The Impact of a Spatial Intervention Program on Students’ Spatial Reasoning and Mathematics Performance

Tom Lowrie, Danielle Harris, Tracy Logan & Mary Hegarty

To cite this article: Tom Lowrie, Danielle Harris, Tracy Logan & Mary Hegarty (2019): The Impact of a Spatial Intervention Program on Students’ Spatial Reasoning and Mathematics Performance, The Journal of Experimental Education, DOI: 10.1080/00220973.2019.1684869

To link to this article: https://doi.org/10.1080/00220973.2019.1684869

Published online: 20 Nov 2019.
Introduction

THE ASSOCIATION BETWEEN spatial reasoning and mathematics has been acknowledged by researchers in education for almost forty years (Bishop, 1989; Clements, 1983; Clements & Battista, 1992; Lean & Clements, 1981). Cognitive psychologists have focused on the benefits of visual and spatial reasoning on problem-solving for much longer (Galton, 1883; Larkin & Simon, 1987; Lohman, 1979; Mayer & Sims, 1994; Piaget & Inhelder, 1956). The relatively recent Science, Technology, Engineering and Mathematics (STEM) agenda has heightened interest in spatial reasoning, since spatial abilities have been shown to predict success in STEM academic pursuits and career involvement in STEM professions (Kell, Lubinski, Benbow, & Steiger, 2013; Wai, Lubinski, & Benbow, 2009). The importance of spatial reasoning on disciplines outside of STEM domains has also become apparent (Newcombe, 2013).

Definitions of spatial reasoning (including spatial thinking, spatial ability and spatial visualization) differ depending on the discipline and focus of the study (Bruce, Sinclair, Moss, Hawes, & Caswell, the Spatial Reasoning Study Group, 2015). In the current investigation, we consider spatial reasoning as an overarching term that includes the skills associated with spatial visualization, mental rotation and perspective taking. More generally, spatial reasoning is a broad term that also includes the utility of these skills to applied situations (such as mathematics).
The present study examined the effectiveness of a practice-based spatial intervention embedded within a pedagogical framework on developing secondary school students’ spatial reasoning skills and mathematics performance. Specifically, this paper considers the extent to which the intervention program was effective: (1) across different aspects of spatial ability; (2a) across different mathematics content areas; and (2b) across problems with different types of graphics (displays).

Components of Spatial Reasoning

Although spatial reasoning is rarely taught explicitly in schools, spatial terms and concepts (e.g., terms such as rotate, orient, and visualize) are mentioned in some discipline-based syllabi (e.g., geography and mathematics, geometry in particular). Recently, researchers have developed an approach to providing explicit spatial instruction that involves developing spatial interventions with classroom teachers (Hawes, Moss, Caswell, Naqvi, & MacKinnon, 2017; Lowrie, Logan, & Ramful, 2017). By relating specific aspects of spatial reasoning to the strategies and processes required to solve problems encountered in syllabi (e.g., the mental rotation of 2D shapes in geometry), these intervention studies have developed training programs that are abreast of the nuances of spatial reasoning. For example, Lowrie et al. (2017) defined three components of spatial reasoning in their training program: namely, mental rotation; spatial orientation (which includes perspective taking); and spatial visualization (which includes imagining complex spatial transformations, such as folding a piece of paper into a three-dimensional form).

Malleability of spatial reasoning

Spatial reasoning can be enhanced through training (see Uttal et al., 2013 for a review) and transfer across different spatial tasks (Wright, Thompson, Ganis, Newcombe, & Kosslyn, 2008). In their meta-analysis of effects of spatial training, Uttal et al. (2013) reported mean effect size improvements of .47 in spatial training across studies. They also noted that the approach to administering the spatial content did not seem to matter—whether with concrete materials, digital tools or a specific scaffolded program. Of note, they found no evidence for differences in the effectiveness of spatial training across age groups, although fewer studies were conducted in secondary school settings than with other cohorts.

A study by Lowrie, Logan, Harris and Hegarty (2018) provided evidence for the effectiveness of a spatial training program across Grades 3-6 of primary school. Effect size improvements in spatial ability for the intervention group (d = .83) were substantially higher than mean improvements reported by Uttal et al. (2013); potentially due to the fact that the spatial content was embedded within a pedagogical framework that mirrored a concept-development learning cycle (the ELPSA framework, described later). Notably, the classroom teachers were able to utilize the learning framework with successful outcomes after five days of intervention training, providing opportunities to implement the program at scale (Lowrie et al., 2018).

Spatial training and its impact on mathematics

The association between spatial reasoning and mathematics performance has been widely acknowledged, even though it has not been the responsibility of mathematics teachers to provide spatial instruction (Bishop, 2008). To date spatial reasoning has been inadvertently promoted through engagement with learning experiences that are used to develop mathematics understanding, such as the use of multiplicative arrays (Clements, Battista, Sarama, & Swaminathan, 1997); the interpretation of maps (Lowrie & Logan, 2007); evoking visual imagery (Presmeg, 1986); proportional reasoning (Möhring, Newcombe, Levine, & Frick, 2016); and geometric reasoning (Battista, 2007). However, recent research has begun to examine effects of explicitly training
Recent studies have begun to demonstrate the impact of spatial training on mathematics performance; (Cheng & Mix, 2014; Hawes et al., 2017; Lowrie et al., 2017; Lowrie, Logan & Hegarty, 2019). The design of these studies included spatial training intervention conditions and control conditions, situated within primary (elementary and middle) school settings. In each study, participants in the intervention condition outperformed the control condition participants on mathematics tasks after students were exposed to some form of spatial training. In the Cheng and Mix (2014) study, the students in the intervention condition (n = 31) were exposed to 40 minutes of mental rotation training while the control group (n = 27) completed crossword puzzles. There were significant group differences on a mental rotation test (but not on other spatial measures) in favor of the intervention group. In addition, there were statistically significant improvements on the intervention group’s mathematics performance associated with tasks that required the completion of missing terms to solve addition and subtraction algorithms. However, there were no differences in the performance of the intervention and control groups in other mathematics tasks associated with number facts and multidigit calculations.

Hawes et al. (2017) conducted a ten-lesson training program with six intervention classes and three active control classes over a school year. The intervention program, implemented by classroom teachers, comprised five geometry-based activities and five spatial-challenge activities that encouraged spatial visualization and the recall of visual-spatial information. As many as 47 hours were devoted to these activities across the 32-week intervention. The active control group undertook an inquiry-based science unit. The intervention group had statistically significant gains on three measures of spatial reasoning; spatial language, visual-spatial geometry and mental rotation. Improvements in mathematics were modest, limited to one of three measures associated with recognition of numerals. There were no gains in a non-symbolic comparison test or a measure of number knowledge.

Lowrie et al. (2017) designed and implemented a 10-week program (20 hours) that focused on mental rotation, spatial orientation and spatial visualization skills, and replaced usual mathematics instruction for ten intervention classrooms. Business-as usual control groups (nine classrooms) received standard mathematics instruction from the Australian curriculum. Participants in the intervention program significantly improved and outperformed those in the control group on spatial measures and this improvement generalized to performance on a mathematics test, which comprised problems on geometry, measurement, and number concepts. The intervention cohorts’ improvements were significant for geometry and measurement items but not number concepts. A subsequent investigation in primary schools (Lowrie et al., 2019) demonstrated that a spatial visualization training program significantly improved an intervention cohort’s performance on geometry problems and content-diverse word-based problems but not on mathematics problems that required the interpretation of visual displays (including bar graphs, maps and number lines).

These studies were conducted with students in the early years of formal schooling (Cheng & Mix, 2014; Hawes et al., 2017) or in the primary grades (Lowrie et al., 2017, 2019), where disciplines are not segregated by the timetabling of subjects. Other intervention studies have examined the extent to which spatial training can improve course outcomes for college-level engineering students (see Sorby, Veurink, & Streiner, 2018). Less is known about the effects of spatial training in the secondary years. The current investigation undertakes a classroom-based intervention into a secondary school setting for the first time, drawing on the theoretical framework from Lowrie et al. (2018) and aligned to the spatial constructs in the mathematics curriculum (Ramful, Lowrie & Logan, 2017). Specifically, we address the impact of spatial training on different aspects of mathematics content performance in the secondary school curriculum; namely, geometry and measurement, number and algebra, statistics and probability.
**Present study**

The first goal of this study was to examine the effects of spatial training on both spatial reasoning and mathematics performance in secondary school. Successful spatial interventions in the early years of education (Cheng & Mix, 2014; Hawes et al., 2017; Lowrie et al., 2017) and at college level (Sorby et al., 2018) have been established to support STEM achievement. However, despite the awareness of the “leaking STEM pipeline” that sees large numbers of students dropping STEM disciplines as content difficulty increases (Ellis, Fosdick & Rasmussen, 2016), little research has been conducted in secondary schools to explore the efficacy of spatial interventions given the shift to more complex mathematical content (Attard, 2010).

The present paper examines the effectiveness of a spatial intervention program set within the five step Experience-Language-Pictorial-Symbolic-Application (ELPSA) pedagogical framework (Lowrie & Patahuddin, 2015; Lowrie et al., 2018) in grade 8 classrooms. The study embedded spatial lessons within this framework to support the development of deep level spatial understandings to foster applications for curriculum content beyond spatial tasks. In the ELPSA framework, the first step (Experience) draws on students’ preexisting knowledge when encountering new tasks, such as class discussions around the utility and scale of different kinds of maps before students produce their own maps of a familiar route. Given that spatial language is a helpful precursor when learning spatial concepts (Newcombe & Stieff, 2012), the next phase (Language) promotes appropriate spatial language to facilitate subsequent spatial learning. Links between pre-existing knowledge and new task demands are introduced during the Pictorial and Symbolic steps of the framework, giving students multiple representations on which to build their understanding. Finally, the Application stage provides students with opportunities to apply spatial reasoning beyond spatial tasks (for a more detailed description of the ELPSA framework see Lowrie & Patahuddin, 2015 and Lowrie et al., 2018). The ELPSA framework has successfully supported spatial intervention programs in primary school (Lowrie et al., 2018, 2019) and in mathematics classes more broadly (Lowrie & Patahuddin, 2015).

The ELPSA pedagogical framework provides a structured way to embed spatial reasoning into classroom practice. It contrasts with spatial training studies that rely on training specific tasks with drill and practice (Cheng & Mix, 2014). Given the success of this approach in enhancing both spatial and mathematics performance in primary classrooms (Lowrie et al., 2017, 2018, 2019), here we test the hypothesis that a comprehensive spatial program designed within the ELPSA framework will also be effective in training spatial reasoning in secondary school classrooms with opportunities to impact on mathematics achievement.

A second goal of this study was to explore which aspects of mathematics are affected by spatial training. Stieff and Uttal (2015) conducted an analysis of spatial training studies. They were cautiously optimistic about the efficacy of spatial training for STEM improvement. They concluded that much of the successful transfer to date has been limited to STEM tasks that are highly spatial, for example, mechanical drawing (Sorby, 1999), suggesting that spatial training might be most effective for improving more spatial aspects of mathematics. There is a sustained body of literature in mathematics education that associates geometric content with spatial reasoning (Battista & Clements, 1988; Clements & Battista, 1992; Lowrie & Logan, 2018). Although there are some studies that point to the importance of spatial reasoning in non-geometric content-related mathematics, including the number-line (Gunderson, Ramirez, Beilock & Levine, 2012) and missing addend subtraction problems (Cheng & Mix, 2014), most of this work is described in relation to early years mathematics only (see Mulligan, Woolcott, Mitchelmore & Davis, 2018; Verdine, Golinkoff, Hirsh-Pasek, & Newcombe, 2017). Moreover, previous intervention studies have not produced improvements in mathematics beyond geometric and measurement items. For example, Lowrie et al. (2017) found significant improvements for geometry and measurement tasks, but not for number problems.
Although previous studies have not found evidence that spatial interventions improve problems that include graphics, it is worth persevering with studies that examine students’ graphic decoding skills. National and international testing bodies are using these types of problems increasingly to assess students’ mathematics understanding (Lowrie & Diezmann, 2009), usually replacing the more traditional word-based problems. Graphic displays such as graphs, maps and diagrams are highly conventional, requiring the multimodal interpretation of several specific graphic conventions such as keys, legends, axis and labels (Bresciani & Eppler, 2009; Mix & Cheng, 2012). In fact, such graphics problems may need to be taught explicitly (Lowrie, Diezmann, & Logan, 2011; Moore, 1993). Consequently, we continue to study the effects of spatial training on problems including graphics displays in the present research. Graphic displays may contain both autonomous information essential to solve the problems successfully and/or auxiliary presentations which may provide cues or distractions in decoding (Gagatsis & Elia, 2004). Specifically, we examined how effects of spatial training differ for: 1) different mathematics content areas, and 2) different problem representations.

Hypotheses

Given the success of interventions using the ELPSA approach in primary schools (Lowrie et al., 2017, 2018, 2019) the following hypotheses were formulated for the present study:

1. A classroom-based spatial training program implemented by teachers would result in greater improvement in spatial reasoning for the intervention group compared with business as usual controls.
2. The intervention would result in increased mathematics scores for the intervention group compared with business as usual controls.
   a. Mathematics performance differences in favor of the intervention cohort would be associated with geometry-measurement concepts.
   b. Mathematics performance differences in favor of the intervention would be associated with problems that contained graphic-rich displays.

Method

Participants

A total of 876 grade 8 students (age range 12 years, 6 months – 14 years, 10 months) from nine secondary schools (six intervention, three control) in Canberra, Australia participated in the study. As the testing was conducted during standard lesson times not all students were able to participate in both the pre- and post-testing. Therefore, only students who completed both tests were included in the analysis (N = 641 students). Thirty-two classes participated in the intervention condition (292 female, 235 male; mean age = 13 years, 10 months) and eight classes acted as business-as-usual controls (64 female, 50 male; mean age = 13 years, 10 months). There was no significant difference in class size between the intervention (mean = 21.6 students, S.D. = 3.1) and control groups (mean = 23.3 students, S.D. = 3.1), t(38) = 1.36, p = .18.

In the absence of random allocation, schools were matched on sociodemographic information. All schools in the study were drawn from within the Australian Capital Territory. In Australia, school demographics are reported as an Index of Community Socio-Educational Advantage (ICSEA; Mean = 1000, S.D = 100). Each school’s ICSEA score is calculated based on Australian Bureau of Statistics data, school location, and proportion of Indigenous students enrolled in the school as well as data on parent’s self-reported income, qualifications and occupation. Thus, a
value on the index corresponds to the average level of educational advantage of the student popu-
lation relative to those of other schools. The ICSEA scores ranging from 989 to 1078 for the
intervention group and 1006 to 1074 for the control group were not significantly different
between intervention and control schools, \( t(7) = .09, p = .93 \).

**Design**

The study employed a quasi-experimental pre-post research design. A request for participation
was sent out through the Government and non-Government Education departments of the
Australian Capital Territory. Department leaders from the intervention schools had to opt-in to
the program. Of these schools half required all grade 8 mathematics teachers to implement the
program to ensure consistency in assessment and reporting while the remaining schools allowed
teachers to self-select into the program. An additional challenge was the recruitment of suitable
control classes. Although some experimental schools offered their nonparticipating classes as con-
trol classes, these classes did not act as controls to preserve the fidelity of the program. Other
schools within the two jurisdictions volunteered to act as control schools—with these schools
offered the incentive of having access to the program after the completion of the study. As ran-
dom allocation was not possible due to ethical restrictions enforced by the governing educational
directorates, the distribution of government and non-government schools within each cohort (3:3
intervention, 2:1 control) were as comparable as possible (see Stieff & Uttal, 2015). Pretest meas-
ures were compared across intervention and control groups to ensure equivalence.

**Spatial reasoning intervention**

**Spatial intervention program**

Eight mathematics department heads attended two days of professional development in 2017
where they were instructed on spatial reasoning, the spatial underpinnings of mathematics, and
the ELPSA framework (Lowrie et al., 2018; Lowrie & Patahuddin, 2015). During a subsequent
two-day workshop they worked with the research team to develop spatial lessons that fit with
curriculum outcomes. These lessons were trialed in their respective schools. Based on feedback
from these trials, the authors finalized a series of lessons following the ELPSA learning cycle that
focused on spatial ability and the spatial reasoning within mathematics across the three constructs
of mental rotation, perspective taking and spatial visualization. These lessons are outlined in
Table 1. The intent of the lessons was not to provide instruction of mathematical content; there-
fore, any noticeable improvements to mathematics were not simply as a result of newly designed
mathematics lessons.

Each lesson was framed around the ELPSA learning framework (Lowrie et al., 2018; Lowrie &
Patahuddin, 2015) depending on the affordances of the lesson. For example, the 2D Mental
Rotation lesson relied heavily on language and pictorial representations while the 3D Mental
Rotation lesson provided opportunities to move towards symbolic understanding and application.
The classroom teachers were encouraged to emphasize visualization processes throughout the
lessons with a focus on having the students visualize spatial transformations to make thoughtful
predictions before observing the results of these transformations. In this way students and teach-
ers were encouraged to develop their spatial vocabulary and explore multiple approaches for tack-
ling problems.

**Professional development workshop**

The 25 teachers\(^1\) in the intervention condition participated in 2 x 2-hour workshops. The first
session familiarized the teachers with spatial reasoning, its components and the ELPSA learning
framework. The second session introduced teachers to the lessons. Both sessions were led by the authors and provided opportunities for teachers to engage with the theory, materials and ask questions. The sessions took place within a month of intervention implementation which allowed time for teachers to review the materials and clarify any areas of concern before implementing the lessons.

**Lesson implementation**

The study ran in term three\(^2\) of the 2018 school year. The spatial program replaced geometry and measurement for the intervention group, whilst the business-as-usual control group continued on with their standard geometry and measurement lessons. The detailed spatial lesson plans and materials were delivered to the intervention schools in the first week of term three once pretesting had taken place. Each teacher delivered twelve hours of the intervention program over the duration of a unit block\(^3\) between pretesting and post-testing. Due to the teaching organizational requirements varying between schools, the intervention was administered across 12 mathematics periods (an additional two periods were set aside for pre- and post-testing).

Geometry and measurement lessons for the control group consisted of a mixture of direct instruction, problem solving and investigation using digital technologies and textbook resources. The learning activities for these units were drawn from the Australian Curriculum guidelines (ACARA, 2015). In Australia, the school curriculum outlines the necessary content to be taught for grade eight students, but the school and classroom teacher determine the structure of the lessons. In grade 8, geometry content includes using units of measurement to find the perimeter of 2D shapes and the volume of rectangular and triangular prisms. Geometric reasoning includes defining congruence of plane shapes using transformations.

### Table 1. Structure of the intervention lessons.

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Spatial construct</th>
<th>Content</th>
<th>Spatial skill targeted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mental Rotation</td>
<td>2D Mental Rotation: Using 2D rotational language and identifying rotations.</td>
<td>Developing 2D rotational language, visualize 2D transformations</td>
</tr>
<tr>
<td>2</td>
<td>Mental Rotation(^*)</td>
<td>3D Mental Rotation: Representing and predicting the rotation of 3-D objects.</td>
<td>Development of 3D physical and mental representational skills</td>
</tr>
<tr>
<td>3</td>
<td>Spatial Orientation</td>
<td>Represent different perspectives of miniature cities.</td>
<td>Understand different perspectives and begin to develop symbolic representations of 3-D space.</td>
</tr>
<tr>
<td>4</td>
<td>Spatial Orientation(^*)</td>
<td>Mapping: Follow and generate a set of directions through visualizing, representing 3-D space on a 2-D plane.</td>
<td>Using visualization to explore routes and directions. Develop and communicate spatial representations at different scales.</td>
</tr>
<tr>
<td>5</td>
<td>Spatial Visualization</td>
<td>Nets and Objects: Constructing shapes from nets, visualize alternate sides of a 3-D object.</td>
<td>Visualization of complex spatial transformations, reflection on spatial processes.</td>
</tr>
<tr>
<td>6</td>
<td>Spatial Visualization</td>
<td>Shape Decomposition: Exploring different ways of breaking down unusual geometric shapes.</td>
<td>Developing awareness of embedded shapes and spatial relationships within polygon shapes (lesson 6) and circular shapes (lesson 7).</td>
</tr>
<tr>
<td>7</td>
<td>Spatial Visualization</td>
<td>Circle Decomposition: Exploring area of unusual circular shapes using visualization and disembedding.</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Spatial Visualization(^*)</td>
<td>Translations: Multi-step transformations using visualization.</td>
<td>Pictorial and symbolic representations of transformations on a 2D plane.</td>
</tr>
<tr>
<td>9</td>
<td>Spatial Visualization</td>
<td>Origami: Activities drawing on spatial skill and language development from previous lessons.</td>
<td>Integrate and apply spatial knowledge through visualizing and communicating about spatial transformations.</td>
</tr>
</tbody>
</table>

\(^*\)Note. One lesson from each construct was delivered over two lessons to allow for further progression towards Application of the spatial construct.
Both intervention and control groups delivered content from the Australian Curriculum associated with number and algebra. No instruction was given to control or intervention group teachers about the delivery of these units. As with the control group geometry and measurement content, these lessons were drawn from the Australian Curriculum guidelines (ACARA, 2015). In grade 8 this includes carrying out the four operations with rational numbers and integers and simplifying algebraic expressions involving the four operations.

**Program fidelity**

Intervention teachers were supported throughout the program, and implementation was gauged in four ways. First, each intervention lesson in the program was observed by members of the research team at multiple sites with different teachers delivering the lesson each time. Second, all schools and 93% of teachers were observed on at least one occasion by one or more members of the research team and a record was taken of adherence to the lesson plan and student engagement with the program. For each site visit observers rated the teachers’ execution of the prescribed lesson on a Likert-type scale from 0 (not observed) to 4 (clearly evident). Observer ratings from 44 site visits produced a mean of 3.61 (S.D. = 1.12). Third, support was offered to teachers throughout the program via the site visits and email to assist with any challenges to implementation. Finally, students’ workbooks were collected at the completion of the program to identify the completed tasks for each lesson. A record was taken of lessons implemented and completed, along with evidence for the application of the ELPSA framework, as students recorded the various elements of the framework. As a result of these steps, we had evidence for implementation of the intended teaching program.

**Materials**

**Mathematics**

The mathematics test (MathT) was developed to reflect appropriate content present in the year 8 mathematics curriculum (ACARA, 2015). These items were modeled on large-scale assessment items and were completed without the aid of a calculator. The MathT included content from three curriculum strands, Geometry and Measurement (12 items; subscale alpha = .52), Number and Algebra (6 items; subscale alpha = .58) and Statistics and Probability (4 items; subscale alpha = .48). Large-scale mathematics assessment in Australia is undertaken as a measure of Numeracy (i.e., real world application of mathematical knowledge). Therefore, the test did not test explicit content taught in either the intervention or control classes, nor the common number-algebra lessons. The MathT contained 22 items across three different problem representation types. The breakdown of the items and examples are presented in Table 2. Each student received the same order of items (static-graphic response, dynamic-graphic response, word problems). The static-graphic and word problems included both multiple choice and text entry questions. Dynamic-graphic response items required different types of interaction with the graphics in the problem in order for students to represent their answers. These included drag and drop answers, moving sliders on an axis, rotating answer graphics and route tracing. Students were given 25 minutes to complete all 22 items. Each item received a maximum score of 1, however some of the dynamic items were given partial marks for being within a reasonable range of the correct answer (MathT total possible score range 0-22). Cronbach’s Alpha for the MathT was .74; while the test-retest reliability was .79.

**Spatial measures**

In light of the scale of the intervention and time restraints on data collection, limited measures of spatial reasoning were available for this age group. Therefore, three separate cognitive tests were used to measure the spatial abilities of interest as indicators of spatial reasoning.
### Table 2. MathT item breakdown.

<table>
<thead>
<tr>
<th>Response Type</th>
<th>Static-Graphic</th>
<th>Dynamic-Graphic</th>
<th>Word Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of items</td>
<td>10</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Example item</td>
<td>Kim throws two standard six-sided dice at the same time. Which point on the number line shows Kim’s chances of throwing an even number on both dice?</td>
<td>A car leaves the garage on the road heading East. It crosses one intersection and then turns right, traveling for 3.5 kilometers, when it reaches the next intersection it turns onto the street heading south-east and continues ahead until it reaches its destination. Trace the route to show which destination the car reaches at the end.</td>
<td>Three coworkers were tasked with wrapping up boxes in preparation for the big Christmas sale. Emma wrapped 11 more boxes than Jake. Jake wrapped 7 more boxes than Matt. In total they wrapped 85 boxes. How many boxes did Matt wrap?</td>
</tr>
</tbody>
</table>

Sample item content strand | Statistics-Probability | Geometry-Measurement | Number-Algebra |
**Mental rotation.** A modified Card Rotation Test (CRT; Ekstrom, French, Harman & Dermen, 1976) presented one reference shape and ten comparison images per screen (Figure 1). Students were instructed to select only the options that were the same (i.e., rotated) as the reference image and complete as many items as possible in the allocated 3 minutes. Scores for the CRT ranged from 0-80 with all correctly selected items and all correctly unselected items given a score of +1. Cronbach’s alpha for the modified test was .94.

![Figure 1. Sample CRT item.](image)

![Figure 2. Sample SOT item.](image)
**Perspective taking.** A digital version of the Perspective Taking Test (SOT; Hegarty & Waller, 2004) was developed using the original 12 items. For each item the static array was presented, and students were instructed to imagine they were standing at one point in the array, facing a second point and then position a dotted line within a circle to indicate where a third point would be located (Figure 2). Students received a score of +2 if they were within 15 degrees either side of the correct answer and a score of +1 if they were outside of this 30 degree zone but within the 90 degree answer range. The range of scores was 0-24. Cronbach’s alpha for the SOT was .81. Students were given 5 minutes to complete the SOT.

**Spatial visualization.** A modified version of the 10-item Paper Folding Test (PFT; Ekstrom et al., 1976) was developed within the online testing platform. The ten original items were redrawn by a graphic designer and presented individually on the screen. Students selected the multiple-choice image they felt was correct before the screen moved to the next item, completing as many items as possible within three minutes. Each correct item was given a score of 1 (the range for the PFT was 0-10). Cronbach’s alpha for the PFT was .57. Although less than ideal, the low reliability is not unexpected due to the small number of items (Field, 2009). Students were given 3 minutes to complete as many items as possible in the PFT. A sample item is presented in Figure 3.

Individual test scores were calculated using raw scores, however, to assess overall growth in spatial reasoning, giving equal weight to each of the tests, scores for each test were converted to give a score out of 10 (where 10 indicated a perfect score on the test and 0 indicated no correct items), thus a total potential spatial score equal to 30. The test-retest reliability (alpha) for the spatial test was .86 across all students (i.e., both intervention and control cohorts).

**Procedures for the pre- and Post-Testing**

Students completed the same assessment at pre- and post-test. The pretest was completed within the first two weeks of term for all classes and was administered by teachers in their classrooms. All assessments were completed during class time in the web browser of students’ personal devices via a secure website. Demographic information was collected before students began the assessment. The order of assessments was randomized by test type (i.e., mathematics, spatial ability tests) but the order of the spatial tests remained the same for all students (i.e., spatial visualization, mental rotation, perspective taking). Students were presented with instructions before each test for an unlimited period of time. Once each test commenced, students were able to move between items using controls in the browser up until the test timed out.

![Figure 3. Sample PFT item.](image-url)
Access to the assessment site was not available during the intervention period. Post-testing was completed in the final three weeks of term after a minimum of six weeks of instruction (either intervention or business-as-usual). As in the pretests, students completed the post-tests on their personal devices under the supervision of their classroom teacher.

**Results**

The results are presented in three sections. The first section presents preliminary analysis to demonstrate group equivalence. The second section reports on the effect of the spatial intervention program on spatial scores. Since the research design contained nesting structures of students within classrooms, a multilevel (hierarchical) modeling approach was adopted to analyze group differences on treatment gains. The third section analyzed the effect of the program in relation to students’ mathematics scores. Hierarchical linear modeling (HLM) was again conducted, within a design that had students nested within classrooms at Level 1. Finally, analysis was conducted of the impact of the program on different mathematics content topics and problems involving different types of graphic displays.

**Preliminary analysis**

Due to the quasi-experimental nature of the research design, screening analysis was performed to ensure equivalence of the intervention and control groups at pretest. Results from the analysis revealed no difference between the mean scores of the intervention and control groups on the spatial pretest \( t(577) = 1.19, p = .26 \), or the MathT pretest \( t(753) = .28, p = .78 \). This speaks to the comparability of the intervention and control groups in terms of preexisting spatial skills and mathematics performance.

There were no significant gender differences found at pretest for spatial reasoning, \( t(580) = 1.07, p = .29 \) or MathT, \( t(749) = 1.64, p = .10 \). No significant gender differences were found in analysis of gain scores when adjusting for multiple comparisons; \( t(83) = 1.59, p = .12 \) (spatial) and \( t(97) = .78, p = .78 \) (MathT) for the control group, and \( t(450) = .19, p = .85 \) (spatial) and \( t(502) = 2.14, p = .03 \) (MathT) for the intervention group. Therefore, the analysis in the present study focused on the effectiveness of the intervention program independent of gender.

**Effectiveness of the spatial reasoning program on spatial performance**

We analyzed group differences (posttest scores – pretest scores) using a two-level HLM model (students within classrooms) with condition dummy coded (1 = intervention and 0 = control) to determine the direct effects of the intervention. Results from the HLM for pretest-posttest gains revealed statistically significant differences between control and intervention groups for spatial reasoning, \( F(33) = 49.69, p < .001, d = .41 \). On average, students in the intervention group gained 13% more than the control group across the intervention period (see Table 3).

Given the significant effects on combined spatial ability, post hoc analyses were conducted to determine the extent of improvement across the respective spatial constructs. There were statistically significant differences for mental rotation, CRT \( F(41) = 7.57, p = .004, d = .32 \); and perspective taking, SOT \( F(42) = 9.08, p = .009, d = .43 \). For spatial visualization, differences were in favor of the intervention group, but were not statistically significant PFT \( F(40) = 3.56, p = .066 \). Descriptive statistics are presented in Table 3.

**Effectiveness of the spatial intervention program on mathematics performance**

Results from the HLM models for pretest-posttest gains revealed gain scores greater than 0 for each group on Math Performance (MathT, see Table 4 for observed mean gains). On average,
students in the intervention group gained 6.5% more than the control group, $F(38) = 15.09, p < .001$, $d = .38$ an indication that, as hypothesized, spatial training impacted on mathematics performance.

**Curriculum Content Analysis**

Additional hierarchical models were conducted to examine specific math content improvement from the treatment program. On average, students in the intervention group had statistically significant increases in performance relative to the control group for both Geometry-Measurement problems $F(39) = 11.45, p = .002, d = .35$ and Number-Algebra problems $F(47) = 10.17, p = .003, d = .26$. With a Bonferroni correction method applied (adjusted alpha = .015) there was no statistically significant difference in performance between the groups for Statistics-Probability problems $F(49) = 5.42, p = .024$ (see Table 5 for improvement by math content categories).

**Problem Representation (Graphic Displays) Analysis**

Further analysis was undertaken to determine whether the intervention program differentially impacted students’ performance on mathematics problems that were displayed in different representations (collapsing over different mathematics content). Students who participated in the intervention group had statistically significant increases in performance relative to the control group for Static-Graphic problems $F(31) = 3.68, p = .001, d = .36$; Dynamic-Graphic problems $F(31) = 6.82, p = .001, d = .37$; and Word problems $F(35) = 2.59, p = .014, d = .31$. Descriptive statistics are reported in Table 6.

**Discussion**

Building on previous research concerning the effectiveness of spatial training programs on students’ mathematics performance in primary school contexts (Cheng & Mix, 2014; Hawes et al.,
2017; Lowrie et al., 2017, 2019), this study examined the impact of spatial training on students in secondary school (Grade 8).

We hypothesized that the intervention groups’ spatial reasoning would increase compared to that of the control group after the implementation of a 12-lesson program that embedded spatial reasoning explicitly into a pedagogical framework. This hypothesis was supported. Specifically, the intervention significantly improved mental rotation and perspective taking performance. These findings support recent studies that found improvements in these spatial abilities with younger students (Lowrie, et al., 2017, 2018). The fact that the intervention group’s spatial visualization scores did not improve significantly compared to that of the control group needs to be considered with caution, given the low reliability of the paper folding instrument.

We further hypothesized that the program would improve the intervention group’s mathematics performance when compared to the control group who received the standard mathematics instruction. This hypothesis was also supported. Although we predicted that effects of the spatial intervention would be most evident in geometry problems and problems with graphics, in fact, its effects extended to number and algebra content, along with graphic problems represented in both static and dynamic form as well as more traditional word-problem representations. Notably, in contrast with results from a younger cohort (Lowrie et al., 2017), the findings from this study did extend beyond spatial-type geometry items to number and algebra.

Although several intervention studies have established significant effects of a spatial training program on school-aged students’ mathematics performance (see Cheng & Mix, 2014; Hawes et al., 2017; Lowrie et al., 2017, 2019), this is the first to demonstrate such results in secondary-school classrooms led by classroom teachers and robust to the differences in classroom practices and environments. The current investigation embedded spatial reasoning with mathematics

| Table 5. Curriculum content pretest, post-test and gain scores. |
| --- | --- | --- | --- |
|  | Control |  | Intervention |
|  | $M$ | $SD$ | $M$ | $SD$ |
| Geometry and Measurement (12 items) |  |  |  |  |
| Pretest | 3.37 | 2.06 | 3.41 | 2.05 |
| Post-test | 3.32 | 1.94 | 4.28 | 2.35 |
| Gain | −.05 | 2.14 | .87 | 2.18 |
| Number and Algebra (6 items) |  |  |  |  |
| Pretest | 1.41 | 1.35 | 1.17 | 1.16 |
| Post-test | 1.37 | 1.30 | 1.58 | 1.36 |
| Gain | −.07 | 1.20 | .42 | 1.22 |
| Statistics and Probability (4 items) |  |  |  |  |
| Pretest | .64 | .79 | .65 | .87 |
| Post-test | .46 | .66 | .67 | .99 |
| Gain | −.20 | .80 | .05 | .92 |

| Table 6. Problem representation pretest, post-test and gain scores. |
| --- | --- | --- | --- |
|  | Control |  | Intervention |
|  | $M$ | $SD$ | $M$ | $SD$ |
| Static – Graphic (10 items) |  |  |  |  |
| Pretest | 2.95 | 1.80 | 2.71 | 1.76 |
| Post-test | 2.70 | 1.70 | 3.09 | 1.92 |
| Gain | −0.19 | 1.73 | 0.38 | 1.73 |
| Dynamic – Graphic (7 items) |  |  |  |  |
| Pretest | 1.38 | 1.25 | 1.49 | 1.27 |
| Post-test | 1.37 | 1.27 | 1.95 | 1.45 |
| Gain | 0.01 | 1.16 | 0.47 | 1.34 |
| Word Problems (5 items) |  |  |  |  |
| Pretest | 1.38 | 1.03 | 1.33 | 1.11 |
| Post-test | 1.33 | 1.05 | 1.61 | 1.23 |
| Gain | −0.07 | 1.18 | 0.23 | 1.31 |
applications into daily teaching practices (Newcombe, 2013) rather than teaching spatial constructs in isolation (e.g., mental rotation skills from the Cheng & Mix, 2014 study). The improvements in the intervention group’s mathematics performance were not restricted to mathematics problems generally associated with spatial reasoning; namely, geometry-content problems. The intervention cohort had significant gains in a range of number- and algebra-content problems not associated with learning activities or content in the intervention program. In addition, performance improvements in favor of the intervention group were attained across both static and dynamic graphic and non-graphic (word) problems. These findings are noteworthy since the intervention students were not taught graphic conventions normally required to improve performance of graphic-rich tasks (Bresciani & Eppler, 2009) or heuristics important for decoding word problems (Eisenmann, Novotná, Přibyl, & Břehovský, 2015). We propose that the intervention cohort’s heightened spatial reasoning provided affordances that supported the challenges associated with solving relatively difficult mathematics tasks. For the graphic-rich tasks this would have included the capacity to decode specific graphic conventions, while for the word problems we suggest that the students would have been more likely to use visual representations to support meaning (Lowrie & Logan, 2018).

**Future directions and limitations**

The present study extends evidence that a spatial training program based on the ELPSA framework (Lowrie & Patahuddin, 2015) and focusing on mental rotation, perspective taking and spatial visualization is effective in enhancing mathematics performance. The utility of this program has now been demonstrated across elementary to middle school classrooms in three separate studies (Lowrie et al., 2017, 2019) including the present study. Future studies are needed to understand the mechanism by which this approach improves mathematics performance. There is a growing body of evidence suggesting that well-developed spatial abilities allow students to utilize supportive “tools” for problem solving, including gesture and spatial language, which are helpful when solving novel mathematics problems (Young, Levine, & Mix, 2018). We speculate that students who develop higher levels of spatial reasoning based on our intervention are able to use a range of representations which add to the flexibility and fluency of novel problem solving. Future studies should tease apart the unique contribution of the ELPSA pedagogy to improvements, separate to the spatial content.

Spatial reasoning skills are also a good predictor of student performance on mathematics problems that contain embedded graphics (Lowrie & Diezmann, 2011) and in the present study spatial training enhanced performance on these types of problems. More focused studies that examine the direct impact of spatial training on performance across different types of mathematical representations are required to better ascertain the affordances of spatial reasoning for novel problems and how spatial training affects students’ use of a range of strategies in problem solving.

Despite best efforts to ensure a representative sample with equivalent intervention and control groups the intervention classrooms greatly outnumbered the control classrooms. As random allocation was not permissible within the ethical restrictions of the educational jurisdiction this was beyond the authors’ control. This may be a consequence of the growing public awareness of spatial reasoning for STEM achievement (Newcombe, 2017) or the support of the educational jurisdiction in conducting such a large-scale program within schools. Future studies may look at ways to enhance experimental control without compromising on issues of ethics or equity. Future designs could move from business-as-usual control groups to active control groups in order to alleviate some of the sampling challenges we have experienced.

The spatial constructs delivered through this intervention were closely aligned with the spatial demands within the corresponding curriculum (Ramful et al., 2017). However, given the efficacy of spatial training for college level courses (Sorby et al., 2018) and life skills more generally (National
Research Council, 2006), it may be argued that limiting the spatial intervention to mathematics classrooms is short-sighted. Future work may benefit from expanding the program to include spatial abilities with applications into other content areas such as science, geography, technology and beyond. Alternatively, the growing STEM agenda beyond disciplines (Lowrie, Leonard & Fitzgerald, 2018) may provide opportunities to incorporate a broader range of spatial reasoning skills.

Conclusion

This paper provides evidence for the effectiveness of a spatial reasoning intervention program with mathematics applications in secondary school classrooms. The program was designed within a pedagogical learning framework (ELPSA) that supported program fidelity—which became crucial given the varied teaching approaches of the intervention cohort. Despite this, the intervention was robust to the effects of classroom implementation across teachers and schools. There were statistically significant differences between intervention and control groups, in favor of the intervention classes exposed to 12 periods of spatial training. Improvements in mathematics performance went beyond the anticipated geometry and measurement concepts to include number and algebra problems. Although specific conventions or problem-solving heuristics were not taught to the intervention group, improvements were observed across problems represented in both graphic and non-graphic (word problem) displays. This study adds to the growing body of evidence that spatial training has implications beyond spatial tasks for supporting mathematics achievement.

Notes

1. Seven teachers taught the intervention lessons in two separate classes.
2. The Australian school year consists of four ten-week terms that run from January to December.
3. Australian Secondary Schools operate in time blocks for content delivery. These blocks vary between schools, so while some schools chose to deliver the entire program in one five-week block to allow for another complete content block within the term, other schools delivered the program for 1-2 lessons per week while delivering an additional mathematics lesson concurrently.
4. Ethical guidelines restricted collection of student demographic information to age, identified gender, teacher name and school.

Acknowledgements

The authors wish to thank Robyn Lowrie, Alex Forndran, Ajay Ramful and Sitti Pattahuddin for assistance with development of the intervention program and classroom observation. Additional thanks to Australian Capital Territory teachers for assistance in developing the intervention program and resources, Destina Winarti for help with classroom observation and Josh Crowley for his work in developing the digital assessment platform.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Australian Research Council (grant number DP150101961).

ORCID

Tom Lowrie http://orcid.org/0000-0003-3042-822X
Danielle Harris http://orcid.org/0000-0002-6189-3522
Tracy Logan http://orcid.org/0000-0001-5283-286X
References


Ellis, J., Fosdick, B. K., & Rasmussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. PLoS One, 11(7), e0157447. doi:10.1371/journal.pone.0157447


