An Exploration of the Role of Spatial Ability and Spatial Anxiety in Gesture Production and Mathematical Thinking

by

Kelsey E. Schenck

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Educational Psychology

at the

UNIVERSITY OF WISCONSIN–MADISON

2023

Date of final oral examination: 5/3/2023

The dissertation is approved by the following members of the Final Oral Committee:

Mitchell J Nathan, Professor, Educational Psychology

Martha W Alibali, Professor, Psychology

Edward M Hubbard, Associate Professor, Educational Psychology

Martina A Rau, Associate Professor, Educational Psychology

Acknowledgments

This endeavor would not have been possible without the tremendous support that I have received on this journey. I would like to thank my advisor, Mitchell Nathan, for his invaluable mentorship, patience, and feedback over the past six years. You have been a tireless source of wisdom and guidance. Thank you for creating an atmosphere that fosters a sense of independence, resilience, and curiosity.

Thank you to Michael Swart for adopting me as an unofficial advisee and imparting your vast design knowledge. Your insights on countless drafts and figures have been invaluable. I am deeply grateful for past and present members and collaborators of the UW MAGIC Lab who have always been willing to provide advice and edits on my numerous projects over the years: Joe Michaelis, Fangli Xia, Doy Kim, Michael Forkner, John McGinty, Ariel Fogel, Matthew Grondin, Hanall Sung, and Candace Walkington. I am also thankful for the network of past and present fellow UW graduate students that have helped support me: Amanda Siebert-Evenstone, Sally Wu, Tiffany Herder, Joel Beier, Roseanne Luu, Jihyun Rho, Xuesong Cang, Indrani Dey, Jaeyoon Choi, Yuanru Tan, and Yeyu Wang.

I am also grateful to my dissertation committee members, who have also served on committees for my other milestones along the way: Mitchell Nathan, Ed Hubbard, Martina Rau, and Martha Alibali. I appreciate your immense feedback and the time you've spent guiding the trajectory of my work.

I would also like to thank the different organizations that have made this work possible. The research on students' fraction knowledge funded by grants from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, and the research on geometric thinking using *The Hidden Village* was funded by grants from Institute of Education Sciences. I have also benefited from the gracious support through the Arvil S. Barr Graduate fellowship.

I am incredibly thankful to my parents Chuck, Shelley and Craig, my siblings Catherine, Keeton and Chloe, and my grandparents Linda and George. I would not have had the courage to make a career change and move across the country to pursue this degree without your encouragement and support. I would also like to thank Maggie, the best emotional support dog. I can't imagine what the last 12 years would have looked like without you. Thank you also to Maycie Schutt, my best friend, who has helped keep me sane and weather many storms over the past six years. I am so grateful to have met you in Madison.

And most importantly, thank you to Rob, my partner and soon to be husband, whose encouragement has meant everything. The last three years would not have been possible without your unwavering love and support and your belief in me. There is no here without you.

Table of Contents

Abstract

Spatial ability is a multifaceted construct with demonstrated ties to success in mathematics and gesture production. Recent empirical work has also begun investigating whether targeted interventions for spatial ability can transfer to mathematics but are often limited to specific populations and specific math outcomes. Spatial anxiety has been shown to be negatively associated with the ability to access and utilize spatial skills. However, little is known about the interplay between spatial ability, spatial anxiety, and gestures in mathematics. This dissertation describes a guiding framework for researching spatial ability in mathematics before describing four studies to investigate the relationships between these constructs. These studies cover a multitude of mathematics domains, populations, and methods of analysis, including an experimental study on adult's performance on standardized mathematics outcomes, an analysis of existing data from a longitudinal study on children's fraction knowledge, and an experimental study on undergraduates' geometric thinking, and a randomized control study on undergraduates' geometric thinking and gesture production following a short spatial intervention. Several findings emerged. First, specific spatial sub-categories could be more critical for success on different math tasks, and spatial anxiety scores predict spatial and math ability scores. Second, spatial anxiety may moderate the relationship between spatial ability and both geometric thinking and representational gesture production. Third, spatial-based Tangram tasks may increase the likelihood of transformational proof and dynamic gesture production. These findings will inform future classroom-based spatial interventions and studies exploring the links between mathematics, gesture, spatial ability, and spatial anxiety.

Introduction

In the last several decades, traditional models of mathematical cognition have seen substantial expansion beyond isolated rule-based processing symbolic systems. Since mathematical thinking is a confluence of how our perceptual, motoric, and cognitive processes interact with our environment and culture (Lakoff & Nuñez, 2000), educational researchers have investigated the ways in which mathematics concepts gain meaning by being grounded in perceptual and motor-based systems, including action and gesture. Researchers have also sought to determine how individual differences in cognitive processes, such as spatial ability, anxiety, and working memory, impact success on mathematics tasks. For this reason, there are a variety of different factors that can impact a student's success on mathematical tasks. Understanding the links between spatial ability, spatial anxiety, and mathematical thinking may be a way to bridge these two lines of often separate lines of research.

Spatial ability can be generally defined as the ability to imagine and manipulate objects and relations. Decades of research have shown that spatial ability can be linked to mathematics success. Findings from numerous studies have demonstrated that spatial ability is critical for many domains of mathematics education, including basic numeracy and arithmetic (Case et al., 1996) and geometry (Battista et al., 2018), as well as more advanced topics such as algebra word problem solving (Oostermeijer et al., 2014), calculus (Sorby et al., 2013), and decoding complex mathematical ideas (Tufte, 2001). Scores on the mathematics portion of the Program for International Student Assessment are significantly positively correlated with scores on tests of spatial cognition (Sorby & Panther, 2020). Spatial ability has also been linked to the entrance into, retention in, and success within STEM fields (e.g., Wolfgang et al., 2003), while deficiencies in spatial abilities have been shown to create obstacles for STEM education (Harris et al., 2013). Combined, the findings suggest that spatial ability serves as a gateway for entry into STEM fields (Uttal & Cohen, 2012) and that educational institutions need to recognize the importance of explicitly training students' spatial thinking skills along with skills in STEM. Many studies do not consider the influences of individual differences in spatial ability on these associations, such as how individuals approach the problems. Moreover, few studies investigate the links between spatial abilities and mathematical reasoning tasks that are not pen-and-paper based.

Gesture production may also be tied to spatial abilities. One function of gestures that cooccur with speech is to express and communicate spatial information from mental representations (Alibali, 2005). For example, speakers often gesture when they describe object locations (Tutton, 2003), spatial patterns (Melinger & Kita, 2007), or motion in space (Kita & Özyürek, 2003). Additionally, the complex motor movements of gesture allow learners to represent dynamic, spatial-relational information (McNeill, 2005) and augment reasoning and problem-solving by connecting sensorimotor experiences with mental representations (Wilson, 2002; Nathan, 2014). For example, Chu and Kita (2011) found that when individuals produce spontaneous gestures during challenging spatial visualization problems, the gestures enhance performance, possibly by improving the internal computation of spatial transformations. However, how other individual differences, such as spatial anxiety, could affect gesture production in these situations is unknown.

Studies across this vast body of work investigating the links between spatial abilities and mathematics performance and gesture production use different spatial taxonomies, employ different spatial measures, and track improvement across many mathematics education topics. This variety makes it difficult for scholars to draw clear causal lines between specific spatial skills interventions and specific mathematics educational improvements and for educators to follow clear guidance on improving mathematics learning through spatial skills development.

As a separate construct, spatial anxiety has been shown to be negatively associated with performance on spatial ability tasks (Ramirez et al., 2012). While some scholars have argued that low anxiety levels are essential for neurocognitive performance, high levels of anxiety, especially domain-trait-specific anxieties, may lead to reduced performance (Derakshan & Eysenck, 2010). Spatial anxiety is a domain-specific trait anxiety that functions similarly to mathematics anxiety and testing anxiety and describes apprehension towards tasks requiring spatial processing (Lyons et al., 2018). Like spatial ability, spatial anxiety appears to be composed of sub-components. One study of nineteen 21-year-old twin pairs identified two components of spatial anxiety: navigation anxiety and rotation/visualization anxiety (Malanchini et al., 2017). Other studies have linked spatial anxiety to spatial orientation skills, specifically, decreased efficiency of orientation strategies and increased errors during navigation tasks (Lawton, 1994; Hund & Minarik, 2006). However, we know little about the effects of spatial anxiety on other components of spatial ability as there are limited measures.

Studies have investigated gender differences within spatial anxiety. Wei et al. (2018) compared groups from China and Russia and found that males outperformed females on spatial ability tasks, with females reporting higher spatial anxiety than males. More specifically, females had higher scores of spatial anxieties on spatial perception and visualization tasks, although the magnitudes of the correlations between anxiety and these spatial abilities were similar between males and females. A different study found that spatial anxiety affected working memory capacity in females but not males and that females performed worse than males on mental rotation tasks (Ramirez et al., 2012). Still, other studies have found no significant sex differences in spatial anxiety (Hund & Minarik, 2006; Saucier et al., 2002). The variety of measures and inconsistency of results could reflect the multifaceted heterogeneity of spatial ability, and participants' experiences of anxiety may only arise during certain spatial activities. In any case, it reinforces the need for further research to clarify how spatial anxiety interacts with spatial ability and in what contexts, and how math educators could leverage these factors to develop effective mathematics instruction addressing these possible gender differences.

The framework and studies in this dissertation are guided by several aspects of research on spatial ability and spatial anxiety, specifically for mathematical thinking and gesture production. However, there are many limitations in the current research. First, there is a lack of consensus among the spatial ability community about the sub-categories that make up spatial ability and how best to select a taxonomy in mathematics education research. Second, much research on the connections between spatial ability and mathematics focuses on paper-based outcomes such as multiple-choice and short-answer measures. While this provides a good foundation for the relationship between these constructs, little is known about how spatial abilities relate to more informal, discourse-based mathematical thinking practices or the interaction between spatial abilities and gestures during these tasks. Additionally, to my knowledge, no studies have tested the potential of spatial interventions to improve mathematical thinking in discourse-based tasks. Finally, spatial anxiety is a relatively unexplored construct, especially outside navigational anxiety. To the best of my knowledge, no studies have examined spatial anxiety's potential associations with mathematics performance.

In this dissertation, I address the limitations described above by creating a guiding framework and a series of four research studies. The overall goals of this work are to support theoretical understanding and practical applications of the cognitive and affective processes of the spatial system (i.e., spatial ability and spatial anxiety) for improving students' mathematical thinking by:

1) summarizing the current state of the spatial ability literature as relevant for mathematics education and creating a guiding framework for mathematics education researchers
2) examining the role of spatial ability, spatial anxiety, and gesture in a variety of domains of mathematics, including overall mathematics ability, fraction knowledge, and geometric discourse, and
3) empirically testing whether a short Tangram task affects geometric thinking and gesture production more than a lower-spatial control task.

Summary of Dissertation Research Studies

This dissertation work is presented in the following four chapters exploring the relationships between spatial ability, spatial anxiety, and mathematical thinking. Chapter 1 first introduces a guiding framework for researchers investigating the connection between spatial ability and mathematics. This chapter summarizes the abundance of spatial ability taxonomies in the literature and provides practical guidance for navigating these taxonomies to select one most appropriate for the research aims with accompanying recommendations for choosing an analytic approach and spatial tasks. The following three chapters present findings from four studies exploring the role of spatial ability and spatial anxiety in gesture production and mathematical thinking. In Chapter 2, I describe two exploratory studies to uncover the specific relationships between spatial abilities, spatial anxiety, and mathematics. In particular, the first study presents findings from an exploratory study on adults' standardized mathematics outcomes previously published as a conference paper at the 2020 AERA Conference. Study 2 focuses on the role of embodied processes across the grounded and embodied learning timescales (e.g., spatial ability, spatial anxiety, number sense, and working memory) in children's symbolic fraction knowledge that was previously published as a poster at the 2021 CogSci Conference. In Chapter 3, I extend these findings to explore the role of spatial anxiety in students' geometric thinking as measured through their verbal proof and gesture. Preliminary findings from this study were published as a conference paper at the 2022 North American Chapter of the International Group for the Psychology of Mathematics Education Conference. These studies are presented with modifications to remove redundancies and provide consistent formatting. Finally, Chapter 4 describes a randomized controlled study testing the potential of a Tangram task which involves spatial thinking for improving students' geometric thinking and gesture production. This study was preregistered[a]. I then concluded with a brief discussion of the overall implications of the framework and studies, general limitations of the studies, and proposed future work.

[a] https://osf.io/dgznu/

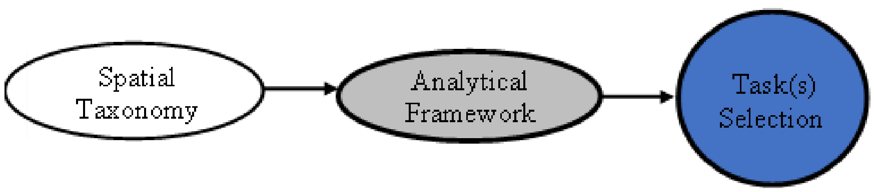
Chapter 1: Navigating Spatial Ability for Mathematics Education

Spatial ability can be broadly defined as imagining, maintaining, and manipulating spatial information and relations. Over the past several decades, researchers have found strong connections between spatial abilities and mathematics performance (e.g., Newcombe, 2013; Young et al., 2018a). However, the sheer plurality of spatial taxonomies and analytical frameworks that scholars use to describe spatial skills, the lack of theoretical spatial taxonomies, and the variety of spatial assessments available make it very difficult for education researchers to make the appropriate selection of spatial measures for their investigations. Researchers also face the daunting task of selecting the ideal spatial skills to design studies and interventions to enhance student learning and the development of reasoning mathematics and STEM (science, technology, engineering, and mathematics) more broadly. To address these needs, we have provided a review that focuses on the relationship between spatial skills and mathematical thinking and learning. Our specific contribution is to offer a practical guide for navigating and selecting among the various major taxonomies on spatial reasoning and various instruments for assessing spatial skills for use in mathematics education research.

The central objective of this review is to describe an organizational framework that acknowledges this muddled picture and strives to operate within it rather than offer overly optimistic proposals for resolving long-standing complexities. This review offers guidance through this complicated state of the literature to help STEM education researchers select appropriate spatial measures and taxonomies in service of their investigations, assessments, and interventions. We review and synthesize several lines of the spatial ability literature and provide researchers exploring the link between spatial ability and mathematics education a guiding framework for research design. To foreshadow, this framework identifies three major design decisions to guide scholars and practitioners seeking to use spatial skills to enhance mathematics education and education research: (1) selecting a spatial ability taxonomy and (2) an analytical frameworks, and (3) choosing spatial tasks (Figure 1). This guiding framework is intended to provide educational researchers and practitioners with a common language and decision process for conducting research and instruction that engage learners' spatial abilities. We hope this may lead to an increased understanding of the associative and causal links between spatial and mathematical abilities to improve educational research and practice.

Figure 1

Major elements of an investigation into the role of spatial reasoning



The Importance of Spatial Reasoning for Mathematics and STEM Education.

Research has shown that there is a positive relationship between spatial skills and various domains of mathematics performance. Deficiencies in spatial abilities can create obstacles in STEM fields (Shea et al., 2001). Some scholars have sought to determine which mathematical concepts engage spatial thinking. For example, studies on specific mathematical concepts found spatial skills were associated with one-to-one mapping (Gallistel & Gelman, 1992), missing-term problems (Cheng & Mix, 2014), mental computation (Verdine et al., 2014), and various geometry concepts (Hannafin et al., 2008). Burte and colleagues (2017) proposed categories of mathematical concepts such as problem type, problem context, and spatial thinking level to target math improvements following spatial invention training. Their study concluded that mathematics problems that included visual representations, real-world contexts, and that involved spatial thinking are more likely to show improvement after embodied spatial training.

Although spatial skills are not typically taught in the general K-12 curriculum, these lines of research have led to recommendations that explicitly teaching children about spatial thinking could increase STEM achievement and retention (Sorby, 2009; Stieff & Uttal, 2015). Several studies have shown that spatial training may positively impact mathematics performance in both experimental and classroom settings. For example, Mix and colleagues (2021) found that training spatial visualization and form perception led to better mathematics performance in first- and sixth-grade students. Similarly, a large, community-based study (n=17,648) conducted by Judd and Klingberg (2021) found enhanced math performance when mathematics training was paired with spatial training. In classrooms, Lowrie and colleagues (2019) found that students who completed a spatial visualization intervention program did significantly better than control classes who received standard mathematics instruction on geometry and word problems. In a meta-analysis of studies exploring the possible effects of spatial training on mathematics, Hawes and colleagues (2022) found that spatial training was effective for increasing success on both spatial and mathematics tasks, moderated by age, use of physical manipulatives, and type of transfer. One limitation of this meta-analysis is that it only surveyed 26 studies.

The Varieties of Approaches to Spatial Reasoning

Difficulties observing spatial reasoning in practice have spurred substantive research focused on uncovering the nature of spatial ability and its sub-components. Factor-analytic studies throughout the last century have sought to determine if spatial ability exhibits a unitary structure or if it is more likely to be composed of various sub-factors such as mental rotation, spatial visualization, navigation, and spatial orientation (e.g., Buckley et al., 2018; Carroll, 1993;

Hegarty & Waller, 2005; Spearman, 1927; Thurstone, 1950). Attempts to define and classify spatial sub-components relate mainly to psychometric indices and the associations between test item performance characterizing spatial skills and are often not clearly grounded in accepted definitions, theoretical bases, or interpretations of findings in the field (Uttal et al., 2013).

Still, the specific nature of these associations is largely unknown. Several lines of research have suggested that shared processing requirements from mathematical and spatial tasks could account for these associations. For example, brain imaging studies have shown similar brain activation patterns in spatial and mathematics tasks (Amalric & Dehaene, 2016; Hubbard et al., 2005; Walsh, 2003). In a study of spatial and mathematical thinking, Mix and colleagues (2016) showed a robust within-domain factor structure and overlapping variance irrespective of task-specificity. They proposed that spatial scaling, visualization, and form perceptions are shared processes required when individuals perform various spatial and mathematical tasks.

Efforts to document the relationship between mathematics performance and spatial skills or to enhance mathematics through spatial skills interventions show significant limitations in their theoretical framing or generalizability of findings. For example, many studies designed to investigate and improve spatial abilities have focused on either a particular spatial sub-component or a particular mathematical skill. Much of the research has primarily focused on measuring only specific aspects of object-based spatial ability, with some scholars focusing exclusively on either small- or large-scale spatial skills. The most commonly studied spatial skills include small-scale skills such as individuals' abilities to mentally rotate objects, visualize objects from different perspectives, and find embedded figures (Linn & Peterson, 1985). An even more restricted body of research has considered large-scale skills involving spatial orientation, such as navigation and map reading (e.g., Hegarty et al., 2018; Hegarty & Waller, 2004; Frick, Möhring & Newcombe, 2014)

Currently, there is no commonly accepted definition of spatial ability or its exact sub-components in the literature (Carroll, 1993; Lohman, 1988; Michael et al., 1957; McGee, 1979; Yilmaz, 2009). This lack of convergence leaves the research and practitioner communities with a muddled picture of how to think about the relationship between spatial thinking and mathematics, designing effective interventions based on clear causal principles, selecting appropriate metrics, and analyzing outcome data. Consequently, there is insufficient guidance for mathematics and STEM education researchers to navigate the vast landscape of spatial taxonomies and analytical frameworks, select the most appropriate measures for documenting student outcomes, and design potential interventions targeting spatial abilities.

One notable exception is the work by Battista and colleagues (2018). They used think-aloud data from individual interviews and teaching experiments with elementary and middle-grade students to investigate the relationship between spatial reasoning and geometric reasoning. Their investigation yielded task-level accounts of spatial reasoning that are described in terms of their role in geometric thinking, such as parallelism and isometries, rather than generalized cognitive processes, such as mental rotation. These "property-based spatial analytic reasoning" processes decompose "objects into their parts using geometric properties to specify how the parts or shapes are related, and, using these relationships, operates on the parts" (Battista et al., 2018, pp. 196-197). As a result, these processes are described in ways that generally fall outside the standardized psychometric assessment instruments. Establishing bridges between education domain-centric analyses of this sort and traditional psychometric accounts about domain-general spatial abilities is central to our review.

Selecting a Spatial Taxonomy

One of the first decisions that mathematics education researchers need to make when designing a study that either aims to explore the role of spatial ability as a covariate or attempts to design an intervention around spatial ability for improving mathematics is selecting a spatial taxonomy that suits the data collected and the ways data will be analyzed. A *spatial taxonomy* is an organizational system for classifying spatial abilities, and thus serves an important role in both the theoretical framework for any inquiry as well as for interpreting and generalizing findings from empirical investigations. In practice, selecting a spatial taxonomy is often difficult for researchers due to the expansive research on spatial ability.

In an attempt to make the vast number of spatial taxonomies more navigable for mathematics researchers and educators, we distinguish between three general types of spatial taxonomies that are reflected in the current literature and help researchers in their selection process through guiding questions. These three spatial taxonomies are ones that classify according to: different specific spatial abilities, different broad spatial abilities, and those that treat spatial abilities as derived from a unitary factor structure. Although this is not a comprehensive account, these three spatial taxonomies were chosen to highlight the main sub-factor dissociations in the literature.

Specific-Factor Structures

Since the earliest conceptualization (e.g., Galton, 1879), the communities of researchers studying spatial abilities have struggled to converge on one all-encompassing definition or provide a complete list of its sub-components. Though the literature provides a variety of definitions of spatial ability that focus on the capacity to visualize and manipulate mental images (e.g., Battista, 2007; Gaughran, 2002; Lohman, 1979; Sorby, 1999), some scholars posit that it may be more precise to define spatial ability as a constellation of skills based on performance on tasks that load on individual spatial factors (Buckley et al., 2018).

Following this charge, studies throughout the first half of the twentieth century attempted to define the many sub-components of spatial ability using psychometric methods. Attempts to dissociate subfactors were often met with difficulty due to differing factor analytic techniques and spatial ability tests (D'Oliveira, 2004). The subsequent lack of cohesion in this field of study led to different camps of researchers adopting inconsistent names for spatial sub-components (Cooper & Mumaw, 1985; McGee, 1979) and divergent factorial frameworks (Hegarty & Waller, 2005; Yilmaz, 2009). For example, McGee (1979) describes factor analytic studies in which Witkin's (1950) field dependence/independence instruments, such as the Rod and Frame Test (RFT) and Embedded Figures Test (EFT) that assess the extent individuals rely on internal or external referents, and Guilford and Zimmerman's (1948) spatial orientation instrument that assessed individuals' ability to orientate in space relative to objects emerge together in a factor, indicating that these tests measure a similar psychological construct. However, Linn and Peterson (1985) replace McGee's spatial orientation factor with a factor called spatial perception, which describes the ability of individuals to determine relationships with respect to their own bodies while ignoring distractions. They also provide evidence that supports the claim that some of Witkin's (1950) field dependence/independence instruments, such as the RFT, align more closely with spatial perception, while other field dependence/independence instruments, such as the EFT, better align with spatial visualization which refers to an individual's ability to mentally rotate and manipulate objects. Such a lack of convergence is clearly problematic for the scientific study of spatial ability.

In light of this, several reviews attempted to better organize the literature in the second half of the twentieth century and into the early part of the twenty-first century. However, these reviews differed in important ways. For example, to provide greater order to the inconsistent and confusing names of spatial ability subfactors, McGee (1979) suggested a simple two-factor taxonomy comprised of spatial visualization and spatial orientation. However, wading into the same literature, Lohman (1979) argued for a three-factor spatial ability taxonomy, including spatial visualization, orientation, and relations. The additional factor, *spatial relations*, was defined as tasks requiring mental rotation with an emphasis on the ability to solve problems quickly and was later (Lohman, 1988) renamed *speeded rotation* to emphasize the relation to processing time. Carroll (1993) subsequently identified five major clusters of spatial ability sub-components. Two sub-components, spatial visualization and spatial relations, do not differ from Lohman's (1979) categories. Carroll's three additional factors included flexibility of closure, closure speed, and perceptual speed, which describe the abilities to detect a known hidden pattern, complete a hidden pattern, and compare figures or symbols.

In the last few decades, several attempts have been made to dissociate sub-components of spatial ability further. Yilmaz (2009) combined aspects of the taxonomies described above with studies identifying dynamic spatial abilities and environmental spatial abilities to create an eight-factor taxonomy that acknowledges several spatial skills (e.g., environmental ability and spatiotemporal ability) needed in real-life situations. More recently, Buckley and colleagues (2018) proposed an extended taxonomy for spatial ability. This taxonomy combines many ideas from the previously described literature and the spatial factors identified in the Cattell-Horn-Carroll theory of intelligence (see Schneider & McGrew, 2012). It currently includes 25 factors that can also be divided into two broader categories of static and dynamic, with the authors acknowledging that additional factors may be added as research warrants. It is unclear how dissociation of these many subfactors could be practically applied in empirical research, which we regard as an important goal for bridging theory and research practices.

As demonstrated above, a variety of spatial taxonomies dissociate factors along different lines through various factor analytic methods. Though specific definitions vary, many scholars agree on making a dissociation between spatial orientation and visualization skills and, more recently, further dissociating mental rotation from other spatial visualization skills. These separations may inform researchers' spatial taxonomy and analytical framework selection and the selection of tasks used in their investigations. We next review several studies that use a specific-factor structure taxonomy and report on its benefits and limitations for education research.

Dissociation Between Spatial Orientation and Spatial Visualization

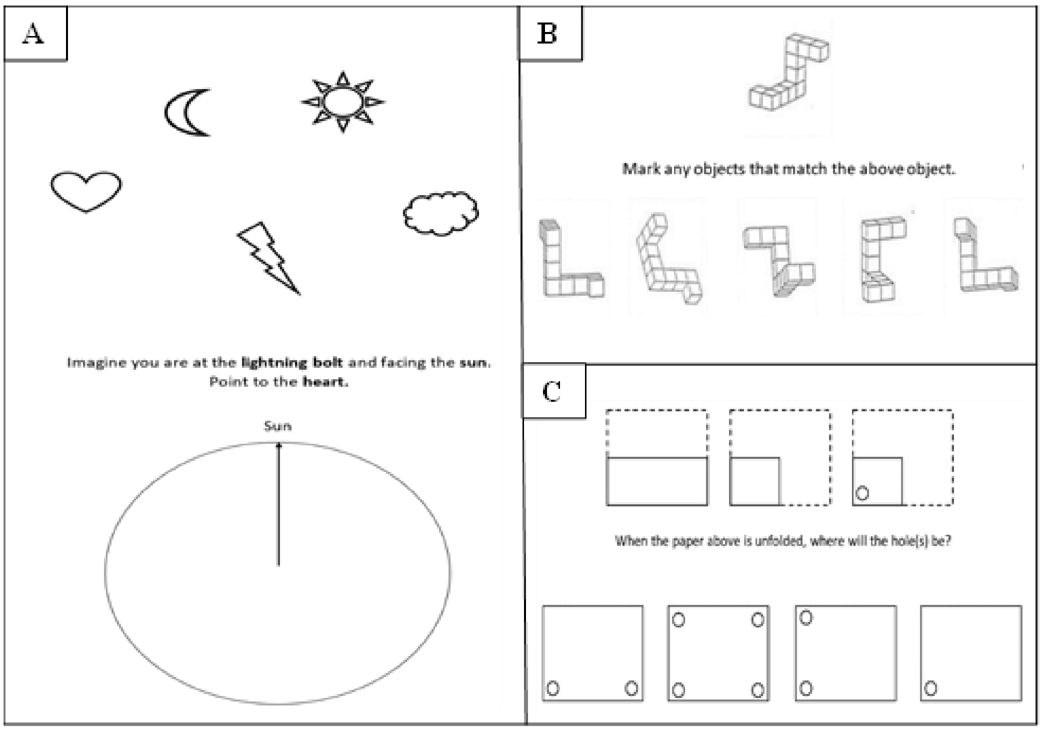
On the surface, perspective-taking (a subfactor of spatial orientation) and spatial visualization may seem equivalent. Measures for these skills often ask participants to anticipate the appearance of arrays of objects after either a rotation (visualization) of the objects or a change in the objects' perspective (perspective-taking), and both depend on a form of rotation. However, several studies have indicated that the spatial orientation subfactor of perspective-taking and the spatial visualization subfactor of spatial visualization processes are psychometrically separate skills (e.g., Huttenlocher & Presson, 1979; Kozhevnikov & Hegarty, 2001; Wraga et al., 2000). Perspective rotation tasks often lead to egocentric errors, such as reflection errors when trying to reorient perspectives, while object rotation task errors are not as systematic (Kozhevnikov & Hegarty, 2001; Zacks et al., 2000). For example, to solve a spatial orientation/perspective-taking task (Figure 2A), participants may imagine their bodies moving to a new position or viewpoint with the objects of interest remaining stationary. In contrast, the objects in a spatial visualization task are often rotated in one's imagination (Figure 2B). Behavioral and neuroscience evidence is consistent with these findings suggesting a dissociation between an object-to-object representational system and a self-to-object representational system (Hegarty & Waller, 2004; Kosslyn et al., 1998; Zacks et al., 1999). Thus, within the specific-factor structure of spatial ability, spatial orientation/perspective-taking can be considered a separate factor from spatial visualization/mental rotation (Thurstone, 1950).

Dissociation Between Mental Rotation and Non-rotational Spatial Visualization

The boundaries between specific factors of spatial ability are often blurred and context dependent. To address this, Ramful and colleagues (2017) have created a three-factor taxonomy that clarifies the distinct differences between spatial visualization and spatial orientation (see section 1.1.1) by treating mental rotation as a separate factor. Their taxonomy is unique in that they used mathematics curricula, rather than solely basing their analysis on a factor analysis, to identify three sub-factors of spatial ability: (1) mental rotation, (2) spatial orientation, and (3) spatial visualization. *Mental rotation* describes how one imagines how a two-dimensional or three-dimensional object would appear after it has been turned (Figure 2B). Mental rotation is a cognitive process that has received considerable attention from psychologists (Bruce & Hawes, 2015; Lombardi et al., 2019; Maeda & Yoon, 2013). *Spatial orientation*, in contrast, involves egocentric representations of objects and locations and includes the notion of perspective-taking (Figure 2A). *Spatial visualization* in their classification system (previously an umbrella term for many spatial skills that included mental rotation) describes mental transformations that do not require mental rotation or spatial orientation (Linn & Peterson, 1985) and can be measured through tasks like those shown in Figure 2C that involve operations such as paper folding and unfolding. Under this definition, spatial visualization may involve complex sequences in which intermediate steps may need to be stored in spatial working memory (Shah & Miyake, 1996). In mathematics, these skills often correlate with symmetry, geometric translations, part to whole relationships, and geometric nets (Ramful et al., 2017).

Figure 2

Exemplars of spatial orientation, mental rotation, and non-rotational spatial visualization tasks



Note. The spatial orientation task (A) is adapted from Hegarty and Waller's (2004) Object Perception/Spatial Orientation Test. The mental rotation task (B) is redrawn and adapted from Vandenberg and Kuse's (1978) Mental Rotation Test. The non-rotational spatial visualization task (C) is redrawn and adapted from Ekstrom and colleagues' (1976) Paper Folding Task.

Recommended Uses, Analytic Frameworks, and Limitations

Specific-factor taxonomies are used in a variety of lines of research, including mathematics education. Studies exploring the association between spatial ability and mathematics often focus on a particular sub-factor. For example, some studies have focused on the association between mental rotation and numerical representations (e.g., Rutherford et al., 2018; Thompson et al., 2013), while others have focused on spatial orientation and mathematical problem solving (e.g., Tartre, 1990). Similarly, those investigating spatial training efficacy often use spatial tasks based on a single factor or a set of factors as pre- and post-test measures and in intervention designs (e.g., Bruce & Hawes, 2014; Gilligan et al., 2019; Lowrie et al., 2019; Mix et al., 2021). However, the sheer number of these specific-factor spatial taxonomies used in education can be overwhelming for researchers, and many are not based in spatial ability theory, so it may be challenging for education researchers to find a suitable, evidence-based taxonomy for their work. The discussion above serves as an evidence-based guide to help researchers to ground their understanding of the various specific factor definitions and narrow their search for an appropriate specific-factor spatial taxonomy.

We recommend that education researchers first decide how many specific factors are relevant for their study by either selecting a taxonomy used in previous research or through exploratory pilot studies to identify the ecologically valid and contextually relevant spatial subfactors. In the design of these pilot studies, we suggest that researchers include several tests of a particular specific spatial sub-factor of interest or of multiple related spatial subfactors that have been used in prior similar studies or may be theoretically relevant to the current study. We also suggest that researchers include measures for constructs (e.g., gender, age, general intelligence, and working memory) that may need to be controlled for in the final study. Then, researchers should select the appropriate analytic framework to either use scores from single sub-factor measures or separate scores from multiple factor-specific measures. As a best practice, we recommend using a composite score for many measures to combat issues such as task-related biases (Moreu & Weibels, 2021). In addition, researchers should be careful to acknowledge the inherent limitations that arise as multiple specific sub-factors may load on any one particular mathematics task (see section 2 for discussion).

Broad-Factor Structures

Studies of mathematical reasoning and learning that rely on specific-factor structures can yield different results depending on their choices of factors (Schenck, Kim, et al., 2022). These differing results present a problem for finding convergence of the role of particular spatial abilities on particular mathematics concepts. An alternative approach relies on much broader distinctions between spatial ability sub-components. Proponents of broad-factor structure approaches argue that many traditional specific-factor structures of spatial ability rely on exploratory factor analysis rather than confirmatory factor analyses informed by a clear theoretical basis of spatial ability (Uttal et al., 2013; Young et al., 2018b). We refer to these as *broad-factor structure* approaches as their categorizations align with combinations of specific spatial ability subfactors.

Some scholars who draw on broad-factor structures have argued for a partial dissociation between large-scale and small-scale spatial abilities (Ferguson et al., 2015; Hegarty et al., 2006; Hegarty et al., 2018; Jansen, 2009; Potter, 1995). Large-scale spatial abilities involve physical navigation through space (e.g., moving one's body through a new environment). Small-scale spatial abilities are defined as those that predominantly rely on mental transformations of shapes or objects (e.g., mental rotation tasks) but do not necessitate large-scale body movement. Hegarty and colleagues (2018) recommend that large-scale abilities can be measured through sense of direction measures and navigation activities. The same study also suggests that small-scale abilities may be measured through typical spatial ability tasks like those discussed in section 2 of this paper, such as mental rotation. A meta-analysis examining the relationship between small- and large-scale abilities provided further evidence that these two factors should be defined separately, with additional evidence that this relationship may be moderated by age but not gender (Wang et al., 2014). Though, another recent meta-analysis investigating gender differences in large- and small-scale spatial ability provided evidence for a high level of gender differences in large-scale ability and a medium level of gender differences in small-scale ability due to differing neural bases and strategy use between the genders (Yuan et al., 2019).

Other lines of broad-factor structures research have drawn on linguistic, cognitive, and neuroscientific findings to develop a 2x2 classification system that distinguishes between intrinsic and extrinsic information and static and dynamic tasks (Newcombe & Shipley, 2015; Uttal et al., 2013). Intrinsic spatial skills involve attention to a single object's spatial properties, while extrinsic spatial skills predominately rely on attention to the spatial relationships between objects. The second distinction defines static tasks as recognizing and thinking about objects and their relations. In contrast, dynamic tasks often move beyond static coding of the spatial features of an object and its relations to imagining spatial transformations of one or more objects.

Uttal and colleagues (2013) describe how this 2x2 broad classification framework combining intrinsic and extrinsic information with static and dynamic reasoning can be mapped onto Linn and Peterson's (1985) three-factor taxonomy, breaking spatial ability into spatial perception, mental rotation, and spatial visualization sub-factors. *Spatial perception* tasks (e.g., water level tasks; see Inhelder & Piaget, 1958) require coding spatial position information between objects or gravity without manipulating the objects and nicely represent the extrinsic-static categorization. Mental rotation tasks (e.g., the Mental Rotations Test of Vandenberg & Kuse, 1978) represent the intrinsic-dynamic category. Spatial visualization tasks fall into the intrinsic classification and can address static and dynamic reasoning depending on whether the objects are unchanged or require spatial transformations. Ekstrom et al.'s (1976) Form Board Test and Paper Folding Test are two examples of spatial visualization tasks that measure the intrinsic-dynamic classification, while the Embedded Figures Test (Witkin et al., 1971) is an example of an intrinsic-static classification. Furthermore, Uttal and colleagues (2013) address a limitation of Linn and Peterson's (1985) taxonomy by including the extrinsic/dynamic classification, which they note can be measured through spatial orientation and navigation instruments such as the Guilford-Zimmerman Spatial Orientation Task (Guilford & Zimmerman, 1948).

Though Uttal and colleagues' (2013) classification provides a helpful framework for investigating spatial ability and its links to mathematics (Young et al., 2018b), it faces several challenges. Some critics posit that spatial tasks often require a combination of spatial sub-components and cannot be easily mapped onto one domain in the framework (Okamoto et al., 2015). For example, a think-aloud task might ask students to describe a different viewpoint of an object. The student may imagine a rotated object (intrinsic-dynamic), moving their body to the new viewpoint (extrinsic-dynamic), or a combination of strategies. Additionally, an experimental study by Mix and colleagues (2018) testing the 2x2 classification framework on children in Kindergarten, 3rd and 6th grade using confirmatory factor analysis failed to find evidence for the static-dynamic dimension at any age or for the overall 2x2 classification framework. This study demonstrates limitations to this framework in practice, and other frameworks with less dimensionality may be more appropriate for understanding spatial abilities in children.

Even in light of these challenges, broad-factor taxonomies may benefit researchers who do not expect specific sub-factors of spatial ability to be relevant for their data or are using a spatial ability measure to control for spatial ability as part of an investigation of a related construct. No valid and reliable instruments have been designed to specifically assess these broad-factor taxonomies. Instead, the developers of these taxonomies suggest mapping existing spatial tasks, which are usually tied to specific sub-factors of spatial ability, to broader categories. For practical applications, we recommend that education researchers first decide whether one or multiple broad factors are relevant for the study by aligning spatial measures used in similar prior studies with a particular broad-factor taxonomy. We suggest selecting an appropriate task and using a single variable for analyses if one broad factor appears relevant. If multiple broad factors are relevant, researchers should select one measure for each relevant factor and use these as separate scores for analyses. As before, we recommend using a composite score from multiple measures and controlling for individual factors that may correlate with spatial ability, such as gender, age, general intelligence, and working memory, where possible.

Unitary-Factor Structure

Many scholars understand spatial ability to be composed of a subset of specific or broad factors. However, there is also empirical support for an additional view: Spatial ability may actually be a *unitary construct*. Spatial ability has been investigated using factor analytical methods that sought to map the structure of general intelligence (Spearman, 1927; Thurstone, 1938). These early studies identified spatial ability as one factor separate from general intelligence that mentally operates on spatial or visual images. Evidence supporting a unitary spatial ability model proposes a common genetic network that supports all spatial abilities (Malanchini et al., 2020). In their study, 16 spatial tests were administered clustered into three main sub-components: *Visualization*, *Navigation*, and *Object Manipulation*. They then conducted a series of confirmatory factor analyses to fit one-factor (Spatial Ability), two-factor (Spatial Orientation and Object Manipulation), and three-factor models (Visualization, Navigation, and Object Manipulation). The best fitting model was the one-factor model -- even when accounting for general intelligence.

This unitary structure could be beneficial for researchers interested in questions about general associations between mathematics and spatial ability or using spatial ability as a moderator in their analyses. However, to date, no valid and reliable instruments have been created to fit within the unitary taxonomy, such as those that include various spatial items. Instead, researchers who discuss spatial ability as a unitary construct often choose one or multiple well-known spatial measures based on a particular sub-factor of spatial ability (e.g., Boonen et al., 2013; Burte et al., 2017). In the absence of a unitary spatial cognition measure designed to assess spatial ability from a unitary perspective, researchers should think critically about selecting measures for their studies and address the limitations of such decisions.

For example, Casey and colleagues (2015) found that early spatial skills were long-term predictors of later math reasoning skills. In their analysis, the authors identified two key spatial skills, mental rotation and spatial visualization, that previous work by Mix and Cheng (2012) found to be highly associated with mathematics performance. To measure these constructs, Casey and colleagues administered three spatial tasks to participants: a spatial visualization measure, a 2-D mental rotation measure, and a 3-D mental rotation measure. The authors were interested in the impact of overall spatial ability and partially replicating previous findings rather than whether or how these two factors impacted mathematics performance. Thus, they combined these three spatial scores into a single composite score. The decision to use a composite score across separate measures aligns with a unitary-factor taxonomy and demonstrates a way to use this taxonomy in the absence of spatial measures designed for a unitary-factor taxonomy.

This example illustrates why we recommend that researchers who select the unitary structure of spatial ability either address possible limitations of their spatial task selection or create a composite score from several tasks that align to various spatial sub-factors and extract the latent factor. Though some confirmatory analyses have shown that there is no clear separation of factors among common spatial ability measures (Colom et al., 2001), this issue motivates the need for an evidence-based, theory-grounded task selection procedure such as the one described in Section 2 and the need for a unitary spatial cognition measure.

An Evidence-Based Procedure for Choosing Spatial Tasks

With so many spatial ability taxonomies, researchers must carefully select tasks and surveys that match their stated research goals and theoretical framework, the spatial ability skills of interest, and the populations under investigation. As mentioned, researchers may select spatial tasks based on practical motivations rather than theoretical ones. These decisions can be complicated by the vast number of spatial tasks, with little guidance for which ones best align with the various spatial taxonomies. To help guide researchers with these decisions, we have compiled a list of spatial instruments referenced in this paper and matched them with their associated spatial sub-components and intended populations (Table 1).

Table 1

Alphabetical List of Spatial Ability Instruments

Instruments	Spatial Ability Factors	Population	Citation
Card Rotation Test	small-scale; intrinsic-dynamic; mental rotation	Adults and adolescents	(Ekstrom et al., 1976)
Children's Embedded Figures Test	small-scale; intrinsic-static; spatial perception; field-independence; spatial visualization	5-12 year old children	(Karp & Konstadt, 1963)
Embedded Figures Test	small-scale; intrinsic-static; spatial perception; field-independence; spatial visualization	Adults	(Witkin et al., 1971)
Foam Board Test	small-scale; intrinsic-dynamic; spatial visualization	Adults and adolescents	(Ekstrom et al., 1976)
Gottschaldt Figures Test	small-scale; intrinsic-static; spatial perception; field-independence; spatial visualization	Adults	(Gottschaldt, 1926)
Group Embedded Figure Test	small-scale; intrinsic-static; spatial perception; field-independence; spatial visualization	Older children (10+), adolescents, and adults	(Oltman et al, 1971)
Guilford-Zimmerman Spatial Orientation Task	small-scale; extrinsic-dynamic; spatial orientation	Adults	(Guilfold & Zimmerman, 1948)
Hidden Figures Test	small-scale; intrinsic-static; spatial visualization	Adults	(Ekstrom et al., 1976)
Make-A-Dice Test	small-scale; intrinsic-dynamic; spatial visualization	Older children (10+), adolescents, and adults	(Burte et al., 2019b)
Mental Cutting Test	small-scale; intrinsic-dynamic; spatial visualization; mental rotation	Adults	(CEEB, 1939)
Mental Rotations Test	small-scale; intrinsic-dynamic; mental rotation	Adults	(Vandenberg & Kuse, 1978)
Navigational Strategy Questionnaire	large-scale; navigation	Adults and adolescents	(Zhong, 2016)
Paper Folding Test	small-scale; intrinsic-dynamic; spatial visualization	Adults	(Ekstrom et al., 1976)

Perspective Taking Test for Children	small-scale; extrinsic-dynamic; spatial orientation; perspective-taking	6-8 year old children	(Frick, Möhring, and Newcombe, 2014)
Perspective Taking/Object Perspective Test	small-scale; extrinsic-dynamic; spatial orientation; perspective-taking	Adults and adolescents	(Hegarty and Waller, 2004)
Picture Rotation Test	small-scale; intrinsic-dynamic; mental rotation	Pre-school aged children	(Quaiser-Pohl, 2003)
Purdue Spatial Visualization Test: Rotations	small-scale; intrinsic-dynamic; mental rotation	Adults and adolescents	(Guay, 1976)
Rod-and-Frame Test	small-scale; extrinsic-static; spatial perception; spatial orientation; field-dependence independence	Adults	(Witkin, 1950)
Rotated Colour Cube Test	small-scale; intrinsic-dynamic; mental rotation	7-11 year old children	(Lütke & Lange-Küttner, 2015)
Santa Barbara Sense of Direction Scale	large-scale; navigation	Adults and adolescents	(Hegarty et al., 2002)
Santa Barbara Solids Test	small-scale; intrinsic-dynamic; spatial visualization; mental rotation	Adults and adolescents	(Cohen & Hegarty, 2012)
SOIVET-Maze	large-scale; extrinsic-dynamic; navigation;	Adults	(da Costa et al., 2018)
Spatial Reasoning Instrument	small-scale; mental rotation, spatial orientation, spatial visualization	Adolescents	(Ramful et al., 2017)
Surface Development Test	small-scale; intrinsic-dynamic; spatial visualization	Adults and adolescents	(Ekstrom et al., 1976)
Test of Spatial Ability	small-scale; spatial visualization; mental rotation; spatial orientation	3-6 year old children	(Verdine et al., 2013)
Virtual SILC Test of Navigation	large-scale; extrinsic-dynamic; navigation	Adults and adolescents	(Weisberg et al., 2014)
Water level tasks	small-scale; extrinsic-static; spatial perception; spatial orientation	Any	(Inhelder & Piaget, 1958)

Comparing the instruments in these ways reveals several important gaps necessary to measure unitary spatial cognition that correlate with mathematics and spatial abilities across the lifespan. In particular, this analysis reveals an over-representation of certain quadrants of the 2 (intrinsic-extrinsic) x 2 (static-dynamic) classification system described in section 1.2. It shows a pressing need for more tasks explicitly designed for extrinsic-static classifications. It also reveals that the slate of available instruments is dominated by tasks that have only been tested on adults. These disparities are important for educational considerations and are taken up in the final section.

Due to the sheer number of spatial tasks and the observations that tasks do not load consistently on distinct spatial ability factors such as spatial visualization or mental rotation due to the tasks' complex nature, it would be impossible in the scope of this review to discuss every task-factor relationship. As a practical alternative, we have decided to group spatial ability tasks into three aggregated categories based on their dissociations, as discussed in the previous section: Spatial orientation tasks, non-rotational spatial visualization tasks, and mental rotation tasks (for examples, see Figure 2). We acknowledge that other scholars may continue to identify different aggregations of spatial reasoning tasks, including those used with mechanical reasoning and abstract reasoning tasks (e.g., Tversky, 2019; Wai et al., 2009). In our aggregated categories, mechanical reasoning tasks would align with either mental rotation or non-rotational tasks depending on the specific task demands, while abstract reasoning tasks would align most closely with non-rotational spatial visualization tasks.

As there are no universally accepted measures of spatial ability for each spatial factor, we have narrowed our discussion in this paper to include exemplars of validated, cognitive, pen-and-pencil spatial ability tasks. These tasks have been historically associated with various spatial ability factors rather than merely serving as measures of general intelligence or visuospatial working memory (Carroll, 1993) and are easily implemented and scored by educators and researchers without specialized software or statistical knowledge. Notably, this discussion of spatial ability tasks and instruments excludes self-report questionnaires such as the Navigational Strategy Questionnaire (Zhong & Kozhevnikov, 2016) and the Santa Barbara Sense of Direction Scale (Hegarty et al., 2002); navigation simulations such as the Virtual SILC Test of Navigation (Weisberg et al., 2014) and SOIVET-Maze (da Costa et al., 2018); and tasks that involve physical manipulation such as the Test of Spatial Ability (Verdine et al., 2014). As such, we were unable to find any published, validated instruments for large-scale spatial orientation, a sub-factor of spatial orientation, that meet our inclusion criteria. The following sections detail the types of tasks and instruments commonly used to measure spatial orientation, non-rotational spatial visualization, and mental rotation.

Spatial Orientation Tasks

Much like spatial ability more generally, spatial orientation skills fit into the broad distinctions of large-scale (e.g., wayfinding, navigation, and scaling abilities) and small-scale (e.g., perspective-taking and directional sense) skills, with small-scale spatial orientation skills being shown to be correlated with larger-scale spatial orientation skills (Hegarty et al., 2002; Hegarty & Waller, 2004). Although few studies have attempted to determine statistical associations between spatial orientation and mathematics, spatial orientation has been correlated with some forms of scholastic mathematical reasoning. One area of inquiry showed associations between spatial orientation and early arithmetic and number line estimation (Cornu et al., 2017; Zhang & Lin, 2015). In another, spatial orientation skills were statistically associated with problem-solving strategies and flexible strategy use during high school level geometric and nongeometric tasks (Tartre, 1990). Studies of disoriented children as young as three years old show that they reorient themselves based on the Euclidean geometric properties of distance and direction, which may contribute to children's developing abstract geometric intuitions (Izard et al., 2011; Lee et al., 2012; Newcombe et al., 2009).

The Guilford-Zimmerman (GZ) Spatial Orientation Test is one of the earliest spatial orientation instruments (Guilford & Zimmerman, 1948). Participants are asked to identify a boat's position that would give a particular view of the landscape. However, psychometric studies of spatial orientation have shown conflicting results on the separability of spatial orientation and spatial visualization. Some researchers conclude that the GZ Spatial Orientation Test and spatial visualization instruments measure the same factors (Lohman, 1979; Schultz, 1991). It is possible that participants' strategies could be to rotate the objects rather than shift their perspectives. Other critics of this instrument believe the GZ Spatial Orientation Test tasks are complicated and confusing for participants. One study reported that 98% of participants commented that their orientation skill test score did not reflect their ability (Kyritsis & Gulliver, 2009).

To combat the GZ Spatial Orientation Test problems, Kozhevnikov and Hegarty (2001) developed the Object Perspective Taking Test, which was later revised into the Object Perspective/Spatial Orientation Test (see Figure 2A; Hegarty & Waller, 2004). Test takers are prevented from physically moving the test booklet, and all items involved an imagined perspective change of at least 90°. Unlike previous instruments, results from the Object Perspective/Spatial Orientation Test showed a dissociation between spatial orientation and spatial visualization factors (though they were highly correlated) and correlated with self-reported judgments of large-scale spatial cognition. A similar instrument, the Perspective Taking

Test for Children, has been developed for younger children. (Frick, Möhring, & Newcombe, 2014). Additionally, simpler versions of these tasks that asked participants to match an object to one drawn from an alternative point of view have also been used, such as those in the Spatial Reasoning Instrument (Ramful et al., 2017).

Non-Rotational Spatial Visualization Tasks

With differing definitions of spatial visualization, measures of this spatial ability sub-component often include tasks that evaluate other spatial ability skills. For example, cross-sectioning tasks such as those in the Mental Cutting Test (CEEB, 1939) and the Santa Barbara Solids Test (Cohen & Hegarty, 2012) ask students to choose the correct cross-section from a criterion figure to be cut with an assumed plane. At first, this may seem like a straightforward spatial visualization task, with some scholars suggesting cross-sectioning tasks are distinct from mental rotation and spatial perception (Ratliff et al., 2010); the nature of the drawings requires participants to visualize the cross-section and then rotate either the object or perspective to match the given answer choices indicating a potential overlap between spatial ability factors. In this section, we focus on tasks that do not overtly require mental rotation.

The three tests for non-rotational spatial visualization come from the Kit of Factor-Referenced Cognitive Tests developed by Educational Testing Services (Ekstrom et al., 1976). These instruments were developed for research on cognitive factors in adult populations. The first instrument is the Paper Folding Test (PFT), one of the most commonly used tests for measuring spatial visualization (see Figure 2C). In this test, participants view diagrams of a square sheet of paper being folded and then punched with a hole. They are asked to select the picture that correctly shows the resulting holes after the paper is unfolded. Though this task assumes participants imagine unfolding the paper without the need to rotate, studies have shown that problem attributes (e.g., number and type of folds and fold occlusions) impact PFT accuracy and strategy use (Burte et al., 2019a).

The second instrument is the Form Board Test. Participants are shown an outline of a complete geometric figure with a row of five shaded pieces. The task is to decide which shaded pieces will make the complete figure when put together. During the task, participants are told that the pieces can be turned but not flipped and can sketch how they may fit together.

The third instrument, the Surface Development Test, asks participants to match the sides of a net of a figure to the sides of a drawing of a three-dimensional figure. Like the PFT, strategy use may also impact accuracy on these two measures. To combat the possible influence of strategy use, the Make-A-Dice test was created (Burte et al., 2019b), which relies on the number of squares in a row and consecutive folding in different directions rather than just increasing the number of folds to increase difficulty. Additionally, none of these three instruments were explicitly designed to test non-rotational spatial visualization but rather a broader definition of spatial visualization that includes mental rotation. Thus, it is possible that some participants' strategies may include mental rotation or spatial orientation.

Other common types of spatial visualization tasks include embedded figures adapted from the Gottschaldt Figures Test (Gottschaldt, 1926). These tasks measure spatial perception, field-independence, and the ability to disembed shapes from a background, which may be a necessary problem-solving skill (Witkin et al., 1977). One instrument, the Embedded Figures Test, originally consisted of 24 trials during which a participant is presented with a complex figure, then a simple figure, then shown the complex figure again with instructions to locate the simple figure within it (Witkin, 1950). Others have used Witkin's (1950) stimuli as a basis to develop various embedded figures tests, including the Children's Embedded Figures Test (Karp & Konstadt, 1963) and the Group Embedded Figure Test (Oltman et al., 1971).

A test of spatial ability based on Ramful and colleagues' (2017) Spatial Reasoning Instrument (SRI) employs a three-factor taxonomy: Spatial orientation, spatial visualization, and mental rotation. Notably, the questions that measure spatial visualization are specifically designed not to require mental rotation or spatial orientation. Unlike previously mentioned instruments, the SRI is not a speed test, though students are given a total time limit of 45 minutes. Additionally, this instrument targets middle school students and was designed to align more closely with students' mathematical curricular experiences rather than a traditional psychological orientation. Mathematical connections in the SRI include visualizing lines of symmetry, using two-dimensional nets to answer questions about corresponding three-dimensional shapes, and reflecting objects.

Mental Rotation Tasks

Mental rotation is a cognitive operation in which a mental image is formed and rotated in space. Mental rotation can be treated as a separate skill from spatial orientation and spatial visualization (Linn & Peterson, 1985; Shepard & Metzler, 1971). This process develops from 3 to 5 years of age with large individual differences (Estes, 1998) and shows varying performance across individuals irrespective of other intelligence measures (Borst et al., 2011). Several studies have also demonstrated significant gender differences, with males typically outperforming women (e.g., Voyer et al., 1995), though this gap may be decreasing across generations (Richardson, 1994), suggesting it is due at least in part to sociocultural factors such as educational experiences rather than exclusively based on genetic factors.

Though there is disagreement on how mental rotation fits within the subset of spatial skills, behavioral and imaging evidence suggests that mental rotation tasks invoke visuospatial representations that correspond to object rotation in the physical world (Carpenter et al., 1999; Shepard & Metzler, 1971). Mental rotation skills are often subsumed under spatial visualization or spatial relations sub-components. Historically, three-dimensional mental rotation ability has fallen under the spatial visualization skill, while two-dimensional mental rotation occasionally falls under a separate spatial relations skill (e.g., Carroll, 1993; Lohman, 1979). Thus, mental rotation measures often include either three-dimensional or two-dimensional stimuli rather than a mixture of both.

As many definitions of general spatial ability include a "rotation" aspect, several studies have investigated the links between mental rotation and mathematics. For young children, cross-sectional studies have shown mixed results. Some studies found significant correlations between mental rotation and calculation and arithmetic (Cheng & Mix, 2014; Gunderson et al., 2012; Hawes et al., 2015). Conversely, Carr et al. (2008) found no significant associations between mental rotation and standardized mathematics performances in similar populations. In middle school-aged children (11-13 years), mental rotation skill was positively associated with geometric knowledge (Battista, 1990; Casey et al., 1999) and problem-solving (Hegarty & Kozhevnikov, 1999; Delgado & Prieto, 2004). Studies on high school students and adults have indicated that mental rotation is associated with increased accuracy on mental arithmetic problems (Geary et al., 2000; Kyttälä & Lehto, 2008; Reukala, 2001).

Three-dimensional Mental Rotation Tasks

In one of the earliest studies of three-dimensional mental rotation, Shepard and Metzler (1971) presented participants with pictures of pairs of objects that were either the same objects in the same or different orientations or different objects because they were mirror images of those objects. This design provided a nice control since the mirror images were truly different objects (i.e., they could not be rotated to match the original). Yet, they shared comparable visual complexity to those that were the same. Participants were asked to answer as quickly as possible whether the objects were the same or different, regardless of the orientation difference. Results found a positive linear association between reaction time and the angular difference in the orientation of objects. In combination with participant post-interviews, this finding illustrated that in order to make an accurate comparison between the object and the answer questions, participants first imagined the object as rotated into the same orientation as the target object and that participants perceived the two-dimensional pictures as three-dimensional objects in order to complete the imagined rotation. Additional studies have replicated these findings over the last four decades (Uttal et al., 2013). Shepard and Metzler-type stimuli have been used in many different instruments, including the Purdue Spatial Visualization Test: Rotations (Guay, 1976) and the Mental Rotation Test (see Figure 2B; Vandenberg & Kuse, 1978). However, recent studies have shown that some items on the Mental Rotation Test can be solved using analytic strategies such as a global-shape strategy to eliminate answer choices rather than relying on mental rotation strategies (Hegarty, 2018).

One common critique of the Shepard and Metzler-type stimuli is that the complex design of the classic cube configurations is inappropriate for younger populations, leading to few mental rotation studies on this population. Studies have shown that children under five years of age have severe difficulties solving standard mental rotation tasks, with children between the ages of 5 and 9 solving such tasks at chance (Frick, Hanson, & Newcombe, 2014). To combat this, studies with pre-school age children often lower task demands by reducing the number of answer choices, removing mirrored and incongruent stimuli, and using exclusively two-dimensional pictures (Krüger et al., 2013; Krüger, 2018). In response, some scholars have begun developing appropriate three-dimensional mental rotation instruments for elementary school students, such as the Rotated Colour Cube Test (Lütke & Lange-Küttner, 2015). In this instrument, participants are presented with a stimulus consisting of a single cube with different colored sides and are asked to identify an identical cube that has been rotated. However, studies on both three-dimensional and two-dimensional rotation have found that cognitive load depends more on the stimulus angle orientation than the object's complexity or dimensionality (Cooper, 1975; Jolicoeur et al., 1985).

Two-dimensional Mental Rotation Tasks

To measure two-dimensional mental rotation, tasks for all populations feature similar stimuli. These tasks, often referred to as spatial relations or speeded rotation tasks, typically involve single-step mental rotation (Carroll, 1993). One common instrument for two-dimensional mental rotation is the Card Rotation Test (Ekstrom et al., 1976). This instrument presents an initial figure and asks participants to select the rotated but not reflected items. Importantly, these tasks can be modified for various populations (Krüger et al., 2013). One standardized instrument for preschool and early primary school-age children, the Picture Rotation Test, demonstrates how easily these two-dimensional stimuli can be modified (Quaiser-Pohl, 2003).

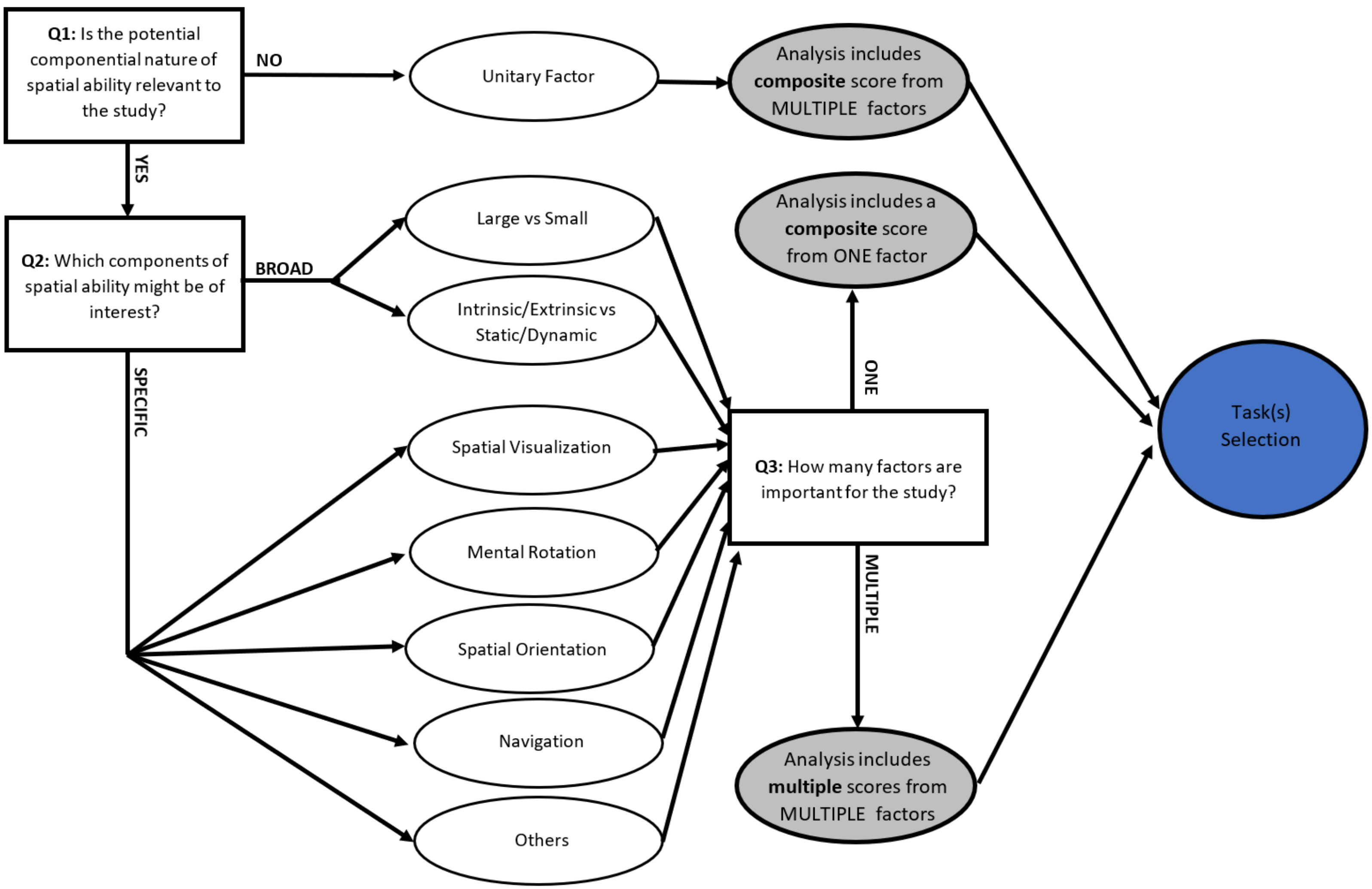
A Guiding Framework

We contend that the decisions made regarding the choice of spatial taxonomies, analytical frameworks, and spatial measures will impact both the results and interpretations of the study's findings. One way these decisions impact study outcomes is that specific sub-factors of spatial ability have been shown to be more strongly associated with specific sub-domains of mathematics than with others (Delgado & Prieto, 2004; Schenck & Nathan, 2020). Thus, selecting a spatial skills instrument poorly suited to the mathematical skills under investigation may fail to show a suitable predictive value and lead the research team to conclude that spatial reasoning overall is not relevant to the domain of interest. These design limitations are often not discussed in the publications we reviewed and, perhaps, may not even be realized by the researchers. However, due to the abundance of spatial taxonomies and lack of consensus in the spatial ability community, it can be difficult for education researchers to select an appropriate framework to match their specific domains of study and foresee the limitations affecting their outcomes.

Due to the various spatial taxonomies and the assumptions and design decisions needed for the accompanying analytical frameworks, we assert that it may be beneficial for some scholars and educators to move away from attempting to create a specific, universally accepted taxonomy of spatial ability. Instead, it may be more useful to consider spatial ability as a web of interconnected sub-components that describe how individuals perceive, generate, relate, and transform spatial information about objects, including themselves. Furthermore, the precise ways individuals interact with spatial information through the various possible sub-components appear to be based on a particular scholar's or educator's perspective, which should inform study design through their choices of spatial taxonomies and analytical frameworks and measures based on their goals.

Figure 3

Flow chart for selecting the appropriate spatial taxonomy and analytic frameworks for one's investigation



To address these shortcomings, we have designed a guide for researchers in the form of a flowchart that helps match spatial taxonomies to analytic frameworks (Figure 3). Our guide does not include every possible spatial taxonomy. Instead, it offers a helpful starting point for incorporating spatial skills into an investigation.

The first question in the flowchart, Q1, draws researchers to the relevance of subcomponents to their investigation. If sub-components do not apply to their goals, we recommend the unitary spatial taxonomy and an analytical framework using a single composite score from multiple tasks.

When sub-components are applicable, Q2 of the flowchart directs the researcher to consider whether these sub-components are specific or broadly defined. The flowchart then guides the investigator to choose a spatial taxonomy that most closely aligns with their research goals. Several analytic frameworks from our review are listed as compatible with specific-factor structures. Broad-factor structures mainly distinguish large-scale from small-scale spatial abilities or intrinsic/extrinsic skills from static/dynamic tasks.

Once an appropriate spatial taxonomy has been selected, Q3 in the flowchart directs researchers to whether one or more sub-components are directly relevant to their investigation. If only one factor is relevant, we suggest an analytic approach that includes a single score from one task among the chosen taxonomies. If multiple factors are relevant to the investigation, we suggest including multiple composite scores from multiple tasks in analyses. Note that the guidance here to use a single composite score from one task as a measure of spatial ability that is conceptualized as made up of sub-components differs from the guidance to use a single composite score from multiple tasks when spatial ability is conceptualized in terms of a unitary-factor structure. For example, if the research goal is to understand the links between mental rotation and a particular domain of mathematics, we suggest the researchers use a single composite score from mental rotation measures. But if the research goal is to understand the general links between spatial ability and a particular domain of mathematics, we suggest that the researcher use a composite score from multiple spatial measures in the absence of a universal spatial cognition measure. Task selection, the final step in the flow chart, will depend on practical considerations such as which spatial sub-components are relevant, population age, and time constraints. Though thousands of spatial tasks are available, the tasks listed in Table 1, which also identify corresponding broad and specific spatial sub-components, can serve as a useful starting point for designing a study.

Conclusions and Lingering Questions

Researchers largely agree that spatial ability is essential for mathematical reasoning and success in STEM fields (National Research Council, 2006). The two goals of this review were, first, to summarize the relevant spatial ability literature, including the various factor structures and measures, in an attempt to more clearly understand the elements of spatial ability that may relate most closely to mathematics education; and second, to provide practical recommendations for education researchers and practitioners for selecting appropriate theoretical taxonomies, analytical frameworks, and specific instruments for measuring, interpreting, and improving spatial reasoning for mathematics education. Our review showed a wide array of spatial taxonomies and analytical frameworks for understanding and measuring spatial reasoning. However, we found no consensus on the nature of this link, no convergence on a definition of spatial ability or agreement regarding the possible sub-components of spatial ability, and no universally accepted set of standardized measures to assess spatial skills. This review exposes several challenges to understanding the relationship between spatial skills and performance in mathematics, which are complicated by divergent descriptions of spatial taxonomies and analytical frameworks scholars use to describe spatial skills and the sheer volume of spatial assessments one encounters as a potential consumer. These challenges are all part of the responsibility of the research community. The lack of progress on these issues impedes progress in designing effective spatial skills interventions for improving mathematics thinking and learning based on clear causal principles, selecting appropriate metrics for documenting change, and analyzing and interpreting the student outcome data.

Our primary contribution in the context of these challenges is to provide a practical guide, well situated in the research literature, for navigating and selecting among the various major spatial taxonomies and various validated instruments for assessing spatial skills for use in mathematics education research and instructional design. In order to anchor our recommendations, we first summarized much of the history and major findings of spatial ability research as it relates to education (Section 1). In this summary, we identified three major types of spatial taxonomies: specific, broad, and unitary, and provided recommendations for associated analytical frameworks. We then discussed the plethora of spatial ability tasks that investigators and educators must navigate (Section 2). To make the landscape more tractable, we divided these tasks into three categories relevant to mathematics education -- spatial orientation, mental rotation, and non-rotational spatial visualization (see Table 1) -- and mapped these tasks to their intended populations. We acknowledge that researchers and educators often select spatial tasks and analytic frameworks for practical rather than theoretical reasons, which can undermine the validity of their own research and assessment efforts. In order to provide educators with a stronger evidence-based foundation, we offered a guiding framework in the form of a flowchart to assist investigators in selecting appropriate spatial taxonomies and analytic frameworks as a precursor to making well-suited task sections to meet their particular needs. A guide of this sort provides some of the best paths forward to utilizing the existing resources for understanding and improving education through the lens of spatial abilities. We focused our efforts on providing a tool to guide the decision-making of investigators and educators seeking to relate spatial skills with mathematics performance based on the existing resources and theoretical frameworks.

Several limitations remain, however. One is that the vast majority of published studies administered spatial skills assessments using paper-and-pencil instruments. In the ensuing years, testing has largely moved online, posing new challenges regarding the applicability of some of the instruments and findings. Updating these assessments will naturally take time until research using online instruments catches up. A second limitation is that studies investigating the associations between spatial ability and mathematics have often focused on a particular spatial or mathematical skill. There are many unknowns for which spatial abilities map to which areas of mathematics performance. This limitation can only be addressed through careful, large-scale studies. A third limitation is that many of the instruments in the published literature were developed for use with adult populations. This greatly limits their applicability to school-aged populations. Here, again, this limitation can only be addressed through more research that extends this work across a broader developmental range. Fourth, many spatial ability instruments reported in the literature include tasks that may be solved using various strategies, thus calling into question whether they measure the specific spatial skills they claim to measure. For example, some tasks, such as the Paper Folding Test, may be solved effectively through counting rather than pure spatial visualization. Thus, there is a pressing need for process-level data, such as immediate retrospective reports and eye tracking (cf. Just & Carpenter, 1985), to accurately describe the various cognitive processes involved and how they vary by age, individual differences, and assessment context. Fifth, there is a need for more tasks and instruments designed specifically for extrinsic-static classifications and unitary frameworks. Though there are currently no studies that provide evidence for a link between mathematics and extrinsic-static spatial abilities, it is possible that a lack of assessments has masked the connection. This limitation can only be addressed with more research into the development of these measures. Perhaps the greatest limitation is that scholarly research on spatial ability still lacks a convergent taxonomy and offers no clear picture of which aspects of spatial thinking are most relevant to STEM thinking and learning.

Overall, research to understand the structure of spatial ability more deeply is at a precipice. Understanding the sub-components of spatial ability and how their structures change with individual differences directly impacts how education researchers understand the relationship between spatial thinking and mathematical ability and how researchers and educational designers develop effective interventions. Synthesizing these lines of research highlighted several unexplored areas that require future work. STEM education and workforce development remain essential for scientific and economic advancements, and spatial skills are important for success and retention in technical fields. Thus, it is critical to further understand the connections between spatial and mathematical abilities as ways to increase our understanding of the science of learning and inform the design of future curricular interventions that transfer skills for science, technology, engineering, and mathematics.

Chapter 2: Exploring Spatial Ability, Spatial Anxiety, and Mathematical Thinking

Portions of this chapter have been published and are reproduced in modified form with permission from:

Schenck, K.E. & Nathan, M.J. (2020, April). *Connecting mathematics, spatial ability and spatial anxiety*. [Paper presentation]. American Educational Research Association Conference, San Francisco, CA, United States.

Schenck, K. E, Hubbard, E. M, Nathan, M., & Swart, M. (2022, July). Expanding Understandings of Embodied Mathematical Cognition in Students' Fraction Knowledge. In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society*, 44. Toronto, Canada.

Study 1: Connecting Mathematics, Spatial Ability, and Spatial Anxiety

Traditional models of mathematical cognition (e.g., Newell & Simon, 1972) have seen substantial expansion beyond individually isolated rule-based processing of symbol systems. While contextual and social influences on mathematical thinking have greatly influenced contemporary theories (e.g., Cobb, 1994; Schliemann & Carraher, 2002), two other classes of processes that extend our theories of mathematical cognition – embodiment and affect – are still in their infancy. One facet of embodied cognition particularly relevant for math is spatial skills (Uttal & Cohen, 2012; Mix et al., 2016; Newcombe, 2013). An aspect of affect relevant to math is anxiety. To further understand the breadth and multifaceted nature of mathematical thinking in terms of expanded notions of cognition, we investigate the relationship of mathematical ability with spatial skills and spatial anxiety. Models of mathematical cognition that include affective and embodied processes contribute to basic theory of human cognition. Such models also inform the design of evidence-based educational practices, including the design of curriculum activities and principles for instruction and assessment.

Theoretical Framework

Spatial Ability

Spatial ability refers to the ability to generate and manipulate spatial objects, images, relationships, and transformations (Battista, 2007). These abilities include memorization and comparison of visual patterns, manipulating mental objects, and performing spatial imagery (Hegarty & Waller, 2005). Spatial imagery, as defined by Hegarty & Waller (2005), refers to "a representation of the spatial relationships between parts of an object, the location of an object in space or their movement" (pg. 144). Spatial ability has been linked to success in mathematics for students as young as three (Verdine et al., 2014), and studies investigating phenomena such as the SNARC effect have the links between spatial-numerical associations and math ability (Berch et al., 1999; Toomarian et al., 2019). As children develop, spatial ability is consistently important. Assessment of spatial skills among elementary-aged children strongly predicts later mathematical capabilities (Casey et al., 2015; Laski et al., 2013).

There is little consensus in the literature about the exact combination of factors and sub-skills critical for spatial ability (Yilmaz, 2009), and naming conventions of the factors vary (Hegarty & Waller, 2005; Lohman, 1988; McGee,1979). We adopted the three-factor framework proposed by Ramful and colleagues (2016): (1) Mental rotation describes how one imagines a 2D or 3D object would appear after it has been turned; (2) spatial orientation involves egocentric representations of objects and locations and includes the notion of perspective taking; and (3) spatial visualization describes mental transformations that do not require mental rotation, spatial orientation, or egocentric reference.

Anxiety in Mathematics and Spatial Reasoning

High levels of anxiety reduce neurocognitive performance (Derakshan & Eysenck, 2010; MacLeod & Donnellan, 1993; Meyers et al., 2013). Mathematics anxiety predicts low mathematics performance even when controlling for other anxiety factors, such as test anxiety (Lukowski et al., 2016). Furthermore, math anxiety appears to be domain specific: When students simply anticipated doing math, those with high math anxiety exhibited greater activity in brain regions associated with threat detection, which was not present when they anticipated doing a reading activity (Lyons & Beilock, 2012).

Spatial anxiety as a psychological construct captures feelings of annoyance, confusion, and frustration when faced with spatial tasks. Spatial anxiety and spatial ability are negatively correlated (Malanchini et al., 2017). Furthermore, spatial anxiety, like spatial ability, appears to be composed of sub-components. One study of twins identified two components of spatial anxiety: navigation and rotation/visualization (Malanchini et al., 2017) or large versus small scale (Hegarty et al., 2018). Other studies have linked spatial anxiety to spatial orientation skills, specifically, a decrease in the efficiency of orientation strategies and an increase in errors on navigation tasks (Hund & Minarik, 2006; Lawton, 1994). However, we know little about the relationship between spatial anxiety on math performance.

Research Questions

There is emerging evidence for connections between the broad constructs of spatial ability, spatial anxiety, and math performance, overall, but little is known about how spatial anxiety relates to math performance, or the associations for their various subcomponents. Thus, our three main research questions are: What is the relationship between spatial ability and overall performance on math tests? (RQ1); What are the relationships between spatial ability sub-categories and performance on math performance sub-categories? (RQ2); What are the relationships between spatial anxiety, spatial ability, and performance on math tests? (RQ3).

Methods and Data Sources

We recruited 153 participants (18 years of age and above) through Amazon's Mechanical Turk service. All participants completed spatial ability, spatial anxiety, and mathematics ability assessments and a demographics survey. For spatial ability, participants completed the *Spatial Reasoning Instrument*, which breaks down into three sub-categories: mental rotation, spatial orientation, and spatial visualization (Ramful et al., 2016). The spatial anxiety measure was a combination of eight questions from the *Spatial Anxiety Scale* (Lawton, 1994) and five questions from the *Child Spatial Anxiety Questionnaire* (Ramirez et al., 2012) that were updated to fit the population. The mathematics ability measure was composed of a subset of 16 questions from the 2012 PISA Mathematics Test, which consisted of four sub-categories of questions: quantity, uncertainty and data, space and shape, and change and relationships (Organization for Economic Co-operation and Development, Programme for International Student Assessment, 2014).

All measures were completed online. In a previous pilot study, participants took an average of 26 minutes to complete the tasks. Based on this data, participants who completed the study in less than 25 minutes were excluded from the analyses. This restriction left a total of 101 participants. See Table 2 for descriptive statistics.

Table 2

Demographic and Descriptive Statistics (N=101)

Variables	**Mean (SD)**	**N (%)**
Age in years	34.05 (10.97)	
Sex, Female		62 (61%)
Native Language, English		86 (85%)
Ethnicity, White		63 (62%)
Total SRI Score	18.34 (6.67)	
Mental Rotation (MR) Subscore	5.89 (2.73)	
Spatial Visualization (SV) Subscore	4.58 (2.60)	
Spatial Orientation (SO) Subscore	7.87 (2.22)	
Total PISA Score	8.99 (3.45)	
Quantity (Q) Subscore	2.35 (1.07)	
Uncertainty and Data (UD) Subscore	2.31(1.14)	
Space and Shape (SS) Subscore	2.37 (1.05)	
Change and Relationships (CR) Subscore	1.92 (1.09)	
Spatial Anxiety Score	23.44 (12.39)	

Results

Reliability for the spatial ability, mathematics, and spatial anxiety measures exceeded the 0.7 threshold for the overall assessments and for each of the sub-scales, with the exception of the mathematics Quantity sub-scale, which had a Cronbach's $\alpha = 0.69$.

For our analyses, we used Ordinary Least Squares (OLS) regression. In each model, we included participant sex and age as fixed effects because these variables have been shown to be significant predictors of spatial ability, spatial anxiety, and mathematics ability (Lawton,1994; Maeda & Yoon, 2013, Voyer, Voyer & Bryden, 1995).

RQ1: Spatial and Math Ability

A multiple linear regression was calculated to predict overall math ability scores based on overall spatial ability scores while controlling for age and sex (RQ1). Spatial ability scores significantly predicted overall math ability scores (t (97) =10.07, $p < .001$). Each one-point increase in spatial ability scores increased math ability scores by 0.38 points. This result indicates that spatial ability may have an important relationship with math ability and gives further weight to previous findings that spatial ability may be an essential factor for success in STEM fields (Davis, 2015; Uttal & Cohen, 2012).

RQ2: Spatial and Math Ability Sub-categories

To investigate the relationships between spatial and math ability sub-categories, we calculated five separate multiple linear regression equations to predict math ability scores based on each of the spatial ability sub-categories (mental rotation, spatial orientation, and spatial visualization) controlling for age and sex (RQ2). The first model predicted overall math ability by the three spatial ability sub-categories. The subsequent models predicted the four sub-categories of mathematics ability (quantity, uncertainty and data, space and shape, and change and relationships) by the three spatial ability sub-categories. Overall, at least one spatial ability sub-category significantly predicted the dependent variable. No sex effects were found in any model.

Mental rotation was significantly predictive of scores for both uncertainty and data questions (UD; t (95) = 2.84, p = .005) and change and relationship questions (CR; t (95) = 2.50, p = .014) but not math ability overall. Spatial orientation was significantly predictive of scores for overall math ability (t (95) = 3.50, p = 0.001), quantity questions (Q; (t (95) = 3.94, p < .000), and uncertainty and data questions (UD; t (95) = 1.99, p = .049). Spatial visualization was significantly predictive of scores for both uncertainty and data questions (UD; t (95) = 2.18, p = .031) and space and shape questions (SS; t (95) = 3.10, p = .003).

RQ3: Spatial Ability, Math Ability, and Spatial Anxiety

The last set of analyses further examined the relationships between spatial ability, spatial anxiety, and math ability (RQ3). Results indicated that spatial anxiety significantly predicted spatial ability scores (t (97) = -8.18, p < .000). As expected, this relationship was negative, which is consistent with previous work on the association between spatial anxiety and performance (e.g., Malanchini et al., 2017).

A second model indicated that spatial anxiety also significantly predicted math ability scores (t (97) = -0.13, p < .000). This result was interesting since the results of the pilot study did not indicate this relationship, and it has not been identified in any previous literature to our knowledge. The relationship between spatial anxiety and math ability is negative, as expected with anxiety relationships. Spatial anxiety may affect students' spatial ability, which leads to lower math ability scores. Spatial anxiety scores may reflect math anxiety, which, though not measured in this study, was not significantly correlated in the pilot study.

To determine the more specific relationship between spatial ability, spatial anxiety, and math ability, the third model in these analyses included spatial ability scores, spatial anxiety scores, and the interaction between these scores. Here, spatial anxiety lost its predictive power for math ability once spatial ability was added to the model (t (95) = -1.26, p = .211). Spatial ability remained highly predictive of math performance (t (95) = 2.51, p = .0134). Additionally, the interaction between spatial ability and anxiety was not significant (t (95) = 1.40, p = .165). This result could mean that spatial anxiety does not directly affect math ability but may indirectly affect it through other factors.

Educational and Scientific Importance

This study indicated three important results. First, spatial ability and math ability are highly related. Second, specific spatial sub-categories could be more critical for success on different types of math tasks. Third, high spatial anxiety scores predict lower spatial and math ability scores.

Taken together, these results reflect the complex nature of spatial and math ability. This research helps uncover the deeper relationships between math ability, spatial ability, and spatial anxiety. The strong correlation between spatial ability and math ability may reflect the highly spatial nature of mathematics.

Additionally, the results illuminate the relationships between the different sub-categories of spatial and math ability that previous research has not identified. Each of the three spatial ability sub-categories (mental rotation, spatial orientation, and spatial visualization) had unique relationships with the four sub-categories of math ability. Spatial visualization was predictive of success on items involving space and shape, suggesting people visualize different views of an object without the need to rotate the object or orient relationships between objects. These differing relationships are consistent with the idea that spatial ability is composed of a variety of factors (e.g., McGee, 1979). Subcomponents of spatial ability factors may be more critical to success on specific math ability tasks than overall spatial ability. Thus, it may be possible to design more specific spatial interventions to improve scores on particular math sub-categories.

Spatial anxiety was negatively associated with both spatial ability and math ability. Spatial anxiety may have a general effect, such as reducing working memory capacity, as posited in Attentional Control Theory (Eysenck & Derakshan, 2011), which would limit resources devoted to mathematical problem-solving. Spatial anxiety may also disrupt specific mathematical skills.

However, there are several limitations to this study that should be addressed in future research. First, this study only included adults, with a majority of participants identifying as white and native English speakers. Second, this study did not include measures of working memory, which may explain why spatial anxiety is negatively associated with both spatial ability and math ability. Additional studies will be needed to see if these relationships extend to other populations and to further the mechanisms behind the relationships between spatial anxiety and both spatial ability and math ability. Third, this study used a composite score for spatial anxiety that consisted of a modified scale designed for children. Though no measures of spatial anxiety designed for adults existed at the time this study was conducted, future studies will be needed to see if these new measures, which are more appropriate for adult populations, have an impact on the results discussed here. These new measures also break spatial anxiety into sub-factors. Further studies should investigate the possible associations between these spatial anxiety sub-factors and both spatial ability and mathematics.

Study 2: Exploring Expanded Notions of Embodiment in Students' Fraction Knowledge

Traditional models of mathematical cognition are expanding beyond individually isolated rule-based processing of symbol systems. While the importance of social factors and contexts influences contemporary theories on mathematical thinking (e.g., Sfard, 2008), two emerging frameworks, embodiment and affect, also extend theories on mathematical cognition. Affective and embodied processes are not only contributing to the basic theory of human cognition, but they are also informing the design of evidence-based educational practices, including the design of curricular activities as well as principles for instruction and assessment. Thus, researchers and educators must understand the multifaceted nature of mathematical thinking, including the relationships between mathematical ability, spatial skills, and spatial anxiety. In the current study, we investigate the relationships between these factors in children with a focus on mathematical fraction operations.

Theoretical Background

Fraction Knowledge and Embodiment

Previous studies on symbolic fraction proficiency have shown that fraction operations are one of the best predictors of later algebra performance (DeWolf et al., 2015; Siegler et al., 2012) and are necessary for advanced learning in STEM (National Mathematics Advisory Panel, 2008). However, studies have shown that children struggle with symbolic fraction skills during the critical period between elementary and early middle school (e.g., Fuchs et al., 2013; Hansen et al., 2017; Siegler & Pyke, 2013). Though research has identified several factors which may explain fraction task performance, including calculation fluency (Hansen et al., 2017), working memory (Fuchs et al., 2013), mathematics anxiety (Starling-Alves et al., 2021), and spatial-numerical associations (Toomarian et al., 2019), research studies on fractions typically focus on only one explanatory domain at a time. Moreover, while other lines of research have shown that spatial ability and spatial anxiety may contribute to mathematical thinking (Schenck & Nathan, 2020), these studies do not focus specifically on fractions knowledge. Thus, little is known about the associations between factors such as number sense, memory, affective states, and spatial ability and their associations with symbolic fraction tasks.

In the embodied cognition community, past research on symbolic fraction understanding has focused on understanding the relationships between fraction knowledge and gestures (e.g., Surahmi & Ekawati, 2018; Swart et al., 2014) and developing digital embodied interventions (e.g., Swart et al., 2016). In this current work, we extend the embodied lens by drawing on the *Grounded and Embodied Learning Framework* (GEL; Nathan, 2021; Figure 4), which draws on Newell (1994) to describe the interconnected nature of learning across timescales, including affective states, memory, spatial ability and number sense processes.

Spatial Ability

One facet of embodied cognition in mathematical thinking is spatial ability (Xie et al., 2020). Spatial ability refers to the skills needed to generate and manipulate mental spatial objects, images, relationships, and transformations (Battista, 2007). Empirical evidence has demonstrated links between numbers and space that give rise to spatial-numerical associations (Dahaene et al., 1993; Hawes & Ansari, 2020) and an automatic shared processing and strategic recruitment of spatial processes (Mix et al., 2016). Furthermore, studies have shown that spatial ability is integrally linked to success in mathematics for children as young as three (Verdine et al., 2014) and that the development of spatial skills among elementary-aged children strongly predicts later mathematical capabilities (Casey et al., 2015).

Spatial ability is often thought to consist of a variety of sub-skills identified through factor analytic methods in the early twentieth century (e.g., Carroll, 1993), although there is little consensus in the literature about the exact combination of critical sub-skills (Yilmaz, 2009). Nonetheless, a framework proposed by Ramful and colleagues (2017) was developed based on skills needed for success in middle grades mathematic classrooms that consists of three sub-skills: *mental rotation, spatial orientation, and spatial visualization* and has been used in prior research seeking to uncover the specific relationships between spatial skills and mathematics (Schenck & Nathan, 2020).

Anxiety

Domain-specific anxieties such as math and spatial anxiety are distinct from general anxiety (Lauer et al., 2018). Prior studies on math anxiety have shown a negative association between math anxiety and math achievement (Ramirez et al., 2013) and that this association may be moderated by working memory (Ashcraft & Krause, 2007). While math anxiety has a well-established correlation with adolescents' math performance (Ma, 1999), spatial anxiety, especially in children, is relatively underexplored. Studies conducted on adults showed that spatial anxiety is negatively correlated with spatial ability, with women reporting greater spatial anxiety (Lyons et al., 2018). Similar results were found in the few studies in children (Lauer et al., 2018; Ramirez et al., 2012). More recent studies have shown a negative correlation between spatial anxiety and mathematics performance in adults (Schenck & Nathan, 2020), despite the fact that spatial anxiety lost its' significance when controlling for spatial ability, and they found no significant interaction between spatial anxiety and spatial ability. Additionally, factor analyses have indicated that spatial anxiety, like spatial ability, is a multifaced construct in adults even though the research community does not agree on the number and nature of sub-factors.

While work by Lyons and colleagues' (2018) indicated that spatial anxiety may consist of three factors (i.e., mental manipulation, navigation, and imagery), other scholars like Malachini and colleagues (2017) have shown that spatial anxiety induced by large-scale navigation tasks can be dissociated from small-scale spatial visualization tasks.

Current Study

The current study extends Schenck and Nathan's (2020) work uncovering the relationships between spatial ability, anxiety, and mathematical thinking by focusing on children's performance, including measures of working memory, and expanding embodied notions of fractions knowledge. The central research questions are: What are the relationships between spatial anxiety, spatial ability, and performance on fraction assessments (RQ1); how are mathematical learning and thinking about fractions rooted in body-based processes such as anxiety and spatial ability? (RQ2)

Methods

The dataset includes middle-grade students (N = 89) recruited from multiple school districts in a Midwestern city in the United States, including a mix of urban and suburban districts of varying demographics. The data was collected as part of a larger longitudinal study on children's fraction knowledge development and comes from the Year 4 cohort of the study. It should be noted that this year of data collection coincided with the 2020 Covid-19 pandemic, which delayed data collection. Consequently, 52% of participants completed at least one session virtually rather than in person, and there was more than a six-month gap between second and first sessions for eight participants. Inclusion criteria for this study were fluent English production and comprehension and completion of all covariate measures of interest. Participants received $10 per hour of participation and a small toy for each study visit as compensation.

Procedure

Participants completed assessments in two one-hour lab visits occurring on separate days, either in-person or virtually. On the first visit, they completed a set of standardized measures, including the Woodcock-Johnson Tests of Abilities, 3rd edition (WJ-III) subtests (Schrank et al., 2001). On the second visit, they completed the rest of the assessments (see Table 3 for demographic information).

Table 3

Demographic Statistics (N=89)

Variables	Mean (SD)	N (%)
Age in years at Session 1	12.02 (1.52)	
Age in years at Session 2	12.27 (1.57)	
Sex, Male		50(56%)
Sex, Female		34(38%)
Sex, Other		5 (6%)

Measures

Participants were given the Fraction Knowledge Assessment (FKA) to measure fraction knowledge, the dependent variable. The FKA is a pencil-and-paper assessment that measures both conceptual and procedural aspects of fraction knowledge. This instrument was constructed using items from national and international assessments, including the National Assessment of Educational Progress and the Trends in International Mathematics and Science Study, and from instruments developed by researchers such as Hallett and colleagues (2012). Separate FKA's were developed for each grade (5th, 6th, and 8th grades for this study) with grade-appropriate items and contained 41 items.

Embodied Covariates. Various measurements were included to explore the relationships between fraction knowledge and embodiment. First, we measured working memory, math anxiety, and spatial anxiety. Working memory was measured using the auditory working memory subtest from WJ-III (Schrank et al., 2001), which is associated with verbal working memory and short-term memory. Math anxiety was measured using the Suinn Mathematics Anxiety Rating Scale, Elementary Form (Suinn et al., 1988). Spatial anxiety was measured using the Child Spatial Anxiety Questionnaire (Ramirez et al., 2012), which correlates with small-scale spatial anxiety, and the Santa Barbara Sense of Direction Scale (Hegarty et al., 2002), which correlates with large-scale spatial anxiety. These scales were used separately and as a composite score in the analyses. We chose to separate these measures into two factors (large- and small-scale) based on Malanchini and colleagues' (2017) work rather than Lyons and colleagues' (2018) three-factor framework because there is not yet an appropriate three-factor measure that has been developed for this age group.

Next, we measured number sense and spatial ability. Number sense was measured using two WJ-III subtests: mathematics fluency and calculation (Schrank et al., 2001). Mathematics fluency is a timed test and involves simple written arithmetic problems, while calculation is untimed and involves more complicated calculations such as fraction operations. Due to overall time constraints of the larger longitudinal study, spatial ability was measured using a truncated spatial reasoning assessment based on the Spatial Reasoning Instrument (SRI; Ramful et al., 2017) consisting of 15-items, split into three subcategories (mental rotation, spatial orientation, and spatial visualization) of 5-items each. Internal reliability for this truncated measure was calculated using all participants from this study who had completed this measure (N = 166) and Cronbach's alphas for the overall composite score ($\alpha = .75$) and the three sub-categories (mental rotation, spatial orientation, and spatial visualization) were .70, .72, and .68, respectively.

Results

RQ1: Spatial Anxiety, Spatial Ability, and Fraction Assessments

To address the first research question about spatial anxiety, we used Ordinary Least Squares regression. Multiple linear regressions were calculated to predict fraction assessment scores while controlling for sex, age at session 2, and working memory by: 1) spatial ability as a composite score and with sub-categories; 2) spatial anxiety as a composite score and with sub-categories, and 3) spatial ability and spatial anxiety composite scores with an interaction term between these constructs.

Fraction Knowledge and Spatial Ability. To investigate the relationship between fraction knowledge scores and spatial ability, we calculated two multiple linear regression equations to predict fraction knowledge scores based on the spatial ability scores while controlling for age at session 2, sex, and working memory. The first model predicted fraction knowledge by the spatial ability composite score. The second model predicted fraction knowledge by three spatial ability sub-categories (mental rotation, spatial orientation, and spatial visualization).

In model 1, spatial ability composite score was significantly predictive of fraction knowledge ($\beta = 0.85$, $p < .001$). In model 2, mental rotation scores ($\beta = 2.30$, $p < .001$) and spatial visualization scores ($\beta = 1.98$, $p = .004$) were significantly predictive of fraction knowledge. Neither age, sex, nor working memory was significant in either model.

Fraction Knowledge and Spatial Anxiety. To investigate spatial anxiety, we calculated two multiple linear regression equations to predict fraction knowledge scores based on spatial anxiety scores while controlling for age at session 2, sex, and working memory. Like the models for spatial ability, the first model predicted fraction knowledge by the spatial anxiety composite score, and the second model predicted fraction knowledge by two spatial anxiety sub-categories (small-scale and large-scale).

In model 1, spatial anxiety composite score was not significantly associated with fraction knowledge ($\beta = 0.06$, $p = .673$). However, working memory was significantly associated with a 0.31-point increase in fraction knowledge scores for each one-point increase in working memory score ($p = .025$). Similarly, in model 2, neither small- nor large-scale spatial anxiety scores were significantly associated with fraction knowledge ($\beta = -0.48$, $p = .004$; $\beta = 0.31$, $p = .004$, respectively), but working memory was significantly associated with a 0.35-point increase in fraction knowledge scores for each one-point increase in working memory score ($p = .012$). Neither age nor sex was significant in either model.

Fraction Knowledge, Spatial Ability, and Spatial Anxiety. To determine the more specific relationship between spatial ability, spatial anxiety, and fraction knowledge, we calculated a multiple regression equation to predict fraction knowledge by spatial ability composite scores, spatial anxiety composite scores, and the interaction between these scores. In this model, spatial ability lost its predictive power for fraction knowledge ($\beta = -1.12$, $p = .211$) and was replaced by spatial anxiety ($\beta = -0.65$, $p = .043$). Furthermore, there was a significant interaction between spatial ability composite scores and spatial anxiety composite scores ($\beta = 0.08$, $p = .009$). Neither age, sex, nor working memory was significant in this model.

RQ2: Embodied Processes and Fraction Knowledge

To address the second research question about how mathematical learning and thinking about fractions are rooted in body-based processes, we calculated Pearson's correlations to estimate the magnitude and direction of the relationship between each variable. Correlations among all variables are shown in Table 3.

Table 3

Pearson Correlations for Embodied Processes and Fraction Knowledge

Variable	FKA	WM	MA	SA	SA-S	SA-L	Math	Calc	SRI	MR	SO
WM	**0.22**										
MA	**-0.25**	-0.17									
SA	-0.02	-0.12	**0.59**								
SA-S	-0.11	-0.04	**0.61**	**0.73**							
LS-S	-0.05	-0.16	**0.41**	**0.92**	**0.42**						
Math	**0.53**	**0.25**	-0.15	0.01	-0.14	0.10					
Calc	**0.69**	0.18	**-0.25**	-0.02	-0.13	0.06	**0.70**				
SRI	**0.58**	**0.27**	**-0.23**	-0.14	-0.20	-0.08	**0.42**	**0.55**			
MR	**0.57**	**0.28**	**-0.27**	-0.14	-0.15	-0.11	**0.40**	**0.49**	**0.87**		
SO	**0.31**	**0.22**	-0.08	-0.13	0.02	-0.15	0.17	**0.35**	**0.72**	**0.51**	
SV	**0.50**	0.18	-0.15	-0.07	**-0.26**	0.04	**0.44**	**0.54**	**0.82**	**0.55**	**0.41**

Note. N = 89. FKA = fraction knowledge score; WM = working memory score; MA = math anxiety score; SA = spatial anxiety score; SA-S = Small-scale spatial anxiety sub-category; SA-L = Large-scale spatial anxiety sub-category; Math = Math fluency score; Calc = Calculation score; SRI = Spatial ability score; MR = Mental Rotation sub-category; SO = Spatial orientation sub-category; SV = spatial visualization sub-category. Two-tailed significance levels are presented. Bolded items are significant at $p \leq 0.05$.

Working Memory and Anxieties. The first measures of interest focused on working memory and mathematic and spatial anxieties. Results indicated that fraction knowledge and working memory were positively associated (r(89) = 0.22, p = .040), while fraction knowledge and math anxiety were negatively associated (r(89) = -0.25, p = .019). There were no significant correlations between fraction knowledge and spatial anxiety as a composite score (r(89) = -0.02, p = .884) or either small-scale or large-scale sub-categories (r(89) = -0.11, p = .293, and r(89) = -0.05, p = .642, respectively). Within this timescale, math anxiety and spatial anxiety as a composite score and both small- and large-scale spatial anxiety were significantly associated (r(89) = 0.59, p < .000, r(89) = 0.61, p < .001, and r(89) = 0.41, p < .001, respectively).

Number Sense and Spatial Ability. The second measures of interest, consisted of two number sense scales, math fluency and calculation, and spatial ability. Fraction knowledge was significantly positively associated with both number sense (math fluency and calculation) scores (r(89) = 0.53, p < .001, and r(89) = 0.69, p < .001, respectively) and the spatial ability composite score (r(89) = 0.58, p < .001). Furthermore, there were significant positive correlations between all three spatial ability sub-categories (mental rotation, spatial orientation, and spatial visualization) and fraction knowledge (r(89) = 0.57, p < .001, r(89) = 0.31, p = .003, and r(89) = 0.50, p < .001, respectively).

Within this timescale, math fluency scores were significantly positively associated with calculation scores (r(89) = 0.70, p < .001), the spatial ability composite score (r(89) = 0.42, p < .001), and mental rotation and spatial visualization sub-categories (r(89) = 0.40, p < .001, and r(89) = 0.44, p < .001, respectively). Calculation scores were also significantly positively associated with the spatial ability composite score (r(89) = 0.55, p < .001), and all three spatial ability sub-categories (mental rotation, spatial orientation, and spatial visualization) scores (r(89) = 0.49, p < .001, r(89) = 0.35, p = .001, and r(89) = 0.54, p < .001, respectively).

Between the Measures. There were also significant associations between the two groups of measures. The two number sense measures were significantly associated with different working memory and anxiety measures. Math fluency was associated with working memory (r(89) = 0.25, p = .021), while calculation scores were associated with math anxiety (r(89) = -0.25, p = .020). The spatial ability composite scores and sub-categories were also associated with differing combinations of working memory and anxiety measures. Spatial ability composite and mental rotation scores were associated with working memory (r(89) = 0.27, p = .011, and r(89) = 0.28, p = .009, respectively) and math anxiety (r(89) = -0.23, p = .031, and r(89) = -0.27, p = .010, respectively). Spatial orientation was only associated with working memory (r(89) = 0.22, p = .043), while the spatial visualization scores were associated with small-scale spatial anxiety (r(89) = -0.26, p = .014).) and math anxiety (r(89) = -0.23, p = .031, and r(89) = -0.27, p = .010, respectively). Spatial orientation was only associated with working memory (r(89) = 0.22, p = .043), while the spatial visualization scores were associated with small-scale spatial anxiety (r(89) = -0.26, p = .014).

Discussion

This study provides evidence of a positive relationship between spatial ability and symbolic fraction knowledge and that spatial anxiety's relationship with fraction knowledge may be moderated by spatial ability (RQ1). Consistent with findings in other domains of mathematics in adults (e.g., Schenck & Nathan, 2020), fraction knowledge was significantly associated with overall spatial ability. This study also identified two specific factors of spatial ability, spatial visualization and mental rotation, associated with fraction knowledge.

The connection between symbolic fraction knowledge could be explained as students accessing previously learned mental representations, such as a mental number line to compare fractional quantities or constructing mental diagrams such as area models or other conceptual models to visualize computations such as addition and multiplication. Though more studies will be needed to assess whether this spatial visualization and fraction knowledge connection is indicative of these nonsymbolic problem-solving strategies, it could provide evidence for how children ground symbolic fractions in a nonsymbolic number sense system such as a ratio processing system (Park & Matthews, 2021). The connection between symbolic fractions and mental rotation is less intuitive. Though mental rotation is one of the most studied spatial abilities, the limited evidence connecting it to mathematics in general and specific domains. While some studies have shown mental rotation is predictive of mental arithmetic which may explain the connection found in this study (e.g., Kyttälä & Lehto, 2008), further studies will be needed to explore the association found in this study.

Further studies may also be needed to address two limitations of the spatial ability measure used in this study. First, the spatial ability measure was a truncated version of a validated instrument. Though metrics indicated that this version was sufficiently reliable for this study, employing the full version in subsequent work may reveal different or more valid associations. Second, the truncated SRI spatial ability measure assumed a specific sub-factor framework of spatial ability. Since there is little agreement on which sub-factors are important for mathematics, let alone whether existing spatial instruments truly measure intended spatial ability sub-skills accurately (Burte et al., 2019a), additional studies may be needed to fully explore the results of this study.

Neither the spatial anxiety composite score nor either of the sub-scores was significantly associated with fraction knowledge in this study. However, when spatial ability was added to the model with spatial anxiety, the interaction and spatial anxiety composite scores were significantly negatively associated with fraction knowledge. These results contradict previous studies on spatial anxiety and mathematics that showed a negative association between spatial anxiety and various mathematics domains with an insignificant interaction between spatial ability and spatial anxiety in adults (Schenck & Nathan, 2020).

This contradiction could have several explanations. First, spatial anxiety may moderate the relationship between spatial ability and fraction knowledge. The relationship may be specific to symbolic fraction knowledge with spatial anxiety interfering with the spatial ability processes such as visualizing the mental representations described earlier. Second, the relationship between spatial anxiety and spatial ability may not be specific to fraction knowledge but rather change with development. As students age, they may develop mechanisms to cope with the impacts of spatial anxiety, or the effects of other domain-specific anxieties, such as math anxiety or testing anxiety, may become more predictive of mathematic performance. Additionally, some studies have suggested that domain-specific anxieties, such as math and spatial anxieties, may not be differentiable from other forms of anxiety in children (Hill et al., 2016; but see Lauer et al., 2018). It could be that the spatial anxiety results seen in this study are indicative of another type of anxiety that was not measured or controlled for in either study. Third, these results could be explained through the development of fraction knowledge specifically. Success on symbolic fraction tasks, in particular, may require a composite of both procedural and conceptual knowledge and bifurcate in terms of the types of experiences that students acquire throughout development. These experiences may provide students with the skill that delays spatial anxieties' impacts from the initial lack of knowing how to manipulate the formalisms of fractions representations. Fourth, the significant relationship between spatial anxiety and mathematics in adults in Schenck and Nathan's (2020) study may result from the lack of a working memory variable. Domain-specific anxieties have been linked to lower performance in complex tasks such as mathematics by creating worries that interfere with working memory (Engle, 2002). It is possible that the spatial anxiety measure associations found in Schenck and Nathan (2020) may change if working memory had been measured and controlled for in their analyses. Additional studies will be needed to test these explanations.

This study also illuminated how embodied processes are correlated with learning and thinking about symbolic fractions (RQ2). Working memory and mathematica anxieties had significant relationships with fraction. Additionally, math anxiety and spatial anxiety were also significantly correlated. These findings support both previous work on symbolic fractions suggesting these are factors of interest in this domain.

Calculation, math fluency, and spatial ability were all significantly correlated with fraction knowledge, which supports prior work on symbolic fractions. Additionally, mental rotation and spatial visualization spatial ability sub-skills were also strongly correlated with fraction knowledge, similar to the results found in the previous analysis of this study. However, unlike the previous models, there was also a medium correlation between spatial orientation sub-skills and fraction knowledge. This relationship could indicate that spatial orientation skills are associated with fraction knowledge, but other sub-skills have more powerful associations. Furthermore, differing correlations between each measure of number sense and the various spatial sub-skills. These correlations may benefit future studies attempting to explore these areas more deeply and may indicate additional variables of interest to researchers.

Overall, this study supports several previous findings on both fraction knowledge and mathematics learning more generally. It also highlights several novel associations. Spatial ability skills, such as spatial visualization and mental rotation, may be specifically relevant for symbolic fraction knowledge learning, and spatial anxiety may moderate that relationship. This study also solidifies the connection between fraction knowledge and embodiment by demonstrating how fraction knowledge is rooted in multiple body-based systems. Though this work is correlational and exploratory, these findings may provide the foundations for future work exploring the mechanisms behind these associations.

Chapter 3: Investigating the Role of Spatial Anxiety in Embodied Geometric Reasoning

Portions of this chapter have been published and are reproduced with permission from:

Schenck, K. E**.,** & Nathan, M. J. (2022, November). *Spatial anxiety moderates the effect of spatial ability in geometric reasoning*. In A. E. Lischka, E. B. Dyer, R.S. Jones, J. N. Lovett, J. Strayer, & S. Drown (Eds.), *Proceedings of the forty-fourth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 598). Nashville, TN, United States.

Embodied processes such as spatial reasoning and gesture production are associated with success on verbal geometric tasks such as proof (e.g., Nathan et al., 2021; Pier et al., 2019). Geometric reasoning often engages spatial reasoning processes for mentally visualizing and transforming geometric objects (e.g., skewing a quadrilateral). Gestures and speech are integral to communicating those transformations for assessments and applications (incl. Alibali & Nathan, 2012). The role of gestures during mathematical reasoning goes well beyond a supporting role in *describing* geometric transformations and has been shown to play a fundamental role in the reasoning process (e.g., Hostetter & Alibali, 2008/2019; Kita et al., 2017). However, far less is known about the impact of spatial anxiety on specific domains of mathematics or the possible interactions between spatial anxiety and spatial ability.

Although previous studies have identified associations between certain domains of mathematics and spatial anxiety (Schenck, Hubbard, et al., 2022; Schenck & Nathan, 2020), there are remaining questions about the exact nature of these associations since neither of these studies focused on a highly spatial domain of mathematics such as geometry nor did they include spatial activities designed to elicit gestures. Thus, the possibility that task demands may explain the connections between spatial anxiety, spatial ability, gesture, and geometric reasoning must be explored.

The central objectives of this paper are, first, to offer an expanded theoretical view of embodied mathematical activity that includes both cognitive and affective processes, and second, to investigate the role of a cognitive and affective spatial system in embodied geometric thinking. Specifically, we examine the role that the cognitive processes of spatial ability and the affective process of spatial anxiety plays in predicting the quality of participants' geometric reasoning and their occurrences of gesture. The main objective is to identify associations between affective spatial processes and geometric thinking to improve our empirical understanding of the embodied and spatial nature of mathematical reasoning to advance our understanding of the role of affective spatial processes. A deeper understanding of the spatial nature of mathematics and the influence of its associations with other cognitive processes are crucial for developing effective, evidence-based approaches to mathematics education that inform pedagogy, teacher professional development, and the design of curricular activities and interventions.

Theoretical Background

Some scholars suggest that reasoning is achieved by using bodily experiences as a grounding mechanism to imagine or simulate perceptions or actions, which can be re-used through mental simulation to make connections between this online multimodal sensorimotor experience and an offline, amodal system of conceptual knowledge (Barsalou, 2008). Scholars have also posited a reliable, causal, bidirectional relationship between body and cognitive states (Nathan, 2014; Shapiro, 2019). Theories of grounded and embodied cognition (GEC) connect reasoning to body-based processes like physical and simulated action (Beilock & Goldin-Meadow, 2010; Hostetter & Alibali, 2008/2019), often through gesture (Alibali & Nathan, 2012).

Grounded and Embodied Mathematical Cognition

GEC acknowledges that bodily experiences are critically important for meaning-making of concrete and abstract concepts. Hostetter and Alibali's (2008/2019) theory of *Gestures as Simulated Action* posits that gestures arise when perceptions push pre-motor activity beyond motor-planning threshold and are expressed as motoric activity that simulates the performance of an action. This threshold can vary by individual depending on current task demand, individual differences, and situational considerations.

Several studies have shown that mathematical ideas are inherently embodied and can be expressed through actions such as gestures. It is thought that the complex motor movements of gesture allow learners to represent dynamic, spatial-relational information (McNeill, 2005) and augment reasoning and problem-solving by connecting sensorimotor experiences with mental representations (Nathan, 2014; Wilson, 2002). For example, Abrahamson and Bakker (2016) explored how hand and arm movements demonstrating proportionality can served as a sensorimotor basis for multiplicative reasoning. Guided by the Mathematics Imagery Trainer for Proportion, Abrahamson and colleagues developed a digital environment in which students constructed new “attentional anchors” representing targeted mathematical concepts.

Though much of the work on embodied mathematical cognition focuses on numbers and operations (e.g., Abrahamson & Bakker, 2016; Ottmar & Landy, 2017), an emerging body of work by Nathan and colleagues has explored GEC in verbal geometric proof processes. Forming proofs is a disciplinary practice that often combines content knowledge with psychological processes such as spatial imagery and logical deduction (Nathan, 2014). During proof generation, students tend to rely on authoritative means (Herbst & Brach, 2006), overgeneralize from specific examples (Knuth et al., 2009), or salient perceptual features (Jones, 2000). These methods align with the emphasis on proof as establishing certainty rather than developing inquiry in high school textbooks (de Villiers, 1998) which is a disconnect from how mathematics uses proof to construct mathematical knowledge. As such, students often struggle to generate generalized proofs (Healy & Hoyles, 2000) and fail to actively investigate the veracity of a statement beyond memorizing theorems to confirm assertions (Hanna, 2000). Thus, scholars have looked at expanding the traditional notion of proof as a product to less formalized proof schemes as a form of disciplinary discourse (Harel & Sowder, 1998; Knuth, 2002).

The theory of *action-cognition transaction* (ACT; Nathan, 2014) posits a relationship between the learner and the environment and between action and perception (in combination with language) as the mechanisms by which gestures can induce cognitive states and vice-versa. In research, Nathan and colleagues (2021) studied how mathematics experts and non-experts solve verbal geometric reasoning. Participants were asked to explain whether the geometric conjectures were true or false. Transcripts of student's explanations (rationales) combined with video recordings of the gestures they made showed that representational gestures, specifically those that dynamically depicted transformations of geometric objects, were strongly associated with the production of correct intuition, mathematical insight, and transformational proof (Nathan et al., 2021). Between groups, experts outperformed non-experts across all performance measures and were more likely to produce representational gestures than non-experts who produced dynamic representational gestures, leading the authors to conclude that dynamic representational gesture may replace expertise. Additional studies working with high school students have shown that students performed better on geometric reasoning tasks when their verbal explanations were accompanied by dynamic representational gestures that were "replays" of actions they were previously directed to perform (Walkington et al., 2022). In effect, these gesture replays demonstrate that learners internalize embodied simulations of a geometric transformation that serve as a primary mechanism supporting geometric insights.

Embodied Components of Spatial Systems

Much of the research to understand geometric reasoning from an embodied lens has focused on the role of gestures (e.g., Nathan et al., 2021; Pier et al., 2019; Walkington et al., 2022). From a cognitive perspective, much research has focused on the roles of spatial abilities (e.g., Hannafin et al., 2008; Newcombe, 2013). Nonetheless, there may be an additional benefit to expanding research on geometric thinking to include affective and cognitive measurements of spatial abilities and spatial anxiety and how they are associated with both verbal mathematical reasoning and gesture production.

Spatial Ability

Spatial ability, the capacity to imagine, retain, and manipulate visuospatial information and relations, is crucial for success in science, technology, engineering, and mathematics (STEM) fields (Shea et al., 2001; Wolfgang et al., 2003). Deficiencies in spatial abilities can hinder progress in STEM disciplines (Harris et al., 2013; Wai et al., 2009), emphasizing the need for institutions to explicitly train students' spatial thinking skills. Evidence supports that spatial thinking training interventions can enhance STEM education performance and retention (Sorby, 2009). Connections between spatial and mathematical task success have been demonstrated in children (Casey et al., 2015; Laski et al., 2013) and adults (Schenck & Nathan, 2020). Research on spatial ability and mathematical skills has shown positive associations (Atit et al., 2021; Xie et al., 2020). Spatial ability is essential for various domains of mathematics education (e.g., Battista et al., 2018; Case et al., 1996; Oostermeijer et al., 2014; Sorby et al., 2013; Tufte, 2001).). Overall, these findings underscore the importance of spatial ability in mathematics education and suggest that fostering spatial thinking skills could enhance success in various mathematical domains.

Anxiety

Though some scholars argue that low anxiety levels can enhance neurocognitive performance, high anxiety levels may impair it. Trait anxieties, relatively stable cognitive and somatic anticipation responses to uncertain situations, can be domain-specific, such as math or spatial anxiety. Math anxiety is a well-established predictor of poor math performance, even when controlling for other factors (Lukowski et al., 2019; Zhang et al., 2019). Similarly, spatial anxiety, an apprehension towards situations requiring spatial abilities, has been found to negatively correlate with spatial ability (Malanchini et al., 2017; Ramirez et al., 2012). However, spatial anxiety has not been as extensively studied as math anxiety.

Spatial anxiety has been linked to lower performance on spatial tasks (Ramirez et al., 2012), reduced sense of direction (Kremmyda et al., 2016; Lawton, 1994), higher math anxiety in adults (Ferguson et al., 2015), and a gender gap in spatial ability and anxiety scores (Lawton, 1994; Maeda & Yoon, 2013; Malanchini et al., 2017; Voyer et al., 1995). Children with higher spatial anxiety levels exhibit reduced spatial skill gains (Gunderson et al., 2013). Moreover, spatial anxiety has been shown to be associated with mathematics performance in both children and adults (Schenck & Nathan, 2020; Schenck, Hubbard, et al., 2022). Existing assessments of spatial anxiety, such as Lawton's Spatial Anxiety Scale (SAS; 1994), the Child Spatial Anxiety Questionnaire (Ramirez et al., 2012), and a more recent measure developed by Lyons et al. (2018), contribute to a better understanding of spatial anxiety. However, these instruments have limitations, such as targeting specific age groups or focusing on particular aspects of spatial anxiety.

Research Questions

We investigate the links between spatial anxiety and geometric reasoning, which drives the following research questions: 1) What is the relationship between spatial anxiety and geometric reasoning as measured by performance on (a) verbal insight and (b) proof; and 2) What is the relationship between spatial anxiety and the occurrence of gestures during geometric reasoning?

Method

Participants

Undergraduate students (N = 94) were recruited from a large Midwestern university in the United States. The data represents a pre-intervention subset of a larger randomized experimental trial investigating students' mathematical intuitions and geometric proofs (Swart et al., 2023). Inclusion criteria required fluent English production and comprehension and completion of all covariate measures of interest. Participants received a $20 gift card as compensation for completing both sessions of the study. Sixty-seven participants (71%) identified as female, while 27 (29%) identified as male. Most participants (81%) identified English as their native language. Responses to race/ethnicity questions were consolidated into three categories (White/Non-Hispanic, Asian, and Other) due to the high number of participants who self-identified as White/Non-Hispanic (63%) and Asian (26%). The "Other" category (12%) represents the remaining options. A full breakdown of the demographics is summarized in Table 4.

Table 4

Demographics and descriptive statistics (N = 94)

Variables	Mean (SD)	N (%)
Average age	20.13 (1.34)	
Sex, Female		67 (71%)
Ethnicity, White/Non-Hispanic		59 (63%)
Ethnicity, Asian		24 (26%)
Ethnicity, Other		11 (12%)
Native English Speaker		76 (81%)
Average spatial ability score	12.68 (2.07)	
Average spatial anxiety score	37.60 (14.98)	
Likelihood of correct insight (per trial)		421 (56%)
Likelihood of correct proof (per trial)		175 (23%)
Likelihood of representational gesture (per trial)		302 (40%)
Likelihood of nondynamic gestures (per trial)		190 (25%)
Likelihood of dynamic gesture (per trial)		112 (15%)

Materials

Conjectures

Eight geometric conjectures that explored the general properties of two-dimensional objects were used in this study. These were selected from a variety of secondary mathematics textbooks and in consultation with a mathematics professor who directed the secondary mathematics teacher education program for their relevance to current high school education. Of the eight conjectures, three involved properties of parallelograms, three involved properties of triangles, and two involved properties of coordinate planes and angles. Five of the eight conjectures were true statements, and three were false. Table 5 includes the text for each conjecture, its veracity, and relevant mathematical insights.

Table 5

Text, truth value, and mathematical concept for each of the eight conjectures

Conjecture	Conjecture Text	Truth	Insight Examples
AAA	Given that you know the measure of all three angles of a triangle, there is only one unique triangle that can be formed with these three angle measurements.	False	Triangles can be scaled to produce similar triangles with proportional sides and equal angles.
Diagonals	The diagonals of a rectangle always have the same length.	True	A rectangle has two pairs of parallel sides, so the distance between two opposite sides much be equal even if it is at a diagonal. The diagonals of a rectangle form two congruent triangles, so they must be the same length.
Doubled Rectangle	If you double the length and the width of a rectangle, then the area is exactly doubled.	False	Doubling the length and height of a rectangle would increase the area by a factor of four.
Opposite Sides and Angles	In triangle ABC, if Angle A is larger than Angle B, then the side opposite Angle A is longer than the side opposite Angle B.	True	As the measure of the angle increases, the side opposite of the angle must also increase so that triangle is closed.
Parallelogram	The area of a parallelogram is the same as the area of a rectangle with the same lengths of base and height.	True	A rectangle is a type of parallelogram, so they have the same formula. A parallelogram can be made into a rectangle by tilting or pushing the sides, which does not change the area.
Triangle Inequality	The sum of the lengths of two sides of a triangle is always greater than the length of the third side.	True	If the sum of the two sides was equal to the third side, it would form a line. If the sum of the two sides was less than the third side, the triangle would not close.
Vertical Angles	The opposite angles of two lines that intersect each other are always the same.	True	The only instance opposite angles would not be the same would be if one of the lines were curved and not a line. Straight angles add up to 180 degrees, and two intersecting lines create four angles with adjacent angles that form straight angles.
X-axis Reflection	Reflecting any point over the x-axis is the same as rotating the point 90 degrees clockwise about the origin.	False	Reflecting a point over the x-axis multiplies the y-coordinate by -1. Rotating a point 90 degrees clockwise multiplies the x-coordinate by -1 and then switches the x- and y- coordinate.

Measures

Though participants completed several covariate measures as part of the larger study, only two measures are of interest for this current study: spatial ability and spatial anxiety. We used a truncated version of the validated Spatial Reasoning Instrument (Ramful et al., 2017) to assess spatial ability. The truncated version was composed of five multiple-choice items for each of the three specific spatial ability subcomponents: mental rotation, non-rotational spatial visualization, and spatial orientation. Each correct answer received one point, and points were summed for a total composite score out of 15 possible points. Internal reliability for this truncated measure was calculated using all participants from the larger study who had completed this measure (N = 151). Though the overall reliability of this measure was acceptable (α = .69), truncating the battery created reliability issues for the individual subcomponents. Thus, we will only be using the composite score for this measure.

Spatial anxiety was measured using the *Novel Spatial Anxiety Scale* (Lyons et al., 2018). This measure consists of 24 5-item Likert scale questions, which break into three subcategories (mental-manipulation, navigation, and imagery) of eight questions each. Point values for answers to each question ranged from 0 (“Not at all”) to 4 (“Very Much”). No questions required reverse scoring. Points were summed for a total composite score as well as three sub-scores. Internal reliability was calculated using all participants from the larger study who had completed the measure (N = 151). The Cronbach’s alpha for the overall composite score was .91, and Cronbach’s alphas for the subcategories of mental manipulation, navigation, and imagery were .87, .91, and .86, respectively. Only the spatial anxiety composite score is used in these analyses.

Coding

For each conjecture, the researcher asked participants to provide a reason why they believed the conjecture was true or false. Videos of the pre-intervention experimental section were organized, and participants' verbal responses were transcribed verbatim using the Transana software system (Woods & Fassnacht, 2020). Timestamps were added to segment the full transcripts into each of the eight conjectures for coding. Segmentation resulted in 752 video clips to be coded (94 participants x 8 conjectures). Two research team members coded each transcript using the coding scheme described below.

Coding Scheme

Typed transcripts and videos were coded for three categories based on the coding scheme developed by Nathan and colleagues (2021): verbal insight, valid transformational proof, and gesture. Verbal insight was coded (1/0) for the presence of key mathematical ideas for each conjecture, as specified by our team of mathematicians and math educators (examples are shown in Table 2). Transformational proof, a type of deductive proof, was coded (1/0) if the verbalized proof (including speech and gesture) met all three criteria set forth by Harel & Sowder, 2005: (1) a logical sequence of reasoning, with conclusions drawn from valid premises; (2) generalizable reasoning, showing the argument holds for an entire class of mathematical objects relevant to the conjecture; (3) operational thinking, with evidence of thought progression through a goal structure, anticipating the outcomes of one's mental operations. Any combination of insights that did not meet all three criteria was coded as a 0.

For gesture, three separate codes were used. The first code identified the occurrence of a representational gesture (i.e., a hand or arm movement that represents or depicts some feature or operation of a mathematical object or idea, per Alibali & Nathan, 2012). The next two codes subdivided representation gestures exclusively into nondynamic depictive or dynamic depictive gestures. *Nondynamic depictive gestures* are representational gestures that only represent mathematical entities (including acts such as tracing or pointing) but do not involve transformations that alter the geometric objects under investigation. *Dynamic depictive gestures* are representational gestures but also demonstrate some type of geometric transformation, such as rotating, dilating, or skewing an object. Investigators (e.g., Garcia & Infante, 2012; Newcombe & Shipley, 2015) have documented the role of dynamic depictive gestures in students' exploration of generalizable mathematical and physical properties of objects, such as invariance of the sum of interior angles of a triangle when skewed or dilated (Pier et al., 2013; Williams-Pierce et al., 2017). Any occurrence of at least one of these types of gestures would lead to assigning a 1 for that respective gesture code. Participant transcripts that did not include any gesture or included non-representational gestures (e.g., beat gestures) were coded 0.

Inter-rater Reliability

To establish inter-rater reliability, a researcher not involved in the initial coding process coded a random selection of 20% of the participants' videos. Overall inter-rater reliability for these codes is $\kappa = 0.89$. Table 5 shows inter-rater reliability for verbal insight, transformational proof, and gestures.

Shaffer's *rho*[b] was also calculated for each code to assess the validity of the inter-rater reliability using a *kappa* threshold of 0.65. (Eagan et al., 2017). Results for each inter-rater reliability measure can be found in Table 6. Overall, Shaffer's *rho* statistics for each code were

[b] The software for calculating rho can be accessed at this link: https://app.calcrho.org/.

less than 0.05, indicating that the sample size used for inter-rater reliability was sufficient to estimate inter-rater reliability at a threshold of at least 0.65.

Table 6

Inter-rater Reliability for Participant Transcript Coding

Code	**Cohen's *kappa***	**Shaffer's *rho***
Omnibus	0.89	0.00
Verbal Insight	0.84	0.00
Transformational Proof	0.86	0.00
Representational Gesture	0.90	0.00
Non-Dynamic Gesture	0.93	0.00
Dynamic Gesture	0.90	0.00

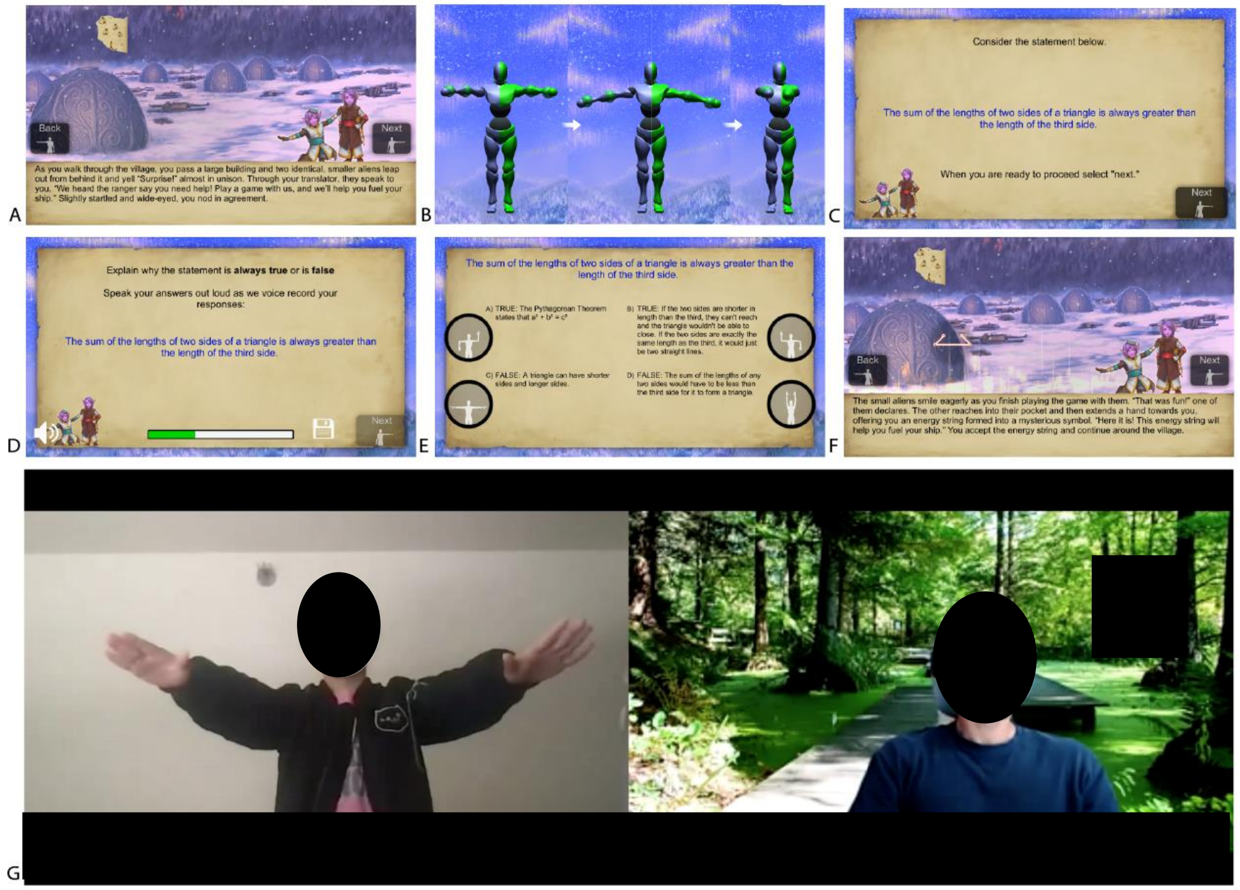
Note. Shaffer's *rho* was calculated with a *kappa* threshold of 0.65.

Procedure

In the first session, participants completed a series of demographic and covariate measures online via Qualtrics. In the second session, participants played an embodied videogame, *The Hidden Village* (THV*;* Nathan & Walkington, 2017; Nathan & Swart, 2021). In the game narrative, players are informed that they have crashed landed on an alien planet and need help from various villagers to retrieve energy strings to refuel the ship and return home. Each villager asks the participant to help them complete an activity in exchange for an energy string and to reveal an additional part of the village map (Figure 4A). During each game activity, participants were prompted to perform a set of relevant directed actions (Figure 4B). The game tracks player movements in real time and determines when an action sequence has been adequately performed. Each action sequence had to be successfully performed three times to advance in the game. For the next step, each player was presented with one of the eight geometry conjectures, with conjecture order determined by a Latin square design to avoid ordering effects on later performance (Figure 4C). The player was then prompted to give an immediate verbal response to a geometry conjecture's veracity (True or False) as a measure of their mathematical intuition. Players were not informed of any association between the actions they were prompted to perform and the current conjecture. They were then prompted to explain their response (Figure 4D). These verbal responses and accompanying gestures were recorded throughout game play. A player then had the conjecture shown again and asked to select the best choice among one of four multiple-choice options that presented a truth value for the conjecture and a justification for selecting that truth option (Figure 4E). The game then concluded their encounter with that particular village member (Figure 4F). This cycle was repeated for each of the eight different conjectures. It should be noted that this data collection period coincided with the 2020 Covid-19 pandemic; consequently, all participants completed questionnaires, surveys, and game play virtually through Zoom facilitated by a research team member (Figure 4G). The gameplay was video- and audio-recorded through Zoom.

Figure 4

Conjecture cycle in The Hidden Village



Note. This figure depicts the cycle for each conjecture in *The Hidden Village*. Participants are introduced to alien villagers that ask for help with a task (A), complete a set of direct actions, (B) and then answer a geometric conjecture (C - E). They are then rewarded with a piece of the map and an energy string to refuel their ship (F). Due to the COVID-19 pandemic, participation took place on Zoom, facilitated by a research team member (G).

Results

A correlation matrix of key variables can be found in Table 7. For our analyses, we used logistic regression for binary outcomes (0/1) on the occurrences of verbal insight, transformational proof, representational gesture, non-dynamic gesture, and dynamic gesture.

Logistic mixed models for insight, transformational proof, and the gesture codes were fit using the *glmer* command in the R software package *lme4* (Bates et al., 2015). We included participant ID and conjecture as random effects. Participant sex and spatial ability scores were included as fixed effects because these variables have been shown to be significant predictors of spatial anxiety and mathematics ability (Lawton, 1994; Maeda & Yoon, 2013; Nathan et al., 2021; Schenck & Nathan, 2020; Voyer et al., 1995). Spatial ability and spatial anxiety scores were centered and scaled. We also added an interaction term between spatial ability and spatial anxiety to test for the potential moderator relationship reported elsewhere (Schenck, Hubbard, et al., 2022; Schenck & Nathan, 2022). Plots for each of the models that include the interaction term can be found in Appendix A.

Table 7

Correlations for Key Factors

Variables	1	2	3	4	5	6	7	8
1.Age								
2.Sex[a]	.078							
3.Spatial Ability Score	**.216**	**-.200**						
4.Spatial Anxiety Score	.050	.060	**-.124**					
5.Verbal Insight	-.001	**-.214**	**.156**	-.001				
6.Transformational Proof	-.067	**-.123**	**.115**	**-.061**	**.425**			
7. Representational Gestures	-0.98	-.032	.007	.000	**.131**	**.101**		
8. Non-Dynamic Gestures	-.074	-.030	-.051	.024	.084	**-.241**	**.710**	
9.Dynamic Gestures	-.045	-.007	.072	-.030	**.077**	**.433**	**.511**	**-.243**

Note. Bolded correlations are significant at $p < .010$.
[a]Female is the reference category.

Odds ratios (*OR*) rather than effect sizes are reported here as our dependent variables are dichotomous. Odd ratios of 1.00 represent no change in the relative odds of the dependent variable. Odds ratios above 1.00 represent an increase in the relative odds of the outcome variable, while odds ratios below 1.00 represent a decrease. We interpreted odds ratios as "small" ($1.68 \leq OR < 3.47$ or $0.29 < OR \leq 0.60$), "medium" ($3.47 \leq OR < 6.71$ or $0.15 < OR \leq 0.29$), or "large" ($OR \geq 6.71$ or ≤ 0.15). These interpretations correspond respectively to Cohen's d = 0.2, 0.5, and 0.8 as "small," "medium," and "large" effect sizes (Chen et al., 2010).

RQ1: Spatial Anxiety and Geometric Reasoning

To examine the role of spatial anxiety on geometric reasoning, we fit separate mixed-effect logistic regression models for the dependent variables of verbal insight and transformational proof. For verbal insight (Model 1, Table 8), gender ($p < .001$) and spatial ability ($p = .006$) were significantly associated with the production of correct verbal insight. Males were associated with an increase in the relative odds of producing correct verbal insight of 3.25. Increased spatial ability scores were associated with an increase in the relative odds of producing correct verbal insight of 3.49. When the interaction between spatial ability and spatial anxiety was added to form Model 2 (Table 8), the interaction was significantly associated with a decrease in the relative odds of producing correct verbal insight ($OR = 0.59$, $p = .045$). Gender ($OR = 3.03$, $p < .001$) and spatial ability scores ($OR = 3.25$, $p = .002$) remained significantly associated with correct verbal insight. These results suggest that spatial anxiety may not be directly associated with verbal insight production but indicate a small moderating interaction between spatial ability scores and verbal insight. Though participants with higher spatial ability scores tend to be more likely to produce verbal mathematical insights, there is a modest decline with increases in spatial anxiety.

Table 8

Results of the Logistic Regression Predicting Verbal Insight

Variable		***B***	**SE**	**OR**	***p***	
Model 0: Null Model						
Random Component: Participant ID variance	0.97		0.99			
Random Component: Conjecture variance	0.93		0.96			
Intercept		0.31	0.37	1.36	.000	***
Model 1: Main Effects						
Random Component: Participant ID variance	0.53		0.72			
Random Component: Conjecture variance	0.93		0.96			
Intercept		1.16	0.85	3.19	.005	**
Male[a]		1.18	0.27	3.25	.000	***
Spatial Ability		1.25	0.06	3.49	.006	**
Model 2: Main Effects with Interaction						
Random Component: Participant ID variance	0.49		0.70			
Random Component: Conjecture variance	0.93		0.96			
Intercept		1.14	1.56	0.70	.005	**
Male[a]		1.20	0.26	3.03	.000	***
Spatial Ability		1.37	0.17	3.25	.002	**
Spatial Anxiety		-0.17	0.05	0.84	.051	
Spatial Ability X Spatial Anxiety		-0.52	0.01	0.59	.045	*

Note. N = 752. OR = odds ratio.
[a]Female is the reference category.
* $p \leq .05$, ** $p \leq .010$; *** $p \leq .001$

The results for the models predicting transformational proof (Model 1, Table 9) showed that gender ($p = .033$) and spatial ability ($p = .014$) were significantly associated with the production of a transformational proof. Males were associated with an increase in the relative odds of producing correct verbal insight of 2.29. Increased spatial ability scores were associated with an increase in the relative odds of producing correct verbal insight of 3.94. When the interaction between spatial ability and spatial anxiety was added to Model 2 (Table 9), the interaction ($OR = 0.60$, $p = .019$) and spatial anxiety scores ($OR = 0.85$, $p = .048$) were significantly associated with a decrease in the relative odds of producing a transformational proof. Gender ($OR = 2.25$, $p = .008$) and spatial ability scores ($OR = 3.32$, $p = .013$) remained significantly associated with transformational proofs. These results suggest that spatial anxiety may not directly affect transformational proof production but shows a small indirect moderating interaction between spatial ability scores and proof production. As with insight, higher levels of spatial anxiety modestly depress the benefits of spatial ability on the likelihood of proof production.

Table 9

Results of the Logistic Regression Predicting Transformational Proof

Variable		*B*	SE	OR	*p*	
Model 0: Null Model						
Random Component: Participant ID variance	1.16		1.08			
Random Component: Conjecture variance	2.71		1.65			
Intercept		-1.96	0.61	0.14	.001	**
Model 1: Main Effects						
Random Component: Participant ID variance	0.82		0.91			
Random Component: Conjecture variance	2.72		1.65			
Intercept		-1.37	0.64	0.25	.033	*
Male[a]		0.83	0.32	2.29	.009	**
Spatial Ability		1.37	0.15	3.94	.014	*
Model 2: Main Effects with Interaction						
Random Component: Participant ID variance	0.69		0.83			
Random Component: Conjecture variance	2.73		1.65			
Intercept		-1.42	0.64	0.24	.027	*
Male[a]		0.81	0.31	2.25	.008	**
Spatial Ability		1.42	0.15	4.14	.006	**
Spatial Anxiety		-0.16	0.14	0.85	.048	*
Spatial Ability X Spatial Anxiety		-0.51	0.14	0.60	.019	*

Note. N = 752. OR = odds ratio.
[a]Female is the reference category.
** $p \leq .05$, ** $p \leq .010$*

RQ2: Spatial Anxiety and Gesture

To examine the role of spatial anxiety on gesture production during geometric reasoning, we fit mixed-effect logistic regression models for representational, nondynamic, and dynamic depictive gestures separately. The results for the model predicting representational gestures (Model 1, Table 10) indicated that spatial ability scores (*OR* = 0.82, *p* = .046) were associated with a decrease in the relative odds of producing a representational gesture. This finding suggests that gesture may have been a more valuable embodied resource for participants with lower spatial abilities. When the interaction between spatial anxiety and spatial ability was added to form Model 2 (Table 10), spatial ability ($OR = 0.87$, $p = .029$) remained significantly associated with a decrease in the relative odds of producing a representational gesture. Increased spatial anxiety scores ($OR = 0.79$, $p = .045$) were also significantly associated with a decrease in the relative odds of producing at least one representational gesture. Furthermore, the interaction between spatial ability and spatial anxiety ($OR = 1.75$, $p = .035$) was significantly associated with producing at least one representational gesture. These results suggest that spatial anxiety may moderate the relationship between spatial ability and the likelihood of producing at least one representational gesture, unlike the associated decrease associated with transformational proof production. There were no significant gender effects in either model for representational gestures.

Table 10

Results of the Logistic Regression Predicting Representational Gestures

Variable		*B*	SE	OR	*p*
Model 0: Null Model					
Random Component: Participant ID variance	2.73		1.65		
Random Component: Conjecture variance	0.59		0.76		
Intercept		-0.64	0.33	1.39	.045 *
Model 1: Main Effects					
Random Component: Participant ID variance	2.72		1.65		
Random Component: Conjecture variance	0.58		0.76		
Intercept		3.88	3.02	48.42	.031 *
Male[a]		0.17	0.43	1.19	.697
Spatial Ability		-0.20	0.21	0.82	.046 *
Model 2: Main Effects with Interaction					
Random Component: Participant ID variance	2.56		1.60		
Random Component: Conjecture variance	0.58		0.76		
Intercept		4.58	3.05	97.51	.003 **
Male[a]		0.16	0.43	1.17	.704
Spatial Ability		-0.14	0.21	0.87	.029 *
Spatial Anxiety		-0.23	0.19	0.79	.045 *
Spatial Ability X Spatial Anxiety		0.56	0.19	1.75	.035 *

Note. N = 752. OR = odds ratio.
[a]Female is the reference category.
* $p \leq .05$, ** $p \leq .010$

For nondynamic depictive gestures (Table 11), the results from Model 1 show that spatial ability scores (*OR* = 0.79, *p* = .049) are significantly associated with a decrease in the relative odds of producing at least one nondynamic depictive gesture. Spatial anxiety scores (*OR* = 1.22, *p* = .065) were not significantly associated with nondynamic depictive gesture production. Model 2 (Table 11) added the interaction term between spatial ability and spatial anxiety. This interaction (*OR* = 1.26, *p* = .051) was marginally significantly associated with nondynamic depictive gesture. Spatial ability scores (*OR* = 0.75, p = .030) also remained significantly associated. However, the odds ratios for spatial ability and the interaction did not meet the criteria for a small effect size. No significant gender effects were found in either model.

Table 11

Results of the Logistic Regression Predicting Non-Dynamic Gestures

Variable		***B***	**SE**	**OR**	***p***	
Model 0: Null Model						
Random Component: Participant ID variance	1.42		1.19			
Random Component: Conjecture variance	1.13		1.06			
Intercept		-1.58	0.41	0.21	.000	***
Model 1: Main Effects						
Random Component: Participant ID variance	1.31		1.14			
Random Component: Conjecture variance	1.13		1.06			
Intercept		2.47	2.43	11.82	.310	
Male[a]		0.26	0.35	1.30	.460	
Spatial Ability		-0.23	0.16	0.79	.049	*
Model 2: Main Effects with Interaction						
Random Component: Participant ID variance	1.25		1.12			
Random Component: Conjecture variance	1.13		1.06			
Intercept		3.14	2.45	23.10	.200	
Male[a]		0.24	0.34	1.27	.477	
Spatial Ability		-0.29	0.17	0.75	.030	*
Spatial Anxiety		-0.15	0.15	0.86	.070	
Spatial Ability X Spatial Anxiety		0.23	0.16	1.26	.051	

Note. N = 752. OR = odds ratio.
[a]Female is the reference category.
* $p \leq .05$, *** $p \leq .005$

The results of the first model (Model 1; Table 12) predicting the occurrence of at least one dynamic depictive gesture show that spatial ability scores (p = .012) were significantly associated with dynamic depictive gesture production. Increased spatial scores were associated with an increase in the relative odds of producing at least one dynamic depictive gesture of 1.28. Decreased spatial anxiety scores (*OR* = 0.89, p = .051) were also marginally associated with dynamic depictive gesture production. However, once the interaction between spatial ability and spatial anxiety was added to the model (Model 2, Table 12), neither spatial anxiety (*OR* = 0.89, p = .067) nor the interaction (*OR* = 1.16, p = .078) were significant. Spatial ability remained significantly associated with dynamic depictive gestures (*OR* = 1.25, p = .045). Gender was not significant in either model.

Table 12

Results of the Logistic Regression Predicting Dynamic Gestures

Variable		*B*	SE	OR	*p*	
Model 0: Null Model						
Random Component: Participant ID variance	1.71		1.31			
Random Component: Conjecture variance	2.17		1.48			
Intercept		-2.79	0.57	0.06	.000	***
Model 1: Main Effects						
Random Component: Participant ID variance	1.60		1.26			
Random Component: Conjecture variance	2.18		1.48			
Intercept		-0.26	2.91	0.77	.059	
Male[a]		-0.14	0.41	0.87	.071	
Spatial Ability		0.25	0.20	1.28	.012	*
Model 2: Main Effects with Interaction						
Random Component: Participant ID variance	1.61		1.27			
Random Component: Conjecture variance	2.20		1.48			
Intercept		0.18	3.08	1.20	.053	
Male[a]		-0.15	0.43	0.86	.072	
Spatial Ability		0.22	0.21	1.25	.045	*
Spatial Anxiety		-0.12	0.19	0.89	.067	
Spatial Ability X Spatial Anxiety		0.15	0.19	1.16	.078	

Note. N = 752. OR = odds ratio.
[a]Female is the reference category.
* $p \le .05$, *** $p \le .005$

Discussion

This investigation sought to understand the role of spatial ability and spatial anxiety in geometric reasoning and gesture production. The evidentiary relationships between spatial anxiety, spatial ability, and gesture suggest that geometric thinking is embodied across both affective and cognitive processes. The current study provides correlational evidence that spatial anxiety may act as a moderator between spatial ability and geometric thinking and is associated with the production of mathematical insights, transformational proofs, and representational gestures. Though this study does not endorse a causal claim about the role of spatial systems in geometric thinking, these findings provide a first step in understanding the complex interactions between embodied learning processes and the role of spatial anxiety in spatial ability, geometric thinking, and gesture production.

Spatial ability and spatial anxiety were negatively correlated, corroborating other accounts (Malanchini et al., 2017; Ramirez et al., 2012; Schenck & Nathan, 2020). While spatial anxiety was not directly associated with geometric reasoning, it was a significant moderator between spatial ability and geometric reasoning measured by verbal mathematical insight and transformational proof performance. For gestures, spatial anxiety was a significant moderator between spatial ability and representational gestures but had no significant associations with nondynamic depictive or dynamic depictive gestures. Though this work is correlational and exploratory, these findings expand the current understanding of spatial anxiety and provide insight into its potential role in geometric reasoning and gesture production.

We measured geometric reasoning through verbal mathematical insights and transformational proof production (RQ1). In both outcome measures, spatial anxiety moderated spatial ability, though the effect sizes were small. As spatial anxiety decreased, the relationship between spatial ability scores and geometric reasoning increased. This connection could be explained as spatial anxiety taking up valuable resources needed to perform spatial tasks during geometric reasoning tasks by forcing students to use their limited working memory capacity to both focus on the tasks and inhibit irrelevant thoughts in accordance with the attentional control theory (Eysenck et al., 2007). Domain-trait anxieties, such as mathematics anxiety (Ashcraft & Kirk, 2001) and testing anxiety (Beilock & Carr, 2005), have also been shown to reduce mathematical performance. Though there are few studies on spatial anxiety, it has been shown that there was an interaction between working memory and spatial anxiety in elementary-aged girls performing mental rotation tasks (Ramirez et al., 2012). This interaction was more pronounced for children with higher working memory scores than those with lower ones.

However, this current study did not include working memory measures. Additionally, it is unclear whether spatial anxiety is more associated with verbal or visuospatial working memory. While visuospatial working memory, rather than verbal working memory, has been shown to mediate the relationship between spatial ability and executive functioning (Wang et al., 2018), studies on domain-trait anxieties occupy often use verbal working memory measures in their designs (e.g., Ramirez et al., 2012) as it is thought that anxiety impacts working memory through verbal components, such as intrusive thoughts (Beilock, 2010). Future studies will be needed to fully understand the interaction between spatial anxiety and spatial ability during geometry reasoning.

For gesture (RQ2), we also found that spatial anxiety moderated the relationship between spatial ability and representational gesture production, though the effect size did not meet the criteria for a small effect size. There were no significant relationships between spatial anxiety and nondynamic depictive gestures or dynamic depictive gestures. The lack of association between spatial anxiety and the two subcategories of representational gesture may be due to the low incidence of these codes in this dataset, as only 25% of cases included a nondynamic depictive gesture, and 15% of cases included a dynamic depictive gesture. For the more inclusive superordinate category of representational gesture, the association between spatial ability and gesture production decreased as spatial anxiety decreased. The direct relationship between spatial ability and representational gesture was negative in this study, suggesting that participants with lower spatial ability were more likely to produce at least one gesture. Combined, these findings suggest that participants with lower spatial ability scores and relatively higher spatial anxiety scores may be more likely to gesture. One explanation for this relationship is that representational gesture may function as a beneficial cognitive offloading mechanism, which may help offset the limited visuospatial and verbal working memory capacity caused by spatial anxiety. This interpretation is supported by research showing that individual differences in working memory capacity determine whether gestures affect cognitive processes (Marsaller & Burianová, 2013). However, it is unclear whether this offloading mechanism of gesture production is due to reducing working memory load through restructuring and externalizing information (Cook et al., 2012) or by lowering the individual's current gesture threshold, allowing even small amounts of motor activation to cause a gesture (Hostetter & Alibali, 2008/2019). It should be noted that previous studies on gestures during geometric reasoning have suggested that increased spatial scores are associated with increased gesture production (Nathan et al., 2021; Pier et al., 2019; Walkington et al., 2019). This difference in the direction of association may be due to the modality of the study. The previous studies have been conducted in person, while this current study was moved online due to the COVID-19 pandemic. Students may gesture differently in virtual environments than in physical environments. Future studies will be needed to uncover the specifics of these relationships.

This research demonstrates further evidence for a significant negative correlation between spatial ability and spatial anxiety (e.g., Malanchini et al., 2017; Ramirez et al., 2012; Schenck & Nathan, 2020). In addition to the evidence of a possible moderation effect detailed above, this correlation leads us to question whether spatial ability and spatial anxiety are independent constructs and how they may influence each other. It is possible that higher spatial anxiety is a result of a person's meta-recognition of their poor spatial ability. Alternatively, higher spatial anxiety may lead a person to seek out fewer spatial experiences, which has been linked to lower spatial ability, especially in women (Baenninger & Newcombe, 1989; Quaiser-Pohl & Lehmann, 2002). Future work will be needed to tease apart these relationships to inform the design of interventions that target spatial ability, spatial anxiety, and geometric thinking.

Implications and Conclusions

Research on embodied mathematical thinking has often focused on observable, conscious behaviors like gestures. Studies have shown that co-speech and co-thought gestures, particularly dynamic gestures, can be used to depict mathematical objects and simulate transformations on those objects (Hostetter & Alibali, 2008/2019). These simulated actions strongly correlate with students' ability to generate correct mathematical insights and construct transformational proofs (e.g., Nathan et al., 2021; Swart et al., 2022; Walkington et al., 2022). Though spatial ability was used as a covariate in previous work and found to be associated with intuition, insight, transformational proof, and gesture (Hostetter & Alibali, 2007, Nathan et al., 2021), it was not singled out as an embodied process. This current study expanded the investigation on mathematical thinking to include affective (e.g., spatial anxiety) and cognitive (e.g., spatial ability) embodied processes, adding to the understandings of how mathematical knowledge is embodied.

This study adds to the limited spatial anxiety literature by providing empirical evidence for spatial anxiety's role in geometric thinking. Prior empirical work on spatial anxiety has focused on spatial anxiety's impact on spatial ability outcomes (Gunderson et al., 2013; Kremmyda et al., 2016; Lawton, 1994: Malanchini et al., 2017; Ramirez et al., 2012), adult's success on standardized mathematics problem (Schenck & Nathan, 2020), and middle-grade student's fraction knowledge (Schenck, Hubbard, et al., 2022). Identifying the significant effect of spatial anxiety on the relationship between spatial ability and geometric thinking and gesture production provides a basis for other studies to further explore this interaction. This finding may lead to investigations into possible spatial anxiety interventions to improve both spatial ability and mathematical reasoning abilities.

Chapter 4: A Tangram Task for Geometric Reasoning during Insight and Proof

Over the last few decades, there has been increasing interest in improving mathematics proficiencies. Several lines of research have investigated the influences of motoric (i.e., gesture) and cognitive (e.g., spatial ability) processes on mathematics abilities. For example, studies have shown that gesture is associated with mathematical reasoning both directly (e.g., Goldin-Meadow et al., 2009; Nathan et al., 2021) or indirectly (e.g., Walkington et al., 2022). It has also been widely demonstrated that spatial abilities are associated with mathematical achievement across different ages and domains of mathematics (e.g., Delgado & Prieto, 2004; Mix et al., 2016). The malleability of spatial ability may be especially useful for building mathematics abilities. Spatial abilities can be improved through both direct and indirect training methods and in and out of educational settings (Baennigner & Newcombe, 1989; Uttal et al., 2013). Spatial training may be effective for increasing student success in STEM fields, as problem solving in STEM areas relies on spatial reasoning (Stieff & Uttal, 2015).

The literature investigating whether interventions that involve spatial abilities to improve math is ambivalent and includes both studies that show support and a lack of support. There is a growing body of causal evidence that spatial-based tasks improve mathematics achievement (e.g., Adams et al., 2022; Gilligan et al., 2020; Hawes et al., 2017; Sorby & Veurink, 2019). Other studies have found no evidence of spatial tasks transferring to mathematics achievement (Hawes et al., 2015). The varied findings could result from the multidimensional and complex nature of spatial ability and mathematics. Furthermore, these studies are often limited by their focus on pen-and-paper mathematics assessments as outcomes (e.g., Hawes et al., 2022, Yang et al., 2020). As discourse tasks such as questioning, proving, and justifying problem-solving are important for fostering conceptual and procedural mathematics knowledge (e.g., Legesse et al., 2020), it is beneficial to understand the potential of a tasks that involves spatial abilities for improving these types of mathematics proficiency.

The central objectives of this paper are to provide evidence for whether a Tangram task is effective for improving geometric reasoning, can impact on students' gesture production, and to explore the associations between working memory and spatial anxiety and geometric reasoning. Specifically, we examine the potential for a short Tangram task to increase students' generation of mathematical insights, transformational proof, and the types of gestures often associated with mathematically valid reasoning in geometry. Understanding the impact of spatial-based tasks in discourse-based geometric reasoning (i.e., proof) and gesture production provides a first step toward understanding how spatial-based tasks can impact discourse-based reasoning tasks and the basis for future intervention designs for use in classrooms, educational technology designs, and teacher professional development.

Theoretical Background

In the following section, I give an overview of the main theories central to this paper. I will first summarize the literature on the different aspects of the spatial system. Next, I discuss the prior work on grounded and embodied mathematical thinking. This background provides the theoretical framing that leads to the current study's research questions.

The Spatial System

Many different components may make up one's system for processing visuospatial information. For this paper, I will address three components of this system: spatial ability, spatial anxiety, and working memory.

Spatial Ability

Spatial ability has been a topic of interest and debate within the scientific community for more than a century, with researchers struggling to agree on a complete definition or list of sub-components. The ability to generate, retain, and manipulate abstract visual images is a recurring theme in many definitions (e.g., Battista, 2007; Gaughran, 2002; Lohman, 1979). Factor-analytic studies have attempted to determine if spatial ability is a unitary structure or comprised of various sub-factors (e.g., Buckley et al., 2018; Carroll, 1993), with some researchers advocating for broader categorical distinctions, such as large- versus small-scale skills (Hegarty et al., 2018) or intrinsic and extrinsic information (Uttal et al., 2013). In contrast, recent work by Malanchini et al. (2020) supports a unitary model of spatial ability, demonstrating the existence of a common genetic network that supports all spatial abilities.

The importance of spatial ability extends to its impact on success in science, technology, engineering, and mathematics (STEM) fields, with deficiencies in spatial abilities posing obstacles to STEM careers (Harris et al., 2013; Shea et al., 2001). Research has found connections between success in spatial tasks and mathematical tasks in both children (e.g., Casey et al., 2015; Schenck, Hubbard et al., 2022) and adults (e.g., Schenck & Nathan., 2020), and several meta-analyses have confirmed positive associations between spatial and mathematics skills (Atit et al., 2021; Xie et al., 2020). Although the specific nature of these associations remains largely unknown, studies have suggested shared processing requirements in both spatial and mathematical tasks (Hubbard et al., 2005; Mix et al., 2016). Additionally, researchers have sought to determine which mathematical concepts engage spatial thinking, with some studies finding associations between specific spatial ability sub-components and mathematics concepts (Burte et al., 2017; Cheng & Mix, 2014; Hannafin et al., 2008; Schenck & Nathan, 2020).

Spatial Anxiety

Spatial anxiety, and its impact on neurocognitive performance, has been a subject of study for researchers. According to the Attentional Control Theory, anxiety negatively affects performance by disrupting working memory capacity (Eysenck et al., 2007). Spatial anxiety, a domain-specific trait anxiety, has been found to be negatively correlated with spatial ability, interfering with processes involved in spatial skill development (Gunderson et al., 2013; Malanchini et al., 2017; Ramirez et al., 2012). Like spatial ability, spatial anxiety is thought to have subcomponents, such as navigation anxiety and rotation/visualization anxiety (Lyon et al., 2018; Malanchini et al., 2017). Various assessments have been developed to measure spatial anxiety but are limited to specific age groups or situations. A recent questionnaire developed by Lyons et al. (2018) provides a relatively comprehensive measure of spatial anxiety, targeting adults and featuring subscales for mental manipulation, navigation, and imagery, making it particularly relevant to the current study.

Studies have also investigated gender differences in spatial anxiety. While some studies have reported females experiencing higher spatial anxiety and performing worse on spatial tasks (e.g., Wei et al., 2018), others have found no significant sex differences (e.g., Hund & Minarik, 2006). The inconsistency in results could be due to the heterogeneous nature of spatial ability and the fact that anxiety may only manifest during certain spatial activities.

Malleability of Spatial Ability. Hundreds of studies have explored whether spatial ability can be improved. Overall, studies have shown that spatial ability is sensitive to training effects (e.g., Cheng, 2016; Peters et al., 1995; Wright et al., 2008) and that spatial training may be beneficial for mathematics outcomes (e.g., Cheng & Mix, 2014; Hawes et al., 2015; Mix et al., 2021). Several meta-analyses summarize the results of these studies and provide comprehensive insight into the malleability of spatial skills.

Baennigner and Newcombe (1989) conducted the first of these large meta-analyses. In two separate analyses, they investigated: (1) the relationship between prior spatial experiences and spatial ability and (2) the effects of spatial training on spatial ability. The first analysis included 11 studies, including 26 samples divided into male and female groups, and provided evidence for a weak link between spatial activity participation and spatial test performance. No significant gender differences were discovered. The second analysis included 19 spatial training studies initially divided into experimental and test-retest groups. These studies were then categorized by gender, the training content (specific, general, or indirect), and the duration of training. Overall, the analyses indicated that spatial training improves performance on spatial tests, and conditions are optimal when specific training is administered in at least three sessions. Furthermore, spatial training improved the scores of both males and females for all groups.

Expanding on Baenninger and Newcombe's (1989) findings, Uttal and colleagues (2013) completed a meta-analysis of 217 studies on spatial ability training. These studies were completed after 1984 (the last date of studies used in the Baenninger and Newcombe (1989) analyses), employed a causally relevant design, and focused on non-clinical populations. Rather than specify studies by spatial skill sub-component, spatial skills were coded based on a 2 x 2 framework with dimensions of intrinsic versus extrinsic and static versus dynamic. Training programs were divided into video games, instructional courses, and direct spatial task training. Overall, three key results were found: (1) Spatial ability is malleable in people of all ages; (2) training is effective, transferable, and long-lasting; (3) interventions are effective both in and out of educational settings. Like the previous meta-analysis, there were no significant differences between types of training or genders. Unlike Baenninger and Newcombe (1989), Uttal and colleagues did not conclude that a particular type of intervention was more effective. Instead, their analysis suggested that the training type may depend on time, resources, individual differences, and study goals. In addition, training was shown to have strong effects of transfer from one assignment to another, and targeted training may increase learners' chances of succeeding or specializing in STEM disciplines.

Spatial Interventions. Studies have shown that spatial training can improve spatial abilities, but some scholars have investigated whether training these skills can improve mathematics abilities. There is ample evidence of a positive relationship between STEM skills and spatial ability and that spatial ability can improve with training. Theoretically, researchers have argued that spatial training may translate into increasing student achievement in STEM fields (Stieff & Uttal, 2015; Uttal & Cohen, 2012). Scholars have proposed that the best way to understand the link between spatial skills and STEM achievement is to build on complementary approaches in the mathematics education and psychology fields by focusing on a better alignment between the sub-factors of spatial ability and STEM reasoning skills in the exploration of how spatial ability inventions influence how students think and learn (Lowrie et al., 2020).

In the last five years, there have been substantive gains in empirical research into the efficacy of spatial training on mathematical ability. At least two recent meta-analyses have focused on understanding the effects of spatial training on mathematics performance. Yang and colleagues (2020) sought to extend Uttal and colleagues' (2013) findings to young children's spatial skills. This meta-analysis included 20 intervention studies for children aged 0 to 8 years. It used Uttal and colleagues' (2013) spatial skill 2 x 2 framework and training type classifications. The videogame training classifier was expanded to include play and hands-on operation to account for the age of the participants in the study. The analysis showed that early spatial skill training is highly effective, with no significant differences in age, training type, or research setting. For gender, however, results revealed that spatial training led to a greater effect for girls than boys.

Hawes and colleagues (2022) also examined the possible effects of spatial training on mathematics performance and under what conditions these effects occur. An examination of 29 pre-post designed studies found that spatial training was effective for increasing success on both spatial and mathematics tasks. Age, use of concrete manipulatives, and type of transfer moderated the effects of spatial training on mathematics. In contrast, spatial gains, type of control group, and training dosage were not significantly associated with training gains. These findings suggest that spatial training that involves concrete materials or short interventions may be beneficial for increasing mathematical thinking. However, this meta-analysis included only two studies on adults.

Overall, evidence supports that spatial skills and mathematics performance could be enhanced through spatial training. That being said, very few studies have explored the impact of spatial training on mathematics in adults and how spatial training may impact mathematics outcomes that are not measured by written tests. Moreover, results remained mixed on a few issues, including whether spatial training can close gender gaps and what design decisions may make training more effective.

Working Memory

It is thought that working memory may be an essential component of aspects of intelligence, including spatial ability (Kyllonen, 1996; Miyake et al., 2001). One of the most widely accepted models of working memory is Baddeley and Hitch's multi-modal model (1974;

see also Baddeley, 2000 for more recent updates). Most crucial to Baddeley and Hitch's framework is the construction of working memory as comprised of three subcomponents: the central executive, the phonological loop, and the visuospatial sketchpad. The visuospatial component of working memory may be particularly critical for mathematics and spatial abilities.

Studies have indicated that specific sub-components of spatial ability (spatial visualization, speeded rotation, and visuospatial perceptual speed) may differ in cognitive load on the working memory system (Shah & Miyake, 1996; Hegarty et al., 2000). Tasks that measure these three spatial ability components require spatial transformation, which cannot be completed in a single eye fixation. Thus, they require temporary visuospatial storage thought to be provided by the visuospatial sketchpad (Miyake et al., 2001). Furthermore, Miyake and colleagues (2001) showed differing degrees of working memory involvement between spatial visualization, speeded rotation, and visuospatial perceptual speed, with spatial visualization tasks showing the high involvement and visuospatial perceptual speeding showing the lowest. Researchers posit that these differences may be due to the maintenance (i.e., cognitive load) of storing the visuospatial representations. These findings could help explain within-subject and between-subject variations in spatial ability measures.

Individual differences in spatial rather than verbal working memory are related more closely to proficiencies in mathematics (Reuhkala, 2001; Kyttälä et al., 2003). The variations in spatial working memory are particularly significant for *number*, *measurement*, and *space*. Researchers have suggested this could be because the visuospatial sketchpad is used to represent visual number forms (Hayes, 1973), non-numeric spatial arrangements (Seron et al., 1992), carrying operations (Heathcote, 1994), counting imagery (Bull et al., 1999) and geometry (Hartje, 1987). Differences in processing speed and spatial working memory capacity could also partly explain variation in mathematics scores. For example, speed-of-processing indices accounted for variance in young children's arithmetic abilities (Bull & Johnston, 1997). Additionally, children with low visuospatial working memory capacity reflect a limited capacity to temporarily store visuospatial information (Geary et al., 2000).

Differences in visuospatial working memory may also impact spatial abilities through problem-solving strategy selection and anxiety. Research on mathematical problem-solving strategies has shown that individuals with higher working memory scores rely on working-memory-demanding strategies, while individuals with lower working memory scores often rely on heuristic strategies, which use a lower working memory demand. (Beilock & DeCaro, 2007). Individuals with high working memory were also found to be most affected by stress-related anxiety, which can disrupt cognitive processes (Beilock & Carr, 2005; Ramirez et al., 2012). Furthermore, other studies have shown that spatial ability explains unique variance in mathematical word problem-solving performance (Blatto-Vallee et al., 2007; Casey et al., 2008; Boonen et al., 2013). Thus, it is possible that individuals with higher working memory scores may be more susceptible to the negative impacts of anxiety on the performance of problem-solving tasks like those on spatial ability measures. As visuospatial working memory has been linked to spatial abilities as well as mathematic abilities, it could be that the differences outlined above could explain the variations in links between spatial ability and mathematics for individual students.

Grounded and Embodied Mathematical Thinking

Grounded and embodied theories of cognition posit that meaning-making, even for abstract concepts, can be achieved through body-based experiences (Barsalou, 2008; Shapiro, 2019). Several theoretical frameworks offer explanatory accounts for how grounded and embodied cognition (GEC) arises from the interaction of cognitive, perceptual, and motor processes. For example, Barsalou (1999) proposed a sensory-motor account of mental representations in which mental processes operate with perceptual symbols, activating a perceptual-motor simulation of properties associated with a particular concept. Glenberg and Robertson (1999) proposed the Indexical Hypothesis, which describes how actions influence comprehension and meaning making during reading tasks. When reading, actions help readers index or ground the abstract symbols in objects and movements. The benefits of grounded reading extend to imagined and physical actions (Glenberg et al., 2004). These benefits seen in reading also extend to include mathematical problem solving, which typically involves a verbal or written prompt (Glenberg et al., 2012).

Gestures provide an important avenue of evidence for the role of the body in complex reasoning due to their role in language production, comprehension, and problem solving (McNeil, 1992). Though gestures are seldom in a one-to-one relation with spoken word or meaning, gestures convey meaning by combining semantic and pragmatic content to simulate action (Hostetter & Alibali, 2008/2019) and typically accompany speech or thought (Goldin-Meadow, 2003; Kita et al., 2017; McNeil, 1992). Gestures serve several functions during communication and thinking, such as affecting cognitive processes, conveying information, and managing cognitive load. Gestures help activate, maintain, manipulate, and package visuospatial and motoric information (Kita et al., 2017). They can keep spatio-motoric information active in working memory, reducing cognitive load during reasoning tasks (Cook et al., 2012; Goldin-Meadow et al., 2001). For example, individuals with lower visuospatial working memory capacity may have higher gesture rates to compensate for their limited memory resources when thinking and speaking about spatial information (Chu et al., 2014; Göksun et al., 2013).

Individuals also have a gesture threshold (Hostetter & Alibali, 2008). This threshold is the level of motor activation needed for a mental simulation to be expressed in action and can vary depending on factors such as individual differences (e.g., spatial and verbal skills, anxiety, working memory), current task demands (e.g., processing spatial imagery), and situational contexts. For example, Hostetter and Alibali (2007) found that individuals with high spatial visualization skills and low verbal skills produce gestures at higher rates than individuals with different spatial and verbal skill combinations.

Recently, several studies have provided empirical evidence for the grounded and embodied nature of geometric thinking, specifically in proof. Constructing proofs is a methodology for generating new mathematical knowledge and is one of the focuses of secondary geometric education (Rav, 1999). The production of mathematical proofs requires students to evaluate universal truths about shape and space and justify their conclusions. Studies have shown that both professional mathematicians and students ground their understandings of mathematical content through embodied practices such as gestures (e.g., Kim et al., 2011; Marghetis et al., 2014; Nathan et al., 2021, Walkington et al., 2019; Walkington et al., 2022). Studies have specifically identified a class of representational gestures, called *dynamic gestures*, as critical for valid, generalizable geometric thinking (Pier et al., 2019; Nathan et al., 2021; Walkington et al., 2019). While non-dynamic gestures (such as pointing and tracing) can effectively simulate and communicate static properties of objects, dynamic gestures depict the invariant characteristics of mathematical objects as they undergo transformations, such as rotation, reflection, dilation, and skewing. Furthermore, prior studies have found strong associations between dynamic gestures, spatial reasoning skills, and geometric reasoning (Göksun et al., 2013; Nathan et al., 2021).

Research Questions

Previous research on the potential of spatial interventions for mathematical outcomes has provided empirical evidence that some tasks that involve spatial abilities may improve mathematics performance. Nevertheless, these studies are limited and typically focus on younger populations and rely on pen-and-paper assessments. Research on embodied geometric reasoning has shown that gestures, particularly dynamic gestures, can influence one's reasoning process. However, it is unclear what the impact of spatial-based tasks on gesture during these reasoning tasks may be. Additionally, previous research has indicated that other factors, such as working memory and spatial anxiety, may be associated with spatial abilities and, in turn, impact the effectiveness of tasks that involve spatial abilities for mathematical thinking. Thus, there is a value in investigating the impacts of spatial-based tasks and how working memory and spatial anxiety are associated with embodied geometric reasoning and gesture production, leading to three theoretically motivated research questions.

R1: Can a short Tangram puzzle task impact students' geometric reasoning?

Tasks that invoke spatial abilities, such as solving Tangram puzzles, are thought to improve access to one's spatial abilities, and spatial abilities have been found to be critical for geometric reasoning. I hypothesize that a short Tangram task will improve geometric reasoning as measured by mathematical insight and transformational proof.

R2: Can a short Tangram puzzle task impact students' tendency to gesture?

Students who produce representational gestures, especially dynamic representational gestures, often exhibit superior mathematics performance because representational gestures can simulate properties and transformations of mathematical objects. Tangram puzzle tasks require students to imagine and manipulate objects. On this basis, I hypothesize that a short Tangram puzzle task will lead to increased gesture production.

R3: How are working memory and spatial anxiety associated with geometric reasoning and gesture production?

This research question is exploratory. Working memory, especially visuospatial working memory, is often associated with mathematical thinking and spatial abilities, and gestures may help reduce the demands on working memory. I hypothesize that visuospatial working memory will be associated with both (a) geometric reasoning and (b) gesture production. Furthermore, spatial anxiety has been shown to be negatively associated with spatial abilities and representational gestures, and it may moderate the relationship between spatial ability and mathematical thinking. Thus, I hypothesize that spatial anxiety will be negatively associated with (c) geometric thinking and (d) representational gesture production.

Methods

Participants

Eighty-four undergraduate students were recruited from a large university in the Midwestern United States. Inclusion criteria required fluent English production and comprehension and completion of all covariate measures of interest. As compensation, participants received a $20 gift card. Participants were randomly assigned to either the control or intervention groups. The control group (n = 42) included 23 students who identified as female, with 90% identifying as native English speakers. The intervention group (n = 42) included 21 students who identified as female, with 95% of students identifying as native English speakers. Further descriptive statistics for both groups can be found in Table 13.

Power Analysis

The a priori power analysis used β = 0.80, α = 0.05, and an effect size of f = 0.25 for the effect of the spatial intervention on proof performance using G*Power's ANOVA Repeated Measures (Faul et al., 2007). The effect size was estimated based on prior studies that have used similar spatial interventions (Cornu et al., 2019; Mix et al., 2021; Siew et al., 2013; Sundberg, 1994). The correlation among participants solving six repeated geometry proofs was estimated at 0.5 based on previous studies' data generating model (see Schenck, Kim, et al., 2022). A minimum sample of 76 (36 per group) is needed. The current sample, which included 42 participants per group, is expected to provide adequate power for the following statistical analyses.

Table 13

Demographics (N = 84)

Variables	**Intervention (n=42)**	**Control (n=42)**
Average Age (SD)	19.29 (1.22)	19.38 (1.24)
% native English speakers	95.24%	90.47%
% Ethnicity- White/Non-Hispanic	57.14%	73.81%
% Ethnicity- Asian	26.19%	14.29%
% Ethnicity- Other	16.67%	11.90%
% Gender- Male	45.24%	42.86%
% Gender- Female	50.00%	54.76%
% Gender- Other	4.76%	2.38%

Materials

Covariate Measures

Participants completed several measures as part of a pretest before the intervention. The first was a demographic information survey via Qualtrics that included self-reported information about age, gender, native language, and ethnicity. After data collection was complete, I collapsed the ethnicity responses into three categories for reporting: White/Non-Hispanic, Asian, and

Other. Gender responses were also collapsed into two categories for analysis: Female and Male/Other.

Spatial anxiety was then assessed using the *Novel Spatial Anxiety Scale* (Lyons et al., 2018) via Qualtrics. This spatial anxiety measure is a 24-question 5-item Likert scale measure that asks students to rate how anxious they would be from “Not at All” to “Very Much” in a variety of situations. These situations can require mental manipulation (e.g., rotating an object), imagery (e.g., recalling specific features of an object), or navigation (e.g., getting around an unfamiliar location). Each question was scored from 0 (“Not at All”) to 5 (“Very Much) to compute a composite score. Internal reliability testing showed Cronbach’s alpha for the overall composite score of 0.85.

To assess spatial ability, I followed the recommendations of Schenck and Nathan (2023). Previous data analysis from a study on geometric reasoning identified mental rotation and non-rotational spatial visualization as significant predictors of verbal geometric reasoning (Schenck, Kim et al., 2022). These two constructs both fall under the specific factor of spatial visualization (Carrol, 1993; Lohman, 1988; McGee, 1979). For task selection, I chose the Paper Folding Task (Ekstrom et al., 1976). This spatial visualization measure has been shown to be predictive of measures of geometric thinking (Nathan et al., 2021; Walkington et al., 2019) and gesture production (Nathan et al., 2021; Hostetter & Alibali, 2007). The Paper Folding Task is a 20-item (2 x 10 items per section), timed, multiple-choice assessment. Participants were given three minutes to complete each section of ten questions. Scores were computed with one point for each correct answer and -0.25 points for each incorrect answer. The internal reliability for my sample was 0.80.

Working memory was assessed using two different span tasks. Both tasks used computer-based software created by Stone and Towse (2015) and were presented using the Tatool platform (von Bastian et al., 2012). The first task was the matrix span which measures visuospatial working memory. In this task, participants were shown a 4 x 4 grid. One at a time, grid locations would light up. After receiving all grid locations for the set, participants were notified of the recall phase and were asked to click the grid locations in the order they recalled seeing them. Each set contained between three and nine grid locations. Each set size was repeated three times in random order for a total of 126 grid locations. The second task was the reading span which measures verbal working memory. In this task, participants were first presented with a two-digit number to be memorized. Immediately following the presentation of the number, the participants were asked to read a sentence as fast as possible and assess whether it made sense. Each set contained from three to seven numbers. At the end of each set, the participants were instructed to recall the numbers in the correct serial order. Each set size was repeated three times in random order for a total of 75 sentences and numbers. At the end of each span task, a partial non-weighted score was calculated without the 85% accuracy criterion on the processing component of the span tasks. Partial, rather than absolute, scores were calculated as partial scores have been shown to have higher internal consistency and convergent validity for these tasks (Đokić et al., 2018). Non-weighted scores were selected as they are more in accordance with standard psychometric procedures (e.g., Conway et al., 2007)

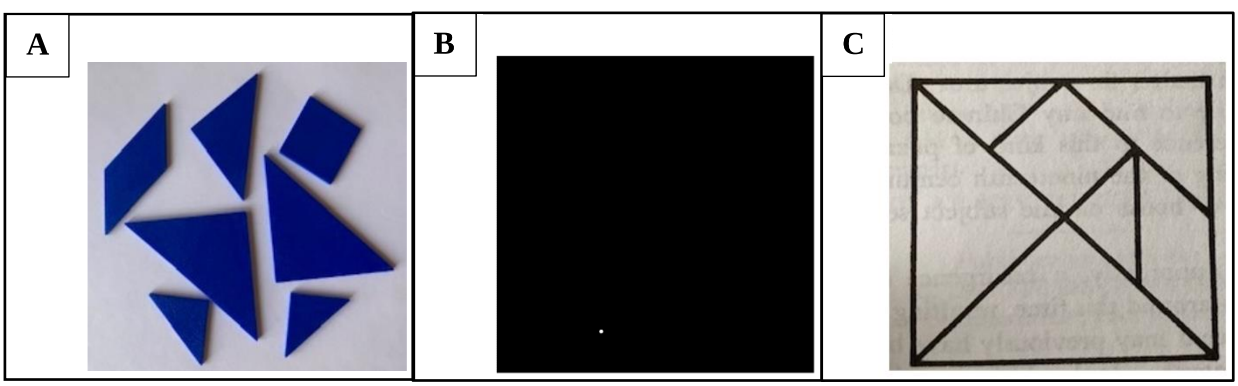
Tangram Task

After completing the pretest covariate measures, participants were randomly assigned into one of two conditions: a control task or a Tangram task. In the control condition, participants were shown a set of six letters arranged in a circle via PowerPoint. Each set of letters included at least two vowels and could be combined to form at least one six-letter English word. They were then instructed to write down all the possible words or common acronyms that could be created from a combination of the letters. Each set was presented for four minutes before a new set was automatically displayed, and a total of ten sets were presented for a total task time of 40 minutes. This task was chosen as the control task as it involves relatively low spatial skills. In a pilot study (N = 32), students were randomly assigned either the control task or the Tangram task condition. Most of the 16 participants in this control condition reported using replacement strategies (i.e., starting with a word and replacing the consonants or vowels) or association strategies (i.e., finding similar words associated with a found word). Additionally, in this pilot, students in the control condition group completed the Paper Folding Test (Ekstrom et al., 1976) before and after the control task. There was no significant difference between the spatial pre- and posttest scores ($t(30) = 1.14$, $p = .262$)

In the Tangram task condition, participants were asked to complete a set of puzzles using physical Tangram manipulatives (Figure 5A). For each puzzle, participants were shown an image of a composite shape via PowerPoint (Figure 5B) and were given 5 minutes to assemble the shape using all seven Tangram pieces. After the participants assembled the composite shape or after 5 minutes, an image of the composite shape with the individual pieces visible was revealed (Figure 5C). Then, participants were instructed to check their assembled solution against the displayed solution and make any necessary changes to their assembled shape before moving on to a new puzzle. Participants completed as many puzzles as possible in 40 minutes to match the control condition's time demands. The mean number of puzzles solved by participants in this condition was eight, with a range of 6 to 13 puzzles. This task was chosen as the spatial intervention for three reasons: (1) Tangram puzzles rely on spatial visualization skills such as disembedding, mental rotation, and visualization, which have been previously linked to geometric reasoning (Moyer-Packenham et al., 2008; Schenck, Kim, et al., 2022); (2) Previous studies with both adults and children have shown that experiences with Tangram puzzles resulted in improves scores on paper-and-pencil geometry tasks (Cornu et al., 2019; Mix et al., 2020; Siew et al., 2013; Sundberg, 1994); and (3) Tangram puzzle pieces are physical manipulatives and include several geometric shapes, indicating a possible close relationship between Tangrams and geometry concepts which Hawes' and colleagues' (2022) suggest may increase the likelihood that a spatial-based task will impact mathematics outcomes. In the pilot study (N = 32), the 16 students in the Tangram spatial intervention condition reported using spatial and geometric strategies such as visualizing how to break up the picture into the relevant pieces, imagining rotating the pieces to fit into the image, and mentally comparing the geometric features (i.e., angles, side lengths, and area) to decide which piece was appropriate, providing further evidence that participant used spatial-based strategies to solve the Tangram puzzles. Additionally, there was a significant increase in difference between the spatial pre- and posttest scores ($t(30) = 2.168$, $p = .048$).

Figure 5

Spatial Intervention Group Task



Conjectures

The six conjectures used in this study were selected from a larger list of conjectures used in another study (Schenck, Walkington, & Nathan, 2022). These conjectures reflect the general properties of two- and three-dimensional objects often included in secondary geometry curriculums. Three conjectures explored the two-dimensional properties of lines, parallelograms, and triangles. The remaining three conjectures explored the three-dimensional properties of prisms, cylinders, and spheres. All six conjectures were true statements. Table 14 includes the text and relevant mathematical insights for each conjecture.

Table 14

Text and mathematical concepts for each of the eight conjectures

Conjecture Name	Conjecture Text	Example Verbal Insights
Triangle	The sum of the lengths of any two sides of a triangle must be greater than the length of the remaining side.	1) References the Triangle Inequality Theorem 2) The shortest distance between two points is a line. 3) If the sum was equal to the third side, it would form a line. If the sum were less than the third side, it would not connect to form a triangle.
Parallelogram	Consecutive angles in a parallelogram add up to 180 degrees. Consecutive angles are two angles inside the shape that are next to each other.	1) Interior angles of a parallelogram sum to 360, and half of 360 is 180. 2) Parallelograms consist of parallel lines and same side interior angles sum to 180. 3) Rectangles and squares are parallelograms with right angles.
Lines	If two parallel lines are cut by a third line, the pairs of corresponding angles are congruent. Corresponding angles are in the same position at each intersection where a straight line crosses two others.	1) References the Corresponding Angles Theorem 2) Transversal cuts both parallel lines at the same angle, so you can use a translation to map the angles to each other 3) Uses vertical angles and alternate interior angle theorems to show a logical train of congruency.
Sphere	A plane can only intersect a sphere at zero, one, or infinite points. A plane is like a sheet of paper that goes on forever in every direction.	1) If a plane does not intersect the sphere, it intersects at zero points. (OR False because intersect implies at least one intersection) 2) If a plane is tangent to the sphere, it intersects at one point. 3) The cross section of a plane is a circle with infinite points.
Prism	If the length, width, and height of a cube are each doubled, then the volume increases by a factor of 8.	1) The volume of a cube is side cubed or length times width times height. 2) Doubling each side multiplies each dimension by 2. 3) Uses examples to find the volume of multiple cubes.
Cylinder	Given a cylinder with radius *r* and height *h,* the cylinder can be unrolled to include a rectangle with dimensions of *h* and 2*pi*r.	1) The rectangle will have the same height as the cylinder. 2) 2* pi* r is the formula for circumference.

Procedure

All participation was completed in a private room with a researcher. Participants began their session by completing the demographics survey and the Novel Spatial Anxiety Scale via Qualtrics on a laptop. Participants then completed the Paper Folding Task by providing answers to the item on the worksheet. Next, participants completed the Matrix and Reading Span tasks on a laptop.

Participants were then randomly assigned to either the control or Tangram task condition. In each condition, the researcher read instructions aloud as participants followed along with the instructions displayed on the PowerPoint software. Participants then completed the word generation (control group) or Tangram puzzle (intervention group) tasks for 40 minutes.

Participants were then asked to stand three feet from the table with the laptop while the researcher introduced the conjecture task. Specifically, participants were instructed to read the geometry conjecture aloud, decide whether the conjecture was always true or ever false, and then provide a justification. All participants completed six conjectures that had their presentation order counterbalanced using a Latin square design. Conjecture responses were videotaped.

Coding

Videos of this section of the experiment were organized, and participants' speech was transcribed verbatim. Initial speech transcription was done using Otter for Business, an online automated transcription service (Otter.ai, 2023), before a researcher reviewed and corrected each transcript. Timestamps were then added to split the full transcripts into the six conjectures for coding, resulting in 504 video clips to be coded (84 participants x 6 conjectures). The occurrence of gestures was then coded for each clip using the videos. One researcher from the research team coded each transcript using the coding scheme described below.

Coding Scheme

Transcripts were coded for mathematical insight, transformational proof, and gesture. Mathematical insight consists of two possible subcategories: verbal insight and gestural insight. Verbal insight was coded (1/0) for the presence of key mathematical ideas for each conjecture in the participant's speech, as specified by our team of mathematicians and math educators (examples are shown in Table 14). Gestural insight (1/0) refers to spontaneous gestures in which a person's gestures exhibit mathematical insights otherwise absent in speech (Xia et al., 2022.). These two categories were combined into a single *Insight* (1/0) category for this analysis.

Transformational proof was coded (1/0) if the verbalized proof showed a logical sequence of generalizable reasoning and included operational thinking as defined by Harel and Sowder (2005). Any verbalized insights that met only part of this criteria were scored a zero.

Three separate gesture codes were used. The first code identified the occurrence of a representational gesture that represents or depicts some feature or operation of a mathematical object or idea (Alibali & Nathan, 2012). Representational gestures were then divided into two subcategories: nondynamic depictive and dynamic depictive gestures. *Nondynamic depictive gestures* are representational gestures that only represent mathematical entities (including acts such as tracing or pointing) but do not involve transformations that alter the geometric objects under investigation. *Dynamic depictive gestures* are representational gestures that demonstrate geometric transformation, such as rotating, dilating, or skewing an object. Any occurrence of at least one of these types of gestures would lead to assigning a 1 for that respective gesture code. Participant transcripts that did not include any gesture or included non-representational gestures (e.g., beat gestures) were coded 0.

Inter-rater Reliability

Inter-rater reliability was established by a researcher not involved in the initial coding process, who coded a random sample of roughly 20% of the participants' videos. I also calculated Shaffer's *rho* statistics to assess the validity of the inter-rater reliability using a *kappa* threshold of 0.65 (Eagan et al., 2017). The overall inter-rater reliability for these codes is κ = 0.84 with a Shaffer's *rho* of 0.00, indicating the sample size was sufficient to estimate the inter-rater reliability at a threshold of at least 0.65. Individual inter-rater reliability measures for the insight, transformational proof, and gesture codes are shown in Table 15.

Table 15

Inter-rater Reliability for Participant Transcript Coding

Code	**Cohen's *kappa***	**Shaffer's *rho***
Omnibus	0.84	0.00
Insight	0.85	0.00
Verbal Insight	0.81	0.00
Gestural Insight	0.87	0.02
Transformational Proof	0.86	0.00
Representational Gesture	0.81	0.00
Non-Dynamic Gesture	0.79	0.00
Dynamic Gesture	0.83	0.00

Note. Shaffer's *rho* was calculated with a *kappa* threshold of 0.65

Results

Descriptive statistics showed no obvious differences in the covariate measures between the control and intervention groups, but there were differences in the percentage correct for proof and all three gesture categories (Table 16). Correlations among all the key factors are presented in Table 17. Verbal and gestural insight categories were combined into one "insight" category for correlations and analyses. For the analyses, mixed effect logistic regression models for binary outcomes (0/1) on the accuracy of insight, proof, representational gesture, non-dynamic gesture, and dynamic gesture were fit using the *glmer* command in the R software package *lme4* (Bates et al., 2015). Participant ID and conjecture were included as random effects in all models. In addition to the covariate measures, gender was added to the models as there are well-documented gender differences in spatial ability (e.g., Maeda & Yoon, 2013) and spatial anxiety (e.g., Malanchini et al., 2017).

As the dependent variables in each model are dichotomous, I will report odds ratios rather than effect sizes. I interpreted odds ratios as "small" ($1.68 \leq OR < 3.47$ or $0.29 < OR \leq 0.60$), "medium" ($3.47 \leq OR < 6.71$ or $0.15 < OR \leq 0.29$), or "large" ($OR \geq 6.71$ or ≤ 0.15) based on the recommendation of Chen and colleagues (2010).

Table 16

Descriptive statistics (N = 84)

Variables	**Intervention (n=42)**	**Control (n=42)**
Average spatial ability score (SD)	13.82 (3.72)	13.61 (3.89)
Average spatial anxiety score (SD)	37.60 (13.59)	35.81 (11.88)
Average verbal working memory accuracy	0.33 (0.17)	0.33 (0.13)
Average visuospatial working memory accuracy	0.65 (0.18)	0.61 (0.16)
Likelihood of correct insight (per trial)	61.11%	57.94%
Likelihood of correct verbal insight (per trial)	56.75%	54.37%
Likelihood of correct gestural insight (per trial)	6.75%	5.95%
Likelihood of correct proof (per trial)	30.16%	11.90%
Likelihood of representational gesture (per trial)	61.11%	50.79%
Likelihood of nondynamic gesture (per trial)	59.52%	47.62%
Likelihood of dynamic gesture (per trial)	32.14%	18.65%

Table 17

Correlations for Key Factors

Variables	1	2	3	4	5	6	7	8	9
1. Gender[a]									
2. Verbal Working Memory	**-.180**								
3. Visuospatial Working Memory	-.114	**.323**							
4. Spatial Ability Score	**-.139**	**.262**	**.325**						
5. Spatial Anxiety Score	.067	**.208**	**-.103**	**-.121**					
6. Insight	**-.206**	**.223**	**.209**	**.277**	**-.002**				
7. Transformational Proof	**-.121**	**.287**	**.205**	**.248**	.047	**.452**			
8. Representational Gestures	-.063	**.177**	.098	.113	-.068	**.148**	**.291**		
9. Non-Dynamic Gestures	-.065	**.181**	**.118**	**.133**	-.066	**.168**	**.295**	**.953**	
10. Dynamic Gestures	-.061	**.184**	.046	.095	-.002	**.155**	.482	**.512**	**.433**

Note. Bolded correlations are significant at $p < .010$.
[a]Female/Other is the reference category.

Spatial Interventions and Geometric Reasoning

I first fit a set of models predicting participants' mathematical insights to test the hypothesis that a short Tangram task will improve geometric reasoning as measured by mathematical insight and transformational proof. The results of Model 1 (Table 18) showed that several of the covariates were significantly predictive of insight. There was a small significant gender effect, with females less likely than non-females to produce a mathematical insight (OR = 0.46; $p = .007$). Spatial ability scores had a small but highly reliable association with the relative odds of producing a correct mathematical insight, significantly associated with a 1.68 increase ($p < .000$). Both non-dynamic (OR = 3.13; $p < .000$) and dynamic gestures (OR = 2.64; $p = .003$) had a small significant association with the production of insight. When the Tangram task was added to form Model 2 (Table 18), it was not significantly predictive of insight (OR = 1.40; $p = .238$). These results suggest that the Tangram task did not affect the participants' odds of producing a correct mathematical insight. However, gender (OR = 0.47; $p = .008$), spatial ability (OR = 1.68; $p < .000$), non-dynamic gestures (OR = 3.16; $p < .000$), and dynamic gestures (OR = 2.69; $p = .002$) continued to have a small significant association with insight production.

Table 18

Results of the Logistic Regression Predicting Insight

Variable		***B***	**SE**	**OR**	***p***
Model 0: Null Model					
Random Component: Participant ID variance	1.61		1.27		
Random Component: Conjecture variance	0.31		0.56		
Intercept		0.53	0.29	1.70	.003 **
Model 1: Main Effects					
Random Component: Participant ID variance	0.58		0.76		
Random Component: Conjecture variance	0.70		0.84		
Intercept		-1.22	0.73	0.30	.049 *
Female[a]		-0.78	0.29	0.46	.007 **
Spatial Ability		0.52	0.16	1.68	.000 ***
Spatial Anxiety		-0.01	0.15	0.99	.099
Verbal Working Memory		1.88	1.13	6.55	.097
Visuospatial Working Memory		1.34	0.90	3.82	.138
Non-Dynamic Gestures[b]		1.14	0.29	3.13	.000 ***
Dynamic Gestures[b]		0.97	0.33	2.64	.003 **
Model 3: Main Effects with Intervention					
Random Component: Participant ID variance	0.55		0.74		
Random Component: Conjecture variance	0.71		0.84		
Intercept		-2.99	0.72	0.05	.001 **
Female[a]		-0.76	0.29	0.47	.008 **
Spatial Ability		0.52	0.15	1.68	.000 ***
Spatial Anxiety		-0.01	0.15	0.99	.093
Verbal Working Memory		1.99	1.12	7.32	.076
Visuospatial Working Memory		1.39	0.89	4.01	.120
Non-Dynamic Gestures[b]		1.15	0.29	3.16	.000 ***
Dynamic Gestures[b]		0.99	0.33	2.69	.002 **
Tangram Task[c]		0.34	0.28	1.40	.238

Note. N = 504. OR = odds ratio.
[a]Male/Other is the reference category.
[b]No Gesture/ Gesture of another type is the reference category.
[c]Control Group is the reference category.
* $p \leq .05$, ** $p \leq .010$; *** $p \leq .001$

I then fit a set of models predicting participants' transformational proof. The first model (Model 1; Table 19) showed a small significant association between both spatial ability scores and non-dynamic gestures and transformational proof. Spatial ability scores were associated with an 80% increase in the relative chance of producing a transformation proof (OR = 1.80; p = .003), while the occurrence of at least one non-dynamic gesture was associated with a 2.34 increase in the relative odds of transformational proof production (p = .039). Additionally, students who had higher visuospatial working memory scores (OR = 11.02; p = .034) or produced at least one dynamic gesture (OR = 11.13; p < .000) were over 11 times more likely to produce a transformational proof than students with lower visuospatial scores or who did not produce dynamic gesture. When the Tangram task was added to the form Model 2 (Table 19), it had a significant medium effect on transformational proof production (OR = 4.26; p = .018). This result means that students in the Tangram task condition were over four times more likely to produce a transformational proof than students in the control condition. In this model, spatial ability scores (OR = 1.70; p = .007), visuospatial working memory scores (OR = 10.28; p = .034), non-dynamic gesture (OR = 2.27; p = .047), and dynamic gesture (OR = 10.38; p < .000) were significantly predictive of transformational proof production. Additionally, in models 1 and 2, spatial anxiety was a marginally negative significant predictor of transformational proof (OR = 0.76; p = .053, and OR = 0.77; p = .052, respectively). Participants with higher spatial anxiety scores were less likely to produce a transformational proof than participants with lower spatial anxiety scores, though these associations did not meet the criteria for a small effect size.

The effect of the Tangram task was also evident in students' verbal explanations. Several participants directly referenced the task while explaining why the statement was always true. For example, one participant explaining the *Parallelogram* conjecture first gave a generalized example that any consecutive angles of a parallelogram must be supplementary due to the parallel lines and the alternate angle theorem. The participant then gave a specific example using the Tangram shapes, stating:

> Another way to think of it is like using the shapes. You could have the one parallelogram and then build an identical version with the two triangles and put them next to each other. Then you can see how the angles match and make a straight line or 180 degrees. I guess you could also do it with the square.

Though the mention of the Tangram task was more common in the two-dimensional conjectures, there were a few cases in the *Cylinder* conjecture. For example, during their explanation, one participant remarked that they were "mentally breaking apart the cylinder to view its' pieces like I did with the [Tangram] puzzles." Combined, these quantitative and qualitative results give clear evidence for the hypothesis that a short Tangram task may positively impact the production of transformation proof but not mathematical insight.

Table 19

Results of the Logistic Regression Predicting Transformational Proof

Variable		***B***	**SE**	**OR**	***p***	
Model 0: Null Model						
Random Component: Participant ID variance	2.42		1.55			
Random Component: Conjecture variance	0.09		0.29			
Intercept		-1.93	0.29	0.15	.000	***
Model 1: Main Effects						
Random Component: Participant ID variance	0.49		0.70			
Random Component: Conjecture variance	0.17		0.41			
Intercept		-5.21	0.91	0.01	.000	***
Female[a]		-0.26	0.34	0.77	.451	
Spatial Ability		0.59	0.20	1.80	.003	**
Spatial Anxiety		-0.28	0.18	0.76	.053	
Verbal Working Memory		2.19	1.13	8.96	.061	
Visuospatial Working Memory		2.40	1.13	11.02	.034	*
Non-Dynamic Gestures[b]		0.85	0.41	2.34	.039	*
Dynamic Gestures[b]		2.41	0.36	11.13	.000	***
Model 3: Main Effects with Intervention						
Random Component: Participant ID variance	0.39		0.63			
Random Component: Conjecture variance	0.16		0.40			
Intercept		-5.23	0.89	0.01	.000	***
Female[a]		-0.26	0.33	0.77	.427	
Spatial Ability		0.53	0.20	1.70	.007	**
Spatial Anxiety		-0.26	0.18	0.77	.052	
Verbal Working Memory		1.94	1.11	6.96	.081	
Visuospatial Working Memory		2.33	1.10	10.28	.034	*
Non-Dynamic Gestures[b]		0.82	0.41	2.27	.047	*
Dynamic Gestures[b]		2.34	0.36	10.38	.000	***
Tangram Task[c]		1.45	0.34	4.26	.018	*

Note. N = 504. OR = odds ratio.
[a]Male/Other is the reference category.
[b]No Gesture/ Gesture of another type is the reference category.
[c]Control Group is the reference category.
* $p \leq .05$, ** $p \leq .010$; *** $p \leq .001$

Spatial Interventions and Gesture

In order to test the hypothesis that a short Tangram task will lead to increased gesture production, I fit three models predicting each of the three gesture categories: representational, non-dynamic, and dynamic. In the model predicting representational gestures (Model 1; Table 20), spatial anxiety ($p = .030$) and visuospatial working memory ($p < .000$) were significantly predictive of representational gesture production. Higher spatial anxiety scores were associated with an increase in the relative odds of producing at least one representational gesture of 1.22 (p = .030), though this association did not meet the criteria for a small effect size. Participants with higher visuospatial working memory scores were more than 11 times more likely to produce a representational gesture, suggesting that visuospatial memory has a large association with this type of gesture production. The Tangram task was not significantly predictive in this model (OR = 1.31; p = .131).

Table 20

Results of the Logistic Regression Predicting Representational Gestures

Variable		***B***	**SE**	**OR**	***p***	
Model 0: Null Model						
Random Component: Participant ID variance	2.90		1.70			
Random Component: Conjecture variance	1.72		1.31			
Intercept		0.42	0.57	1.52	.046	*
Model 1: Main Effects						
Random Component: Participant ID variance	2.92		1.51			
Random Component: Conjecture variance	1.71		1.31			
Intercept		-0.40	0.53	0.67	.045	*
Female[a]		-0.08	0.19	0.92	.697	
Spatial Ability		0.17	0.03	1.19	.382	
Spatial Anxiety		0.20	0.01	1.22	.030	*
Verbal Working Memory		0.16	0.61	1.17	.794	
Visuospatial Working Memory		2.41	0.73	11.13	.000	***
Tangram Task[b]		0.27	0.19	1.31	.131	

Note. N = 504. OR = odds ratio.
[a]Male/Other is the reference category.
[b]Control Group is the reference category.
* $p \leq .05$, ** $p \leq .010$; *** $p \leq .001$

The model predicting non-dynamic gestures (Model 1; Table 21) produced similar results to the representational gestures model. Higher spatial anxiety scores were significantly associated with an increase in the relative odds of producing non-dynamic gestures of 1.22 (p = .042). Increased visuospatial working memory scores had a large association with non-dynamic gesture production and were significantly associated with an increase in the relative odds of 9.68 (p = .002). The Tangram task was not significantly predictive in this model (OR = 1.40; p = .075).

Table 21

Results of the Logistic Regression Predicting Non-Dynamic Gestures

Variable		*B*	SE	OR	*p*	
Model 0: Null Model						
Random Component: Participant ID variance	3.14		1.77			
Random Component: Conjecture variance	1.76		1.33			
Intercept		0.29	0.58	1.34	.046	*
Model 1: Main Effects						
Random Component: Participant ID variance	3.15		1.70			
Random Component: Conjecture variance	1.74		1.34			
Intercept		-0.77	0.53	0.46	.014	**
Female[a]		-0.08	0.19	0.92	.696	
Spatial Ability		0.03	0.03	1.03	.228	
Spatial Anxiety		0.20	0.01	1.22	.042	*
Verbal Working Memory		0.37	0.61	1.45	.560	
Visuospatial Working Memory		2.27	0.72	9.68	.002	**
Tangram Task[b]		0.33	0.19	1.40	.075	

Note. N = 504. OR = odds ratio.
[a]Male/Other is the reference category.
[b]Control Group is the reference category.
* $p \leq .05$, ** $p \leq .010$; *** $p \leq .001$

The results for the model predicting dynamic gestures (Model 1, Table 22) showed that dynamic gestures were significantly associated with an increase in visuospatial working memory scores (p = .017). Higher visuospatial working memory scores were associated with an increase in the relative odds of producing at least one dynamic gesture of 28.79, exceeding the threshold for a large effect size. Additionally, the Tangram task condition significantly predicted dynamic gesture production (p = .039). Participants in the Tangram task condition were 2.32 times more likely to produce at least one dynamic gesture than participants in the control condition, meeting the threshold for a small effect size. These results support the hypothesis that Tangram tasks may positively impact participants' gesture production, at least for this type of gesture.

Table 22

Results of the Logistic Regression Predicting Dynamic Gestures

Variable		***B***	**SE**	**OR**	***p***	
Model 0: Null Model						
Random Component: Participant ID variance	2.15		1.47			
Random Component: Conjecture variance	0.67		0.82			
Intercept		-1.64	0.40	0.19	.000	***
Model 1: Main Effects						
Random Component: Participant ID variance	1.64		1.28			
Random Component: Conjecture variance	0.68		0.82			
Intercept		-2.45	1.18	0.08	.037	*
Female[a]		-0.11	0.41	0.90	.794	
Spatial Ability		0.03	0.06	1.03	.563	
Spatial Anxiety		-0.02	0.02	0.98	.397	
Verbal Working Memory		-0.89	1.31	0.41	.499	
Visuospatial Working Memory		3.36	1.41	28.79	.017	*
Tangram Task[b]		0.84	0.41	2.32	.039	*

Note. N = 504. OR = odds ratio.
[a]Male/Other is the reference category.
[b]Control Group is the reference category.
* $p \leq .05$, ** $p \leq .010$; *** $p \leq .001$

Discussion

Spatial systems and gesture production each demonstrate reliable associations with mathematical reasoning. While gestures have often been considered to be an effective malleable factor for improving mathematical reasoning and problem solving, there has been far less research exploring the potential for changes in spatial abilities to alter math performance or to consider how such interventions would impact gesture production. Spatial abilities impact our ability to complete geometric tasks. Due to the malleability of spatial skills, students engaging in geometric thinking may benefit from engaging in spatial activities prior to completing geometric tasks. The central goal of this paper was to investigate the potential of a brief Tangram task for improving students' geometric reasoning and to understand the relationships between geometric reasoning, spatial anxiety, gesture production, and working memory. One additional contribution of this study is to document the role of spatial anxiety in understanding the relationships between geometric reasoning and spatial systems.

To address the first research question of whether a short Tangram task impacts students' geometric reasoning, I looked for the main effects of completing the invention on students' ability to generate correct mathematical insights and transformational proofs. Even when controlling for prior spatial ability, working memory, spatial anxiety, and gesture production, this study found statistically solid support favoring the Tangram task for improving geometric reasoning through the increased production of transformational proof, supporting my hypothesis. However, there was no statistical evidence for the impact of a Tangram task on mathematical insight production. Thus, it appears that the Tangram task does not lead students to be more likely to generate key mathematical insights, but -- most importantly -- it does benefit students' ability to produce mathematically valid proofs in high school level geometry. These transformational proofs require students to move beyond simply stating mathematical insights to incorporating those insights into argumentation that is logical, generalizable, and operational (Harel & Sowder, 2005), which may be more important for mathematics educational curriculum standards).

One explanation for how Tangram tasks directly benefit the production of mathematically valid transformational proofs is that the spatial strategies used in Tangram tasks activities may transfer due to shared domain-general and domain-specific processes. Participants often recruit spatial skills and representations when solving mathematics problems to model, simulate, and manipulate mathematical relations and to ground the meaning of mathematical symbols (Mix et al., 2019; Seron et al., 1992). This study utilized a physical spatial-based task that was likely to engage students' spatial and motoric systems in ways that are congruent with the task demands for geometry reasoning. This prior engagement may allow students to access and employ those systems more readily during the geometric reasoning task, increasing the likelihood that students will successfully construct accurate, transformational geometric proofs.

For example, after practicing spatial visualization right before geometric reasoning, students may transfer spatial visualization strategies such as generating visualization and transforming the geometric objects to facilitate reasoning through their proof, leading to more generalizable and operative thoughts than strictly recalling relevant mathematical theorems and properties. As seen in the qualitative examples, students also incorporated specific examples from the Tangram task into their verbal explanations to demonstrate their more generalized reasoning concretely. Students could be grounding their understandings of geometric properties in their physical experiences with the Tangram shapes. These explanations align with findings from Hawes and colleagues' (2022) suggestions that spatial interventions with concrete materials (such as Tangrams) used with older children may be more effective.

For the second research question addressing whether a short Tangram task impacts students' tendency to gesture, I looked for the main effects of completing the invention on students' ability to generate representational gestures. When controlling for spatial ability, working memory, and spatial anxiety, this study found that students in the Tangram task condition were more likely to produce dynamic gestures (a sub-category of representational gesture that are particularly important for mathematical reasoning) than students in the control condition, supporting my hypothesis. There were no significant differences between the intervention and control conditions in the likelihood that participants would produce at least one non-dynamic or representational gesture.

These results suggest that the spatial and motoric activation of the Tangram task uniquely impacts dynamic production rather than representational gestures. The movements that commonly occur during the task include picking up, rotating, flipping, and sliding the Tangram pieces. These actions are used to transform physical objects. Since dynamic gestures depict transformations on mental representations of objects, students in the spatial condition may be replaying the actions required in the Tangram task as spontaneous actions in the geometric tasks when they align with their mental transformations. As suggested by the model of action-cognition transduction (Nathan, 2014), students in the Tangram task condition may benefit from a body-based experience which is then internalized and incorporated into their situation model of the geometric concepts as they manage the feedforward and feedback processes that monitor and regulate the influences of one's goal-directed actions on the world (Nathan et al., 2014). This updated situation model can then be invoked during future geometric reasoning and displayed through dynamic representational gestures. This study and prior research have shown that increased dynamic gesture production (similar to that described by Garcia & Infante, 2012) is specifically tied to the increased likelihood of transformational proof production in geometric reasoning tasks (Nathan et al., 2021; Schenck et al., 2021; Walkington et al., 2019; Williams-Pierce et al., 2017). Thus, the Tangram task may also indirectly impact proof production by increasing the likelihood of dynamic gesture production.

The third research question was exploratory and investigated the associations between working memory and spatial anxiety and geometric reasoning and gesture production. For working memory, I hypothesized that visuospatial working memory would be associated with both geometric reasoning and gesture production. The results provided evidence in support of this hypothesis. Results indicated that visuospatial working memory, but not verbal working memory, was significantly associated with transformational proof with large odds ratios (OR = 11.02 across conditions; and OR = 10.28, with intervention condition in the model). These odds ratios were even greater than those for spatial ability and non-dynamic gesture and on par with the factors coding for the presence of dynamic gestures. Though the geometry reasoning tasks in this study had both spatial and verbal components, these results confirm that visuospatial working memory may play a uniquely important role in geometric reasoning (Giofrè et al., 2013; Giofrè et al., 2018). It could be that although the task is presented verbally, the use of geometry-specific vocabulary invokes spatial representations needed for spatial visualization as both visuospatial working memory and spatial abilities may overlap as constructs (Miyake et al., 2001).

Visuospatial working memory was also significantly associated with all three gesture categories, again with large odds ratios (OR = 11.13 for representational gesture; OR = 9.68 for non-dynamic gestures; and OR = 28.79 for dynamic gestures). As visuospatial working memory and gesture were both associated with geometric thinking, these results may support the argument that one function of gestures is to active visuospatial working memory resources since visuospatial working memory is thought to support the imagery needed for spatial tasks (Cattaneo et al., 2006; Smithson & Nicoladis, 2014).

A contribution of this study is the incorporation of a spatial anxiety measure to better understand the relationship between mathematical reasoning, spatial systems, gesture production, and working memory. I hypothesized that spatial anxiety would be negatively associated with geometric reasoning and gesture production. Spatial anxiety can impair both spatial abilities and working memory (Ramirez et al., 2012). Spatial anxiety was significantly associated with representational and non-dynamic gesture production. Participants with higher spatial anxiety were more likely to produce representational and non-dynamic gestures. It could be that students are using representational gestures as an offloading mechanism to overcome their increased working load and inability to fully access their spatial abilities. Additionally, spatial anxiety was marginally significantly associated with proof production but not mathematical insight. This result suggests that spatial anxiety may be directly associated with students' ability to produce logical, generalizable, operative proofs but not their ability to articulate mathematical insights. As with gesture, the reduction in performance could be due to spatial anxiety's interference with spatial abilities and working memory.

Limitations and Future Directions

These findings should be interpreted in light of several limitations. First, no post-test measures for spatial ability were included in this study. The spatial posttest was not included due to concerns about the length of time of the intervention. Though a pilot study showed that students had immediate spatial gains on identical measures after the Tangram task, it is possible that there are limits to these gains or that the Tangram task did not increase spatial ability in the participants in this study. Without delayed measures, it is impossible to determine how long any potential gains will last after the Tangram task. Future studies will need to investigate the potential for lasting effects of the short Tangram task or whether longer or repeated interventions provided additional gains in geometric reasoning and gesture production.

Second, the geometry conjectures were selected from a larger set of conjectures used in a separate study that selected conjectures from secondary mathematics textbooks. Furthermore, this study included both two- and three-dimensional properties in the conjectures. It is possible that these decisions could lead to task differences that were not accounted for in the results. For example, three-dimensional geometry concepts are not focused on as deeply as two-dimensional concepts in high school curriculums. Though some undergraduates in this study have had additional geometry curricula that focused on three-dimensional concepts, it is likely that the majority of participants are less familiar with reasoning about the three-dimensional conjectures. Additionally, the Tangram task, primarily a two-dimensional task, may provide more substantial benefits for the two-dimensional conjectures due to task congruency. Future studies will need to account for participants' expertise and explore the impact of conjecture task differences.

Third, in this study, I considered the Tangram puzzle task to be more aligned with a spatial intervention task rather than a geometry knowledge intervention task based on the interviews of pilot participants and prior studies that have specified that Tangram puzzle tasks are spatial interventions (e.g., Cornu et al., 2019; Mix et al., 2020; Siew et al., 2013; Sundberg, 1994). These studies were also included in the meta-analyses by Hawes and colleagues (2022) that categorized the potential for spatial interventions to improve mathematical thinking. Though the pieces within Tangram contain geometric shapes, this study provided no direct instructions to the participants the linked the shapes to specific geometric properties or knowledge. It is possible that in addition to invoking participants' spatial and motor systems, the Tangram task also transferred geometry properties and knowledge to the subsequent geometric reasoning task. Future studies will need to investigate whether tasks such as the Tangram puzzle task impact spatial systems in order to be characterized as a spatial intervention. Studies will also need to investigate which, if any, geometric properties were transfer by students from the Tangram task to the geometric reasoning task.

Finally, the third research question of this study was exploratory. Due to the complicated analyses used in this study and the focus on the spatial intervention questions, I did not complete an *a priori* power analysis for the effects of working memory and spatial anxiety on geometric reasoning and gesture production. It is possible that the analyses performed for this research question are underpowered in this study. Future studies should address this issue by doing a focused study on the role of working memory and spatial anxiety in geometric reasoning and gesture production with a larger sample to thoroughly investigate the claims made in this paper.

Implications and Conclusion

In light of these findings, I consider the implications of this study on emerging theories of spatial-based tasks for improving mathematics and grounded and embodied cognition (GEC). Recent meta-analyses have shown that spatial intervention experiences may be effective for improving mathematics outcomes (Hawes et al., 2022, Yang et al., 2020). These include studies that indicate that even short spatial interventions can be beneficial (e.g., Cheng & Mix, 2014; Hawes et al., 2015). However, the spatial interventions included in these two recent meta-analyses are limited in scope and often focus on elementary-aged participants and paper-based assessments. The randomized controlled experimental design of the current study allows for causal conclusions that strengthen the endorsement for a causal relationship that short spatially based tasks such as Tangram puzzles play in geometric reasoning and gesture production. This study corroborates prior research that spatial interventions can affect mathematical reasoning and extends the prior research by including undergraduate participants completing discourse-based geometric reasoning tasks and gesture production. These results demonstrate that the skills learned or accessed in spatial interventions may transfer to more informal, verbal reasoning tasks often seen in classrooms and used as formative assessments by instructors. This study also added to the limited literature on the role of spatial anxiety in mathematical thinking, though further research is necessary.

This study also contributes to a series of empirical studies on the reciprocal relationships between the mind and body posited by many GEC scholars (Nathan, 2021; Shapiro, 2019). This study reinforces prior studies showing that students' gestures and spatial skills benefit geometric proof production (Nathan et al., 2021; Pier et al., 2019; Walkington et al., 2019). Moreover, this study extends this work by demonstrating that spatial tasks may directly and indirectly impact proof production. The Tangram task led to an increase in the production of dynamic gestures, specifically those indicative of successful proof production. Gestures often convey information about how students think and what they understand that may not be apparent in their speech. Spatial tasks may be a way to increase students' gesture production, supporting their mathematical reasoning and providing teachers insight into that reasoning.

Synthesis of Conclusions Across All Studies

Mathematics education research is expanding away from a focus on traditional formalisms and abstract rules into understanding the ways in which mathematical thinking is grounded in our cognitive, motoric, and perceptual systems (e.g., Nathan et al., 2021). By deepening our understanding of the relationship between spatial anxiety and spatial ability, researchers can develop targeted interventions and educational strategies to address these issues, ultimately fostering skill development and performance improvement for individuals across different age groups and genders. This dissertation aims to deepen the understanding of the cognitive and affective components of the spatial system (i.e., spatial ability and spatial anxiety) and its relation to mathematics and explore the potential of a spatial task to improve success on discourse-based geometry tasks.

Summary of Findings

Chapter 1 first reviews the many spatial ability taxonomies and spatial tasks created through decades of research. This review illuminated several challenges in understanding the relationship between spatial ability and mathematics. One, the spatial taxonomies that describe spatial skills are often not based on well evidenced theoretical and analytical frameworks. Two, the sheer variety of spatial tasks and assessments are often untethered to evidence based spatial taxonomies. Three, there are critical gaps in spatial tasks and assessments, including a pressing need for tasks explicitly designed for a unitary spatial taxonomy, extrinsic-static classifications of spatial ability, and for use with children. These challenges can impede progress in designing effective spatial ability interventions for improving mathematics and learning, selecting appropriate metrics for documenting change, and analyzing and interpreting student outcome data. In light of these gaps, the chapter concludes by offering a practice guide for navigating and selecting among the various spatial taxonomies and spatial tasks for use in mathematics education research.

In Chapter 2, I presented two studies that explored the role of spatial ability and spatial anxiety in mathematical thinking. Study 1 provided evidence from adults that spatial ability and spatial anxiety is associated with performance on standardized mathematics outcomes. Evidence corroborated findings that spatial ability and mathematics ability are highly correlated (Davis, 2015; Uttal & Cohen, 2012). Moreover, specific spatial ability sub-categories were associated differently with various mathematic tasks, providing a first step for identifying particular relationships between spatial ability sub-categories and mathematics domains. Finally, the results showed that high spatial anxiety scores predict lower spatial and mathematics ability scores. Study 2 partially built on this work by investigating the role of grounded embodied learning processes, such as spatial ability and spatial anxiety, in children's symbolic fraction knowledge. Like in study 1, particular sub-categories of spatial ability (i.e., mental rotation and spatial visualization) were associated with fraction knowledge scores. Additionally, spatial anxiety appeared to moderate the relationships between spatial ability and fraction knowledge. These two studies demonstrated that distinct spatial ability sub-categories may be associated with various mathematical domains and that spatial anxiety impair the relationship between spatial ability and mathematics ability.

Chapter 3 expands the work on spatial ability, spatial anxiety, and mathematics to a discourse-based geometry task and gesture production. This study's findings built on Chapter 2 to provide correlational evidence that spatial anxiety may moderate the relationship between spatial ability and geometric thinking and between spatial ability and representational gesture production. When spatial anxiety decreased, the effect of spatial ability on the likelihood of mathematical insight and transformational proof production increased. Additionally, students with higher spatial anxiety scores and lower spatial ability scores were more likely to produce a representational gesture. In this study, I demonstrate the links between spatial ability, spatial anxiety, and gesture production in an area of mathematics that was previously unexplored.

Finally, Chapter 4 aimed to test whether a short Tangram task would improve geometric thinking outcomes and impact gesture production using guidance from the framework. Evidence from this study showed that not only does a Tangram task directly improve geometric thinking through the increased likelihood of producing a transformational proof, but it may also indirectly improve geometric thinking by increasing the likelihood of dynamic gesture. Though there were no delayed measures in this study, it is encouraging to see that a Tangram task may at least have an immediate impact on geometric thinking. This study built on the work in Chapters 2 and 3 to further explore the role of spatial anxiety and working memory in geometric thinking and gesture production. Findings showed that, like in previous studies, participants with higher spatial anxiety were less likely to produce transformational proofs but more likely to make representational gestures. For working memory, visuospatial working was significantly associated with transformational proof and gesture, with large odds ratios. These results demonstrate that spatial anxiety and visuospatial working memory may be additional considerations in designing spatial interventions and mathematics education research.

The studies confirm that spatial ability is essential for various mathematics domains, including more formalized and discourse-based mathematics tasks. Additionally, two of the studies provide foundational evidence that particular sub-categories of spatial ability may be tied explicitly to different mathematical domains. Across all four studies, findings suggest that spatial anxiety may influence mathematic outcomes, possibly moderating the relationship between spatial ability and mathematical thinking. It may also play a role in gesture production, a significant predictor of geometric thinking in several prior studies (e.g., Nathan et al., 2021; Pier et al., 2019; Walkington et al., 2019). Together, the review of the literature and studies presented in this dissertation demonstrate the interconnectedness of spatial systems, gesture production, and mathematical thinking.

Limitations and Future Directions

Several factors limit the findings of these studies. First, the majority of the studies in this dissertation are exploratory, and two studies use existing data from more extensive studies. As such, there may be power issues with the sample sizes, and I am unable to make strong generalizable claims. Future work with larger samples, *a prior* power analyses, and studies designed to test the specific claims can focus on quantitative data collection for generalizable claims. Additional studies with larger sample sizes are especially vital for understanding the role of spatial anxiety, as the effect sizes seen in this study rarely meet the threshold for a small effect. A larger sample size may be able to more robustly measure spatial anxiety's actual associations with spatial and mathematic ability.

Second, all the studies used spatial anxiety as a variable of interest. There is relatively limited literature on spatial anxiety and its cognitive impacts, especially in mathematics. Additionally, very few measures of spatial anxiety exist. In fact, to the best of my knowledge, only one measure is appropriate for use with adults and includes spatial anxiety factors other than navigational anxiety. Thus, these studies are limited by this lack of research. Future studies should focus on refining measurement tools and exploring the contexts in which spatial anxiety manifests.

Third, three studies were limited to adults due to limited access to school-age populations during the COVID-19 pandemic. Though this population was appropriate for the goals of this dissertation and for providing foundational knowledge of the relationships between spatial systems, gesture production, and mathematics, it does limit the practical applications of this work. In future work, studies should explore other age groups to understand how these relationships occur across development. A longitudinal study may also inform structural models of the long-term impacts of spatial ability and spatial anxiety on mathematics achievement. In a similar vein, all studies were completed in a laboratory setting. To better understand the role of spatial systems in mathematics and the potential of spatial tasks for improving mathematical thinking, future work should include classroom-based research to explore the benefit of integrating spatial tasks into existing mathematics curricula and reveal possible lasting effects of spatial training in authentic settings.

I also plan to continue to develop and test spatial interventions to facilitate learning in STEM contexts. Specifically, I am interested in exploring the affordances of collaboration and immersive technologies such as virtual, augmented, and mixed reality. The few studies comparing the use of manipulatives such as Tangrams in virtual reality and physical settings have shown an increase in collaborative problem solving in the virtual reality groups but more opportunities for mirroring gestural imagery in the physical group due to the direct manipulation of the objects (Evan et al., 2011). This work suggests that virtual reality may have affordances for increasing collaborations at the expense of opportunities for students to make and mirror relevant gestures. However, the technology used in this study required a handheld controller, limiting the movements that could be made in the virtual reality condition. To the best of my knowledge, no studies have investigated collaboration or immersive technologies that do not require handheld controllers for spatial interventions. These technologies provide high motoric engagement, gestural congruency, and immersion (Johnson-Glenberg et al., 2014) and may facilitate mathematics learning by increasing engagement (Buentello-Montoya et al., 2021).

Implications for Mathematics Education Research

Given these limitations, the work presented in this dissertation has several implications for spatial ability, grounded and embodied learning, and mathematics education. Other researchers can utilize this work to support future work on the intersection of spatial ability and mathematics. Chapter 1 illuminates many limitations of the current spatial ability literature that could inspire future work in the field, including developing unitary spatial ability measures. It also provides a practice guide for researchers to choose appropriate spatial taxonomies, analytic frameworks, and spatial tasks. This guide could make incorporating spatial ability into mathematics education research more approachable for researchers, especially those who are not as familiar with the decades of work in this area. Chapter 4 provides an example of how this guide could be used in research on spatial interventions for increasing mathematics outcomes. The studies in Chapter 2 identified specific spatial ability sub-categories distinctly associated with different mathematics domains. In line with the guide's recommendations, the findings in Chapter 2 could be used as a basis for selecting a specific-factor spatial taxonomy.

These findings also contribute to the theoretical understanding of how spatial systems (i.e., spatial ability and spatial anxiety) are associated with gesture production and mathematical thinking. Though decades of research have demonstrated a connection between spatial ability and mathematics, this work extends the existing literature in three ways. First, the studies in

Chapter 2 provide evidence that different mathematical tasks may rely on specific spatial sub-categories. Though other studies have focused on the role of particular spatial sub-categories in mathematics, such as mental rotation (e.g., Hawes et al., 2015; Thompson et al., 2013), these studies often do not include other tasks for other spatial sub-categories, limiting their use as a basis for other studies wishing to identify associations between mathematics and spatial sub-categories. While spatial abilities are not explicitly taught in classrooms, there is a growing consensus that these skills should be included in STEM curricula (Gilligan-Lee et al., 2022). This work may provide the basis for designing curricula to target specific spatial skills relevant to mathematics. Second, Chapter 3 extends the work on spatial ability and mathematics into discourse-based informal tasks. Tasks such as these are relatively underrepresented in spatial ability literature. As these tasks often occur in mathematics classrooms and are essential for students to build mathematical argumentation, researchers must continue exploring the cognitive and affective components that affect these tasks. Third, Chapter 4 adds to the limited work on spatial tasks for mathematical thinking, especially in discourse-based tasks. Research on spatial training and interventions suggests that short interventions can improve spatial ability (Uttal et al., 2013). The results of this study suggest that spatially-involved tasks, such as a Tangram puzzle task, may extend to improved geometric thinking and increased gesture production.

This work informs the limited work on spatial anxiety. Spatial anxiety is a multifaceted and complex construct that significantly impacts neurocognitive performance (Lauer et al., 2018; Lyons et al., 2018) and is negatively correlated with spatial ability (Malanchini, 2017). This research continues to support these theories. Moreover, the findings in this work indicate that spatial anxiety affects mathematical thinking and gesture production in children and adults.

Though the work here is largely exploratory, the contributions provide a basis for understanding spatial anxiety beyond the contexts of navigation and its association with cognitive processes.

Finally, these studies also help extend the theories of grounded and embodied learning (GEL) and gesture. While GEL theories typically focus on observable behaviors, including gesture and language, recent frameworks, such as the Grounded and Embodied Learning Timescales (Nathan, 2021), have expanded to include unobservable behaviors, including affective and cognitive processes. The studies apply this expansion by positing that spatial ability and spatial anxiety are inherently embodied processes that help individuals ground their mathematical knowledge. Furthermore, there is evidence that spatial anxiety moderates the relationship between spatial ability and mathematics outcomes and that individuals with higher anxiety levels may use gesture as a cognitive offloading mechanism. These studies also add to the gesture literature by providing evidence that spatial anxiety may impact gesture production by modifying an individual's gesture threshold (Hostetter & Alibali, 2019). Thus, it is beneficial to include spatial ability and spatial anxiety in embodied processes to understand their role in grounding students' learning.

Conclusion

The work in this dissertation explores the connections between spatial systems, gesture production, and mathematics. My findings inform future mathematics educational researchers and instructional practices with spatial ability and spatial anxiety using a grounded and embodied perspective. I propose that these constructs are interrelated, and it is critical to continue conducting research that examines them simultaneously. Spatial ability and spatial anxiety impact various domains of mathematics, and it may be possible to develop spatial interventions that can target mathematics performance, including discourse-based tasks and gesture production during these tasks.

References

Abrahamson, D., & Bakker, A. (2016). Making sense of movement in embodied design for mathematics learning. *Cognitive Research: Principles and Implications,1(*1), 1–13.

Adams, J., Resnick, I., & Lowrie, T. (2022). Supporting senior high-school students' measurement and geometry performance: Does spatial training transfer to mathematics achievement? *Mathematics Education Research Journal*.

Alibali, M. W. (2005). Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation, 5*(4), 307 – 331.

Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences*, 21(2), 247–286.

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Sciences*, *113*(18), 4909-4917.

Ashcraft, M. H., & Krause, J. A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review, 14,* 243 – 248.

Atit, K., Power, J. R., Pigott, T., Lee, J., Geer, E. A., Uttal, D. H., Ganley, C. M., & Sorby, S. A. (2021). Examining the relations between spatial skills and mathematical performance: A meta-analysis. *Psychonomic Bulletin & Review.*

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Science, 4*, 417-423.

Baddeley, A. D. & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). Academic Press.

Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles, 20*, 327 – 344.

Banich, M. T., & Compton, R. J. (2018). Cognitive neuroscience. Cambridge University Press.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, *22*(4), 577-660.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, 59(1), 617–645.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4.* R package version 1.1-7. http://CRAN.R-project.org/package=lme4.

Battista, M. T. (1990). Spatial visualization and gender differences in high school geometry. *Journal for Research in Mathematics Education, 21*(1), 47-60.

Battista, M. T. (2007). The development of geometric and spatial thinking. In F. K. Lester Jr. (Ed.), Second handbook of research on mathematics teaching and learning (pp. 843-908). Charlotte, NC: Information Age Publishing

Battista, M.T., Frazee, L. M., & Winer, M. L. (2018). Analyzing the relation between spatial and geometric reasoning for elementary and middle school students. In K. S. Mix & M. T. Battista (Eds.), *Visualizing Mathematics. Research in Mathematics Education* (pp. 195 – 228). Springer, Cham.

Beilock, S. L. (2010). *Choke: What the secrets of the brain reveal about getting it right when you have to*. Free Press, Simon & Schuster.

Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science, 16*(2), 101 – 105.

Beilock, S. L. & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory & Cognition, 33*(6), 983–998.

Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science, 21*(11), 1605–1610.

Berch, D. B., Foley, E. J., Hill, R. J., & Ryan, P. M. (1999). Extracting parity and magnitude from Arabic numerals: Developmental changes in number processing and mental representation. Journal of Experimental Child Psychology, 74, 286–308.

Blacker, K. J., Weisberg, S. M., Newcombe, N. S., & Courtney, S. M. (2017). Keeping Track of Where We Are: Spatial Working Memory in Navigation. *Visual Cognition*, *25*(7-8), 691–702.

Beilock, S. L. & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory & Cognition, 33*(6), 983–998.

Boonen, A. J. H., van der Schoot, M., van Wesel, F., de Vries, M. H. & Jolles, J. (2013). What underlies successful world problem solving? A path analysis in sixth grade students. *Contemporary Educational Psychology, 38*, 271 – 279.

Borst, G., Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2011). Representations in mental imagery and working memory: Evidence from different types of visual masks. *Memory & Cognition*, *40*(2), 204-217.

Bower, C., Zimmermann, L., Verdine, B., Toub, T. S., Islam, S., Foster, L., Evans, N., Odean, R., Cibischino, A., Pritulsky, C., Hirsh-Pasek, K., & Golinkoff, R. M. (2020). Piecing together the role of a spatial assembly intervention in preschoolers' spatial and mathematics learning: Influences of gesture, spatial language, and socioeconomic status. *Developmental Psychology*, *56*(4), 686–698.

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

Buckley, J., Seery, N. & Canty, D. (2018) A heuristic framework of spatial ability: a review and synthesis of spatial factor literature to support its translation into STEM education. *Educational Psychology Review, 30*, 947–972.

Buentello-Montoya, D. A., Lomelí-Plascencia, M. G., & Medina-Herrera, L. M. (2021). The role of reality enhancing technologies in teaching and learning of mathematics. *Computers and Electrical Engineering, 94*, 107287.

Bull, R. & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, identification, and short-term memory. *Journal of Experimental Child Psychology, 65*.

Bull, R., Johnston, R. S. & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology, 15*(3), 421-442.

Burte, H., Gardony, A. L., Hutton, A., & Taylor, H. A. (2017). Think3d!: Improving mathematical learning through embodied spatial training. *Cognitive Research: Principles and Implications, 2*(13).

Burte, H., Gardony, A. L., Hutton, A., & Taylor, H. A. (2019a). Knowing when to fold 'em: Problem attributes and strategy differences in the Paper Folding Test. *Personality and Individual Differences, 146*, 171 – 181.

Burte, H., Gardony, A. L., Hutton, A., & Taylor, H. A. (2019b). Make-A-Dice test: Assessing the intersection of mathematical and spatial thinking. *Behavior Research Methods, 51*(2), 602-638.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W., & Thulborn, K. (1999). Graded functional activation in the visuospatial system with the amount of task demand. *Journal of Cognitive Neuroscience, 11*(1), 9–24.

Carr, M., Steiner, H. H., Kyser, B., and Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences, 18*(1), 61–75.

Carroll, J. B., (1993). *Human cognitive abilities: A survey of factor-analytic studies*. University of Cambridge Press.

Case, R., Okamoto, Y., Griffin, S., McKeough, A., Bleiker, C., Henderson, B., Stephenson, K. M., Siegler, R. S., & Keating, D. P. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development, 61*(1/2), i-295.

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

Casey, M. B., Nuttall, R. L., & Pezaris, E. (1999). Evidence in support of a model that predicts how biological and environmental factors interact to influence spatial skills. *Developmental Psychology, 35*(5), 1237-1247.

Casey, M. B., Pezaris, E., Fineman, B., Pollock, A., Demers, L., & Dearing, E. (2015). A longitudinal analysis of early spatial skills compared to arithmetic and verbal skills as predictors of fifth-grade girls' math reasoning. *Learning and Individual Differences, 40*, 90-100.

Cattaneo, Z., Fastame, M. C., Vecchi, T., & Cornoldi, C. (2006). Working memory, imagery and visuo-spatial mechanisms. In T. Vecchi & G. Bottini (Eds.), *Imagery and spatial cognition: Methods, models and cognitive assessment* (pp. 101 – 137). John Benjamins Publishing Company.

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics – Simulation and Computation, 39*(4), 860 – 864.

Cheng, Y. L. (2016). The improvement of spatial ability and its relation to spatial training. In M. Khine (Ed.). *Visual-spatial ability in STEM education* (pp. 143 –172). Springer.

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

Chu, M. & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General, 137*, 706 – 723.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, *140*(1), 102–116.

Cohen, C. A., & Hegarty, M. (2012). Inferring cross sections of 3D objects: A new spatial thinking test. *Learning and Individual Differences, 22*(6), 868 -874.

College Entrance Examination Board. (CEEB, 1939). *Spatial Aptitude Test in Spatial Relationships.*

Colom, R., Contreras, M.J., Botella, J., & Santacreu, J. (2001). Vehicles of spatial ability. *Personality and Individual Differences, 32*(5), 903-912.

Cobb, P. (1994). Where Is the Mind? Constructivist and Sociocultural Perspectives on Mathematical Development. *Educational Researcher, 23*(7), 13–20.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O. & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769 - 786.

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten memory load when explaining math. *Language and Cognitive Processes, 27*(4), 594 – 610.

Cooper, L. A. (1975). Mental rotation of random two-dimensional shapes. *Cognitive Psychology, 7*(1), 20-43.

Cooper, L. A., & Mumaw, R. J. (1985). Spatial aptitude. In R. F. Dillman (Ed.). *Individual differences in cognition* (2nd. Ed., pp.67-94). Academic Press.

Cornu, V., Hornung, C., Schiltz, C., & Martin, R. (2017). How do different aspects of spatial skills relate to early arithmetic and number line estimation? *Journal of Numerical Cognition, 3*(2).

da Costa, R., Pompeu, J. E., de Mello, D. D., Moretto, E., Rodrigues, F. Z., Dos Santos, M. D., Nitrini, R., Morganti, F., & Brucki, S. (2018). Two new virtual reality tasks for the assessment of spatial orientation Preliminary results of tolerability, sense of presence and usability. *Dementia & Neuropsychologia*, *12*(2), 196–204.

Davis, B. (Ed.). (2015*). Spatial reasoning in the early years: Principles, assertions, and speculations*. Routledge.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371.

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

Derakshan, N., & Eysenck, M. W. (2010). Introduction to the special issue: Emotional states, attention, and working memory. *Cognition and Emotion*, *24*(2), 189–199.

DeWolf, M. Bassok, M., & Holyoak, K. J. (2015). From rational numbers to algebra: Separable contributions of decimal magnitude and relational understanding of fractions. *Journal of Experimental Child Psychology, 133*, 72 – 84.

de Villiers, M. (1998). An Alternative Approach to Proof in Dynamic Geometry. In R. Lehrer & D. Chazan (Eds.), *Designing Learning Environments for developing Understanding of Geometry and Space* (pp. 369–393). Mahwah, NJ: Lawrence Erlbaum Associates.

Đokić R, Koso-Drljević M, Đapo N (2018) Working memory span tasks: Group administration and omitting accuracy criterion do not change metric characteristics. PLOS ONE 13(10): e0205169.

D'Oliveira, T. (2004). Dynamic spatial ability: An exploratory analysis and a confirmatory study. *The International Journal of Aviation Psychology, 14*(1), 19–38.

Eagan, B., Rogers, B., Serlin, R., Ruis, A. R., Arastoopour Irgens, G., & Shaffer, D. W. (2017, June). Can we rely on IRR? Testing the assumptions of inter-rater reliability. [Paper Presentation] In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.), *Proceedings of 12th International Conference on Computer Supported Learning,* Vol. 2. Philadelphia, PA.

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

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19 -23.

Estes, D. (1998). Young children's awareness of their mental activity. The case of mental rotation. *Child Development, 69* (5), 1345–1360.

Evans, M. A., Feenstra, E., Ryon, E., & McNeill, D. (2011). A multimodal approach to coding discourse: Collaboration, distributed cognition, and geometric reasoning. *International Journal of Computer-Supported Collaborative Learning, 6,* 253 – 278.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336-353.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149-1160.

Ferguson, A. M., Maloney, E. A., Fugelsang, J., & Risko, E. F. (2015). On the relation between math and spatial ability: The case of math anxiety. *Learning and Individual Differences, 39,* 1-12.

French, J. W. (1951). *The description of aptitude and achievement tests in terms of rotated factors.* University of Chicago Press.

Frick, A., Hanson, M.A., & Newcombe, N. S. (2014) Development of mental rotation in 3- to 5-year-old children. *Cognitive Development, 28*(4), 386-399.

Frick, A, Möhring, W. and Newcombe, N. S. (2014). Picturing Perspectives: Development of Perspective-Taking Abilities in 4- to 8-Year-Olds. *Frontiers in Psychology, 5,* 386.

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., Jordan, N. C., Siegler, R., Gersten, R., & Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology*, *105*, 683 – 700.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*(1/2), 43–74.

Galton, F. (1879). Generic images. *The Nineteenth Century*, *6*(1), 157–169.

Garcia, N., & Infante, N. E. (2012). Gestures as facilitators to proficient mental modelers. In L. R. Van Zoest, J.-J. Lo, & J. L. Kratky (Eds.), *Proceedings of the 34th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 289–295). Kalamazoo, Michigan.

Gaughran, W. (2002). *Cognitive modelling for engineers* [Paper Presentation]. American Society for Engineering Education annual conference and exposition, Montréal, Canada.

Geary, D. C., Saults, S. J., Liu, F., and Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology*, *77*(4), 337 – 353.

Gilligan, K. A., Thomas, M. S. C., & 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).

Gilligan, K. A., Thomas, M. S., & Farran, E. K. (2020). 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.

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*(1), 1-4.

Giofrè, D., Donolato, E., & Mammarella, I. C. (2018). The differential role of verbal and visuospatial working memory in mathematics and reading. *Trends in Neuroscience and Education*, *12*, 1-6.

Giofrè, D., Mammarella, I. C., Ronconi, L., & Cornoldi, C. (2013). Visuospatial working memory in intuitive geometry, and in academic achievement in geometry. *Learning and Individual Differences*, *23*, 114-122.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*(1), 1-26.

Glenberg, A., Willford, J., Gibson, B., Goldberg, A., & Zhu, X. (2012). Improving reading to improve math. *Scientific Studies of Reading*, *16*(4), 316-340.

Göksun, T., Goldin-Meadow, S., Newcombe, N. & Shipley, T. (2013). Individual differences in mental rotation: What does gesture tell us? *Cognitive Processing, 14*(2), 153 – 162.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Harvard University Press.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological science, 20*(3), 267-272.

Gottschaldt, K. (1926). Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren. *Psychologische Forschung, 8*, 261-318.

Guay, R. B. (1976). *Purdue Spatial Visualization Test*. Purdue Research Foundation.

Guilford, J. P. & Lacey, J. I. (Eds) (1947). *Printed Classification Tests* (No. 5). U.S. Government Printing Office.

Guilford, J. P. & Zimmerman, W. S. (1948). The Guilford-Zimmerman Aptitude Survey. *Journal of Applied Psychology, 32*(1), 24–34.

Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Developmental Psychology, 48*(5), 1229–1241.

Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2013). Teachers' spatial anxiety relates to 1st- and 2nd graders' spatial learning. *Mind, Brain, and Education, 7*(3), 196-199.

Guttman, R., Epstein, E. E., Amir, M., & Guttman, L. (1990). A structural theory of spatial abilities. *Applied Psychological Measurement, 14*(3), 217 – 236.

Hanna, G. (2000). Proof, explanation, and exploration: An overview. *Educational Studies in Mathematics, 44*(1), 5-23.

Hannafin, R. D., Truxaw, M. P., Vermillion, J. R., & Liu, Y. (2008). Effects of spatial ability and instructional program on geometry achievement. *Journal of Educational Research, 101*(3), 148–157.

Hansen, N., Jordan, N. C., & Rodrigues, J. (2017). Identifying learning difficulties with fractions: a Longitudinal study of student growth from third through sixth grade. *Contemporary Educational Psychology, 50*, 45 – 59.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. *Research in collegiate mathematics education III*, 234-283.

Harel, G., & Sowder, L. (2005). Toward comprehensive perspectives on the learning and teaching of proof. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 805-842). Reston, VA: National Council of Teachers of Mathematics.

Harris, J., Hirsh-Pasek, K., & Newcombe, N. S. (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.

Hartje, W. (1987). The effect of spatial disorders on arithmetical skills. In G. Deloche & X. Seron (Eds.), *Mathematical disabilities A cognitive neuropsychological perspective.* (pp. 121 - 135). Lawrence Erlbaum Associates.

Hawes, Z., & Ansari, D. (2020). What explains the relationship between spatial and mathematical skills? A review of evidence from brain and behavior. *Psychonomic Bulletin & Review, 27*, 465 – 482.

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.

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.

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.

Hayes, J. R. (1973). On the function of visual imagery in elementary mathematics. In W. G. Chase (Ed.), *Visual information processing* (pp. 177-214). Academic.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428.

Heathcote, D. (1994). The role of visuo-spatial working memory in the mental addition of multi~ digit addends. *Current Psychology of Cognition, 13*, 207-245.

Hegarty, M. (2018). Ability and sex differences in spatial thinking: What does the mental rotation test really measure? *Psychonomic Bulletin & Review, 25,* 1212-1219.

Hegarty, M., Burte, H., & Boone, A. P. (2018). *Individual differences in large-scale spatial abilities and strategies.* In D. R. Montello (Ed.), *Handbook of behavioral and cognitive geography* (p. 231–246). Edward Elgar Publishing.

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

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial layout learning. *Intelligence, 34*(2), 151–176. https://doi.org/10.1016/j.intell.2005.09.005

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence, 30*, 425–447.

Hegarty, M., Shah, P. & Miyake, A. (2000). Constraints on using the dual-task methodology to specify the degree of central executive involvement in cognitive tasks. *Memory and Cognition, 28*, 376 – 385.

Hegarty, M. & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence, 32*(2), 175 – 191.

Hegarty, M. & Waller, D. A. (2005). Individual differences in spatial abilities. In P. Shah, & A. Miyake (Eds.), *The Cambridge Handbook of Visuospatial Thinking* (pp. 121-169). Cambridge University Press.

Herbst, P., & Brach, C. (2006). Proving and doing proofs in high school geometry classes: What is it that is going on for students? *Cognition and Instruction, 24*(1), 73-122.

Hill, F., Mammarella, I. C., Devine, A., Caviola, S., Passolunghi, M. C., & Szűcs, D. (2016). Maths anxiety in primary and secondary school students: Gender differences, developmental changes and anxiety specificity. *Learning and Individual Differences, 48*, 45 – 53.

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture, 7*(1), 73 – 95.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review, 15*(3), 495 – 514.

Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action: Revisiting the framework. *Psychonomic Bulletin & Review, 26*, 721 – 752.

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6*(6), 435–448.

Hund, A. M. & Minarik, J. L. (2006). Getting From Here to There: Spatial Anxiety, Wayfinding Strategies, Direction Type, and Wayfinding Efficiency. *Spatial Cognition & Computation,* 6, 179–201.

Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking: From childhood to adolescence.* Basic Books.

Izard, V., Pica, P., Spelke, E.S., & Dehaene, S. (2011). Flexible intuitions of Euclidean geometry in an Amazonian indigene group. *Proceedings of the National Academy of Sciences, 108*(24), 9782 – 9787.

Jansen, P. (2009). The dissociation of small-and large-scale spatial abilities in school-age children. *Perceptual and motor skills, 109*(2), 357-361.

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology, 106(*1), 86 – 104.

Jolicoeur, P., Regehr, S., Smith, L. B. J. P., Smith, G. N. (1985). Mental rotation of representations of two-dimensional and three-dimensional objects. *Canadian Journal of Psychology, 39*(1), 100-129.

Jones, K. (2000). Providing a foundation for deductive reasoning: Students' interpretations when using dynamic geometry software and their evolving mathematical explanations. *Educational Studies in Mathematics, 44*(1), 55-85

Judd, N., & Klingberg, T. (2021). Training spatial cognition enhances mathematical learning in a randomized study of 17,000 children. *Nature Human Behaviour, 5*(11), 1548-1554.

Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, *92*(2), 137.

Karp, S. A., & Konstadt, N. L. (1963). *Manual for the Children's Embedded Figures Test.* Cognitive Tests.

Kim, S., Jiang, Y., & Song, J. (2015). The effects of interest and utility value on mathematics engagement and achievement. In K. A. Renninger, M. Nieswandt, & S. Hidi (Eds.), *Interest in mathematics and science learning* (pp. 63-78). Washington, DC: American Educational Research Association.

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review, 124*(3)., 245 – 266.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48,* 16-32.

Knuth, E. J. (2002). Secondary school mathematics teachers' conceptions of proof. *Journal for Research in Mathematics Education, 33,* 379-405.

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students’ production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K–16 perspective* (pp. 153–170). Routledge.

Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 723 – 735.

Kozhevnikov, M. & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition, 29*(5), 745-756.

Kremmyda, O., Hüfner, K., Flanagin, V. L., Hamilton, D. A., Linn, J., Strupp, M., Jahn, K. &Brandt, T. (2016). Beyond dizziness: Virtual navigation, spatial anxiety and hippocampal volume in bilateral vestibulopathy. *Frontiers in Human Neuroscience, 10*.

Krüger, M. (2018). Mental rotation and the human body: Children's inflexible use of embodiment mirrors that of adults. *British Journal of Developmental Psychology, 23(*3),

Krüger, M., Kaiser, M., Mahler, K., Bartels, W., & Krist, H. (2013). Analogue mental transformations in 3-year-olds: Introducing a new mental rotation paradigm suitable for young children*. Infant and Child Development, 23,* 123-138.

Kyllonen, P. C. (1996). Is working memory capacity Spearman's g? In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 49-75). Erlbaum.

Kyritsis, M. & Gulliver, S. R. (2009). *Guilford-Zimmerman orientation survey: A validation* [Paper Presentation]. The 7th International Conference on Information, Communications and Signal Processing, Macau.

Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J. & Hautamäki, J. (2003). Visuospatial working memory and early numeracy. *Educational and Child Psychology, 20*(3)*,* 65–76.

Kyttälä, M. & Lehto, J. E. (2008). Some factors underlying mathematical performance: the role of visuospatial working memory and non-verbal intelligence. *European Journal of Psychology of Education, 23*(77).

Lakoff, G. & Nuñez, R. (2000). *Where Mathematics Come From*. Basic Books.

Laski, E. V., Casey, B. M., Yu, Q., Dulaney, A., Heyman, M. & Dearing, E. (2013). Spatial skills as a predictor of first grade girls' use of higher level arithmetic strategies. *Learning and Individual Differences*, *23*, 123–130.

Lauer, J. E., Esposito, A. G., & Bauer, P. J. (2018). Domain-specific anxiety relates to children's math and spatial performance. *Developmental Psychology, 54,* 2126 – 2138.

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex roles,* 30(11-12), 765-779.

Lee, S.A., Sovrano, V.A., & Spelke, E.S. (2012). Navigation as a source of geometric knowledge: Young children's use of length, angle, distance, and direction in a reorientation task. *Cognition, 1*, 144 – 161.

Legesse, M., Luneta, K., Ejigu, T. (2020). Analyzing the effects of mathematical discourse-based instruction on eleventh-grade students' procedural and conceptual understanding of probability and statistics. *Studies in Educational Evaluation, 67*, 100918.

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

Lohman, D. F. (1979). *Spatial ability: A review and re-analysis of the correlational literature (*Technical Report No. 8). Stanford, CA: Aptitudes Research Project, School of Education, Stanford University.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 181–248). Lawrence Erlbaum.

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

Lowrie, T., Logan, T. & Hegarty, M. (2019). The influence of spatial visualization training on students' spatial reasoning and mathematics performance. *Journal of Cognition and Development, 20*(5), 729 – 751.

Lowrie, T., Resnick, I., Harris, D., & Logan, T. (2020) In the search of the mechanisms that enable transfer from spatial reasoning to mathematics understanding. *Mathematics Education Research Journal, 32*, 175 -188.

Lukowski, S. L., DiTrapani, J., Jeon, M., Wang, Z., Schenker, V. J., Doran, M. M., . . . Petrill, S. A. (2019). Multidimensionality in the measurement of math-specific anxiety and its relationship with mathematical performance. *Learning and Individual Differences*, *70*, 228-235.

Lütke, N. & Lange-Küttner, C. (2015). Keeping it in three dimensions: Measuring the development of mental rotation in children with the Rotated Colour Cube Test (RCCT). *International Journal of Developmental Science, 9*(2), 95-114.

Lyons, I. M., Beilock, S. L. (2011). Mathematics anxiety: Separating the math from the anxiety. *Cerebral Cortex*, 22, 2102–2110.

Lyons, I. M., Ramirez, G., Maloney, E. A., Rendina, D. N., Levine, S. C., & Beilock, S. L. (2018). Spatial anxiety: a Novel questionnaire with subscales for measuring three aspects of spatial anxiety. *Journal of Numerical Cognition, 4,* 526-553.

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education, 30,* 520 – 540.

Malanchini, M., Rimfeld, K., Shakeshaft, N. G., McMilan, A., Schofield, K. L., Rodic, M., Rossi, V., Kovas, Y., Dale, P. S., Tucker-Drob, E. M., & Plomin, R. (2020). Evidence for a unitary structure of spatial cognition beyond general intelligence. *Science of Learning*, *5*(9).

Malanchini, M., Rimfeld, K., Shakeshaft, N. G., Rodic, M., Schofield, K., Selzam, S., . . . Kovas, Y. (2017). The genetic and environmental aetiology of spatial, mathematics and general anxiety. *Scientific reports*, *7,* 42218.

Marghetis, T., Edwards, L. D., & Núñez, R. (2014). More than mere handwaving. *Emerging perspectives on gesture and embodiment in mathematics*, 227-246.

Marstaller, L., & Burianová, H. (2013). Individual differences in the gesture effect on working memory. *Psychonomic Bulletin & Review, 20*, 496 - 500.

MacLeod, C., & Donnellan, A. M. (1993). Individual differences in anxiety and the restriction of working memory capacity. *Personality and Individual Differences*, *15*(2), 163–173.

Maeda, Y. & Yoon, S. Y. (2013). A meta-analysis on gender differences in mental rotation ability measured by the Purdue spatial visualization tests: Visualization of rotations (PSVT: R). *Educational Psychology Review, 25*(1), 69 – 94.

McGee, M. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal and neurological influences. *Psychological Bulletin, 86*(5), 889–918.

McGee, M. G. (1979). *Human spatial abilities: sources of sex differences*. New York, NY: Preager.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago: The University of Chicago Press.

McNeill, D. (2005). *Gesture and thought.* University of Chicago Press.

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes, 22*, 473 – 500.

Meyers, J. E., Grills, C. E., Zellinger, M. M., & Miller, R. M. (2013). Emotional distress affects attention and concentration: The difference between mountains and valleys. *Applied Neuropsychology: Adult*, *21*(1), 28–35.

Michael, W. B., Guilford, J. P., Fruchter, B. & Zimmerman, W. S. (1957). The description of spatial-visualization abilities. *Educational and Psychological Measurement, 17*, 185-199.

Mix, K. S. (2019). Why are spatial skills and mathematics related? *Child Development Perspectives*, *13*(2), 121 – 126.

Mix, K. S., Hambrick, D. Z., Satyam, V. R., Burgoyne, A. P., & Levine, S. C. (2018). The latent structure of spatial skill: A test of the 2 x 2 typology. *Cognition, 180*, 268 – 278.

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.

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.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P. & Hegarty, M. (2001). How are visuo-spatial working memory, executive functioning, and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology: General, 130*, 621-640.

Mohler, J. L. (2008). A review of spatial ability research. *Engineering Design Graphics Journal, 72*(2), 19 – 30.

Moreau, D., & Wiebels, K. (2021). Assessing change in intervention research: The benefits of composite outcomes. *Advances in Methods and Practice in Psychological Science, 4*(1), *2515245920931930*.

Moyer-Packenham, P. S., Salkind, G., & Bolyard, J. J. (2008). Virtual manipulatives used by K-8 teachers for mathematics instruction: Considering mathematical, cognitive, and pedagogical fidelity. *Contemporary Issues in Technology and Teacher Education, 8*(3), 202 – 218.

Nathan, M.J. (2014). Grounded mathematical reasoning. In L. Shapiro (Ed.), *The Routledge Handbook of Embodied Cognition* (1st ed., 171 -183). Routledge.

Nathan, M. J. (2021). *Foundations of embodied learning: a Paradigm for education*. Routledge.

Nathan, M. J., Schenck, K. E., Vinsonhaler, R., Michaelis, J. E., Swart, M. I., & Walkington, C. (2021). Embodied geometric reasoning: Dynamic gestures during intuition, insight, and proof. *Journal of Educational Psychology, 113*(5), 929–948.

Nathan, M. J., & Swart, M. I. (2021). Materialist epistemology lends design wings: Educational design as an embodied process. *Educational Technology Research and Development, 69,* 1925 – 1954.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. U.S. Department of Education.

National Research Council (2006). *Learning to think spatially*. Washington, DC: The National Academies Press.

Newcombe, N. S. (2013). Seeing Relationships: Using Spatial Thinking to Teach Science, Mathematics, and Social Studies. *American Educator, 37*, 26–40.

Newcombe, N. S. & Shipley, T. F. (2015). Thinking about spatial thinking: new typology, new assignments. In J. S. Gero (Ed), *Studying Visual and Spatial Reasoning for Design Creativity* (pp. 179 – 192). Springer, Dordrecht.

Newcombe, N.S., Ratliff, K. R., Shallcross, W. L., & Twyman, A.D. (2009). Young children's use of features to reorient is more than just associative: Further evidence against a modular view of spatial processing. *Developmental Science, 13*(1), 213 – 220.

Newell, A. (1994). *Unified theories of cognition.* Harvard University Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Prentice-Hall.

Okamoto, Y., Kotsopoulos, D., McGarvey, L., & Hallowell, D. (2015). The development of spatial reasoning in young children. In B. Davis (Ed), *Spatial Reasoning in the Early Years: Principles, Assertions, and Speculations.* (pp. 25-38). Routledge.

Oltman, P. K., Raskin, E., & Witkin, H. A. (1971). *Group Embedded Figure Test*. Consulting Psychologists Press.

Oostermeijer, M., Boonen, A. J. H., & Jolles, J. (2014). The relation between children's constructive play activities, spatial ability, and mathematical work problem-solving performance: a mediation analysis in sixth-grade students. *Frontiers in Psychology, 5*, 782.

Otter.ai (2023). *Otter for Business* [Computer Software].

Ottmar, E., & Landy, D. (2017). Concreteness fading of algebraic instruction: Effects on learning. *Journal of the Learning Sciences, 26*(1), 51 – 78.

Organization for Economic Co-operation and Development, Programme for International Student Assessment. (2014) PISA 2012 Released Mathematics Questions. Retrieved from Organization for Economic Co-operation and Development website: http://www.oecd.org/pisa/test/PISA%202012%20items%20for%20release_ENGLISH.pdf

Park, Y., & Matthews, P. G. (2021). Revisiting and refining relations between nonsymbolic ratio processing and symbolic math achievement. *Journal of Numerical Cognition, 7*, 328 – 350.

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse Mental Rotations Test- Different versions and factors that affect performance. *Brain and Cognition, 28*(1), 39 – 58.

Pier, E. L., Walkington, C., Clinton, V., Boncoddo, R., Williams-Pierce, C., Alibali, M. W., & Nathan, M. J. (2019). Embodied truths: How dynamic gesture and speech contribute to mathematical proof practices. *Contemporary Educational Psychology, 58*, 44 – 57.

Potter, L. E. (1995). Small-scale versus large-scale spatial reasoning: Educational implications for children who are visually impaired. *Journal of Visual Impairment & Blindness, 89*(2), 142-152.

Quaiser-Pohl, C. (2003). The mental cutting test "Schnitte" and the Picture Rotation Test – Two new measures to assess spatial ability. *International Journal of Testing, 3*(3), 219-231.

Ramful, A., Lowrie, T., & Logan, T. (2017). Measurement of spatial ability: Construction and validation of the Spatial Reasoning Instrument for Middle School Students. *Journal of Psychoeducational Assessment*, *35*(7), 709 - 727.

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2012). Spatial anxiety relates to spatial abilities as a function of working memory in children. *Quarterly Journal of Experimental Psychology, 65*(3), 474–487.

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development, 14,* 187 – 202.

Ratliff, K.R. McGinnis, C.R., & Levine, S. C. (2010, August). *The development and assessment of cross-sectioning ability in young children,* [Paper presentation]. The 32nd Annual Meeting of the Cognitive Science Society, Portland, OR.

Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica*, *7*(1), 5-41.

Reuhkala, M. (2001). Mathematical skills in ninth-graders: relationship with visuo-spatial abilities and working memory. *Educational Psychology, 21*, 387–399.

Richardson, J. T. E. (1994). Gender differences in mental rotation. *Perceptual and Motor Skills, 78*(2), 435–448.

Richardson, F. C., & Suinn, R. M. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology, 19*(6), 551-554.

Rutherford, T., Karamarkovich, S. M. & Lee, D. S. (2018). Is the spatial/math connection unique? Associations between mental rotation and elementary mathematics and English achievement. *Learning and Individual Differences, 62,* 180 – 199.

Saucier, D. M., Green, S. M., Leason, J., MacFadden, A., Bell, S., Elias, L. J. (2002). Are sex differences in navigation caused by sexually dimorphic strategies or by differences in the ability to use the strategies? *Behavioral Neuroscience, 116* (3), 403 - 410.

Schenck, K. E., Hubbard, E. M., Nathan, M. J., & Swart, M. I. (2022) Expanding understandings of embodied mathematical cognition in students' fraction knowledge. In J. Culbertson, A., Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society, 44*. Toronto, Canada.

Schenck, K. E., Kim, D., Swart, M. I., & Nathan, M. J. (2022, April). *With no universal consensus, choice of spatial frameworks can affect model fitting and interpretation*. [Paper Presentation]. The annual American Educational Research Association Conference, San Diego, CA.

Schenck, K. E. & Nathan, M. J. (2020, April). *Connecting mathematics, spatial ability, and spatial anxiety*. [Paper Presentation]. The annual American Educational Research Association Conference, San Francisco, CA.

Schenck, K. E**.,** & Nathan, M. J. (2022, November). *Spatial anxiety moderates the effect of spatial ability in geometric reasoning*. In A. E. Lischka, E. B. Dyer, R.S. Jones, J. N. Lovett, J. Strayer, & S. Drown (Eds.), *Proceedings of the forty-fourth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 598). Nashville, TN, United States.

Schenck, K. E., & Nathan, M. J. (2023). *Navigating spatial ability for mathematics education: A review and roadmap for researchers.* Manuscript in preparation.

Schenck, K. E., Nathan, M. J., & Walkington, C. (2021). Groups that move together, prove together: Collaborative gestures and gesture attitudes among teachers performing embodied geometry. In S. Macrine & J. Fugate (Eds.), *Movement Matters: How Embodied Cognition Informs Teaching and Learning.* MIT Press.

Schenck, K. E., Walkington, C., & Nathan, M. J. (2022). Mathematical Reasoning with Dynamic Geometry. https://doi.org/10.17605/OSF.IO/R9CUV

Schliemann, A. D., & Carraher, D. W. (2002). The Evolution of Mathematical Reasoning: Everyday versus Idealized Understandings. *Developmental Review*, *22*(2), 242–266.

Schneider, J., & McGrew, K. (2012). The Cattell–Horn–Carroll model of intelligence. In D. Flanagan & P. Harrison (Eds*.), Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99–144). Guilford Press.

Schrank, F. A., McGrew, K. S., & Woodcock, R. W. (2001). *Technical Abstract.* [Woodcock-Johnson III Assessment Service Bulletin No. 2.]. Itasca, IL: Riverside Publishing.

Schultz, K. (1991). The contribution of solution strategy to spatial performance. *Canadian Journal of Psychology*, *45*(4), 474–491.

Seron, X., Pesenti, M., Noël, M., Deloche, G. & Cornet, J. (1992). Images of numbers, or "When 98 is upper left and 6 sky blue.". *Cognition*, *44*, 159-196.

Sfard, A. (2008) *Thinking as communicating: Human development, the growth of discourses, and mathematizing.* Cambridge University Press.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: an individual differences approach. *Journal of experimental psychology: General*, *125*(1), 4 - 27.

Shapiro, L. (2019). *Embodied cognition*. Routledge.

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, *93*(3), 604 - 614.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701-703.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., ... Chen, M. (2012). Early predictors of high school mathematics achievement. Psychological Science, 23, 691– 697.

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49*, 1994 – 2004.

Siew, N. M., Chong, C. L., & Abdullah, M. R. (2013). Facilitating students' geometric thinking through Van Heile's phased-based learning using Tangram. *Journal of Social Science, 9(3), 101 – 111.*

Smithson, L., & Nicoladis, E. (2014). Lending a hand to imagery? The impact of visuospatial working memory interference upon iconic gesture production in a narrative task. *Journal of Nonverbal Behavior, 38*, 247 – 258.

Sorby, S. (1999). Developing 3-D spatial visualization skills. *Engineering Design Graphics Journal*, *63*(2), 21–32

Sorby, S. A. (2009). Educational research in developing 3-D spatial skills for engineering students. *International Journal of Science Education*, *31*(3), 459-480.

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

Sorby, S. A. & Panther, G. C. (2020). Is the key to better PISA math scores improving spatial skills? *Mathematics Education Research Journal, 32*(2), 213 – 233.

Sorby, S., & Veurink, N. (2019). Preparing for STEM: Impact of spatial visualization training on middle school math performance. *Journal of Women and Minorities in Science and Engineering, 25*(1), 1-23.

Spearman, C. (1927). *The abilities of man, their nature and measurement*. Macmillan.

Starling-Alves, I., Wronski, M. R., & Hubbard, E. M. (2021). Math anxiety differentially impairs symbolic, but not nonsymbolic, fraction skills across development. *Annals of the New York Academy of Sciences.*

Stieff, M., & Uttal, D. (2015). How much can spatial training improve STEM achievement? *Educational Psychology Review, 27*(4), 607 – 615.

Stone, J. M, & Towse, J. N. (2015). A working memory test battery: Java-based collection of seven working memory tasks. *Journal of Open Research Software, 3* (1), e5.

Suinn, R. M., Taylor, S., & Edwards, R. W. (1988). Suinn mathematics anxiety rating scale for elementary school students (MARS-E): Psychometric and normative data. *Educational and Psychological Measurement, 48,* 979 – 986.

Sundberg, S. E. (1994). *Effect of spatial training on spatial ability and mathematical achievement as compared to traditional geometry instruction*. University of Missouri-Kansas City.

Surahmi, E., Budayasa, I. K., & Ekawati, R. (2018). The role of gesture to mathematics' fraction base concept instruction. *Journal of Physics Conference Series, 1108.*

Swart, M.I., Friedman, B., Vitale, J.M., Kornkasem, S., Hollenberg, S., Lowes, S., Nankin, F., Sheppard, S. and Black, J.B.B. (2014). Mobile Movement Mathematics: Exploring the gestures students make while explaining FrActions. *Proceedings of the annual conference of the American Educational Research Association*. Philadelphia, PA.

Swart, M. I., Friedman, B., Kornkasem, S., Lee, A., Lyashevsky, I., Black, J. B., Vitale, J. M., & Sheppard, S. (2016). A design-based approach to situating embodied learning of mathematical fractions using narratives and gestures in a tablet-based game. *Proceedings of the annual conference of the American Educational Research Association.* Washington, D.C.

Swart, M. I., Schenck, K. E., Xia, F., Kim, D., Grondin, M., Nathan, M. J., & Walkington, C. (2023, May). *Embodying students' geometric thinking in an interactive narrative game.* [Paper Presentation]. American Educational Research Association Conference, Chicago, IL, United States.

Tam, Y. P., Wong, T. T., & Chan, W. W. L. (2019). The relation between spatial skills and mathematical abilities: The mediating role of mental number line representation. *Contemporary Educational Psychology, 56*, 14 – 24.

Tartre, L. A. (1990). Spatial orientation skill and mathematical problem solving. *Journal for Research in Mathematics Education, 21*(3), 216 -229.

Thompson, J. M., Nurek, H. C., Moeller, K., & Kardosh, R. C. (2013). The link between mental rotation ability and basic numerical representations. *Acta Psychologica, 144,* 324–331.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago University Press.

Thurstone, L. L. (1950). *Some primary abilities in visual thinking (Report number. 59).* Chicago, IL: Psychometric Laboratory, University of Chicago.

Toomarian, E.Y., Meng, R., & Hubbard, E.M. (2019). Individual Differences in Implicit and Explicit Spatial Processing of Fractions. *Frontiers in Psychology, 10(569).*

Tufte, E.R. (2001). *The visual display of quantitative information* (2nd ed.). Graphics Press.

Tutton, M. (2013). A new approach to analysing static locative expressions. *Language and Cognitive Processes*, *22*, 25 – 60.

Tversky, B. (2019). Transforming Thought. In *Mind in motion* (pp. 85 – 106). Basic Books.

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how? In B. Ross (Ed.), *Psychology of learning and motivation* (Vol. 57, pp. 147-181). Academic Press.

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.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599–604.

Viarouge, A., Hubbard, E. M., & McCandliss, B. D. (2014). The cognitive mechanisms of the SNARC effect: an individual differences approach. *PloS one, 9*, e95756.

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., & Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Development, 85*(3), 1062–1076.

von Bastian, C. C., Locher, A., & Ruflin, M. Tatool: A Java-based open-source programming frameworking for psychological studies. *Behavior Research Methods, 45*, 108-115.

Voyer, D., Voyer, S., & Bryden, M.P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250 -270.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817.

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space, and quantity. *Trends in Cognitive Sciences, 7*, 483–488.

Walkington, C., Chelule, G., Woods, D., & Nathan, M.J. (2019). Collaborative gesture as a case of extended mathematical cognition. *Journal of Mathematical Behavior, 55*, 1-20.

Walkington, C., Nathan, M. J., Wang, M. & Schenck, K.E. (2022). The effect of relevant directed arm motions on gesture usage and proving of geometry conjectures. *Cognitive Science*, *46*(9).

Wang, L., Bolin, J., Lu, Z., & Carr, M. (2017). Visuospatial Working Memory Mediates the Relationship Between Executive Functioning and Spatial Ability. *Frontiers in Psychology*.

Wang, L., Cohen, A. S., & Carr, M. (2014). Spatial ability at two scales of representation: A meta-analysis. *Learning and Individual Differences, 36*, 140 -144.

Wei, W., Budakova, A., Bloniewski, T., Matsepuro, D., & Kovas, Y. (2018). Spatial anxiety and spatial performance in university students in Russia and China. In *The European Proceedings of Social & Behavioural Sciences*, 771–781.

Weisberg, S. M., Schinazi, V.R., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2014). Variations in cognitive maps: Understanding individual differences in navigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 669 – 682.

Williams-Pierce, C., Pier, E. L., Walkington, C., Boncoddo, R., Clinton, V., Alibali, M. W., & Nathan, M. J. (2017). What we say and how we do: action, gesture, and language in proving. *Journal for Research in Mathematics Education*, *48*(3), 248-260.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636.

Witkin, H. A. (1950). Individual differences in ease of perception of embedded figures. *Journal of Personality, 19*, 1-15.

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research, 47*(1), 1 – 64.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *A manual for the embedded figures test.* Consulting Psychologist Press.

Wolfgang, C., Stannard, L., & Jones, I. (2003). Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics. *Early Child Development and Care*, *173*(5), 467-475.

Woods, D., & Fassnacht, C. (2020). *Transana Multiuser* (Version 3.32d) [Computer software]. Retrieved from http://transana.com.

Wraga, M., Creem, S. H., & Profitt, D. R. (2000). Updating displays after imagined object and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*(1), 151-168.

Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., & Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychonomic Bulletin & Review, 15*(4), 763 – 771.

Xia, F., Schenck, K.E**.,** Swart, M.I., & Nathan, M.J. (2022, June). Directed actions scaffold gestural insights in geometric reasoning. In *Proceedings of the 2022 International Conference of the Learning Sciences-ICLS2022* (pp. 1982-1983). Hiroshima, Japan: International Society of the Learning Sciences.

Xie, F., Zhang, L., Chen, X., & Xin, Z. (2020). Is spatial ability related to mathematical ability: a Meta analysis. *Educational Psychology Review, 32*, 113 -155.

Yang, W., Liu, H., Chen, N., Xu, P, & Lin, X. (2020). Is early spatial skills training effective? A meta-analysis. *Frontiers in Psychology, 27.*

Yilmaz, B. (2009). On the development and measurement of spatial ability. *International Electronic Journal of Elementary Education*, *1*(2), 1–14.

Young, C. J., Levine, S. C., & Mix, K. S. (2018a). The connection between spatial and mathematical ability across development. *Frontiers in Psychology, 9*, 775.

Young, C. J., Levine, S. C., & Mix, K. S. (2018b). What processes underlie the relation between spatial skill and mathematics? In K. S. Mix & M. T. Battista (Eds.), *Visualizing Mathematics. Research in Mathematics Education* (pp. 195 – 228). Springer, Cham.

Yuan, L., Kong, F., Luo, Y., Zeng, S., Lan, J., & You, X. (2019). Gender differences in large-scale and small-scale spatial ability: A systematic review based on behavioral and neuroimaging research. *Frontiers in Behavioral Neuroscience*, *13*, 128.

Zacks, J., Rypma, B., Gabrieli, J. D. E., Tversky, B., & Glover, G. H. (1999). Imagined transformations of bodies: An fMRI investigation. *Neuropsychologia, 37*(9), 1029 – 1040.

Zacks, J. M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation*, 2, 315–332.

Zhang, X., & Lin, D. (2015). Pathways to arithmetic: The role of visual-spatial and Language skills in written arithmetic, arithmetic word problems, and nonsymbolic arithmetic. *Contemporary Educational Psychology, 41*, 188-197.

Zhang, J., Zhao, N., & Kong, Q. P. (2018). The Relationship Between Math Anxiety and Math Performance: A Meta-Analytic Investigation. *Frontiers in Psychology.*

Zhong, J. Y., & Kozhevnikov, M. (2016). Relating allocentric and egocentric survey-based representations to the self-reported use of a navigation strategy of egocentric spatial updating. *Journal of Environmental Psychology*, *46*, 154-175.

Appendix A

Interaction Plots from Chapter 3

Figure A1

Probability of Verbal Insight by Spatial Ability for Three Levels of Spatial Anxiety

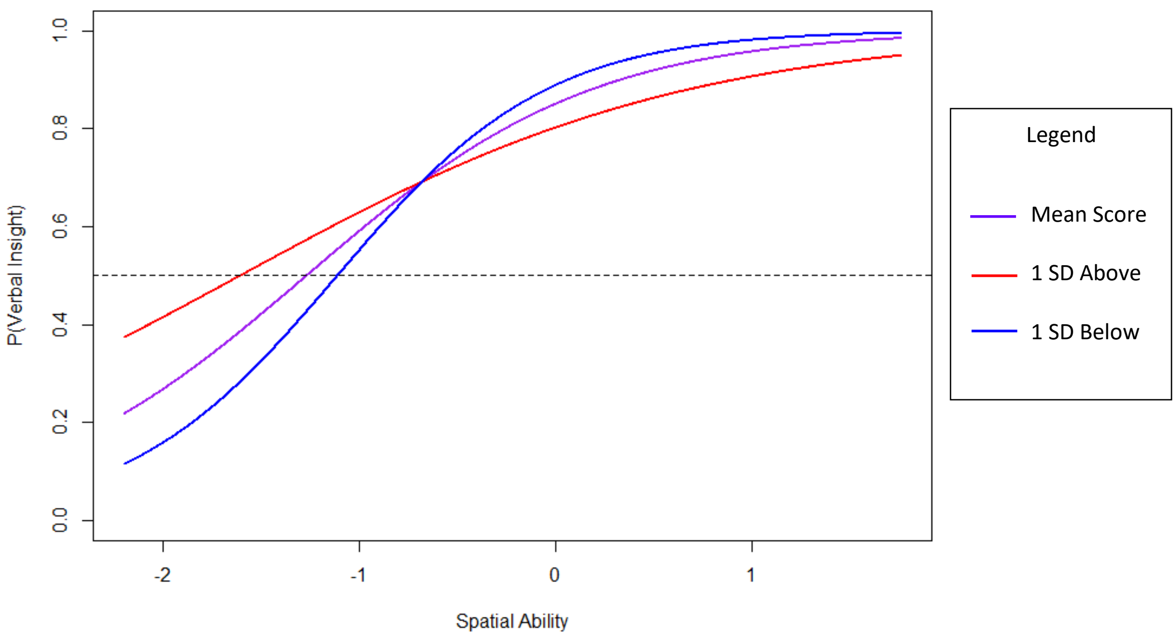


Figure A2

Probability of Transformational Proof by Spatial Ability for Three Levels of Spatial Anxiety

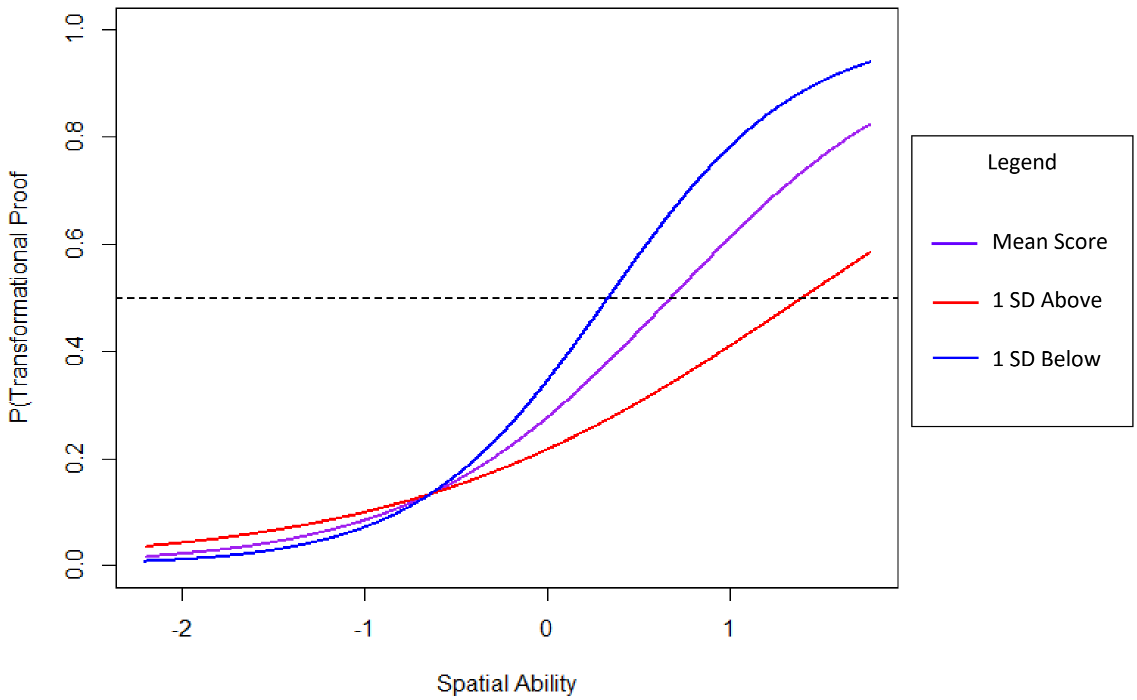


Figure A3

Probability of Representational Gesture by Spatial Ability for Three Levels of Spatial Anxiety

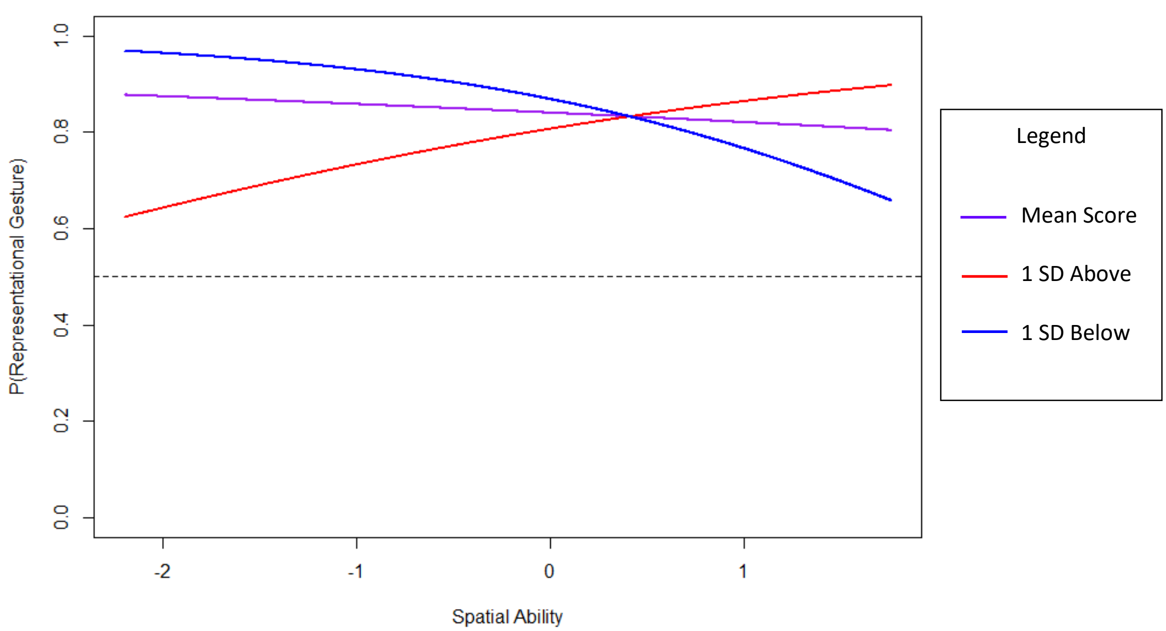


Figure A4

Probability of Non-Dynamic Gesture by Spatial Ability for Three Levels of Spatial Anxiety

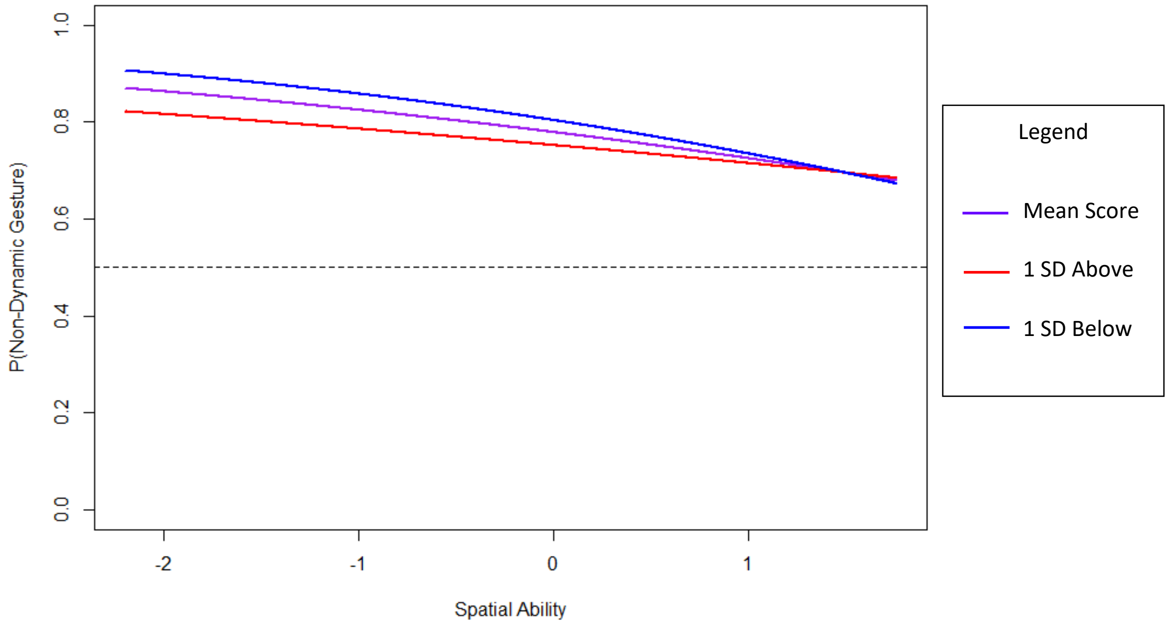


Figure A5

Probability of Non-Dynamic Gesture by Spatial Ability for Three Levels of Spatial Anxiety