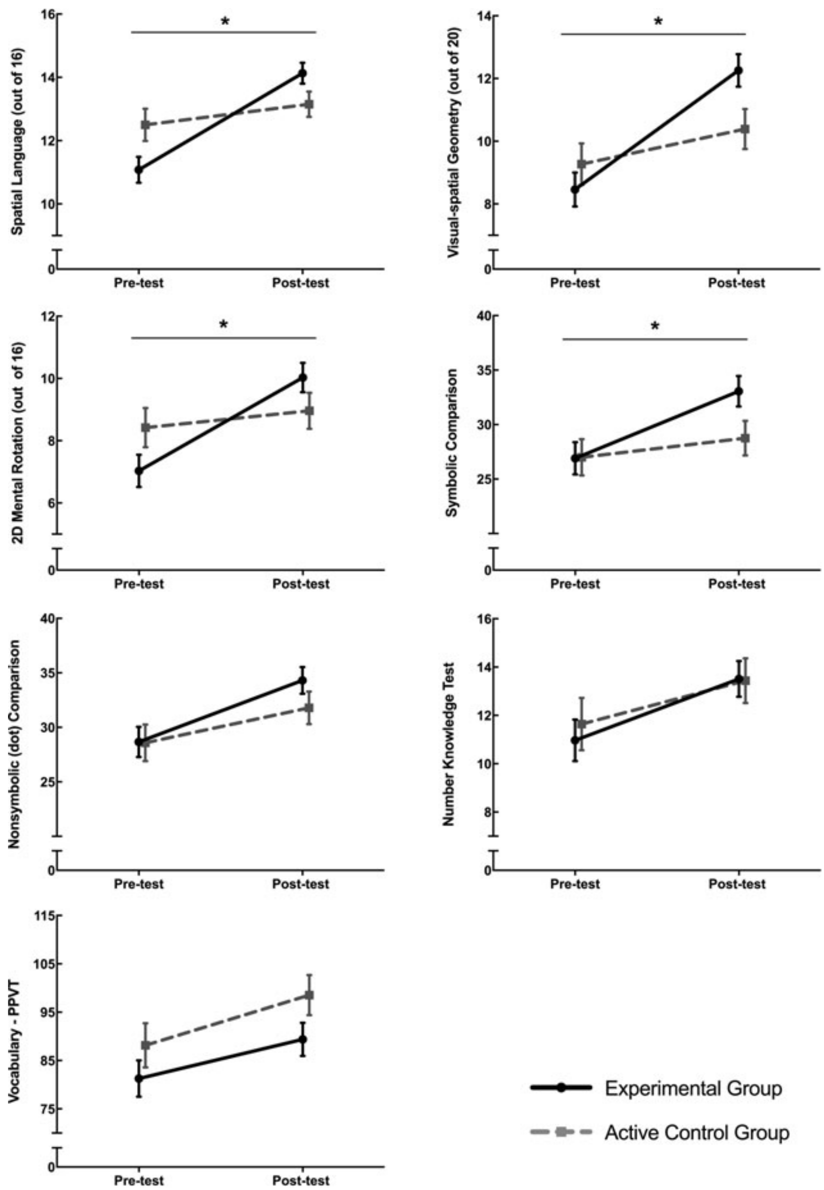


Figure 2. Display of group by time interactions across all seven measures. Error bars represent standard error of the mean. *$p < .05$.

provides a summary of the main findings. For each measure, there was a significant effect of time (all ps $< .001$). These results are not surprising, given the considerable time lag between pre and post (i.e., ~ 32 weeks). Thus, of primary interest was whether improvements from pre- to post- were moderated by, or dependent on, the group factor (experimental vs. control). As shown in Figure 2, significant group-by-time interactions were present on all three measures of spatial thinking: spatial language: $F(1, 63) = 12.39$, $p = .001$, $\eta_p^2 = .16$.; visual-spatial geometry: $F(1,63) = 14.79$, $p < .001$, $\eta_p^2 = .19$; 2D mental rotation: $F(1,63) = 11.53$, $p = .001$, $\eta_p^2 = .16$. Each one of these effects are considered large (i.e., $\eta_p^2 > .14$; Cohen, 1988[7]) . To further evaluate whether the two groups statistically differed at posttest on any of the three spatial measures, independent samples t-tests were conducted (Figure 2 a-c). Results revealed a marginally significant group difference on the spatial language measure, $t(63) = 1.87$, $p = .066$, a significant difference on the visual-spatial geometry task, $t(63) = 2.28$, $p = .026$, and no significant difference on the 2D mental rotation task, $t(63) = 1.43$, $p = .159$. Thus, although the control group significantly outperformed the experimental group on the spatial language measure at pretest, the opposite trend was found at posttest; the experimental group not only caught up to but marginally outperformed the control group at posttest.

Analyses on the three number-based assessments, indexes of far transfer, revealed a significant interaction on the symbolic comparison task (numerals), $F(1,55) = 6.34$, $p = .015$, $\eta_p^2 = .10$, but not the nonsymbolic comparison task (dots), $F(1,57) = 1.45$, $p = .234$, $\eta_p^2 = .03$, or the number knowledge test, $F(1,62) = .58$, $p = .45$, $\eta_p^2 = .01$. The effect size for the interaction on the symbolic comparison task (numerals) is considered medium to large (Cohen, 1988). Note that due to regression toward the mean, the results are even more favorable for the experimental group when the analyses included children who scored zero at pretest (e.g., they could not perform the task), but were able to perform the task at posttest. This is due to larger number of children in experimental group who could not perform the task at pre- compared to the control. Analyses were also carried out on the vocabulary measure, our repeated measures control task. As expected, there was no group-by-time interaction, $F(1,50) = .58$, $p = .57$, $\eta_p^2 = .01$. These analyses were also carried out with age (months) and vocabulary (PPVT raw scores) at pretest entered as covariates. The time-by-group interactions reported above remained significant (ps $\leq .006$), but the time-by-age and time-by-vocabulary interactions failed to reach statistical significance across measures (i.e., ps $\geq .09$). These results suggest that the intervention effects were not moderated by children's age or vocabulary scores at pretest, but rather group membership. Finally, it should also be mentioned that highly similar conclusions are met when one-way ANCOVAs are used to analyze the data and pretest scores are entered as covariates and posttest scores are entered as the dependent variable. Overall, the large effect sizes coupled with convergent findings across different statistical analyses and despite reductions in sample size (i.e., due to listwise deletion), suggest that the experimental intervention was an effective means for improving children's performance on the three spatial measures and the symbolic comparison task. In this next section, we examine whether the effects were specific to grade or the amount of time teachers spent implementing the program.

Implementation time and grade as potential contributing factors to performance gains

Did the amount of time spent implementing the program contribute to differences in children's gain scores? To examine this question correlational analyses were carried out between children's gain scores and the amount of time teachers spent implementing the intervention in their classroom. Correlation values were all close to zero (rs –.23 to .09), suggesting that implementation time had little to no impact on improving performance. Said differently, there was no evidence of a does-response effect. In part, this is likely due to the fact that with the exception of one teacher, teachers' spent a similar amount of time implementing the intervention (40–55 hr). The one teacher that spent considerable less time implementing the intervention only had four children in her class that were tested. Thus, her children's gain scores—even if substantially lower than the other groups (which they were not)—provided little predictive power due to the small sample size.

Next, we examined the possibility that the intervention effects might have differed according to grade level. As it was not advisable to carryout an analysis with teacher and group entered as between-subjects factors, due to extremely small sample sizes in some cases, we opted to carry out a series of repeated measures ANOVAs with grade (SK, first grade, second grade) and group (experimental, control) as the between-subjects factors and time (pre- and postperformance) as the within-subjects factor. This approach was used to provide a reasonable estimate of potential teacher effects. For these analyses, we were interested in time-by-group-by-grade interactions, whereby statistically significant results would suggest that pre-post performance varied as a function of both grade and group. Results revealed no such three-way interactions (ps $\geq .09$); a somewhat unsurprising finding given that entering age as a covariate in the prior analyses did not alter the pattern of results.

These analyses, along with a visual inspection of gain scores across classrooms, suggests that the intervention effects were relatively consistent across the experimental classrooms. Said differently, improvements in children's performance could not be explained by one or two extremely strong teachers in the experimental condition, or, vice versa, one or two extremely weak teachers in the control group. Additionally, the time devoted to the intervention program did not appear to contribute to children's gain scores in the experimental group.

Discussion

Summary of main findings

This study investigated the effects of a 32-week teacher-led intervention aimed at developing young children's (4- to 7-year-olds) geometrical and spatial thinking. Results revealed that, compared to an active control group, children in the spatial intervention demonstrated gains on three separate measures of spatial thinking; spatial language, visual-spatial geometry, and 2D mental rotation. Children in the intervention also demonstrated significant gains relative to the control group on a symbolic number comparison task. This is a novel finding and one that provides preliminary evidence that spatial training might, indeed, transfer to basic numerical skills. Furthermore, this study demonstrates the efficacy of implementing a comprehensive *dynamic spatial* approach to early geometry instruction; an approach that targets the training of multiple spatial skills across multiple time points and varied learning contexts.

On improving children's spatial thinking

Our first research objective was to determine the extent to which children's spatial thinking could be improved through the implementation of a wide variety of in-class spatial activities and lessons. In comparison to the control group, children in experimental classrooms demonstrated broad improvements across three separate measures of spatial thinking—spatial language, visual-spatial geometric reasoning, and 2D mental rotation—suggesting that improvements were not specific to a particular task or set of highly familiar stimuli. Indeed, improvements were demonstrated on untrained spatial tasks. At no point were children ever training for the test as teachers were unaware of the specific requirements of the spatial tests. For these reasons, it seems fair to interpret our findings as evidence that the intervention was effective in bringing about generalized improvements in children's spatial thinking. Our findings add to a growing body of research that suggest spatial thinking is not only amenable to training but, perhaps more important, that the effects of training are generalizable to novel stimuli and untrained spatial tasks (e.g., see Hawes, Moss, Caswell, & Poliszczuk, 2015; Uttal et al., 2013; Wright, Thompson, Ganis, Newcombe, & Kosslyn, 2008). The implications of such findings are potentially significant and far reaching, and provide reasons to be optimistic about improving STEM performance through developing children's spatial skills (Newcombe, 2010; Stief & Uttal, 2015).

This study makes an important contribution to research on spatial thinking and, more specifically, approaches to spatial training. Despite widespread calls to place much greater emphasis on early spatial instruction in our schools (e.g., see Newcombe, 2010; NRC, 2006), there are relatively few examples of such efforts (but see Casey et al., 2008 and Clements & Sarama, 2007). A major goal of this study was to bridge this divide and move out of the laboratory and into the classroom to study the effects of spatial instruction. Our approach differs in important ways from typical approaches to training. To this point, the majority of spatial training interventions have been relatively short in duration (e.g., less than 10 hr), involve repeated practice with a single training task (e.g., playing Tetris or repeatedly practicing a particular spatial test), and are conducted by trained experimenters (Uttal et al., 2013). This intervention, on the other hand, was extensive and long in duration (i.e., distributed throughout the school year), involved repeated and novel exposure to a wide variety of spatial tasks (with a strong focus on spatial visualization), and was implemented by classroom teachers as part of regular mathematics instruction. Identifying these differences helps elucidate key features of the current intervention that may have contributed to the observed findings.

Intriguingly, growing evidence suggest that the most successful attempts at improving core aspects of cognition, such as working memory, involve approaches to training that integrate complexity, novelty, and diversity to maximize ecological validity (Moreau & Conway, 2014).

Our study capitalized on these features and demonstrates the potential utility of an ecologically valid approach to school-based spatial instruction. By offering children many and varied opportunities to develop their spatial thinking skills, it is likely that individual children were differentially influenced by the various lessons, activities, and specific geometric and spatial skills targeted in each. For example, the intervention tasks differentially targeted skills related to the mental composition/decomposition of 2D shapes and 3D figures, mental transformations (e.g., rotations, translations, reflections), and visual-spatial memory. As hypothesized, this comprehensive spatial approach to early geometry instruction, with a strong emphasis on developing spatial visualization skills, appears to support the learning of the pre-k to grade 2 geometry standards outlined by NCTM (see http://www.nctm.org/standards/content.aspx?id = 26846).

Furthermore, what was an effective activity for one child need not be an effective activity for another. Similarly, certain spatial tasks were likely more effective than others at bringing about the intended outcomes, at both the individual and group level. Evidence for both these possibilities comes from the intervention teachers themselves. During our regular conversations together, teachers often expressed surprise in their observations of individual differences in how children took to the various activities. For example, one teacher mentioned how a student demonstrated an exceptional ability to build/draw a structure/image from memory but regularly struggled to articulate how he approached the task. This same child demonstrated difficulties expressing himself in other contexts, as well. As this example points out, by targeting multiple spatial skills both within and between activities, children were provided with opportunities to both build on their strengths (e.g., visual-spatial memory) and work toward improving underdeveloped skills (e.g., spatial language). Furthermore, teachers were often in agreement that some activities worked better than others at both addressing the specific learning goals but also in maintaining student engagement. This speaks to the point made earlier that some tasks were likely more effective than others at inducing change at the group level. Taken together, it seems likely that the strength of an intervention such as this one is due to the cumulative and potentially synergistic effects afforded by reinforcing spatial skills within and across multiple contexts and presentation modes. In other words, the whole is greater than the sum of the individual parts.

Importantly, the overall patterns of student growth were consistent across the experimental classrooms. Similar improvements were made regardless of the teacher and/or grade of student. This is an important finding and one that indicates that results were not merely a result of one or two highly talented teachers in the experimental group. Rather, all six teachers were similarly successful in boosting their students' spatial thinking performance, despite classroom characteristics unique to the teacher (e.g., differences in grades taught and student needs). Indeed, in working with these teachers, it was clear although each one had its own obstacles to deal with the in the classroom, all members remained motivated throughout the intervention and successfully implemented the activities and lessons on a regular basis. Importantly, teachers were encouraged to adapt the intervention in ways that best suited their own needs as well as the needs of their students. For example, although both kindergarten teachers and the grade 2 teachers carried out the same quick challenge activities, specific modifications were required to best suit the learning needs of these two diverse age groups. That is, the challenges and points of discussion for the kindergarten students looked much different than those of the second-grade students.

Lastly, it is worth pointing out that the findings could not be attributed to general cognitive differences between the experimental and control group. Both groups demonstrated similar improvements in vocabulary, a proxy for general intellectual functioning (Bonny & Lourenco, 2013). Furthermore, when controlling for the influence of children's vocabulary skills and age, group differences remained statistically significant. Taken together, our findings indicate that children's gains in spatial and symbolic comparison skills were widespread and consistent across the experimental classrooms.

On improving children's numerical skills through spatial learning

Our second research objective was to investigate the extent to which spatial learning would support children's numerical development. Results revealed that children in the intervention, but not the control group, made significant gains on a measure symbolic comparison. This is a meaningful finding insofar as researchers have identified symbolic (i.e., digits) numerical comparison skills to be highly related to concurrent and later mathematics achievement (De Smedt, Verschaffel, & Ghesquiere, 2009; Durand, Hulme, Larkin, & Snowling, 2005; Nosworthy et al., 2013). For example, De Smedt and colleagues (2009) demonstrated that numerical comparison skills in first grade were predictive of general mathematics achievement in second grade (i.e., number knowledge, understanding operations, simple arithmetic, word problems and measurement), even after controlling for general processing speed and intelligence. A recent review of the literature suggests that data are consistent and robust across studies and populations: Weak performance in numerical comparison relates with low mathematics achievement and dyscalculia (De Smedt, Noël, Gilmore, & Ansari, 2013). For these reasons, symbolic numerical comparison skills have been identified as a foundational building block for mathematics learning (Durand et al., 2005). As such, early interventions aimed at improving this skill are highly desirable (De Smedt et al., 2013).

There are several reasons why the current intervention may have benefited children's symbolic numerical comparison skills. The most probable explanation, however, involves the intertwined nature and unavoidable overlap between spatial and numerical processing skills inherent in many of intervention activities (for a review see Newcombe, Levine, & Mix, 2015).

Although the focus of the spatial intervention was primarily aimed at developing spatial visualization skills, there were several tasks that explicitly dealt with number and many tasks that likely implicitly dealt with number. For example, in the area measurement lesson (see Table 1), children were asked to visualize the number of square tiles required to fill a given area. In this same activity, children were required to engage in proportional reasoning, using the number of square units as a way of comparing areas of various proportional relations to one another (e.g., "The square is twice as big as the rectangle because the square takes up 8 squares and the rectangle takes up 4"). In the majority of activities, however, the use of numbers were less explicit but nonetheless an integral component. For example, in the pentomino lesson (see Table 1), children were required to build as many unique configurations as possible using sets of five square tiles. Children were thus provided with an opportunity to experience, conserve, and compare *fiveness* across multiple representations (i.e., 12 unique configurations). In short, the intervention tasks provided children with a meaningful context to develop an understanding of number concepts and their various relations. It is possible that these repeated experiences of relating number words with various spatial representations proved helpful in the acquisition of symbolic number understanding (e.g., reading and writing numerals) in learning contexts outside the current study.

However, there are at least two other ways in which the intervention may have contributed to improvements in children's symbolic comparison skills. First, from a more theoretical standpoint, it is possible that the spatial intervention resulted in an improved spatial representation of number. It has been suggested, and indeed there is some empirical support for the claim, that spatial thinking skills are related to the strength and precision of one's internal representation of symbolic number via the *mental number line* (e.g., see Gunderson et al., 2012; LeFevre et al., 2013). For example, Viarouge et al. (2014) recently proposed that a strong ability to manipulate mental images (e.g., mental rotation) is likely related to a strong ability to move attention along the mental number line. Additionally, findings from neuroimaging studies indicate neighboring and overlapping brain regions, namely, the intraparietal cortex, that are activated during both basic numerical and spatial tasks (Hubbard et al., 2009). Taken together, there is evidence that numerical and spatial skills are fundamentally linked. Improvements in spatial thinking might provide one means of strengthening symbolic number skills.

A second way in which the intervention may have led to improvements in symbolic comparison was through strengthening children's executive function skills, including visual-spatial working memory. Indeed, research has revealed a close connection between executive functioning, spatial skills, and mathematics (Hubber, Gilmore, & Cragg, 2014; Kyttälä & Lehto, 2008; Verdine et al., 2014). An especially close relationship has been found between visual-spatial working memory and mathematics, including basic numerical processing competencies (Mix & Cheng, 2012). Given that many of the quick challenge activities placed heavy demands on children's visual-spatial memory, attention, and effortful control (i.e., key features of executive functioning), it is possible that intervention also worked towards strengthening children's executive functions. In turn, improvements in executive functioning would be expected to contribute to improved test performance, especially on measures placing heavy demands on children's executive functions. For children, symbolic comparison might be one such task that demands strong executive function skills.

To summarize, it is possible that improvements in symbolic comparison were a consequence of: (a) engaging in tasks that involved both spatial and numerical processing demands, (b) improvements in children's spatial representation of symbolic magnitude (e.g., development of a more precise mental number line), and (c) through gains in executive functioning. A combination of these mechanisms is equally plausible. Given the well-documented association between spatial and numerical skills, a clear avenue of future research involves the development of a more sophisticated and theoretically-driven account of this relationship. Furthermore, more training studies—utilizing both experimental and ecologically valid approaches—are needed to test proposed mechanisms of transfer.

There are several reasons why the spatial intervention may not have resulted in improvements on the nonsymbolic magnitude comparison task and the number knowledge test. According to a recent study, first-grade children demonstrated steeper learning trajectories in symbolic versus nonsymbolic comparison tasks over the course of the school year (Matejko & Ansari, 2016). Although recent research suggest that the Approximate Number System (ANS) is malleable (Park & Brannon, 2013) and develops over time (Halberda & Feigenson, 2008), the learning constraints may be greater for nonsymbolic than symbolic numerical representations. Our data support this view, as indicated by larger gains by the experimental group on both comparison tasks, but with a noticeably steeper trajectory on the symbolic task (see Figure 2).

The absence of a group difference on the number knowledge test is potentially a result of the mathematics learning that occurred in the control classrooms. Although the number knowledge test is a developmentally sensitive test, it also contains many items that are likely practiced in a typical mathematics classroom (e.g., counting, word problems, basic arithmetic). Future research efforts of this sort should look to include measures of mathematics that provide both an index of novel mathematical problem solving as well as more conventional school-based mathematics (e.g., arithmetic). In this way, we might begin to disentangle the differential effects of spatial training across a wide assortment of mathematical tasks.

Limitations

There are several limitations in the study design worth noting. First, the classrooms were not randomly assigned to either the experimental or control conditions. Thus, it is impossible to rule out the influence of any systematic differences between groups at the outset of the study which may have contributed to the study's outcomes. With that said, both groups appeared well matched. In fact, except for performance on the spatial language measure, both groups of children began the study at similar starting points. Furthermore, the teachers in both groups were comparable in terms of years of teaching experience. Another limitation, which ironically is perhaps the study's greatest strength, was the decision to employ a comprehensive approach to spatial instruction. Teachers varied not only in the amount of time they devoted to the various activities, but also in how they chose to implement and adapt them. Thus, it is impossible to isolate any one mechanism responsible for the student gains observed. Instead, we can only speculate that it was the combination of spatial curricula that brought about significant change. Tightly controlled research studies are needed to test the activities and lessons in isolation or in various combinations to determine their relative effectiveness.

Last, this study could be improved by enlisting an active control group that engages teachers in mathematics PD and intervention lessons/activities that does not explicitly target children's spatial thinking. It remains possible that changes in student thinking were specific to our approach to mathematics PD and not necessarily a consequence of the targeted spatial instruction per say. Future research is needed to test the effects of the spatial intervention in the absence or minimization of researcher involvement.

Conclusion: Implications of findings for early mathematics education

Overall, this study suggests that there are many benefits to be had in offering young children a challenging yet engaging assortment of spatial activities. Whereas geometry is often underemphasized in the early years curriculum (Clements, 2004), this study made it a priority to attend to geometry as a critical component of early years mathematics and one that provides plenty of opportunities to develop spatial thinking skills. To this point, teachers carried out the intervention as part of the regular mathematics curriculum and worked toward spatializing the geometry curriculum throughout the course of the school year. The central focus of the intervention activities was the development of children's spatial visualization skills. Across all lessons and quick challenge activities, children were asked to visualize and anticipate certain outcomes (e.g., "How many square tiles will it take to cover the rectangle?"). Furthermore, children were provided with opportunities to engage in the type of dynamic spatial thinking that is typically ignored in early geometry. For example, students were asked to build and transform shapes and figures with their mind's eye, mentally cut and unfold pieces of paper, and predict and reason about geometrical transformations using the shape transformer. Not only did children show high levels of engagement with these types of activities, but our findings also suggest that through participating in such activities, children demonstrated improvements in a number of important geometry standards outlined by NCTM (2000; 2006). The spatial intervention also appeared to benefit children's numerical comparison skills, a foundational skill key to both current and future mathematics achievement (De Smedt et al., 2013; De Smedt et al., 2009). In summary, this research indicates the potential significance of attending to and developing young children's spatial thinking skills as a regular component of early years mathematics instruction.

Notes

1. Note that throughout this article, *spatial thinking* and *spatial reasoning* will be used interchangeably.
2. First Nation is a term used to describe the descendants of the first peoples of Canada who lived in Canada for many thousands of years pre-colonization.
3. Note that in our model, research findings can, and often do, play an important role throughout all stages. For example, not only does the literature help identify areas of spatial thinking to work towards developing in students (e.g., mental rotation), but also provides a new lens on which to view student learning in the context of lesson implementation and reflection (e.g., noticing the importance of mental rotation in area measurement).
4. Decisions around which schools we could work with and serve as the experimental and control groups were carried out at the district school board level. Even before our involvement in this study, it was part of the school board's strategic action plan, and allocation of finances and resources, to improve the mathematics program at the three selected schools. In an effort to work towards this objective, we were approached by the numeracy instructional leader (an employee of the school board and the Ontario Ministry of Education, as well as the fifth author of this article) who had previously read and heard about our team's approach to teacher PD in other areas of the province. She asked whether we might be interested in providing our teacher PD in her district and in working with the three lowest-performing schools. Given our prior interests in working with schools and children from underserved urban communities, we saw this as an opportunity to test our research in rural region of Ontario. Furthermore, we were made aware of the strong First Nation population in this area of Ontario, and saw this offering as an opportunity to work and learn from First Nation communities, especially with respect to cultural specific practices in mathematics teaching and learning. However, this aspect of our work lies outside the scope of this study and is not reported on here.
5. A much more detailed account of our pedagogical approach can be found in Moss, Bruce et al., 2016 and Moss, Hawes et al., 2015. To gain a better understanding of how the lessons/activities typically playout in actual classrooms, please see: Bruce, Sinclair, Moss, Hawes, & Caswell, 2015; Moss, Bruce, et al., 2016; Moss, Bruce, Bobis, 2015; Moss et al., 2015. These resources provide detailed descriptions of the lessons/activities, including transcripts and examples of teacher and student interactions. Video samples of lessons/activities that were created and implemented in the lead-up to the present project are available here:
http://www.oise.utoronto.ca/robertson/1/Math_For_Young_Children/index.html.
6. Analyses were initially conducted using a repeated measures MANOVA with time as the within-subjects variable and group as the between-subjects variable. The overall time-by-group interaction was significant, $\lambda = .61$, $F(7,36) = 3.28$, $p = .008$, with a very large effect size, $\eta_p{}^2 = .39$. That is, approximately 39% of the variance in pre-post performance could be explained by the group factor (experimental vs. control). As a follow-up to this analysis, we opted to carry out a series of independent repeated measures ANOVAs to maximize sample sizes and statistical power. This approach allowed us to deal with missing data on various measures at either time point. It should be noted, however, that even with a reduced sample size (experimental $n = 26$; control $n = 18$), the univariate analyses conducted as a follow-up to the MANOVA, yielded identical patterns of statistically significant findings as reported above in the main text.

7. Partial eta squared ($\eta_p{}^2$) values are used to index effect sizes, where .01, 0.6., .14 denote small, medium, and large effects (Cohen, 1988).

Acknowledgments

We extend our heartfelt gratitude to the teachers, students, schools, and First Nation communities for opening their doors to us and making this project possible. We are also indebted to the teachers at the Dr. Eric Jackman Institute of Child Study (Ontario Institute for Studies in Education at the University of Toronto) for their tireless support. Finally, we acknowledge the generous support of the N.S. Robertson Foundation.

References

Battista, M. (1990). Spatial visualization and gender differences in high school geometry. *Journal for Research in Mathematics Education*, *21*, 47–60.

Bishop, A. J. (1980). Spatial abilities and mathematics education—A review. *Educational Studies in Mathematics*, *11*(3), 257–269.

Bonny, J.W., & Lourenco, S.F. (2013). The approximate number system and its relation to early math achievement: evidence from the preschool years. *Journal of Experimental Child Psychology*, *114*(3), 375–388.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *42*(1), 189–201.

Bruce, C. D., & Hawes, Z. (2015). The role of 2D and 3D mental rotation in mathematics for young children: what is it? Why does it matter? And what can we do about it? *ZDM Mathematics Education*, *47*(3), 331–343.

Bruce, C. D., Moss, J., & Ross, J. (2012). *Survey of JK to Grade 2 teachers in Ontario, Canada: Report to the literacy and numeracy secretariat of the ministry of education*. Toronto: Ontario Ministry of Education.

Bruce, C., Sinclair, N., Moss, J., Hawes, Z., & Caswell, B. (2015). Spatializing the mathematics curricula. In B. Davis and Spatial Reasoning Study Group. (Eds.). *Spatial reasoning in the early years: Principles, assertions and speculations* (pp. 85–106). New York, NY: Routledge.

Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., & Copley, J. (2008). The development of spatial skills through interventions involving block building activities. *Cognition and Instruction*, *26*(3), 269–309.

Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of Cognition and Development*, *15*(1), 2–11.

Clarke, D., Clarke, B., & Roche, A. (2011). Building teachers' expertise in understanding, assessing and developing children's mathematical thinking: the power of task-based, one-to-one assessment interviews. *ZDM–International Journal on Mathematics Education*, *43*(6-7), 901–913.

Clements, D. H. (2004). Geometric and spatial thinking in early childhood education. In D. Clements, J. Sarama, & M. A. DiBaise (Eds.), *Engaging young children in mathematics: Results of the conference on standards for pre-school and kindergarten mathematics education* (pp. 83–90). Mahwah, NJ: Erlbaum Associates.

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education*, *38*, 136–163.

Clements, D. H., & Sarama, J. (2011). Early childhood teacher education: The case of geometry. *Journal of Mathematics Teacher Education*, *14*, 133–148.

Clements, D. H., Wilson, D. C., & Sarama, J. (2004). Young children's compositions of geometric figures: A learning trajectory. *Mathematical Thinking and Learning*, *6*(2), 163–184.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Connolly, A. J. (2007). *KeyMath diagnostic assessment* (3rd ed.). Minneapolis, MN: Pearson Assessments.

Cross, C. T., Woods, T. A., & Schweingruber, H. (Eds.). (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Academies Press.

Delgado, A. R., & Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence*, *32*(1), 25–32.

De Smedt, B., Noël, M. P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*(4), 469–479.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/0012-1610.1037/0012-1649.43.6.1428

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7-to 10-year-olds. *Journal of Experimental Child Psychology*, *91*(2), 113–136.

Ehrlich, S., Levine, S., & Goldin-Meadow, S. (2006). The importance of gestures in children's spatial reasoning. *Developmental Psychology*, *42*, 1259–1268.

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*(6), 1872–1878.

Frick, A., Möhring, W., & Newcombe, N. S. (2014). Development of mental transformation abilities. *Trends in Cognitive Sciences*, *18*(10), 536–542.

Frick, A., Hansen, M. A., & Newcombe, N. S. (2013). Development of mental rotation in 3- to 5-year-old children. *Cognitive Development*, *28*(4), 386–399.

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference*. Boston, MA: Pearson.

Ginsburg, H.P. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. New York, NY: Cambridge University Press.

Ginsburg, H. P., Kaplan, R. G., Cannon, J., Cordero, M. I., Eisenband, J. G., & Galanter, M., et al. (2006). Helping early childhood educators to teach mathematics. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 171–202). Baltimore, MD: Paul H. Brookes.

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report of the Society for Research in Child Development*, *22*(1), 1–23.

Goldsmith, L. T., Winner, E., Hetland, L., Hoyle, C., & Brooks, C. (2013). Relationship between visual arts learning and understanding geometry. Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.

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*(5), 1229–1241.

Hadamard, J. (1945). *The psychology of invention in the mathematical field*. Princeton, NJ: Princeton University Press.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457–1465.

Harris, J., Newcombe, N. S., & Hirsh-Pasek, K. (2013). A new twist on studying the development of dynamic spatial transformations: Mental paper folding in young children. *Mind, Brain, and Education*, *7*(1), 49–55.

Hawes, Z., LeFevre, J. A., Xu, C., & Bruce, C. D. (2015). Mental rotation with tangible three-dimensional objects: A new measure sensitive to developmental differences in 4-to 8-year-old children. *Mind, Brain, and Education*, *9*(1), 10–18.

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

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

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, *91*(4), 684–689.

Hsi, S., Linn, M. C., & Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. *Journal of Engineering Education*, *86*(2), 151–158.

Hubbard E.M., Piazza, M., Pinel, P., & Dehaene, S. (2009). Numerical and spatial intuitions: A role for posterior parietal cortex? In L. Tommasi, L. Nadel & M. A. Peterson (Eds.), *Cognitive Biology: Evolutionary and Developmental Perspectives on Mind, Brain and Behavior* (pp. 221–246). Cambridge, MA: MIT Press.

Hubber, P.J., Gilmore, C., & Cragg, L. (2014). The roles of the central executive and visuospatial storage in mental arithmetic: A comparison across strategies. *The Quarterly Journal of Experimental Psychology*, *67*(5), 936–954.

Kell, H. J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2013). Creativity and technical innovation: Spatial ability's unique role. *Psychological Science*, *24*(9), 1831–1836.

Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., & Hautamaki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology*, *20*, 65–76.

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

Lee, J. (2010). Exploring kindergarten teachers' pedagogical content knowledge of mathematics. *International Journal of Early Childhood*, *42*(1), 27–41.

LeFevre, J. A., Lira, C. J., Sowinski, C., Cankaya, O., Kamawar, D., & Skwarchuk, S. L. (2013). Charting the role of the number line in mathematical development. *Frontiers in Psychology*, *4*, 641. https://doi.org/10.3389/fpsyg.2013.00641

Lehrer, R., & Chazan, D. (Eds). (2012). *Designing learning environments for developing understanding of geometry and space*. New York, NY: Routledge

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

Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science*, *16*(11), 841–845.

Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational Researcher*, *35*(3), 3–14.

Lohman, D. F. (1996). Spatial ability and G. In I. Dennis, & P. Tapsfield (Eds.), *Human abilities: Their nature and assessment* (pp. 97–116). Hilldale, NJ: Erlbaum.

MacDonald, A., Davies, N., Dockett, S., & Perry, B. (2012). Early childhood mathematics education. In B. Perry (Ed.), *Research in mathematics education in Australasia 2008–2011* (pp. 169–192). Boston: Sense Publishers.

Mast, J. V., & Ginsburg, H. P. (2010). Child study/lesson study: Developing minds to understand and teach children. In N. Lyons (Ed.), *Handbook of reflection and reflective inquiry: Mapping a way of knowing for professional reflective inquiry* (pp. 257–271). New York: Springer Publishing Co.

Matejko, A. A., & Ansari, D. (2016). Trajectories of symbolic and nonsymbolic magnitude processing in the first year of formal schooling. *PloS one*, *11*(3), e0149863.

Miller, D. I., & Halpern, D. F. (2013). Can spatial training improve long-term outcomes for gifted STEM undergraduates?. *Learning and Individual Differences*, *26*, 141–152.

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

Moreau, D., & Conway, A. R. (2014). The case for an ecological approach to cognitive training. *Trends in Cognitive Sciences*, *18*(7), 334–336.

Moss, J., Bruce, C., & Bobis, J. (2015). Young children's access to powerful mathematics ideas: A review of current challenges and new developments in the early years. In L. D. English & D. Kirshner (Eds), *International handbook on mathematics education* (pp. 153–190). New York, NY: Routledge.

Moss, J., Bruce, C. D., Caswell, B., Flynn, T., & Hawes, Z. (2016). *Taking shape: Activities to develop geometric and spatial thinking: Grades K-2*. Toronto, ON: Pearson Canada.

Moss, J., Hawes, Z., Naqvi, S., & Caswell, B. (2015). Adapting Japanese Lesson Study to enhance the teaching and learning of geometry and spatial reasoning in early years classrooms: A case study. *ZDM Mathematics Education*, *47*(3), 377–390.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence*. Reston, VA: National Council of Teachers of Mathematics.

National Research Council (2006). *Learning to think spatially: GIS as a support system in the K-12 curriculum*. Washington, DC: National Academic Press.

Natural Curiosity: Building children's understanding of the world through environmental inquiry/A resource for teachers (2011). Toronto: The Laboratory School at the Dr. Erick Jackman Institute of Child Study.

Naqvi, S., Hawes, Z., Chang, D., & Moss, J. (2013) Exploring pentominoes in 7 diverse Pre-k/K classrooms. In M. Martinez & A. Castro Superfine (Eds.), *Proceedings of the 35th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Chicago, IL: University of Illinois at Chicago.

Newcombe, N. S. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *American Educator*, *34*(2), 29–35.

Newcombe, N. S., Levine, S. C., & Mix, K. S. (2015). Thinking about quantity: the intertwined development of spatial and numerical cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(6), 491–505.

Newcombe, N. S., & Shipley, T. F. (2012). Thinking about spatial thinking: New typology, new assessments. In J. S. Gero (Ed.), *Studying visual and spatial reasoning in design creativity*. New York, NY: Springer.

Nosworthy, N., Bugden, S., Archibald, L., Evans, B., & Ansari, D. (2013). A two-minute paper-and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS ONE*, *8*(7). e67918. https://doi.org/10.1371/journal.pone.0067918

Okamoto, Y., & Case, R. (1996). Exploring the microstructure of children's central conceptual structures in the domain of number. *Monographs of the Society for Research in Child Development*, *61*(1–2), 27–58.

Ontario Ministry of Education (2014). *Paying attention to spatial thinking:Support document for paying attention to mathematical education*. Toronto, ON: Queen's Printer for Ontario.

Orion, N., Ben-Chaim, D., & Kali, Y. (1997). Relationship between earth-science education and spatial visualization. *Journal of Geoscience Education*, *45*, 129–132.

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science*, *24*(10), 2013–2019.

Pietsch, S., & Jansen, P. (2012). Different mental rotation performance in students of music, sport, and education. *Learning and Individual Differences*, *22*(1), 159–163.

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. New York, NY: Routledge.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428–44.

Small, M. Y., & Morton, M. E. (1983). Research in college science teaching: Spatial visualization training improves performance in organic chemistry. *Journal of College Science Teaching*, *13*(1), 41–43.

Sorby, S., Casey, B., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences*, *26*, 20–29.

Statistics Canada (2004). *The number knowledge assessment. National longitudinal survey for children and youth—cognitive measures*. Human Resources and Skills Development Canada
Stieff, M., & Uttal, D. (2015). How much can spatial training improve STEM achievement? *Educational Psychology Review*, *27*(4), 607–615.
Taylor, H. A., & Hutton, A. (2013). Think3d!: Training spatial thinking fundamental to STEM education. *Cognition and Instruction*, *31*(4), 434–455.
Thompson, J. M., Nuerk, H. C., Moeller, K., & Cohen Kadosh, R. (2013). The link between mental rotation ability and basic numerical representations. *Acta Psychologica*, *144*(2), 324–331.
Tolar, T. D., Lederberg, A. R., & Fletcher, J. M. (2009). A structural model of algebra achievement: computational fluency and spatial visualisation as mediators of the effect of working memory on algebra achievement. *Educational Psychology*, *29*(2), 239–266.
Tzuriel, D., & Egozi, G. (2010). Gender differences in spatial ability of young children: The effects of training and processing strategies. *Child Development*, *81*(5), 1417–1430.
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.
Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., & Newcombe, N. S. (2014a). Finding the missing piece: Blocks, puzzles, and shapes fuel school readiness. *Trends in Neuroscience and Education*, *3*(1), 7–13.
Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014b). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development*, *85*(3), 1062–1076.
Verdine, B. N., Irwin, C. M., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Contributions of executive function and spatial skills to preschool mathematics achievement. *Journal of Experimental Child Psychology*, *126*, 37–51.
Viarouge, A., Hubbard, E. M., & McCandliss, B. D. (2014). The cognitive mechanisms of the SNARC effect: An individual differences approach. *PloS one*, *9*(4), e95756.
von Károlyi, C. (2013). From tesla to tetris: Mental rotation, vocation, and gifted education. *Roeper Review*, *35*(4), 231–240.
Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over fifty years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*, 817–835.
Wei, W., Yuan, H., Chen, C., & Zhou, X. (2012). Cognitive correlates of performance in advanced mathematics. *British Journal of Educational Psychology*, *82*(1), 157–181.
Wheatley, G. (1996). *Quick draw: Developing spatial sense in mathematics*. Tallahassee: Florida Department of Education.
Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychonomic Bulletin & Review*, *15*(4), 763–771.
Zimmerman, W. (1991). Visual thinking in calculus. In W. Zimmerman, & S. Cunningham (Eds.), *Visualization in teaching and learning mathematics* (pp. 127–137). Washington, DC: Mathematical Association of America.