DOI: 10.1111/cdev.13963

CHILD DEVELOPMENT

REGISTERED REPORT

Hands-On: Investigating the role of physical manipulatives in spatial training

Katie A. Gilligan-Lee[1,2,3] | Zachary C. K. Hawes[4] | Ashley Y. Williams[2] | Emily K. Farran[2,3] | Kelly S. Mix[5]

[1]School of Psychology, University College Dublin, Dublin, Ireland
[2]School of Psychology, University of Surrey, Guildford, UK
[3]Centre for Educational Neuroscience, Birkbeck, University of London, London, UK
[4]Department of Applied Psychology & Human Development, Ontario Institute for Studies in Education, University of Toronto, Toronto, Ontario, Canada
[5]Department of Human Development and Quantitative Methodology, University of Maryland, College Park, Maryland, USA

Correspondence
Katie A. Gilligan-Lee, School of Psychology, University College Dublin, Dublin, Ireland.
Email: katie.gilligan-lee@ucd.ie

Funding information
British Academy, Grant/Award Number: SRG20\201210; University of Surrey Faculty Research Support Fund

Abstract
Studies show that spatial interventions lead to improvements in mathematics. However, outcomes vary based on whether physical manipulatives (embodied action) are used during training. This study compares the effects of embodied and non-embodied spatial interventions on spatial and mathematics outcomes. The study has a randomized, controlled, pre-post, follow-up, training design (N=182; mean age 8 years; 49% female; 83.5% White). We show that both embodied and non-embodied spatial training approaches improve spatial skills compared to control. However, we conclude that embodied spatial training using physical manipulatives leads to larger, more consistent gains in mathematics and greater depth of spatial processing than non-embodied training. These findings highlight the potential of spatial activities, particularly those that use physical materials, for improving children's mathematics skills.

Spatial thinking requires the perception of the location and dimension of objects and their relations with one another. It is fundamental to independent living as it is required for navigating through the world, manipulating and using tools, and communicating using spatial language and gesture (Newcombe, 2018). Spatial thinking is one aspect of cognition that is malleable through intervention and meta-analysis findings show that this malleability is particularly high in children (Hedges's g=0.61) (Uttal et al., 2013). This demonstrates that spatial skills are a suitable intervention target for the primary (elementary) school classroom. Spatial thinking is also fundamental to mathematics performance. Spatial-mathematical relations have been reported across a multitude of cross-sectional and longitudinal studies in childhood (e.g., Gilligan et al., 2017; Mix et al., 2016, 2017; Verdine et al., 2014).

Recent studies have extended correlational work to show that spatial intervention leads to improvements in mathematics outcomes (for a meta-analysis see Hawes et al., 2022). However, despite its potential in the development of childhood mathematics skills, spatial thinking is often absent from primary school mathematics curricula. This is due, in part, to a lack of clarity on how best to introduce spatial skills into the classroom. One prominent feature by which previous studies differ is whether they include embodied action in spatial training. In this study, we establish whether embodied spatial training using physical manipulatives (*Hands-On* training) leads to a greater depth of processing in the spatial domain

Abbreviations: ANCOVA, Analysis of Covariance; ANOVA, Analysis of Variance; BF, Bayes Factor; DV, dependent variable; Math, mathematics; VSWM, visuospatial working memory.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2023 The Authors. *Child Development* published by Wiley Periodicals LLC on behalf of Society for Research in Child Development.

and greater transfer of gains to mathematics, compared to non-embodied training (*Hands-Off* training) and active control training.

What is embodied cognition?

Embodied action is the interaction of a learner with some aspect of their physical environment, that is, through using physical manipulatives, concrete materials, or models (Barsalou, 2008, 2010). Theories of embodied cognition propose that individuals conceptualize ideas with grounded representations, that is, representations that involve sensory-motor encoding, rather than amodal representations (Barsalou, 2008; Glenberg & Kaschak, 2002; Pecher & Zwaan, 2005). Once formed, perception-action representations can be re-activated in the same neural circuits that originally experienced the movements or physical sensations, even without the presence of the original physical stimuli (e.g., Pulvermüller, 2005).

There is broad evidence to suggest that embodied action offers an advantage to learning in educational contexts (see Alibali & Nathan, 2018, or Glenberg, 2008 for reviews). For example, when 8- to 12-year-olds were asked to solve arithmetic problems without using their fingers to calculate, the neural regions associated with these finger movements were nonetheless activated (Berteletti & Booth, 2015). Prior research has also demonstrated that making physical movements during learning improves memory for specific content, such as physically spinning a wheel before learning about angular momentum in physics (Kontra et al., 2015) or acting out a story with toys while reading (Glenberg et al., 2004).

The same mechanism may confer an advantage for spatial training paradigms that offer practice at physically manipulating concrete materials, such as rotating a card cut-out to choose which picture shows a rotated object in a mental rotation task. Such physical actions might help learners form a multimodal cognitive representation of spatial relations that is rooted in action (Barsalou, 2008; Glenberg, 2008), and this embodiment could make the underlying spatial representations better grounded and more easily accessed. Indeed, such effects have been demonstrated for infants performing mental rotation tasks after handling the same objects prior to testing (Mohring & Frick, 2013). If asked to solve a mathematics problem following embodied spatial training, even in the absence of any concrete support (physical materials), participants may more readily activate the same neural circuits and recruit a mental simulation of these spatial movements for use in problem-solving. As we have argued elsewhere (e.g., Mix, 2019), spatial circuits may be activated automatically because of the processing overlap between spatial thought and mathematical thought, or strategically as a way to ground the mathematics symbols in an embodied context that helps learners choose an appropriate algorithm (e.g., imagining objects moving together to form a set as a metaphor for addition). In either case, there is reason to believe that spatial training based on embodied experiences may be particularly powerful because it results in stronger, more accessible spatial representations.

No known studies have explicitly investigated the role of embodied action in cognitive training in the spatial domain. Here we will address this gap in the literature. From a theoretical perspective, understanding the role of embodiment in spatial training will help to develop our mechanistic understanding of not only *if* but *why* there is a causal effect of spatial thinking on mathematics outcomes in childhood. Practically speaking, our findings will help to refine the optimal design of spatial training as a means of developing mathematics skills. This will have substantial implications for the spatialization of mathematics learning and instruction.

Spatial interventions using concrete materials

In the past 10 years, several studies have provided convincing evidence that spatial training can confer benefits to mathematics domains (for a meta-analysis see Hawes et al., 2022). These studies have succeeded in demonstrating a causal effect of spatial skill on mathematics (Pearl, 2009). However, the cognitive mechanisms that underpin this effect remain unknown and there is no clear consensus on the optimal design of spatial training interventions. Most notably, with relevance to the current study, there are mixed findings on whether embodied action should be used in the delivery of spatial training. In the aforementioned meta-analysis (Hawes et al., 2022) training delivery method (concrete materials vs. non-concrete) was identified as one of the only significant moderators of the effectiveness of spatial training. Spatial training that included concrete materials led to larger gains in mathematics (Hedges's $g = 0.416$) compared to training that had no concrete component (Hedges's $g = 0.052$). However, the individual features of studies within these two groups differed substantially. In the next sections, we will review and compare specific studies that include embodied action (concrete materials) in the delivery of spatial training to those that do not.

Studies that use concrete materials in the delivery of spatial training predominantly report positive effects, that is, transfer of spatial training gains to math outcomes. In a seminal study, Cheng and Mix (2014) found that 6- to-8-year-olds who completed 40 min of mental rotation training showed significant improvement on a calculation task compared to controls. Spatial training was delivered using concrete manipulatives (card cut-outs) and participants were instructed to move the shapes provided to check their answers. This work prompted a series of subsequent studies in this domain.

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

For example, Lowrie et al. (2017) investigated a 10-week teacher-led intervention in 10-to-11-year olds that focused on training spatial visualization, mental rotation, and spatial orientation skills using drawing, navigating, folding, and cutting activities, that is, activities requiring concrete materials. The study found that children in the spatial intervention group had significant improvements in spatial and mathematics performance (geometry and arithmetic) compared to a business-as-usual control group. Similarly, in a longer, 32-week classroom intervention with 4-to-7-years-olds, Hawes et al. (2017) reported that spatial visualization training using concrete materials such as tiles, multi-link cubes, and magnetic shapes, led to gains in mathematics (symbolic number comparison) and spatial skills compared to an active control group. Taken together, these findings suggest that spatial training using physical manipulatives leads to positive effects on mathematics.

Spatial interventions that do not use concrete materials

Studies that do not use concrete manipulatives in the delivery of spatial training report more varied results in terms of transfer to mathematics. Some studies with non-embodied designs have demonstrated far transfer to mathematics performance. For example, Cheung et al. (2019) found that 6- to-7-year old children who completed a computer-based, mental rotation training program in their own homes had gains in arithmetic performance compared to an active control group who completed computer-based literacy activities. The training lasted 50 min overall and was entirely online with no concrete materials provided. Similarly, Gilligan, Thomas, and Farran (2019) investigated computer-based spatial training in 8-year-olds and found that simply watching a short video outlining spatial processes led to improvements in number line skills (spatial scaling video) and missing term problems (both rotation and spatial scaling videos) compared to a control group who watched a video on spelling. Finally, Bower et al. (2022) investigated the effects of a digital spatial training program where children completed 50 min of computer-based spatial puzzle activities. They found that this spatial training led to improvements in applied mathematics problems in 3-year-olds compared to a business-as-usual control group. However, the gains were only for children from low-socio-economic-status backgrounds. These findings suggest that in some instances non-embodied spatial training does transfer to mathematics domains, and these computer-based paradigms may offer a more convenient and accessible way of delivering spatial training. Computer-based training is not dependent on additional resources such as blocks or shapes, can be easily administered online in either a school or home setting, and does not necessarily require teacher or parent training.

However, other studies with non-embodied spatial training paradigms have reported conflicting findings. Hawes et al. (2015) delivered approximately 4.5 h of tablet-based mental rotation training to children aged 6 to-8 years old, and despite reporting gains in spatial skills, found no improvements in children's calculation skills compared to an active control group. Similarly, Cornu et al. (2019) delivered 20 visuospatial training sessions (20-min each), including tasks such as mental rotation, embedded figures, and shape closing, to children aged 4 to-7 years old using iPads. Although gains were reported in the spatial domain, there was no transfer to mathematics performance, when the spatial training group was compared to a business-as-usual control group. It is noteworthy that in each of these examples, improvements in spatial skills were reported following spatial training. This suggests that the spatial training delivered was effective, however, the effects did not immediately transfer to mathematics performance.

Overall, the literature summarized above suggests that spatial training using concrete manipulatives may offer an advantage over spatial training that does not use concrete manipulatives. However, these conclusions are based on broad comparisons across different studies with different training paradigms and populations. The closest evidence to date comparing the effects of spatial training with and without concrete manipulatives in a single study was completed by Mix et al. (2021). In their paradigm, Mix and colleagues provided different types of spatial training to two groups of children aged 7 and 12 years. They compared their maths performance to an active control group who engaged in language activities. The spatial visualization training group used concrete materials including card and puzzle pieces to complete three training tasks: a part-whole completion task, a mental rotation task, and tangram puzzles. The form perception and visuospatial working memory (VSWM) spatial training group did not use concrete materials. Instead, they used iPads and pencil and paper workbooks to complete three tasks: a VSWM task, a Corsi block tapping test, and a figure copying task. Both spatial training conditions showed significant performance increases in their broad maths measures, however, the spatial visualization training that included concrete materials produced larger effect sizes than the non-concrete form perception and VSWM training. However, interpreting these findings is complicated by the fact that these differences could also be due to differences in the spatial skills being trained across training groups.

Taken together, no single study to date has used an experimental design to investigate the impact of physical manipulatives in spatial training when all other elements of training are kept uniform. Here, we fill this gap in the literature and explicitly compare embodied (*Hands-On*) to non-embodied (*Hands-Off*) using the same spatial training paradigm.

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Other factors in cognitive training

Despite its importance for the implementation of spatial training in the classroom, no known study of spatial training and mathematics has investigated the durability of gains, that is, how long gains in mathematics persist following spatial training. All known studies in this domain have completed post-testing no later than 10 days after delivering training (Hawes et al., 2022). However, for spatial training to be optimally beneficial in the classroom, it should not only lead to gains in mathematics, but these gains should persist over time. For the first time, we will investigate the durability of gains in mathematics following spatial training, 6 weeks after training. In this way, our study will compare embodied and non-embodied spatial training approaches based on both the *size* of transfer effects from spatial training to mathematics outcomes and the *durability* of any gains reported.

This study also controls for motivational factors including expectation and engagement effects. In doing so, we will strengthen the causal inferences made from this cognitive training work (Boot et al., 2013). As outlined by Green et al. (2019), expectation (placebo) effects occur when the expectation that training will be effective leads to cognitive gains, irrespective of the training delivered. In contrast, engagement effects occur when differences in engagement across training conditions lead to differences in post-test outcomes that are not attributable to training content.

Current study

To date no known studies have explicitly investigated the role of embodied action (use of physical-manipulatives) in spatial training, controlling for other variables, for example, spatial training design and motivational factors. This is a critical next step in establishing the optimum design of spatial training paradigms for the mathematics classroom. The current study investigates whether embodied training using physical manipulatives (*Hands-On* training) leads to larger, more durable gains in spatial and maths outcomes compared to non-embodied training (*Hands-Off* training) and active control training. We also investigate the durability of spatial and math gains following both *Hands-On* and *Hands-Off* spatial training.

METHODOLOGY

Participants

We recruited participants from Year 3 (aged 8.0 ± 0.48 years) from state primary schools in the UK. This age group was chosen as correlational findings show that spatial-mathematical relations are particularly strong at this age (Gilligan, Hodgkiss, et al., 2019; Mix et al., 2016), and previous studies have demonstrated transfer of spatial training gains to mathematics outcomes using both embodied and non-embodied approaches in this age group (e.g., Mix et al., 2021). The sample size was determined using GPower for the largest proposed analysis in the study, that is, for an analysis of covariance (ANCOVA) with one independent variable and three covariates. To achieve power of 80% with an alpha of 0.05 and a medium effect size of 0.25, it was determined that 158 participants were required. To account for drop-off through the study, the target sample size was increased by 15%, and 182 participants were recruited. A medium effect size was chosen as this was the estimated effect for spatial training with concrete materials on mathematics (g = 0.41) reported in a recent meta-analysis (Hawes et al., 2022). A medium effect is also the minimum desirable effect for educational interventions (Hattie, 2009).

The final sample (N = 182) comprised of 49% females. The sample was predominantly white (83.5%), with lower proportions of children who were mixed race (11.5%), Pakistani or Bangladeshi (1.6%), Black, or Black British (1.6%), Indian (1%) and Asian (0.5%). We measured mother's highest level of education as a proxy for socioeconomic status: 17% of mothers had a postgraduate degree or equivalent; 47% of mothers had a university degree or equivalent (undergraduate); 18% had further vocational training; 16% had a school leaving certificate; only 2% had no formal education.

Study design

The study has a randomized, controlled, pre-post, follow-up, training design with school-based data collection. All data was collected between May 2021 and May 2022. Participants were randomly assigned to one of three training groups: *Hands-On* spatial training (n = 63), *Hands-Off* spatial training (n = 59), or *Control* training (n = 60). Note that the mother's highest level of education (SES) was comparable across training groups, χ^2 (10, 182) = 8.93, p = .539. Each group completed four 30-min intervention sessions across a 2-week period. The spatial training groups completed spatial visualization activities either with concrete objects, that is, physical manipulatives made of foam board (*Hands-On* spatial training) or without concrete objects (*Hands-Off* spatial training). The control group completed a vocabulary intervention. To investigate the effects of training, participants completed a test battery of spatial, mathematics, and vocabulary measures (spanning two 30-min sessions) one-week pre-intervention (Time 1) and one-week post-intervention (Time 2). They also completed 1 session of mathematics measures 6 weeks post-intervention, that is, at follow-up (Time 3). At each time point, participants

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

completed mathematics measures prior to spatial measures to avoid possible improvements in mathematics due to spatial training effects. Beyond this stipulation, task order was randomized across participants. For every session (test and intervention), participants were taken out of their classroom in groups of four children supervised by one researcher. Children completed testing and intervention sessions in these groups but worked independently separated by dividers. All training materials and measures are available on our Open Science Framework (OSF) page (https://osf.io/mer9t/).

Spatial training (Hands-On and Hands-Off training)

This study included two spatial training conditions. The procedures for *Hands-On* and *Hands-Off* spatial training were identical with the exception that those in the *Hands-On* spatial training group completed all training sessions with physical manipulatives, while the *Hands-Off* spatial training group was not given any physical manipulatives to support learning. For both groups, each training session included three activities presented in a randomized order. While the activities were the same across sessions, the specific items included in the activities varied. This ensured that participants remained engaged with training. Feedback was provided to both groups during training. For all activities, participants selected their answer(s) by placing a finger on them. All participants were then shown the correct answer(s). For the *Hands-On* spatial training group, participants were prompted to check their answers by first imagining manipulating the shapes and then by using physical cut-outs made from foam board to check their answers, that is, by manipulating the object(s) provided. For the *Hands-Off* spatial training group, participants were prompted to check their answers, that is, "Can you imagine turning this shape in your mind so that it is the same as the target shape?" but were not provided with any physical manipulatives.

Activity 1. Mental rotation

In each training session, participants completed 5 trials of mental rotation. Each trial required participants to determine which two, out of four animals, positioned above a horizontal line, were rotated versions of a target animal positioned below the line. As such, participants completed two rotations in each trial. In each training session, participants completed two rotations at 0°, 45°, 90°, 135°, and 180°, respectively (including equal numbers of clockwise and anticlockwise rotations). The order of trial presentation was randomized across sessions. For each trial, the distractor images included two mirror images of the target animal, rotated to the same degree as the correct answers for that trial. The position of the correct answers on the horizontal axis, and the choice of animal stimuli used for different rotations were counterbalanced. The task stimuli were taken from Neuburger et al. (2011).

Activity 2. Mental transformation

In each session, participants completed 6 trials of mental transformation. The procedure was taken from Ehrlich et al. (2006). In each trial, participants were required to choose which of 4 shapes could be created by joining two target shapes together. Trials varied systematically by rotation and translation. In each session, participants completed two trials at each of 0°, 45°, and 90°. For each degree of rotation, participants completed 1 trial that required translation, that is, the two pieces are not displayed on the same plane, and 1 trial that does not require translation, that is, the two pieces are displayed on the same plane. Trial types were presented in a randomized order. The position of the correct answer was counterbalanced.

Activity 3. Object completion

In each session, participants completed 4 items of object completion taken from Thurstone's Part-Whole Object Completion Task (Thurstone, 1974). For each item, a target image of a square was shown with a portion missing. Participants were required to choose which of four shapes could be rotated and joined to the target shape to create a perfect square. Each session included two trials that required 45° and 90° rotations, respectively. The position of the correct answer was counterbalanced.

Control training (vocabulary training)

The control group completed a vocabulary intervention developed at the University of Oxford as part of the LiFT project http://www.education.ox.ac.uk/research/lift-learning-for-families-through-technology/ (Booton & Murphy, in prep). The intervention aims to improve understanding and use of homonyms, that is, words with dual meanings. The aim of Session 1 is to raise awareness that words can have multiple meanings (homonyms), to make students aware when they do not know the meanings of words, and to introduce strategies for working out the meaning of words in context. In Session 2, children practiced using strategies for working out the meaning of homonyms in context. They also made and evaluated inferences to choose the best guesses. Session 3 aims to reinforce the strategies introduced in Session 2 and to help children evaluate their answers. Session 4 aims

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to bring together all the skills acquired in the previous sessions. Across the four sessions, children completed activities guessing the meaning of words in context. This included simple verbal examples, sorting words in sentences, and activity worksheets. Feedback was provided for all activities.

Measures

The test battery included nine measures. Spatial tasks assessing mental rotation, mental transformation, and object completion were included to measure the near transfer of gains following training, that is, gains in the spatial tasks trained. A mental folding task assessed the intermediate transfer of gains following training, that is, transfer of gains to untrained spatial domains. The battery includes four mathematics tests measuring missing term problems, word problems, calculation, and place value concepts. These measures assessed far transfer of gains, that is, transfer of gains to an untrained cognitive domain. A math composite (an average of the percentage accuracy on the four math tasks) was also calculated to capture children's overall math performance. A composite was not created in cases where participants had missing data for more than 2 out of 4 tasks. A vocabulary task was administered to measure the impact of *Control* training. All measures were computer-based and all instructions were delivered through the computer using earphones. All tasks are openly available at https://gorilla.sc/openmaterials/163616. To measure motivational factors, a paper-based participant expectation measure was completed at the beginning of the first training session and a measure of engagement was given at the end of each training session.

Mental rotation

The procedure for the Mental Rotation Task closely resembles the mental rotation activity described for spatial training activity 1. However, in each trial, participants were only required to identify which one of two animal images (one rotation of the target and one mirror image of the target) located above a horizontal line matched a target image (Gilligan, Hodgkiss, et al., 2019). Participants choose an answer using the computer mouse. They completed four practice trials (with feedback) and 40 experimental trials with no feedback. Practice trials were included to ensure participants understood how to answer the trials. The experimental trials included equal numbers of clockwise and anti-clockwise rotations at 45°, 90°, and 135° (eight trials for each degree of rotation), and eight trials at 180° and 0°. The order of trials was randomized. Percentage accuracy was recorded.

Mental transformation

The procedure for the Mental Transformation Task closely resembles the mental transformation activity described for spatial training activity 2. Participants choose (by clicking on it) which of 4 shapes could be created by joining two target shapes together. Participants completed 3 practice trials with feedback and 16 experimental trials with no feedback. Trials differed systematically by rotation (45° or 0°) and by translation (presented on split planes or the same plane). The three practice trials were added to ensure that participants understand the task in a computer-based format. The position of the correct answer was counter-balanced across trials and the order of item presentation was fixed. Overall performance accuracy was recorded.

Object completion

The procedure resembles the object completion activity described for spatial training activity 3. For each item, a target image of a square was shown with a portion missing. Participants choose (by clicking on it) which of four shapes could be rotated and joined to the target shape to create a perfect square. Participants completed two practice trials at 0°. They then completed 12 experimental trials (equal numbers requiring 45° and 90° rotations) with no feedback. The position of the correct answer was counterbalanced across trials. Performance accuracy was recorded.

Mental folding

The Mental Folding Task requires participants to imagine folds made to a piece of paper, without a physical representation of the folding (Harris et al., 2013). In each trial, participants were shown a shape at the top of the screen with a dotted line and arrow representing where a fold should be made. Participants were required to click on one of four images, shown at the bottom of the screen, to select which image would be created after the fold was made. To ensure that participants understood that the dotted lines represent folding, they completed two practice trials in which they were given a physical card to check their answers, and feedback was provided. Participants then completed 14 experimental trials with no feedback. The position of the correct answer was counterbalanced. Performance accuracy was measured.

Missing term problems

The items were modified from Hawes et al. (2015) and Cheng and Mix (2014). In each trial, participants were asked to complete the missing number(s)

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in a mathematical calculation, for example, 3+_=7. Participants completed two practice items with feedback, and a further 14 test items where no feedback was provided. Items were presented in order of increasing difficulty. Approximately equal numbers of addition versus subtraction items, and single versus multi-digit numbers were included. Percentage accuracy was recorded.

Word problems

Participants completed 2 practice and 12 test items modified from Mix et al. (2021). For each item, participants were asked to type the correct answer to a mathematical word problem that was shown onscreen. The question was also read aloud to participants through their earphones to ensure that reading ability did not influence performance. As this task was originally designed for use in the United States, some of the language, for example, names have been changed for use with children in the UK. No mathematical content has been altered. As the task was originally presented on paper, two practice items were added to ensure that children understood how to answer on the computer. The trials were presented in a fixed order. Percentage accuracy was recorded.

Place value concepts

This task was modified from Mix et al. (2021) who delivered a paper-based version. The task has been shortened from 20 to 12 items. In each item, participants were asked to compare, order, and interpret multi-digit numerals, for example, "Which number has a 4 in the tens place?". They were also asked to order numbers from smallest to largest and to match multi-digit numerals to their expanded notation equivalents (543: 500+40+3). Items were presented in a fixed order. Percentage accuracy was measured.

Calculation task

This task was designed for use in this study. For each of the 14 items, participants were asked to answer a mathematical calculation in a prototypical format, for example, 4+3=X. Equal numbers of addition and subtraction items, and single and multi-digit items were presented in a fixed order. Percentage accuracy was measured.

Vocabulary

Participants completed a vocabulary task assessing their knowledge of homonyms. This task was modified from Booton et al. (2022). In each of the 16 trials, participants were shown a target word on the screen and were asked to choose which one of four images represents the word shown. This word was also read aloud to participants through their earphones. Performance accuracy was recorded.

Expectation and engagement measures

These measures were taken from Gilligan, Hodgkiss, et al. (2019). At the beginning of the first training session, participants' expectations of training were measured using the question, "We are going to be playing some games. How much do you think the games will help you with your maths?". Participants responded by drawing a line on the rating scale provided. Responses were coded as 1–12 based on the position of the line where 1 indicates low and 12 indicates high expectations. After each training session, a participant engagement questionnaire (4 questions; see Gilligan, Hodgkiss, et al., 2019) was administered. Participants responded by drawing a line on the rating scales provided. As described above, responses were coded as 1–12 where 1 indicates low and 12 indicates high engagement with training. Participants were awarded a mean engagement score.

Analysis plan

Our analysis plan including further details on exclusions and missing data are available on our OSF page (https://osf.io/mer9t). This paper is a registered report and the Stage 1 manuscript can be found on the OSF (https://osf.io/mer9t).

Exclusions and missing data

Only participants who completed both test sessions at Time 1 received training. Only participants who completed at least 75% of the intervention sessions (3 sessions) were included. Participants scoring higher than 85% on a given task at pre-test Time 1, were deemed to have reached "ceiling level" performance on the task and were excluded from training analysis for that task only. No data were imputed at Time 2 or Time 3. To achieve the minimum target sample size (n=158), additional participants were recruited (where possible).

Analyses

All analyses were conducted using R. Pearson correlations were completed between all dependent variables (DV's) at Time 1. To investigate the effects of our intervention, we used t-tests to determine whether performance in each group improved significantly, and ANCOVA to determine whether some groups showed

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Correlations between mathematics and spatial measures at pretest (Time 1) reported as Pearson's r

	Calculation	Place value	Word problems	Missing term problems	Mental rotation	Mental transformation	Object completion	Mental folding
Mathematics composite	.87*** (n = 180)	.76***	.89*** (n = 182)	.88***	.42***	.51***	.31*** (n = 182)	.35***
Calculation		.53*** (n = 180)	.69*** (n = 180)	.70*** (n = 179)	.37*** (n = 180)	.49*** (n = 180)	.31*** (n = 180)	.30*** (n = 180)
Place value			.55	.61*** (n = 180)	.26***	.35***	.24***	.28***
Word problems				.71***	.41***	.46***	.26*** (n = 182)	.36***
Missing term problems					.36*** (n = 180)	.42*** (n = 180)	.27***	.24*** (n = 180)
Mental rotation						.37***	.31***	.37***
Mental transformation							.39***	.58***
Object completion								.47***

Note: Note that p < .001 for all correlations (indicated with ***) and all BF_{10} values were greater than 25.17. Unless otherwise stated n = 181.

more improvement than others. We completed separate one-way t-tests (from pre-to-post-test) for each spatial and math DV, for each group. The tests were one-tailed as we hypothesized that training would improve task performance. We next completed ANCOVAs for each spatial and math DV, which allowed us to control for pre-test differences while also comparing post-test outcomes across training groups. A Tukey adjustment was applied to control for between-group comparisons. To investigate the durability of gains in math skills 6 weeks after training, the same analyses were completed using follow-up scores (i.e., at Time 3) instead of post-training scores (i.e., Time 2). We also computed Bayes Factors for all analyses with the BayesFactor R package, (default Bayes factor with a wide Cauchy distribution, scale of effect = 0.707). An advantage of using Bayes factors is that they provide a quantitative estimate of the strength of evidence for the alternative hypothesis BF_{10} (i.e., there are group differences at post-test, controlling for pre-test scores) *and* the null hypothesis BF_{01} (i.e., there are no group differences at post-test, controlling for pre-test scores). In addition, Bayes factors can be used to indicate when there is insufficient power (more data are needed) to make claims about the presence or absence of an effect. Here, we report Bayes factors as they correspond to evidence in favor of the alternative hypothesis compared to the null hypothesis. The following guidelines for interpreting the strength of Bayes factors have been recommended (e.g., see Jarosz & Wiley, 2014): Bayes factors between 1 and 3 = weak/ anecdotal support (not enough evidence to make any substantial claims either for or against the predicted relation); Bayes factors between 3 and 10 = substantial support (enough evidence to make moderate claims about effect); Bayes factors between 10 and 100 = strong evidence (enough evidence to make moderate to strong claims about effect); Bayes factors greater than 100 = very strong or decisive evidence (enough evidence to make strong claims about effect).

RESULTS

Preliminary analysis and correlations

Preliminary analyses found no floor or ceiling effects for any measure at Time 1. Therefore, no tasks were excluded from subsequent analyses. However, across the tasks, several children achieved above 85% at Time 1, which was our cut-off for inclusion in intervention analyses. Therefore, the sample size, and by extension power, was reduced for some analyses. The exact sample size (n) for each analysis is listed in Table 2. Although violations of normality were reported for some measures, parametric analyses were used as all groups were large enough (N > 30) for the central limit

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Descriptive statistics at pretest, posttest and follow-up for the Hands-On (Table 2a), Hands-Off (Table 2b) and Control (Table 2c) groups

Table 2a	Hands-On Group														
Measure	**Pre**	**Post**	***t***	***p***	***d***	***n***	BF_{10}	BF_{01}	**Follow-up**	***t***	***p***	***d***	***n***	BF_{10}	BF_{01}
Maths measures															
Maths composite	63.64 ± 16.55	68.46 ± 17.60	3.01	.002***	.418	52	8.13	0.12	64.45 ± 18.11	0.48	.683	.069	48	0.17	5.72
Calculation	51.60 ± 21.29	59.18 ± 22.30	2.45	.009**	.349	49	2.28	0.44	51.99 ± 27.24	0.16	.563	.024	43	0.17	5.99
Place value	66.13 ± 17.78	69.35 ± 16.02	1.82	.037*	.231	62	0.66	1.52	68.43 ± 16.28	1.76	.042*	.245	52	0.64	1.57
Word problems	59.15 ± 21.79	72.36 ± 26.71	3.23	<.001***	.505	41	13.13	0.07	64.58 ± 29.30	1.05	.151	.175	36	0.30	3.37
Missing term problems	64.10 ± 17.81	64.79 ± 25.77	0.21	.419	.033	39	0.18	5.68	65.42 ± 23.55	0.40	.346	.070	32	.20	4.91
Spatial measures															
Mental rotation	66.96 ± 12.09	82.16 ± 18.05	4.73	<.001***	.777	37	661.18	<0.01							
Mental transformation	62.62 ± 14.98	72.72 ± 17.81	5.51	<.001***	.763	52	13627.56	<0.01							
Object completion	56.32 ± 18.09	63.98 ± 20.06	2.78	.004***	.354	62	4.63	0.22							
Mental Folding	49.18 ± 17.35	54.57 ± 20.46	2.47	.008**	.316	61	2.28	0.44							
Expectation and engagement															
Expectation of training	8.63 ± 3.0														
Overall engagement with training	8.53 ± 1.95														

Table 2b	Hands-Off Group														
Measure	**Pre**	**Post**	***t***	***p***	***d***	***n***	**BF10**	**BF01**	**Follow-up**	***t***	***p***	***d***	***n***	**BF10**	**BF01**
Maths measures															
Maths composite	52.42 ± 22.20	55.14 ± 23.24	1.35	.092	.185	53	0.35	2.84	50.64 ± 24.62	0.94	.823	.136	47	0.24	4.19
Calculation	43.28 ± 22.38	42.86 ± 27.77	0.15	.561	.022	51	0.15	6.49	39.06 ± 28.14	1.75	.957	.256	47	0.65	1.54
Place value	57.04 ± 23.04	59.62 ± 22.12	1.07	.146	.140	58	0.25	4.07	59.61 ± 25.37	0.88	.192	.122	52	0.22	4.60
Word problems	43.52 ± 28.31	56.48 ± 28.97	3.32	<.001***	.495	45	17.25	0.06	48.98 ± 34.82	1.37	.089	.214	41	0.40	2.50
Missing term problems	55.61 ± 21.27	55.76 ± 25.52	0.06	.477	.009	44	0.16	6.12	57.78 ± 26.82	0.11	.457	.017	42	0.17	5.96
Spatial measures															
Mental rotation	64.64 ± 11.91	79.64 ± 19.70	4.70	<.001***	.725	42	739.88	0.00							
Mental transformation	62.23 ± 15.08	70.35 ± 19.61	3.58	<.001***	.522	47	34.73	0.03							
Object completion	55.71 ± 14.82	64.35 ± 17.98	3.04	.002***	.413	54	8.66	0.12							
Mental folding	48.02 ± 19.35	48.68 ± 20.40	.23	.409	.031	54	0.15	6.57							
Expectation and engagement															
Expectation of training	9.08 ± 3.09														
Overall engagement with training	8.41 ± 2.21														

(Continues)

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 (Continued)

Table 2c	Control Group														
Measure	**Pre**	**Post**	***t***	***p***	***d***	***n***	***BF$_{10}$***	***BF$_{01}$***	**Follow-up**	***t***	***p***	***d***	***n***	***BF$_{10}$***	***BF$_{01}$***
Maths measures															
Maths composite	60.67 ± 18.44	58.95 ± 21.37	0.63	.734	.090	49	0.19	5.34	59.65 ± 18.27	0.93	.820	.138	45	0.24	4.14
Calculation	51.59 ± 18.95	48.41 ± 24.83	0.99	.836	.148	45	0.26	3.91	46.60 ± 22.38	1.36	.909	.209	42	0.39	2.56
Place value	63.60 ± 17.01	64.47 ± 19.70	0.32	.377	.042	57	0.15	6.59	66.83 ± 20.31	1.59	.059	.222	51	0.49	2.04
Word problems	52.93 ± 26.66	53.15 ± 32.24	0.04	.484	.006	37	0.18	5.65	49.38 ± 33.38	1.27	.107	.224	32	0.39	2.56
Missing term problems	57.37 ± 20.59	62.63 ± 26.05	1.19	.121	.193	38	0.34	3.00	63.64 ± 20.76	1.43	.081	.249	33	0.47	2.13
Spatial measures															
Mental rotation	67.14 ± 14.29	69.29 ± 16.89	0.98	.166	.166	35	0.28	3.53							
Mental transformation	64.67 ± 15.65	66.33 ± 18.08	0.71	.241	.101	49	0.20	5.08							
Object completion	56.85 ± 14.49	56.25 ± 15.67	0.28	.611	.038	56	0.15	6.60							
Mental folding	52.34 ± 19.33	59.07 ± 21.41	2.81	.004**	.389	52	5.00	0.20							
Expectation and engagement															
Expectation of training	9.04 ± 2.97														
Overall engagement with training	9.43 ± 1.54														

Note that the null results reported for t-tests with fewer than 52 participants may be due to reduced power to detect medium (d = 0.4) effects. However, the Bayes Factors reported for both BF_{10} and BF_{01} also support the t-test findings presented. Also note that the pretest means reported here are based on the participants included in the pre-post test analyses. The pretest means used for comparisons at follow-up varied slightly due to different participants being included and excluded for the two different analyses.

*** p<.005; ** p<.01; * p<.05.

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

theorem to apply (Field, 2013). The Pearson correlations reported in Table 1 show significant associations between all spatial and mathematics measures at Time 1 (p < .001; BF_{10} > 25.17 for all). This demonstrates that the associations between spatial and mathematics skills that have been found in previous studies, and that form the rationale for the training paradigm used in this study, are present within this data set.

Measuring pre to post-test gains

For each DV, we measured changes in children's performance from pre-to post-testing using one-tailed t-tests. The results of these t-tests, alongside means and standard deviations for each time point, are reported in Table 2. We also completed ANCOVAs with Time 2 scores as the dependent variable, training mode (*Hands-On, Hands-Off*, and *Control* training) as a between-participant variable and Time 1 scores as a covariate. Where there was a significant main effect of group, we completed follow-up Tukey pairwise comparisons.

Far transfer of gains: Performance on mathematics measures

Aligned with the main aim of this study, we investigated the far transfer of spatial training to an un-trained domain, mathematics.

Mathematics composite measure

Only children in the *Hands-On* condition improved significantly after training (see Table 2). The ANCOVA found an effect of group on time 2 scores, $F(2, 150)=3.10$, $p=.048$, $\eta_p^2 = .040$, $BF_{10}=0.91$, $BF_{01}=1.09$. Post hoc tests showed higher performance for *Hands-On* compared to *Control* ($p=.040$) but no difference between *Hands-Off* and *Control* ($p=.670$), or between *Hands-On* and *Hands-Off* ($p=.257$) groups.

Calculation

Only the *Hands-On* condition significantly improved after training (see Table 2). Within the ANCOVA, there was a significant effect of group on Time 2 scores, $F(2, 141)=4.41$, $p=.013$, $\eta_p^2 = .058$, $BF_{10}=2.89$, $BF_{01}=0.35$. Post hoc tests found significantly higher performance for the *Hands-On* compared to *Control* ($p=.028$) and *Hands-Off* groups ($p=.033$), with no significant difference between *Hands-Off* and *Control* ($p=.994$).

Place value

Only the *Hands-On* group improved significantly with training (see Table 2). However, the effect of group within the ANCOVA was not significant, $F(2, 173)=1.30$, $p=.276$, $\eta_p^2 = .015$, $BF_{10}=0.19$, $BF_{01}=5.37$.

Word problems

Both *Hands-On* and *Hands-Off* groups improved significantly with training while the *Control* group did not (see Table 2). The ANCOVA found an effect of group on Time 2 scores, $F(2, 119)=3.59$, $p=.031$, $\eta_p^2 = .057$, $BF_{10}=1.58$, $BF_{01}=0.63$. There was higher performance for *Hands-On* compared to *Control* ($p=.023$) but no differences between *Hands-On* and *Hands-Off* ($p=.373$), or *Hands-Off* and *Control* ($p=.352$) groups.

Missing term problems

None of the groups improved significantly with training (see Table 2). There was also no significant effect of group within the ANCOVA, $F(2, 117)=0.70$, $p=.498$, $\eta_p^2 = .012$, $BF_{10}=0.14$, $BF_{01}=7.04$. This may be attributable to the low power for this analysis. However, the findings are supported by the Bayes factors reported.

Near and intermediate transfer of gains: Performance on spatial measures

As outlined above all means, standard deviations, and results of t-tests are reported in Table 2.

Near transfer

As a manipulation check, we investigated whether spatial training led to improvements in the spatial skills targeted (near transfer). For all three near-transfer measures, children in both the *Hands-On* and *Hands-Off* conditions improved significantly after training, but the *Control* condition did not (see Table 2). This was supported by the ANCOVA analyses where there was a significant effect of group on Time 2 scores for Mental rotation, $F(2, 110)=6.19$, $p=.002$, $\eta_p^2 = .101$, $BF_{10}=11.62$, $BF_{01}=0.09$, Mental transformation, $F(2, 144)=3.93$, $p=.021$, $\eta_p^2 = .052$, $BF_{10}=1.76$, $BF_{01}=0.57$, and Object completion, $F(2, 168)=4.38$, $p=.014$, $\eta_p^2 = .050$, $BF_{10}=2.50$, $BF_{01}=0.40$. Tukey post hoc tests were completed for each task. For both Mental rotation and Object completion respectively, there was significantly higher performance for both *Hands-On* ($p=.005$, $p=.034$) and *Hands-Off* ($p=.011$, $p=.026$) groups compared to *Control*, but no significant difference between *Hands-On* and *Hands-Off* ($p=.937$, $p=.980$) groups. For Mental transformation, the *Hands-On* group performed better than *Control* ($p=.020$) but there were no significant differences between *Hands-Off* and either *Control* ($p=.127$) or *Hands-On* ($p=.759$) groups.

Intermediate transfer

We also investigated intermediate transfer to other non-trained spatial skills (Mental Folding). Children in both *Hands-On* and *Control* conditions showed significant improvement after training (see Table 2). There was no significant effect of group on Time 2 scores, $F(2, 163)=2.75$, $p=.067$, $\eta_p^2 = .033$, $BF_{10}=0.64$, $BF_{01}=1.57$.

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Measuring gains at follow-up (durability of gains)

T-tests between Time 1 and Time 3 scores were completed for all groups, for all mathematics measures. Means, standard deviations, and the results of these *t*-tests are reported in Table 2. The only significant improvement reported was for the *Hands-On* group for the Place Value task. There were no other significant differences in performance between pre-test and follow-up testing (6 weeks after training). The ANCOVA analyses largely supported these findings as there was no significant effect of group for any task: Composite measure, $F(2, 136)=1.25$, $p=.289$, $\eta_p^2 = .018$, $BF_{10}=0.22$, $BF_{01}=4.59$; Calculation, $F(2, 128)=0.98$, $p=.378$, $\eta_p^2 = .015$, $BF_{10}=0.17$, $BF_{01}=5.82$; Place value, $F(2, 151)=1.09$, $p=.338$, $\eta_p^2 = .014$, $BF_{10}=0.17$, $BF_{01}=6.01$; Word problems, $F(2, 105)=0.43$, $p=.651$, $\eta_p^2 = .008$, $BF_{10}=0.13$, $BF_{01}=7.66$; Missing term problems, $F(2, 103)=1.07$, $p=.347$, $\eta_p^2 = .020$, $BF_{10}=0.18$, $BF_{01}=5.71$. ANCOVAs with fewer than 158 participants were underpowered to find medium effects (0.25) and, therefore, null results should be interpreted accordingly, and in the context of the corresponding Bayes Factors. Levene's test was violated for the mathematics composite at follow-up. Applying a Greenhouse Geisser adjustment did not alter the findings, and so the unadjusted results are reported here.

Expectation and engagement effects

We tested for motivational effects using Analysis of Variance (ANOVA) with the group as the independent variable and engagement score and expectation score as the DV's, respectively. There was no difference in expectation of training across groups, $F(2, 179)=0.42$, $p=.658$, $\eta_p^2 = .005$, $BF_{10}=0.08$, $BF_{01}=12.40$. However, there was a significant difference in children's reported engagement with training across groups, $F(2, 179)=5.11$, $p=.007$, $\eta_p^2 = .054$, $BF_{10}=4.56$, $BF_{01}=0.22$. Post hoc tests revealed that the children in the *Control* condition reported higher engagement than both *Hands-On* ($p=.026$) and *Hands-Off* training ($p=.011$) (see Table 2). We did not repeat our main ANCOVA analyses with engagement as a covariate because (a) some of our ANCOVA analyses were already underpowered, (b) differences in engagement favored the control group and we proposed that adding this to our models would only strengthen, and not change the pattern of results supporting our hypotheses.

DISCUSSION

Spatial training is effective at improving spatial and mathematics skills

Comparing intervention and control groups, spatial training, including both *Hands-On* and *Hands-Off* paradigms, was effective at eliciting change in the spatial domain. Spatial training led to significant near-transfer gains in the spatial skills targeted compared to the control condition, highlighting spatial thinking as one aspect of cognition that appears to be particularly susceptible to improvement through training. These findings align with the existing literature supporting the malleability of spatial cognition in both children and adults (Uttal et al., 2013; Yang et al., 2020). For intermediate transfer to an untrained spatial task, there was an unexpected improvement for the *Control* group that cannot be explained by practice effects. We propose that some element of control training inadvertently led to an advantage on the mental folding task, rendering it an ineffective comparison group, and leading to inconclusive findings for intermediate transfer.

For mathematics, spatial training also led to gains in performance when compared to the control group which had no significant gains in mathematics for any measure (see next section for details on differences between *Hands-On* and *Hands-Off* training). These findings add to a growing body of literature showing a causal effect of spatial training on mathematics (Hawes et al., 2022). From a theoretical perspective, given that spatial training had differing effects across mathematics measures, our findings also support the theory that there is not a linear coupling between all spatial and mathematical skills (Mix et al., 2016). Consequently, there may be several explanations or different underlying mechanisms underpinning spatial-mathematical relations (Hawes et al., 2023). First, spatial and mathematical thinking may rely on similar brain regions and shared neural mechanisms (Hawes, Sokolowski, et al., 2019). Second, spatial processing or spatial visualization may provide children with a "mental blackboard" for modeling and solving mathematical problems (Lourenco et al., 2018; Mix, 2019), a tool that may be particularly useful for novel or unfamiliar mathematical content (Hawes, Moss, et al., 2019). Third, spatial skills may be beneficial in mathematics as space is used to organize, represent, and communicate meaning in mathematics, for example, number lines, rulers, diagrams, graphs, place value, positioning of digits, etc. (Mix, 2010). Finally, several sub-domains of mathematics including geometry and measurement domains are inherently spatial, and for this reason, success in these domains relies heavily on spatial reasoning (Hawes et al., 2023). Given the various mechanisms through which spatial processing may influence mathematics, it is reasonable to expect differing effects of spatial training for different mathematics sub-domains. In short, different spatial-mathematical relations (and the mechanisms underpinning them) may be differently affected by spatial training paradigms.

Embodied action in spatial training leads to larger gains

Our findings suggest that *Hands-On* spatial training is more effective than *Hands-Off* training in eliciting

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

spatial and mathematical gains. As outlined above, both spatial training paradigms led to gains in spatial skill with no significant differences between the paradigms when they were directly compared. However, examining the findings in more detail, only *Hands-On* training was significantly better than *Control* for all near transfer spatial measures, that is, including the mental transformation measure, and the gains (size of Cohen's *d* effect sizes) following spatial training were larger for the *Hands-On* compared to the *Hands-Off* group for two out of the three near transfer spatial measures. Furthermore, only the *Hands-On* group had improvements in the intermediate transfer (mental folding) task. This pattern suggests a slight *Hands-On* advantage for embodied spatial training in improving spatial skills, which should be viewed as nuanced as the *Hands-Off* training was also effective in eliciting spatial gains. This *Hands-On* advantage was more apparent for mathematics such that *Hands-On* spatial training using physical manipulatives led to significant improvements in all mathematics measures except for missing term problems. By comparison, the *Hands-Off* paradigm only improved on word problems. Furthermore, all group effects reported were driven by improved performance of the *Hands-On* group compared to the *Control* (and in some cases *Hands-Off* training). In short, although children in the *Hands-On* group did not outperform the *Hands-Off* group on every measure, when there were significant differences, they consistently favored the *Hands-On* group, particularly when compared to the *Contro*l condition.

Our findings showing larger and broader effects for *Hands-On* compared to *Hands-Off* training are consistent with the Hawes et al. (2022) meta-analysis which demonstrated that spatial interventions that used physical manipulatives were more effective than those that did not. However, the comparisons made by Hawes et al. (2022) were between studies and hence this finding was confounded by differences in not only the use of physical materials across studies, but also differences in the spatial paradigms used. Here, we provide an explicit comparison of embodied versus non-embodied training for the first time, where all other elements of training were kept constant. Our findings for a *Hands-On* advantage also align with evidence from the broader educational literature that favors the use of physical materials in learning (Alibali & Nathan, 2018, or Glenberg, 2008 for reviews) and studies on the use of embodied action for learning mathematics specifically (Berteletti & Booth, 2015).

Here we propose that the advantage conferred by *Hands-On* training is due to the role of embodied cognition in the *Hands-On* training paradigm. Aligned with theories of embodied cognition, embodied action elicited through interaction with physical manipulatives enabled our *Hands-On* participants to conceptualize ideas with grounded representations based on sensory-motor encoding (Barsalou, 2008; Glenberg & Kaschak, 2002; Pecher & Zwaan, 2005). We suggest that by enabling *Hands-On* learners to form these multimodal cognitive representations of spatial relations rooted in action, we promoted the generation of more deeply grounded, easily accessed spatial representations. The *Hands-On* condition may afford a combination of visual and haptic cues and feedback that are not present in either the *Hands-Off* or *Control* conditions. Thus, when asked subsequently to solve spatial and mathematics questions, the *Hands-On* learners may more readily activate shared neural circuits and recruit spatial processes useful for mathematical reasoning. This can be compared to our *Hands-Off* spatial training group whose representations were not based on perception-action representations and therefore, we propose, were less well-grounded and more difficult to access. This is one likely explanation for the superior gains of the *Hands-On* compared to the *Hands-Off* training group.

However, our findings also show that *Hands-Off* training is somewhat effective. Spatial processes were evolved to enable humans to navigate through spatial environments and to manipulate three-dimensional objects (Newcombe, 2018). It, therefore, makes sense that *Hands-On* experiences are more likely to activate spatial processes. However, this does not mean that *Hands-On* action is necessarily required to activate these processes, as there are many examples where humans have applied these processes to understand navigation and object manipulation in digital environments as well. Additionally, there are several examples of spatial training paradigms (including our *Hands-Off* paradigm) that have elicited spatial and mathematics gains with non-embodied spatial paradigms, that is, without using concrete materials (Bower et al., 2022; Cheung et al., 2019; Gilligan, Thomas, & Farran, 2019). Thus, while it is possible to activate and train spatial thought without embodied action, we conclude here that the most direct and effective way of activating spatial processes is through physical movement and the adoption of embodied action approaches.

The durability of spatial training gains

This is the first study to investigate the durability of spatial training gains. We found that while the *Hands-On* group had significant gains in place value between pretest and follow-up, there were no other significant improvements between these time points, suggesting that gains in mathematics elicited through our spatial training paradigm do not persist over time. These findings can be interpreted to offer insights into the mechanistic relations between spatial and mathematics skills, such that there are at least two possible reasons for these findings.

First, it is possible that the short-term gains in mathematics following spatial training reflect spatial priming effects. Spatial training may lead to improvements

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in mathematics not necessarily due to changes in spatial cognition, but rather due to training-induced shifts in spatial attention and strategy use. Many mathematics problems, including basic arithmetic problems, afford multiple solution strategies, varying in the extent to which verbal and spatial processes are recruited. Moreover, spatial reasoning is a highly effective problem-solving tool, offering the learner with new insights and understanding of the problem at hand (e.g., see Casey & Fell, 2018; Hawes et al., 2022). Thus, it is plausible that spatial training may encourage individuals to adopt the use of more effective and spatially-based problem-solving approaches. For example, the gains observed on word problems may have been due to the adoption of spatially-based "schematic" solutions (see Hegarty, & Kozhevnikov, 1999), encouraged through the repeated training of spatial visualization processes. Relatedly, spatial training may encourage individuals to pay greater attention to the spatial features of given mathematics problems. For example, in the case of place value problems, spatial training may correspond to a shift in the spatial attention paid to digit position and the mathematical meaning derived from digit position.

This account offers an explanation for two other findings in the literature at large. First, training-induced improvements in mathematics are not necessarily related to the amount of spatial change observed (e.g., see Hawes et al., 2022; Mix et al., 2021), including the lack of intermediate spatial transfer observed in the present study. Second, it offers a reason for the absence of a dose–response relation in the space-mathematics training literature to date; longer training does not necessarily result in larger short-term mathematics gains (Hawes et al., 2022). However, all previous studies in this domain have pre- to post-test designs, and none have tested durability. Therefore, alternative explanations for our findings, that should be investigated in future work, are that the quantity of training (2h) delivered in this study was insufficient to elicit long-term gains in mathematics, or that the limited durability found in this study may be due to the relatively small magnitude of spatial gains between pre and post testing. That is, while spatial training was effective, the magnitude of gains was small and therefore even a small decrement means that the gains are no longer present. As such, future studies should also investigate whether spatial training that elicits larger gains leads to more durable transfer of training effects. Taken together, this study adds to a growing body of evidence that suggests spatial training may be effective at improving mathematics due to spatial priming effects (shifts in spatial attention and strategy use) (Uttal et al., 2013). Empirical studies are needed to test this possibility. If priming effects are at play, this might explain why only limited long-term gains were observed in the present study. This too needs to be studied moving forward. It is possible that the extent to which mathematics gains are sustained depends on the sustained use of spatially-based problem-solving strategies.

A second reason for the limited durability of gains observed may have to do with the training design employed. Our findings could be interpreted as showing that isolated bouts of spatial training, that is, discrete spatial training paradigms over a short period of time, do not elicit stable, long-term improvement in mathematics and alternative integrated approaches to training should be adopted. Practically speaking, we know that spatial training can improve mathematics outcomes in the short term, which is beneficial to students and should be monopolized on (e.g., Hawes et al., 2022). Therefore, if gains elicited through discrete spatial training paradigms are not durable, an alternative is to regularly integrate elements of spatial training into classroom lessons. This integration should be based on evidence from previous work that demonstrated high spatial transfer effects to math, which anecdotally could be argued to be more likely to produce durability. First, based on the current study and the findings of Hawes et al. (2022), spatial training using concrete materials is likely to render larger, more durable transfer to math. Second, spatial training appears to be more beneficial for math outcomes that have closer alignment to the training delivered, for example, origami training and geometry outcomes (Hawes et al., 2022). Therefore, when integrating spatial training into the math classroom, it may be important to align the spatial skills trained with the math learning objectives being taught. Using an integrated approach to spatial training would ensure that students frequently engage and improve their spatial thinking skills, and these skills are thus maintained at an optimal level for supporting mathematics. Another possible benefit of this training design is that when spatial skills are activated in a mathematical context, students may more easily recognize strategies and reasoning employed in the spatial context as useful to the mathematical context (Hawes et al., 2022).

Unfortunately, no other studies in this domain have completed follow-up testing longer than 10 days after spatial training, which is similar to our post-testing (after 7 days) and not our follow-up (after 6 weeks). Therefore, there are no other studies from which we can gain meaningful insights on the relative likelihood of each of the two explanations outlined above. While these findings provide the first evidence on the durability of spatial training gains in mathematics and insights into the mechanistic relations between spatial and mathematics skills, future studies are needed to unpack the outstanding questions pertaining to spatial priming and training design outlined above.

Considerations and limitations

The findings should be interpreted in the context of study limitations and considerations. Across all mathematics measures the *Hands-On* group had higher starting points (i.e., time 1 scores). This occurred even though

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

participants were randomly allocated to groups by their classroom teacher, and these groups were randomly allocated to a training condition by the primary investigator who was blind to which participants were in each group. Although this higher initial mathematics skill might be seen as limiting the possible growth from spatial training, the results argue against that interpretation because growth was actually greater in the *Hands-On* condition nonetheless.

A second limitation is that we have reduced power for some analyses due to small sample sizes which means that we are less likely to detect small effect sizes (gains) in this study. Prior to the commencement of the study, all measures were piloted ($n=56$) and appeared suitable for this age group. Despite this, a relatively high proportion of children did not meet the inclusion criteria for individual analyses in the current study as they scored above our pre-registered cut-off for Time 1 performance, that is, they scored 85% or higher during pre-testing. Throughout the results, we have highlighted null results within our analyses that should be interpreted with particular caution due to low power. We have also reported Bayes factors for all findings. As Bayesian statistics are influenced to a lesser degree by issues of power, interpreting frequentist and Bayesian statistics in combination provides a better insight into true effects. They have also allowed us to generate estimates of the evidence for not only the alternative hypothesis BF_{10} (i.e., there are group differences at post-test, controlling for pre-test scores) but also the null hypothesis BF_{01} (i.e., there are no group differences at post-test, controlling for pre-test scores).

Finally, it is possible that the spatial tasks measured here have some contribution from fluid intelligence. However, evidence from both behavioral (Hawes et al., 2019) and neural (Ebisch et al., 2012) studies suggest that spatial tasks tap into unique abilities beyond fluid intelligence. For example, Hawes, Sokolowski, et al. (2019) demonstrated that the association between spatial ability and maths remained after controlling for non-verbal reasoning. This suggests that, while we cannot rule out contributions from fluid intelligence, the gains in ability observed in the current study relate to improved spatial ability.

Implications

Here, we have shown that taking an embodied (*Hands-On*) approach is the optimal design of spatial training as a means of developing mathematics skills. This has substantial implications for the spatialization of mathematics learning and instruction, and there are clear pathways for translating this evidence into the classroom. As outlined in previous literature (Gilligan-Lee et al., 2022) spatial thinking is often absent from primary school mathematics curricula, despite convincing evidence that supports a causal effect of spatial training on mathematics (Hawes et al., 2022). In many cases when spatial thinking is implemented in curricula, it is in the context of learning shape names and properties, and not in practicing spatial visualization skills.

Our findings show that our embodied spatial training paradigm has a valuable place in the primary school classroom as a means of improving mathematics. It is a simple-to-administer paradigm, and as evidenced by the engagement scores, the training is enjoyable for the children who take part. All the tools for delivering our intervention can be found on our OSF page (https://osf.io/mer9t). Furthermore, our findings can be interpreted in a wider context. Beyond this specific spatial training paradigm, we propose that all embodied approaches to spatial instruction are likely to be beneficial to children, over non-embodied approaches. Therefore, we suggest that if educators plan to include spatial activities in their lessons, then we encourage them to incorporate embodied action using physical manipulatives (objects). Finally, these findings should not be interpreted to mean that non-embodied spatial training has no value in the classroom. Indeed, our *Hands-Off* group showed significant improvement in spatial performance, and in instances where no physical tools are available then non-embodied approaches may also offer a pathway to spatial gains, albeit less-optimal gains.

CONCLUSION

The goal of this study was to refine the optimal design of spatial training for use in mathematics learning and instruction. We conclude that embodied spatial training using physical manipulatives (*Hands-On* training) leads to larger, more consistent gains in mathematics and greater depth of spatial processing than *Hands-Off* training. We thus encourage greater use of spatial activities in the classroom, more specifically activities that use physical materials and by extension embodied action.

ACKNOWLEDGMENTS

We would like to thank all of the children and schools who participated. We would like to acknowledge the support of Megan Davies, Charlotte Bradley, Rebecca Watt, Georgia Tomlin and Sarah Jones, for their assistance with data collection and the administration of training sessions.

FUNDING INFORMATION

This study was funded by the British Academy Small Research Grant SRG20\201210 and the University of Surrey Faculty Research Support Fund.

CONFLICT OF INTEREST STATEMENT

We have no conflict of interest to disclose.

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

DATA AVAILABILITY STATEMENT

Open Science Statement: The data, analytic code and materials associated with this manuscript are publicly accessible on the Open Science Framework (https://osf.io/mer9t/).

ORCID

Katie A. Gilligan-Lee https://orcid.org/0000-0002-5406-2149
Zachary C. K. Hawes https://orcid.org/0000-0003-4055-7607
Emily K. Farran https://orcid.org/0000-0001-6409-0875
Kelly S. Mix https://orcid.org/0000-0002-4402-2557

REFERENCES

Alibali, M. W., & Nathan, M. J. (2018). Embodied cognition in learning and teaching: Action, observation, and imagination. In *International handbook of the learning sciences* (pp. 75–85). Routledge.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science*, *2*(4), 716–724. https://doi.org/10.1111/j.1756-8765.2010.01115.x

Berteletti, I., & Booth, J. R. (2015). Perceiving fingers in single-digit arithmetic problems. *Frontiers in Psychology*, *6*(226), 1–10. https://doi.org/10.3389/fpsyg.2015.00226

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science*, *8*(4), 445–454. https://doi.org/10.1177/1745691613491271

Booton, S. A., Hodgkiss, A., Mathers, S., & Murphy, V. A. (2022). Measuring knowledge of multiple word meanings in children with English as a first and an additional language and the relationship to reading comprehension. *Journal of Child Language*, *49*(1), 164–169. https://doi.org/10.1017/S0305000921000052

Booton, S. A., & Murphy, V. M. (in prep). A randomised controlled trial of inference training for homonyms: impact and transfer for first and second English language learners.

Bower, C. A., Zimmermann, L., Verdine, B. N., Pritulsky, C., Golinkoff, R. M., & Hirsh-Pasek, K. (2022). Enhancing spatial skills of preschoolers from under-resourced backgrounds: A comparison of digital app vs. concrete materials. *Developmental Science*, *25*(1), e13148. https://doi.org/10.1111/desc.13148

Casey, B. M., & Fell, H. (2018). Spatial reasoning: A critical problem solving tool in children's mathematics strategy tool-kit. In K. Mix & M. Battista (Eds.), *Visualizing mathematics: The role of spatial reasoning in mathematical thought* (pp. 47–75). Springer Nature Publishing. https://doi.org/10.1007/978-3-319-98767-5_3

Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves Children's mathematics ability. *Journal of Cognition and Development*, *15*(1), 2–11. https://doi.org/10.1080/15248372.2012.725186

Cheung, C. N., Sung, J. Y., & Lourenco, S. F. (2019). Does training mental rotation transfer to gains in mathematical competence? Assessment of an at-home visuospatial intervention. *Psychological Research*, *84*, 2000–2017. https://doi.org/10.1007/s00426-019-01202-5

Cornu, V., Schiltz, C., Pazouki, T., & Martin, R. (2019). Training early visuo-spatial abilities: A controlled classroom-based intervention study. *Applied Developmental Science*, *23*, 1–21. https://doi.org/10.1080/10888691.2016.1276835

Ebisch, S. J., Perrucci, M. G., Mercuri, P., Romanelli, R., Mantini, D., Romani, G. L., … Saggino, A. (2012). Common and unique neuro-functional basis of induction, visualization, and spatial relationships as cognitive components of fluid intelligence. *NeuroImage*, *62*(1), 331–342. https://doi.org/10.1016/j.neuroimage.2012.04.053

Ehrlich, S. B., Levine, S. C., & Goldin-Meadow, S. (2006). The importance of gesture in Children's spatial reasoning. *Developmental Psychology*, *42*(6), 1259–1268. https://doi.org/10.1037/0012-1649.42.6.1259

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Gilligan, K. A., Flouri, E., & Farran, E. K. (2017). The contribution of spatial ability to mathematics achievement in middle childhood. *Journal of Experimental Child Psychology*, *163*, 107–125. https://doi.org/10.1016/j.jecp.2017.04.016

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

Gilligan, K. A., Thomas, M. S., & Farran, E. K. (2019). First demonstration of effective spatial training for near transfer to spatial performance and far transfer to a range of mathematics skills at 8 years *Developmental Science*, *23*(4), e12909. https://doi.org/10.1111/desc.12909

Gilligan-Lee, K. A., Hawes, Z. C. K., & Mix, K. S. (2022). Spatial thinking as the missing piece in mathematics curricula. *npj Science of Learning*, *7*, 10. https://doi.org/10.1038/s41539-022-00128-9

Glenberg, A. M. (2008). Embodiment for education. In *Handbook of cognitive science* (pp. 355–372). Elsevier.

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology*, *96*, 424–436. https://doi.org/10.1037/0022-0663.96.3.424

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin and Review*, *9*, 558–565. https://doi.org/10.3758/BF03196313

Green, S. C., Bavelier, D., Kramer, A., Vinogradov, S., Ansorge, U., Ball, K., & Witt, C. (2019). Improving methodological standards in behavioral interventions for cognitive enhancement. *Journal of Cognitive Enhancement*, *3*(1), 2–29. https://doi.org/10.1007/s41465-018-0115-y

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. https://doi.org/10.1111/mbe.12007

Hattie, J. A. C. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Routledge.

Hawes, Z. C. K, Gilligan-Lee, K. A., & Mix, K. S. (2023). Infusing Spatial Thinking Into Elementary and Middle School Mathematics: What, Why, and How? In *Mathematical Cognition and Understanding: Perspectives on Mathematical Minds in the Elementary and Middle School Years* (pp. 13–33). Cham: Springer International Publishing.

Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: Effects of a 32-week intervention. *Cognition and Instruction*, *35*(3), 236–264. https://doi.org/10.1080/07370008.2017.1323902

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. https://doi.org/10.1016/j.tine.2015.05.001

Hawes, Z., Moss, J., Caswell, B., Seo, J., & Ansari, D. (2019). Relations between numerical, spatial, and executive function skills and mathematics achievement: A latent-variable approach. *Cognitive Psychology*, *109*, 68–90. https://doi.org/10.1016/j.cogpsych.2018.12.002

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hawes, Z., Sokolowski, H. M., Ononye, C. B., & Ansari, D. (2019). Neural underpinnings of spatial and numerical cognition: An *f*MRI meta-analysis of brain regions associated with symbolic number, arithmetic, and mental rotation. *Neuroscience & Biobehavioral Reviews*, *103*, 316–336. https://doi.org/10.1016/j.neubiorev.2019.05.007

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

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

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*(1), 2–9. https://doi.org/10.7771/1932-6246.1167

Kontra, C., Lyons, D. J., Fischer, S. M., & Beilock, S. L. (2015). Physical experience enhances science learning. *Psychological Science*, *26*(6), 737–749. https://doi.org/10.1177/0956797615569355

Lourenco, S. F., Cheung, C. N., & Aulet, L. S. (2018). Is visuospatial reasoning related to early mathematical development? A critical review. In A. Henik & W. Fias (Eds.), *Heterogeneity of function in numerical cognition* (pp. 177–210). Academic Press.

Lowrie, T., Logan, T., & Ramful, A. (2017). Visuospatial training improves elementary students' mathematics performance. *British Journal of Educational Psychology*, *87*(2), 170–186. https://doi.org/10.1111/bjep.12142

Mix, K. S. (2010). Spatial tools for mathematical thought. In K. S. Mix, L. B. Smith, & M. Gasser (Eds.), *Space and language* (pp. 41–66). Oxford University Press.

Mix, K. S. (2019). Why are spatial skill and mathematics related? *Child Development Perspectives*, *13*(2), 121–126. https://doi.org/10.1111/cdep.12323

Mix, K. S., Levine, S. C., Cheng, Y.-L., Stockton, J. D., & Bower, C. (2021). Effects of spatial training on mathematics in first and sixth grade children. *Journal of Educational Psychology*, *113*(2), 304–314. https://doi.org/10.1037/edu0000494

Mix, K. S., Levine, S. C., Cheng, Y.-L., Young, C., Hambrick, D. Z., Ping, R., & Konstantopoulos, S. (2016). Separate but correlated: The latent structure of space and mathematics across development. *Journal of Experimental Psychology General*, *145*(9), 1206–1227. https://doi.org/10.1037/xge0000182

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

Mohring, W., & Frick, A. (2013). Touching up mental rotation: Effects of manual experience on 6-month-old infants' mental object rotation. *Child Development*, *84*(5), 1554–1565. https://doi.org/10.1111/cdev.12065

Neuburger, S., Jansen, P., Heil, M., & Quaiser-Pohl, C. (2011). Gender differences in pre-adolescents' mental-rotation performance: Do they depend on grade and stimulus type? *Personality and Individual Differences*, *50*(8), 1238–1242. https://doi.org/10.1016/j.paid.2011.02.017

Newcombe, N. S. (2018). Three kinds of spatial cognition. In J. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed., pp. 1–31). John Wiley & Sons. https://doi.org/10.1002/9781119170174.epcn315

Pearl, J. (2009). *Causality: Models, reasoning, and inference*. Cambridge university press.

Pecher, D., & Zwaan, R. A. (2005). *Grounding cognition: The role of perception and.action in memory, language, and thinking*. Cambridge University Press.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews: Neuroscience*, *6*(7), 576–582. https://doi.org/10.1038/nrn1706

Thurstone, T. G. (1974). *PMA readiness level*. Science Research Associates.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352–402. https://doi.org/10.1037/a0028446

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. https://doi.org/10.1016/j.jecp.2014.02.012

Yang, W., Liu, H., Chen, N., Xu, P., & Lin, X. (2020). Is early spatial skills training effective? A Meta-Analysis. *Frontiers in Psychology*, *11*, 1938. https://doi.org/10.3389/fpsyg.2020.01938

How to cite this article: Gilligan-Lee, K. A., Hawes, Z. C. K., Williams, A. Y., Farran, E. K., & Mix, K. S. (2023). *Hands-On*: Investigating the role of physical manipulatives in spatial training. *Child Development*, *00*, 1–17. https://doi.org/10.1111/cdev.13963

14678624, 0, Downloaded from https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.13963 by University Of Toronto Mississauga, Wiley Online Library on [07/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License