DEVELOPING SPATIAL ABILITIES THRU COMPUTER-AIDED DESIGN SOFTWARE EXPERIENCES: ENHANCING STEM EDUCATION

A Dissertation

by

ASHLEY M. WILLIAMS

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Mary M. Capraro
Co-Chair of Committee,	Jamaal R. Young
Committee Members,	Christopher G. Thompson
	Julie Singleton
Head of Department,	Joanne Olson

May 2022

Major Subject: Curriculum and Instruction

ABSTRACT

The research I conducted during this dissertation focuses on improving students' spatial abilities as measured by mental rotation through the implementation of computer-aided design (CAD) projects in an elementary classroom. I choose the option of submitting three professional journal articles rather than the traditional five-chapter doctoral dissertation product. The purpose of article one was to synthesize empirical research on CAD software as an intervention in various classrooms. There were 19 studies that met selection criteria to best represent the impact CAD software had on students' spatial ability. The calculated weighted Hedges' *g* effect size of 0.37 for the group-based studies and the weighted *d* effect size of 0.35 for the single-group studies indicates the CAD interventions had a positive impact on students' spatial ability. In article two I assessed the reliability of my final study instrument, the redrawn mental rotations test (MRT-A) and calculated a weighted averaged Cronbach Alpha of .86 to provide a generalized reliability coefficient for the MRT-A. Additionally, I also identified how researchers often fail to report reliability based on their sample, which is concerning given reliability is inherently sample-dependent. Finally, in article three, I explored how experiences with computer-aided design (CAD) software in engineering design projects can enhance elementary students' spatial ability. Quantitative data were collected before and after participants spent a week working on CAD design projects. A paired sample *t* test and 95% confidence intervals indicated a statistically significant difference between observed pre and posttest scores. The calculated Cohen's *d* effect size of 0.55 indicated that the CAD intervention had a positive impact on students' mental rotation skills. It can be concluded that utilizing technologies such as CAD software can aid in developing and improving spatial abilities.

DEDICATION

This dissertation is dedicated to my loving husband, Stephen J. Williams, who has been a source of strength, support, patience, and motivation for me throughout this entire experience. I am truly blessed to have such a wonderful husband and father as my partner in life.

I would also like to dedicate this to my colleague and mentor, Dr. Rachael M. Welder, who has academically and emotionally supported me through this journey. Dr. Welder has taught me so much about well-being and always advocated for me to take care of myself. I am so grateful we had the opportunity to work together and for the life-long friendship we have developed.

ACKNOWLEDGMENTS

First I would like to thank my committee chair, Dr. Mary Margaret Capraro, co-chair, Dr. Jamaal Young, and committee members, Dr. Thompson and Dr. Singleton for their support and guidance throughout the course of this research.

Thanks also go to my friends, colleagues, students, and the department faculty and staff for making my time at Texas A&M University a valuable experience. As well as Aggie STEM for providing me with so many opportunities to learn, instruct, and conduct research.

Finally, many thanks to my husband, parents, siblings, and in-laws for their continuous assistance, encouragement, patience, and love.

CONTRIBUTORS AND FUNDING SOURCES CONTRIBUTORS

Contributors

This work was supported by a dissertation committee consisting of Dr. Mary Margaret Capraro, Dr. Jamaal Young, and Dr. Julie Singleton of the Department of Teaching, Learning, and Culture and Dr. Christopher Thompson of the Department of Educational Psychology.

The data analyzed for Chapter IV were provided by Aggie STEM. The data analyzed in Chapter II were double coded and reviewed by Dr. Thompson. The data analyzed in Chapter III were double coded and reviewed by Dr. Young. All other work conducted for the dissertation was completed by the student independently.

Funding Sources

All work was completed without outside financial support.

NOMENCLATURE

2D	Two-dimensional
3D	Three-dimensional
CAD	Computer-aided Design
DBL	Design-based Learning
MRT	Mental Rotations Test
MRT-A	Mental Rotations Test redrawn
RG	Reliability Generalization
STEM	Science, Technology, Engineering, and Mathematics

TABLE OF CONTENTS

Page

ABSTRACT II

DEDICATION IV

ACKNOWLEDGMENTS V

CONTRIBUTORS AND FUNDING SOURCES CONTRIBUTORS VI

NOMENCLATURE VII

TABLE OF CONTENTS VIII

LIST OF FIGURES XI

LIST OF TABLES XII

1. INTRODUCTION 1

  1.1. Research Questions 2
  1.2. Methods 3
  1.3. Journal Selection 4

2. EFFECTS OF THREE-DIMENSIONAL MODELING OR DESIGNING SOFTWARE INTERVENTIONS ON STUDENTS' SPATIAL ABILITIES: A META-ANALYSIS 6

  2.1. Introduction 6
  2.2. Background 7
    2.2.1. Spatial Ability 7
    2.2.2. CAD Software 9
    2.2.3. CAD and Spatial Ability 10
  2.3. Methods 12
    2.3.1. Literature Search 12
    2.3.2. Data Collection 12
    2.3.3. Coding Procedures 14
    2.3.4. Independent Variable 14
    2.3.5. Outcome Variable 14
    2.3.6. Moderator Variables 15
    2.3.7. Statistical Analysis 15
  2.4. Results 18

2.4.1. Standardized Mean Difference Analysis Results 18
2.4.2. Standardized Mean Change Analysis Results 22
2.5. Discussion 25

3. REVIEW OF THE REDRAWN MENTAL ROTATION TEST: A RELIABILITY GENERALIZATION META-ANALYSIS OF COEFFICIENT ALPHA 28

3.1. Background 29
3.1.1. The Mental Rotation Test 29
3.1.2. Reliability 32
3.1.3. Reliability Generalization 33
3.2. Methods 35
3.2.1. Literature Search 35
3.2.2. Data Collection 35
3.2.3. Coding Procedures 37
3.2.4. Moderator Variables 37
3.2.5. Transformation of Cronbach's Alpha 37
3.2.6. Statistical Analysis 38
3.3. Results 40
3.4. Discussion 43

4. FOSTERING ELEMENTARY STUDENTS' SPATIAL ABILITY DEVELOPMENT WITH COMPUTER-AIDED DESIGN SOFTWARE 45

4.1. Introduction 45
4.2. Conceptual Framework 46
4.3. Literature Review 49
4.3.1. Design-Based Learning 49
4.3.2. Engineering Design Process 51
4.3.3. CAD Software 52
4.3.4. Spatial Ability 53
4.4. Methods 56
4.4.1. Participants 56
4.4.2. Instrument 56
4.4.3. CAD Software Used for Instruction 58
4.4.4. Intervention 58
4.4.5. Data Analysis 59
4.5. Results 60
4.6. Discussion 62
4.7. Conclusion 64

5. CONCLUSIONS 66

5.1. Intellectual Merit 66
5.2. Broader Impacts 70

REFERENCES..73

LIST OF FIGURES

Page

Figure 2.1 Flowchart of inclusion and exclusion decisions 13

Figure 2.2 Forest plot for standardized mean difference effect sizes.............................. 17

Figure 2.3 Forest plot for standardized mean change effect sizes 17

Figure 2.4 Funnel plot of SMD effect sizes. Filled circles denote observed effect sizes. Open circles denote Trim-and-Fill imputed effect sizes......................... 19

Figure 2.5 Funnel plot of SMC effect sizes 23

Figure 3.1 Practice Item Included on MRT-A 30

Figure 3.2 Flowchart of inclusion and exclusion decisions 36

Figure 3.3 Forest plot for transformed alpha coefficients 40

Figure 3.4 Funnel plot of transformed alpha coefficients. Filled circles denote observed alpha coefficients. Open circles denote Trim-and-Fill imputed coefficients... 42

Figure 4.1 Conceptual Framework for Improving Spatial Abilities Through CAD Projects 48

Figure 4.2 Example Item from MRT-A 57

Figure 4.3 Q plots of pre-post MRT-A observed difference scores................................ 60

Figure 4.4 Box plots of pre-post MRT-A performance 61

Figure 4.5 95% CI plots by pre-post MRT-A observed scores....................................... 61

LIST OF TABLES

Page

Table 1.1 Articles and Journals. 4

Table 2.1 Description of Moderator Variables 15

Table 2.2 Primary Study Identification 18

Table 2.3 Descriptive Statistics and Effect Sizes for Spatial Ability SMD Analysis 19

Table 2.4 SMD Moderator Analysis 19

Table 2.5 Descriptive Statistics and Effect Sizes for Spatial Ability SMC Analysis 23

Table 2.6 SMC Moderator Analysis 24

Table 3.1 Formulas for Transformation of Cronbach Alpha. 38

Table 3.2 Primary Study Characteristics 40

Table 3.3 Moderator Analysis Results 42

Table 4.1 Intervention Project Descriptions 59

1. INTRODUCTION

The purpose of this dissertation study was to investigate how engaging students in experiences with 3D modeling software, more specifically computer-aided design (CAD) software, impacts their spatial abilities. First, I synthesized previous findings in order to examine the relationship between interventions which utilized CAD or CAD like software and students' spatial abilities. This synthesis allowed for the analysis of the overall effects of CAD or CAD like software on various students' spatial abilities across studies. I investigated sample and study characteristics in order to explain variations in effects across studies and identify if there is a gap in the research literature examining the effects of CAD or CAD like software specifically on elementary students' spatial abilities.

Second, I examined the reliability of the instrument that was be utilized in the third study to measure students' spatial abilities via mental rotation skills. Several instruments exist to measure mental rotation skills, however, the widely used Mental Rotation Test (MRT) by Vandenburg and Kuse (1978) was among the first paper-and-pencil group test created and validated to measure mental rotation skills. The MRT was redrawn by Peters et al. (1995) due to image quality degradation and was renamed the MRT-A. Although the MRT is a well-established valid instrument and is known to be highly reliable with previous participants, there exist the need to assess the reliability of the MRT-A for a number of reasons. First, Peters et al. (1995) simply inducts reliability estimates from Vandenburg and Kuse (1978), failing to report score reliability for the newly redrawn version of the instrument in various settings. Second, although it was stated that the MRT-A was exactly redrawn from the original MRT, inconsistencies were found in the literature concerning the number of items on the instruments. The MRT-A consists of 24 items while the original MRT consists of only 20 items and although an earlier researcher (Bryant, 1982) used the MRT, citing only Vandenburg and Kuse (1978), also include 24 items, however, the source of the four additional items has not been found. Synthesizing MRT-A reliability estimates across studies provided an expected range of score reliability for the instrument as well as allow for the examination of characteristics that may impact reliability estimates across studies.

Finally, applying findings and results of the aforementioned syntheses, I designed an exploratory research study to investigate how utilizing TinkerCAD software as the primary tool for engineering design projects influences elementary students' spatial abilities.

1.1. Research Questions

Article 1:

1. What is the overall effect for CAD, or CAD like, software interventions comparing spatial abilities of a CAD treatment groups to a non-CAD group?
2. What is the overall effect for CAD, or CAD like, software interventions comparing pre to posttest spatial abilities scores of a CAD intervention group?
3. What is the risk of publication bias?
4. How heterogenous are results from studies on the overall effectiveness of CAD or CAD like interventions?

5. To what degree do study characteristics, including country, educational level, length of intervention, instrument, and software type moderate the effect?

Article 2:

1. What is the overall reliability, using Cronbach's alpha, of the redrawn mental rotation test scores?
2. What is the risk of publication bias?
3. How heterogenous are reliability estimates across various applications of the redrawn mental rotation test?
4. To what extent do primary study characteristics including region, mean age, scoring procedures, instrument modifications, and percent female moderate the average weighted Cronbach's alpha reliability estimate?

Article 3:

1. How does utilizing TinkerCAD software as the primary tool for engineering design projects foster growth and development of elementary students' spatial abilities as measured by mental rotation?

1.2. Methods

The methodological approach utilized while researching and writing the three articles of this dissertation was different according to the research question(s) for each study and type of data that was be collected. Quantitative analyses were used in all three articles. For the first and second article I applied meta-analytic techniques using secondary data from primary studies to analyze effect sizes from previous CAD or CAD like interventions, and reliability estimates from previous applications of the MRT-A. In addition, for both studies, meta-regressions were used to analyze possible sample and study characteristics that could explain variability across effect sizes and reliability estimates in primary studies. For the third study, I analyzed pre- and post-intervention mental rotation skills using descriptive statistics, a paired sample t-test, 95% confidence intervals (CI), and Cohen's *d* effect sizes.

1.3. Journal Selection

Two prospective journals were selected for publication of each manuscript by inclusion criteria as follow: (1) journals which include articles cited in the literature review of this dissertation, and (2) impact factors (*SCImago Journal Rank* [SJR] and *Source Normalized Impact per Paper* [SNIP]), acceptance and invited rates, and prestige of the editorial board. The impact factors were found on the Scimago Journal & Country Rank website, and Academic Accelerator SNIP database website. Acceptance rates, invited rates, type of review, and manuscript length were gathered from *Cabell's Directory of Publishing Opportunities* and were referenced to choose the journals (see Table 1.1).

Table 1.1 Articles and Journals.

Proposed Article	Proposed Journal #1	Proposed Journal #2
Effects of three-dimensional modeling or designing software interventions on students' spatial abilities: A meta-analysis	*Journal of Engineering Education* Acceptance rate: 10% (SJR/SNIP): 3.032/7.53 Editor in chief: Lisa C. Benson Publisher: Wiley-Blackwell Publishing Type of review: Double Blinded Peer Review Length: 8,000-10,000 words	*Journal of Educational Psychology* Acceptance rate: 11% (SJR/SNIP): 3.4559/3.12 Editor in chief: Steve Graham Publisher: American Psychological Association Type of review: Double Blinded Peer Review Length: 12,000 words

Table 1.2 Continued

Proposed Article	Proposed Journal #1	Proposed Journal #2
Review of the redrawn mental rotation test: A reliability generalization meta-analysis	*Educational and Psychological Measurement* Acceptance rate: 10-12% (SJR/SNIP): 1.747/1.664 Editor in chief: George A. Marcoulides Publisher: Sage Publications Type of review: Double Blinded Peer Review Length: 26-30 pages	*Journal of Psychoeducational Assessment* Acceptance rate: 30% (SJR/SNIP): 0.702/0.941 Editor in chief: Donald H. Saklofske Publisher: Sage Publications Type of review: Blinded Peer Review Length: 6,000 words
Integrating computer-aid design projects in elementary classrooms to foster spatial ability development	*Computers & Education* Acceptance rate: 24% (SJR/SNIP): 2.323/3.797 Associate editors: Shelly Heller, M. Nussbaum, Chin-Chung Tsai Publisher: Elsevier, Inc. Type of review: Double Blinded Peer Review Length: no limit	*Educational Psychology Review* Acceptance rate: 11-20% (SJR/SNIP): 4.055/4.271 Editor in chief: Fred Paas Publisher: Springer Nature Type of review: Double Blinded Peer Review Length: 30 pages

2. EFFECTS OF THREE-DIMENSIONAL MODELING OR DESIGNING SOFTWARE INTERVENTIONS ON STUDENTS' SPATIAL ABILITIES: A META-ANALYSIS

2.1. Introduction

Improving science, technology, engineering, and mathematics (STEM) achievement in early education is of much interest to educators, researchers, and communities. There are several ways to improve STEM achievement; spatial ability training is one that receives little emphasis. Spatial abilities have been shown to be a strong predictor of those who do and do not enter STEM fields and, in some fields, have been shown to contribute more unique variance than SAT scores in the prediction of STEM achievement and attainment (Uttal & Cohen, 2012). Several researchers have linked spatial abilities to STEM achievement (Buckly et al., 2018; Höffler, 2010; McConnell, 2015; Smith, 2018), interest, and future success (Shea et al., 2001; Wai et al., 2009). Investment in students' spatial abilities, especially in early childhood education, have not held value in the past but as researchers and educators have gained awareness of the positive links between spatial abilities and STEM it has become an area of interest.

Spatial ability development is vital and can be promoted through experience with three-dimensional modeling programs such as computer-aided design (CAD) software (Matthews & Geist, 2002; Onyancha et al., 2009). The connection between three-dimensional modeling or design software and spatial abilities have become more widely investigated in educational settings such as engineering (i.e., Contero et al., 2006), architecture (i.e., Falcón, 2011), and interior design (i.e., Zavotka, 1986) but have yet to establish a strong presence in early childhood education. Incorporating experiences with three-dimensional modeling or design software (i.e., CAD software) has the potential to improve spatial abilities as well as increase STEM interest and achievement.

2.2. Background

2.2.1. Spatial Ability

Spatial abilities play an important role in mathematical understanding and have been linked to STEM achievement and future success. Spatial ability has been defined as the general capability to represent, transform, generate, and recall symbolic information (Linn & Petersen, 1985). Spatial skills involve understanding, manipulating, and re-organizing or interpreting relations visually (Tartre, 1990) and are highly essential to solve various issues in daily life (Erkoc et al., 2013). These skills are employed during problem solving activities especially when manipulating and processing visual information (Rafi et al., 2005). Spatial abilities allow people to use concepts of shape in concrete and abstract ways to make and use things in the world navigate and communicate (Basham & Kotrlik, 2008), and involves tasks such as mentally transforming objects, knowing left from right when viewing objects from different perspectives, playing Tetris, or packing the trunk of your car (Shavalier, 2004). Spatial ability may also be described as the ability to generate, retain, retrieve, and transform well-structured visual images (Lohman, 1996). Spatial abilities play an important role in STEM achievement and are highly utilized in everyday life.

Spatial ability can be categorized into three categories: (1) spatial perception, (2) spatial visualization, and (3) mental rotation (Linn & Peterson, 2004). All categories play significant roles in the learning and understanding of mathematics and can improve students' problem solving and reasoning skills.

2.2.1.1. Spatial Perception

Spatial perception is comprised of several specific components and is important for physical movement and orientation. Spatial relationships, one of the components of spatial perception, is the ability to establish spatial relationships from visual information with respect to personal orientation despite distracting information (Linn & Petersen, 1985; Seabra & Santos, 2008). The representation of real-world physical conditions is facilitated by spatial perception (Matthews & Geist, 2002). Therefore, spatial perception plays an essential role in the learning of mathematics, especially for two- and three-dimensional geometry.

2.2.1.2. Spatial Visualization

Spatial visualization refers to the ability to manipulate complex spatial information and visual problems imagining the relative movement of an image or object (Linn & Petersen, 1985; Seabra & Santos, 2008). Spatial visualization has also been defined as the process of capturing, translating, and mentally manipulating three-dimensional spatial forms (Carroll, 1993) and involve tasks such as connecting two-dimensional representations to three-dimensional models (Uttal & Cohen, 2012). Spatial visualization has been a basic skill for understanding primary mathematical skills and a gateway to advanced problem solving (Augustynaik et al., 2005). Advanced problems solving is enhanced by spatial visualization because students have to complete tasks that typically require them to make complex decisions about orientation and measurements such as length, width, and angles. The teaching and learning of spatial visualization are inextricably interwoven with problem solving and foundational mathematical skills.

2.2.1.3. Mental Rotation

Mental rotation is the ability to rotate images or objects into specific orientations. Advanced mental rotation is characterized by the ability to quickly and accurately rotate two- or three-dimensional figures mentally (Linn & Petersen, 1985) to verify how they would look from a different angle or perspective (Moé, 2018). It involves the visual examination and mental simulation of an object's rotation in space (Hegarty & Waller, 2005). Mental rotation is involved in performing many everyday tasks and is especially useful in STEM disciplines. The ability to retain complex ideas and to think flexibly is one characteristic requisite in being successful in retaining and converting text into solvable problems.

2.2.2. CAD Software

Computer-aided design software is used to create detailed three-dimensional models and two-dimensional drawings. Most commonly, engineers and architects use CAD software to design and model constructions. Using CAD software in the classroom can increase motivation and STEM achievement because it combines real-life problem solving, reflection, and critical thinking (Huleihil, 2016) through technology-based design learning. In 3D printing and design classes, where the use of CAD software is necessary, improvements in student's interests, motivation, mathematics skills, and real-life skills were seen (Kwon, 2017). There are several types of CAD software on the market such as SketchUp, SolidWorks, AutoCAD, TinkerCAD, and Maya. All of these programs are used to create 2-dimentional renderings of 3-dimentional design. Three-dimensional modeling technology such as CAD software can be incorporated into the classroom to target spatial training in students.

2.2.3. CAD and Spatial Ability

Computer-aided design software is an excellent way to use technology to develop and improve students' spatial abilities. Enhancing the creative design process (Chang, 2014), CAD software has been a vital and necessary tool in engineering education and recently has become a fundamental part of technology education as well (Chester, 2007). Spatial abilities are key skills associated with CAD (Johnson & So, 2015) which are fundamental for engineering students (Contero et al., 2006; Kösa & Karakus, 2018). Several researchers have found that spatial ability, which has shown to positively correlated with retention and achievement in STEM disciplines, has been improved through experience and training with CAD software (Kinsey et al., 2008; Onyancha et al., 2009; Sorby & Baartmans, 2000). Several studies have examined this relationship between the use of CAD software and spatial ability, noting that spatial abilities are key skills associated with CAD (Johnson & So, 2015). Three-dimensional modeling programs are increasingly becoming more common in education and researchers have linked students' abilities to design 3D objects in CAD software to their spatial abilities (Contero et al., 2006; Company et al., 2004) and finding improvements in spatial abilities and motivation (Martín-Dorta et al., 2008). The use of CAD software in classrooms can positively impact students' spatial abilities and has the potential to increase STEM engagement, retention, and achievement.

The utilization of CAD software in educational settings have increasingly become more popular. To encourage the use of CAD software in the classroom it is critical that educators are aware of the positive outcomes and impact they have on students' spatial abilities leading to possible increase in students' STEM attainment and achievement. It is important to understand the general consensus of prior research on the effectiveness of using CAD software on the development of spatial abilities in students from primary to post-secondary. The purpose of this study was to synthesize empirical research on the implementation of CAD software as an intervention in various educational settings as well as investigate how the effectiveness of the intervention may be impacted by select study characteristics. The following research questions guided this meta-analytic study: (1) What is the overall effect for CAD, or CAD like, software interventions comparing spatial abilities of a CAD treatment groups to a non-CAD group? (2) What is the overall effect for CAD, or CAD like, software interventions comparing pre to post-test spatial abilities scores of a CAD intervention group? (3) What is the risk of publication bias? (4) How heterogenous are results from studies on the overall effectiveness of CAD or CAD like interventions? (5) To what degree do study characteristics, including country, educational level, length of intervention, instrument, and software type moderate the effect?

2.3. Methods

2.3.1. Literature Search

The first step in the process was to identify primary studies which examined the relationship between students' spatial ability and CAD, or CAD like, software. The literature search was conducted through a library search engine of a tier one research university in which the following databases were searched simultaneously: Academic Search Ultimate, Applied Science & Technology Source Ultimate, APA PsycInfo, Art & Architecture Source, British Education Index, CINAHL Complete, Complementary Index, Computer Source, Directory of Open Access Journals, Education Source, Engineering Source, ERIC, Professional Development Collection, Science Citation Index, ScienceDirect, Science & Technology Collection, Social Sciences Citation Index, and Vocational Studies Complete. The following keywords and phrases were used in a Boolean search to locate related available studies: "*computer-aided design*" or "*CAD*" or "*3D modeling*" or "*3D printing*" and *"mental rotation" or "spatial perception" or "spatial visualization"* or *"spatial ability."* The search was limited to studies listed or published from 1980 to 2021 and resulted in an initial pool of 222 works. Covidence (Veritas Health Innovation, 2021) was utilized to organize the data collection processes.

2.3.2. Data Collection

Following the literature search, all results were imported to Covidence to further examined for possible inclusion. After the removal of duplicates and unrelated articles based on titles and abstracts, the full text of the remaining possible primary studies was examined. The final selection phase assed articles based on the following inclusion criteria: (1) reported in English, (2) reported the use of CAD, CAD like, or 3D modeling software, (3) reported outcome measures of spatial ability, spatial perception, spatial visualization, and/or mental rotations, and (4) reported sufficient statistical data to calculate an effect size comparing differences between a CAD and non-CAD intervention groups or pre-test to post-test gains of a CAD intervention group. Figure 2.1 provides a detailed flowchart of the inclusion and exclusion decisions that lead to the final dataset.

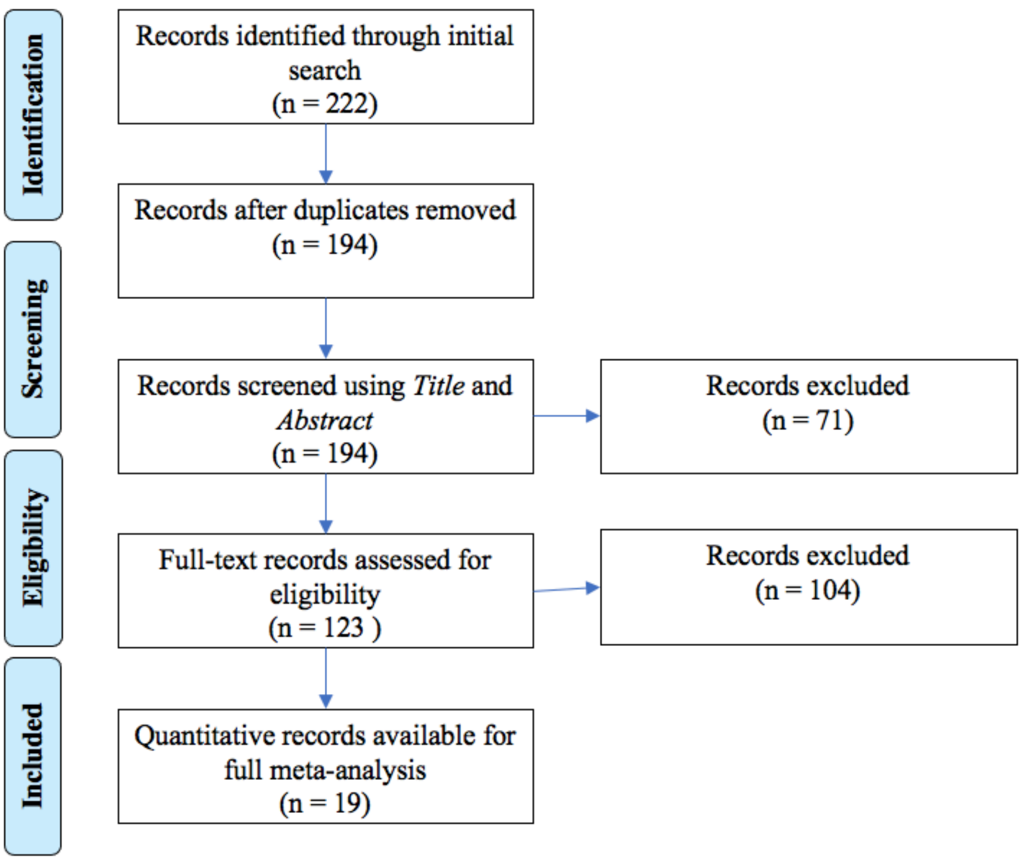

Figure 2.1 Flowchart of inclusion and exclusion decisions

2.3.3. Coding Procedures

After identifying the 19 primary studies to be included a 26-item coding form was utilized to gather information on important variables that helped to calculate effect size and characterize the study and sample. The form was used separately by two coders to collect the following information: author(s), year, article type, study design, country, educational setting (formal vs informal), educational level, length of intervention, instrument, reported score reliability for instrument, spatial ability variable (i.e., mental rotation, spatial visualization), software type, percent female, percent white, sample size (CAD and non-CAD groups), group means (pre and post), and group standard deviations (pre and post). The initial overall agreement rate between the two coders was 97%. All discrepancies were examined, and complete consensus was reached.

2.3.4. Independent Variable

For the purpose of this meta-analysis the independent variable is defined as any intentional hands on experiences with three-dimensional modeling programs, such as computer-aided design (CAD) software, by participants. In our coding procedures, we noted the length of the experience with CAD software and which type of software was utilized.

2.3.5. Outcome Variable

Spatial abilities were assessed as the outcome variable. Spatial ability is often measured through various constructs including mental rotation, spatial perceptions, and spatial visualization. In our coding procedures, we noted which spatial ability construct was measured and the instrument used to measure the spatial ability construct.

2.3.6. Moderator Variables

The study characteristics assed as moderators included country, educational level, length of intervention, instrument, and software type. Sufficient data were not provided by including studies to examine percent female and percent white as potential moderators. See Table 2.1 for description of, and codes used for, moderator variables.

Table 2.1 Description of Moderator Variables

Variable	Description
Country	categorical variable indicating whether the study was conducted in United States (0) or other (1)
Educational Level	categorical variable indicating whether the study was conducted with elementary school (0), middle school (1), high school (2), post-secondary (3), or mixed (4) aged participants
Length of Intervention	categorical variable indicating the intervention took place in 6 weeks or less (0), or more than 6 weeks (1)
Instrument	categorical variable indicating if the instrument used was the PSVT (0), the MRT (1), or Other (2)
Software Type	categorical variable indicating if the software used for the CAD intervention was SketchUp (0), AutoCAD (1), Solidworks (2), or Other (3)

Note: PSVT = Purdue spatial visualization test; MRT = mental rotation test

2.3.7. Statistical Analysis

All statistical analyses were conducted using the statistical software RStudio (RStudio Team, 2019) using the package *metafor* (Viechtbauer, 2010). From the primary studies included, two separate meta-regression analyses were conducted to analyze (1) mean difference effect size of spatial ability differences between CAD and non-CAD groups, and (2) mean change effect size of spatial ability differences between pre and post of CAD groups.

First for the primary studies comparing a CAD intervention group to a non-CAD group an unbiased standardized mean difference (Hedges' *g*) effect sizes were calculated from means and standard deviations provided in the primary studies with the exception of one study (Toptas et al., 2012) which did not include standard deviations. However, for this study *p*-value's for comparisons were provided and therefore a mean standard deviation was calculated using formulas to calculate a *t*-value from the *p*-value, a standard error from the *t*-value, then a standard deviation from the standard error provided by Higgins and Greens (2008) in section 7.7.3.3.

For primary studies in which change mean scores were compared pre and post intervention, unbiased standardized mean change effect sizes and their variances were calculated, using formulas reported in Becker (1988), from summary statistics provided. One article (Onyancha et al., 2009) did not provide mean standard deviations therefor a pooled standard deviation was calculated from the means and Cohen's *d* effect size provided. The variance of the standardized mean change is a function of the population mean change and the population pre-test-post-test correlation. Because pre-test-post-test correlations were not provided an array of correlations (.20, .40, .60, .80) were assessed. The variance of the standardized mean change did not vary drastically between the various pre-test-post-test correlation estimates tested and therefore .60 was used as the pre-test-post-test correlation estimate.

Homogeneity of effect sizes was assessed by visually examining forest plots (see Fig. 2.2 and Fig. 2.3) as well as examining the Q statistic, I^2, and $\hat{\tau}^2$ values provided. After assessing effect-size homogeneity, we elected to use a random-effects model to compute overall effect size to account for within-study sampling error and between-studies heterogeneity. A categorical moderator analysis was conducted to examine the potential relationships of study-level characteristics on the heterogeneity of effect sizes using ANOVA-like models. Next, publication bias was visually assessed using funnel plots which included imputations for potentially missing effects with the trim-and-fill method (Duval & Tweedie, 2000). Lastly, Egger's regression test was also used to assess funnel plot asymmetry (Egger et al., 1997).

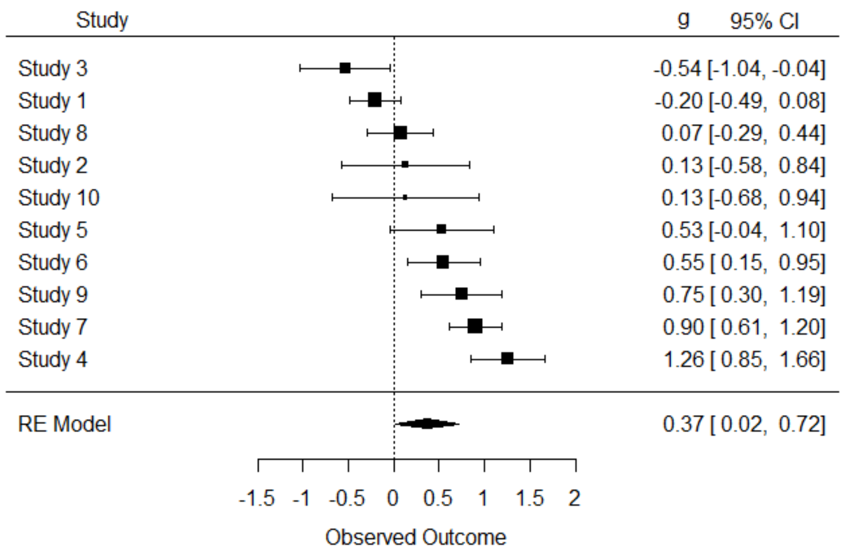

Figure 2.2 Forest plot for standardized mean difference effect sizes

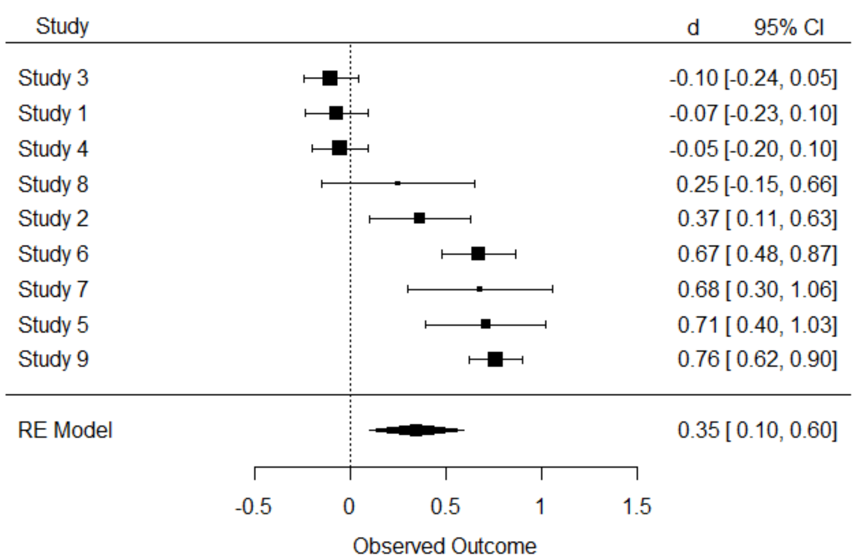

Figure 2.3 Forest plot for standardized mean change effect sizes

2.4. Results

Our meta-analysis included 19 studies published between 1980 and 2021. Of the 19 studies, 10 were group-based studies (Study# 1-10) and 9 were single-group studies (Study# 11-19); see Table 2.2 for primary study identification.

Table 2.2 Primary Study Identification

Study#	Authors	Year	N	Educational Level
1	Basham & Kotrlik	2008	199	High School
2	Duesbury & O'Neil	1996	33	Mixed
3	Erkoç, Gecü, & Erkoç	2013	62	Middle Grades
4	Kösa & Karakuş	2018	116	Post-Secondary
5	Kurtuluş & Uygan	2010	48	Post-Secondary
6	Rafi, Anuar, Samad, Hayati, & Mahadzir	2005	98	Post-Secondary
7	Šafhalter, Vukman, & Glodež	2016	196	Mixed
8	Shavalier	2004	116	Elementary
9	Toptaş, Çelik, & Karaca	2012	82	Middle Grades
10	Workman & Zhang	1999	22	Post-Secondary
11	Bairaktarova	2017	115	Post-Secondary
12	Budinoff & McMains	2018	47	Post-Secondary
13	Hilton, Paige, Williford, Li, Hammond, & Linsey	2017	146	Post-Secondary
14	Johnson & Yoon	2015	143	Post-Secondary
15	Martín-Dorta, Saorín, & Conerto	2008	40	Post-Secondary
16	Marunić & Glažar	2014	104	Post-Secondary
17	Onyancha, Derov, & Kinsey	2009	27	Post-Secondary
18	Rodriguez & Rodriguez	2016	19	Post-Secondary
19	Tumkor	2018	217	Post-Secondary

2.4.1. Standardized Mean Difference Analysis Results

We analyzed 10 standardized mean difference (SMD) effect sizes ranging from -0.20 to 1.25 (See Table 2.3). Based on results from the effect-size heterogeneity test ($Q(9) = 66.79$, $p < .0001$; $I^2 = 85.08\%$), we assumed there to be heterogeneity among the 10 effect sizes. This result was consistent with our visual inspection of the forest plot. In our assessment for publication bias, the funnel plot and Trim-and-Fill results indicated potential plot asymmetry (see Figure 2.4). However, the results from Egger's regression test ($z = -0.5362$, $p = .5918$) are not statistically significant which would suggest that publication bias is not of concern. The overall random-effects weighted effect-size estimate (g = 0.3726, p = .0367; See Table 2.4) was statistically significant.

Table 2.3 Descriptive Statistics and Effect Sizes for Spatial Ability SMD Analysis

	Experimental Group			Control Group						
Study	N	Mean	SD	N	Mean	SD	SD_{pooled}	Hedges' g	lower bound	upper bound
1	116	11.37	5.87	83	12.66	6.83	6.287	-0.204	-0.486	0.0770
2	22	11.41	4.85	11	10.82	3.49	4.457	0.129	-0.578	0.836
3	31	14.03	3.16	31	15.81	3.4	3.282	-0.536	-1.0358	-0.0352
4	72	23.07	4.95	44	15.86	6.75	5.696	1.257	0.852	1.663
5	24	18.833	6.162	24	15.875	4.73	5.493	0.530	-0.0365	1.0959
6	49	64.8	11.27	49	57.14	16.04	13.862	0.548	0.148	0.948
7	95	10.32	2.83	101	7.39	3.57	3.233	0.903	0.610	1.196
8	55	11	9.0697	61	10.312	9.305	9.194	0.0744	-0.288	0.437
9	42	5.86	2.86	40	3.7	2.86	2.86	0.748	0.304	1.192
10	10	43	16.62	12	41	12.83	14.657	0.131	-0.677	0.939

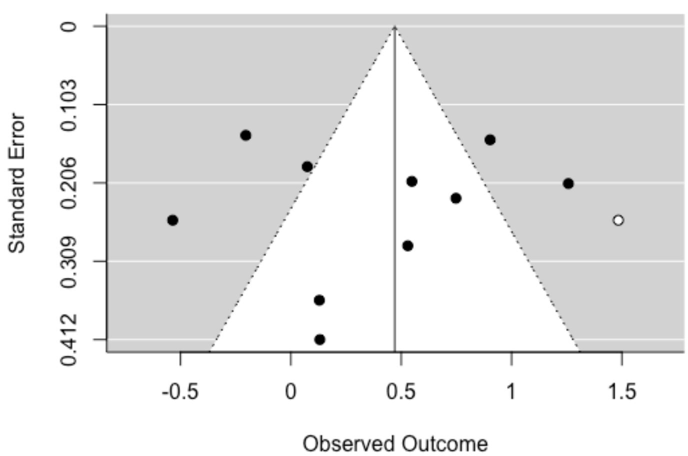

Figure 2.4 Funnel plot of SMD effect sizes. Filled circles denote observed effect sizes. Open circles denote Trim-and-Fill imputed effect sizes

Table 2.4 SMD Moderator Analysis

Moderator [Q_b]	K_j	Mean(SE)	95% CI	Q_{wj}
Overall Model	**10**	**0.37(0.18)**	**[0.02, 0.72]**	
Country [$Q_b(1) = 31.16$ $p < .0001$]				
United States	3	-0.079(0.11)	[-0.29, 0.13]	1.78
Non-United States	6	0.69(0.084)	[0.52, 0.85]	33.44**
Educational Level [$Q_b(4) = 38.89$, $p < .0001$]				
Elementary	1	0.074(0.18)	[-0.29, 0.44]	0
Middle	2	0.18(0.17)	[-0.15, 0.52]	14.16**
High School	1	-0.20(0.14)	[-0.49, 0.08]	0

Table 2.5 Continued

Moderator [Q_b]	K_j	Mean(SE)	95% CI	Q_{wj}
Post-Secondary	4	0.76(0.12)	[0.52, 1.00]	9.81*
Mixed	2	0.79(0.14)	[0.52, 1.06]	3.92*
Length of Intervention [$Q_b(1) = 7.32$, $p = .0068$]				
Equal to or less than 6 weeks	4	0.31(0.12)	[0.07, 0.55]	16.33**
Greater than 6 weeks	3	0.74(0.10)	[0.54, 0.94]	20.41**
Instrument [$Q_b(2) = 11.83$, $p = .0027$]				
PSVT	2	0.27(0.12)	[0.04, 0.50]	33.67**
MRT	3	0.13(0.12)	[-0.11, 0.38]	14.34**
Other	5	0.65(0.10)	[0.45, 0.86]	6.95
Software Type [$Q_b(2) = 25.93$, $p < .0001$]				
SketchUp	4	0.58(0.10)	[0.38, 0.78]	24.35**
AutoCAD	2	0.98(0.18)	[0.63, 1.33]	7.36*
Other	4	0.058(0.096)	[-0.13, 0.25]	9.14*

Note. *$p < 0.05$ and **$p < 0.001$.

2.4.1.1. SMD Moderator Analysis Results

For our SMD data set we examined five categorical moderators using ANOVA-like methods; results are summarized in Table 2.4.

2.4.1.1.1. Country

The categorical moderator, country, explained a statistically significant amount of effect-size heterogeneity ($Q_b(1) = 31.16$ $p < .0001$). Thus, whether or not the study took place in the United States (US) had an impact on students' spatial ability outcomes. Specifically, the mean effect size for the non-US group ($d = 0.69$, $SE = 0.084$) was much greater in magnitude than the US group ($d = -0.079$, $SE = 0.11$). However, there remained a statistically significant portion of effect-size heterogeneity left unexplained; the Non-US group demonstrated significant within-group variability ($Q_w(5) = 33.44$, $p < .0001$).

2.4.1.1.2. Educational Level

The categorical moderator, educational level, explained a statistically significant amount of effect-size heterogeneity ($Q_b(4) = 38.89$ $p < .0001$). Thus, the educational level of the study sample had an impact on students spatial ability outcomes. The mean effect size for the post-secondary ($d = 0.76$, $SE = 0.12$) and mixed ($d = 0.79$, $SE = 0.14$) populations was much greater in magnitude than the elementary ($d = 0.074$, $SE = 0.18$), middle ($d = 0.18$, $SE = 0.17$), and high ($d = -0.20$, $SE = 0.14$) school groups. Yet, there remained a statistically significant portion of effect-size heterogeneity left unexplained by the educational level moderator. The middle ($Q_w(1) = 14.16$, $p = .0002$), post-secondary ($Q_w(3) = 9.81$, $p = .0203$), and mixed ($Q_w(1) = 3.92$, $p = .0475$) groups exhibited significant within-group variability.

2.4.1.1.3. Length of Intervention

The categorical moderator, length of intervention, explained a statistically significant amount of effect-size heterogeneity ($Q_b(1) = 7.32$, $p = .0068$). Thus, the length of the intervention (6 weeks or less verse greater than 6 weeks) had an impact on students spatial ability outcomes. Particularly, the mean effect size for the greater than 6 weeks group ($d = 0.74$, $SE = 0.10$) was greater in magnitude than the equal to or less than 6 weeks group ($d = 0.31$, $SE = 0.12$). However, there remained a statistically significant portion of effect-size heterogeneity left unexplained; the equal to or less than group demonstrated significant within-group variability ($Q_w(3) = 16.33$, $p = .0010$).

2.4.1.1.4. Instrument

The categorical moderator, instrument, explained a statistically significant amount of effect-size heterogeneity ($Q_b(2) = 11.83$, $p = 0.0027$). Thus, the spatial ability instrument used had an impact on students' spatial ability outcomes. The mean effect size for the group of studies that used the PSVT ($d = 0.27$, $SE = 0.12$) and the MRT ($d = 0.13$, $SE = 0.12$) was lower in magnitude than the studies that used other instruments ($d = 0.65$, $SE = 0.10$). Yet, there remained a statistically significant portion of effect-size heterogeneity left unexplained by the instrument moderator. The PSVT ($Q_w(1) = 33.66$, $p < .0001$) and MRT ($Q_w(2) = 14.34$, $p = .0008$) groups exhibited significant within-group variability.

2.4.1.1.5. Software Type

The categorical moderator, software type, explained a statistically significant amount of effect-size heterogeneity ($Q_b(2) = 25.93$, $p < .0001$). Thus, the CAD software type used had an impact on students' spatial ability outcomes. The mean effect size for the studies that used SketchUp ($d = 0.58$, $SE = 0.10$) and AutoCAD ($d = 0.98$, $SE = 0.18$) was much higher in magnitude than studies that used other software types ($d = 0.058$, $SE = 0.096$). Yet, there remained a statistically significant portion of effect-size heterogeneity left unexplained by the software type moderator. SketchUp ($Q_w(3) = 24.35$, $p < .0001$), AutoCAD ($Q_w(1) = 7.36$, $p = .0067$), and other ($Q_w(3) = 9.14$, $p = .027$) groups exhibited significant within-group variability.

2.4.2. Standardized Mean Change Analysis Results

We analyzed 9 standardized mean change (SMC) effect sizes ranging from -0.04 to 0.76 (See Table 2.5). Based on results from the effect-size heterogeneity test ($Q(8) =$

137.21, $p < .0001$; $I^2 = 92.69\%$), we assumed there to be heterogeneity among the 9 effect sizes. This result was consistent with our visual inspection of the forest plot. In our visual assessment for publication bias, the funnel plot indicated potential plot asymmetry (see Figure 5). However, the trim-and-fill and not statistically significant Egger's regression test (z = 1.0445, $p = 0.2963$) would suggest publication bias is not indicated. The overall random-effects weighted effect size estimate ($d = 0.3488$, $p = .0056$; See Table 2.6) was statistically significant.

Table 2.6 Descriptive Statistics and Effect Sizes for Spatial Ability SMC Analysis

		Pre		Post				
Study	N	Mean	SD	Mean	SD	*d*	lower bound	upper bound
11	115	0.77	0.15	0.76	0.12	-0.06645	-0.22961	0.096712
12	47	20.57	5.37	22.57	5.72	0.369395	0.105184	0.633606
13	146	23.29	4.97	22.8	5.13	-0.09834	-0.24348	0.046808
14	143	24.6	4	24.4	4.6	-0.04987	-0.19619	0.096455
15	40	19.03	7.6	24.5	8.5	0.712794	0.397703	1.027885
16	104	11.1	4	13.8	3.9	0.672539	0.478562	0.866517
17	27	48.7	14.49	58.7	14.49	0.680129	0.30263	1.057628
18	19	22.84	4.48	24	4.58	0.253497	-0.14807	0.655064
19	217	60	17	73	16	0.763378	0.624622	0.902133

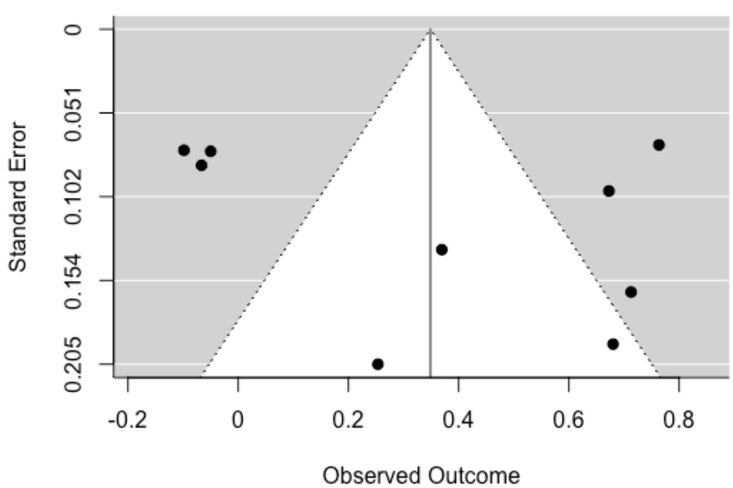

Figure 2.5 Funnel plot of SMC effect sizes

Table 2.7 SMC Moderator Analysis

Moderator [Q_b]	K_j	Mean(SE)	95% CI	Q_{wj}
Overall Model	**9**	**0.35(0.13)**	**[0.10, 0.60]**	
Country [$Q_b(1) = 28.31$, $p < .0001$]				
United States	7	0.20(0.035)	[0.13, 0.27]	108.86**
Non-United States	2	0.68(0.084)	[0.52, 0.85]	0.046
Instrument [$Q_b(1) = 112.40$, $p < .0001$]				
PSVT	6	0.014(0.040)	[-0.07, 0.09]	24.25**
MRT	3	0.73 (0.054)	[0.62, 0.84]	0.57
Software Type [$Q_b(3) =17.89$, $p = .0005$]				
SketchUp	1	0.71(0.16)	[0.40, 1.03]	0
AutoCAD	1	0.25(0.20)	[-0.15, 0.66]	0
Solidworks	2	0.054(0.071)	[-0.09, 0.19]	7.57*
Other	5	0.31(0.038)	[0.23, 0.38]	111.76**

Note. *$p < 0.05$ and **$p < 0.001$.

2.4.2.1. SMC Moderator Analysis Results

For our SMC data set we examined three categorical moderators using ANOVA-like methods; results are summarized in Table 2.6. The educational level and length of intervention moderators were not examined for our SMC data set because they were the homogeneous across studies.

2.4.2.1.1. Country

The categorical moderator, country, explained a statistically significant amount of effect-size heterogeneity ($Q_b(1) = 28.31$ $p < .0001$). Thus, whether or not the study took place in the United States (US) had an impact on students' spatial ability outcomes. Specifically, the mean effect size for the non-US group ($d = 0.68$, $SE = 0.084$) was greater in magnitude than the US group ($d = 0.20$, $SE = 0.035$). However, there remained a statistically significant portion of effect-size heterogeneity left unexplained; the Non-US group demonstrated significant within-group variability ($Q_w(5) = 108.86$, $p < .0001$).

2.4.2.1.2. Instrument

The categorical moderator, instrument, explained a statistically significant amount of effect-size heterogeneity ($Q_b(1) = 112.40$, $p < .0001$). Thus, the spatial ability instrument used had an impact on students spatial ability outcomes. The mean effect size for the PSVT ($d = 0.014$, $SE = 0.040$) group was much lower in magnitude than the MRT ($d = 0.73$, $SE = 0.054$) group. There remained a statistically significant portion of effect-size heterogeneity left unexplained; the PSVT group exhibited significant within-group variability ($Q_w(5) = 24.25$, $p = .0002$).

2.4.2.1.3. Software Type

The categorical moderator, software type, explained a statistically significant amount of effect-size heterogeneity ($Q_b(2) = 17.89$, $p = .0005$). Thus, the CAD software type used had an impact on students spatial ability outcomes. The mean effect size for the studies that used SketchUp ($d = 0.71$, $SE = 0.16$) was much higher in magnitude than studies that used AutoCAD ($d = 0.25$, $SE = 0.20$), Solidworks ($d = 0.054$, $SE = 0.071$), and other software types ($d = 0.31$, $SE = 0.038$). There remained a statistically significant portion of effect-size heterogeneity left unexplained by the software type moderator. The solidworks ($Q_w(1) = 7.57$, $p = .0067$) and other ($Q_w(4) = 9.14$, $p = .027$) groups exhibited significant within-group variability.

2.5. Discussion

Spatial ability is an important factor in mathematical understanding and achievement and can be developed or improved through various technologies. Computer-aid design software are an excellent way to increase technology integration in the mathematics classroom and have been shown to have positive effects on students' spatial ability. The purpose of this study was to synthesize empirical research on the implementation of CAD software as an intervention in various educational settings as well as investigate how the effectiveness of the intervention may be impacted by study characteristics. From the primary studies included, two separate meta-regression analyses were conducted to analyze the mean difference effect size of spatial ability differences between CAD and non-CAD groups, and the mean change effect size of spatial ability differences between pre & post of CAD groups. There were 19 studies (10 group-based and 9 single-case) that met selection criteria to best represent the impact CAD software had on students' spatial ability.

The calculated weighted effect size of 0.37 for the group-based studies and 0.35 for the single-group studies indicates the CAD interventions had a positive impact on students' spatial ability. Many mathematical concepts and skills can be learned through geometrical construction and modeling in both two and three dimensions (Huleihil, 2016) and the use of CAD systems were shown to improve students' drawing capabilities of two and three-dimensional objects (Martin & Velay, 2010). These drawing activities showed improvement of spatial skills of learners (Erkoc et al., 2013) which correlates with retention and achievement (Onyancha et al., 2009). Chang (2014) found that CAD greatly enhanced creative performance and inspired spatial creativity.

Falcón (2011) indicated that the use of CAD systems engages the students and improves attitudes towards mathematics. Researchers found that a semester long course using CAD software improved students' spatial abilities (Kösa & Karakus, 2018; Onyancha et al., 2009) and even in just a six hours course (Contero et al., 2006). Overall, the research on CAD systems implies that the use of CAD systems in the classroom enhances students' spatial abilities and may increase student engagement. Throughout all 19 articles the concept of spatial visualization, geometric knowledge, and student engagement is persistent. CAD systems are an excellent teaching tool. With CAD systems students are learning several mathematical concepts but also engaging in advanced technologies. Using CAD systems can increase spatial ability, which could have a positive impact on STEM achievement.

3. REVIEW OF THE REDRAWN MENTAL ROTATION TEST: A RELIABILITY GENERALIZATION META-ANALYSIS OF COEFFICIENT ALPHA

Spatial abilities play an essential role in mathematical understanding and achievement. Spatial abilities allow individuals to use concepts of shape in concrete and abstract ways to create and utilize objects in the world, to navigate, and to communicate (Basham & Kotrlik, 2008). The mental rotation of objects is a fundamental spatial ability that affects several aspects of life. Mental rotation is associated with the ability to mentally rotate images or objects into various orientations. Advanced mental rotation is depicted by the ability to quickly and accurately mentally rotate two- or three-dimensional figures (Linn & Petersen, 1985) to confirm how they would look from different angles or perspectives (Moé, 2018). Mental rotation abilities have been identified as a predictor of selection and success in science, technology, engineering, and mathematical (STEM) career paths (Chang, 2014; Von Károlyi, 2013). Since the mid-20th century, researchers in various fields have been studying mental rotation abilities at all stages of life, resulting in the creation of several instruments to measure this ability, the most widely used of which is the Mental Rotation Test (MRT) developed by Vandenburg and Kuse in 1978. The MRT was among the first paper-and-pencil group tests to measure mental rotation abilities, thus arguably one of the premier tools used to measure spatial ability. Given the implications of spatial ability presented above, it is imperative that the measurement of spatial ability is effective, efficient, and most of all reliable. Hence, the aim of the present study was to conduct a reliability generalization meta-analysis of the Cronbach's alpha coefficients reported in extant literature using the

MRT. Furthermore, to our knowledge this is the first reliability generalization conducted on the MRT instrument.

In the remaining sections, we first provide a brief background on the development and evolution of the MRT. Then, we make a case for the importance of the current study by reviewing the literature on reliability measurement and reliability generalization meta-analysis, before presenting the procedures used to conduct the present study in the methods section. We conclude by unpacking the results of the present study and providing implications for future work and next steps in the discussion.

3.1. Background

3.1.1. The Mental Rotation Test

The Vandenburg and Kuse MRT was based off of Shepard and Metzler's (1971) mental rotation task using three-dimensional objects, is one of the most widely used measures of mental rotation abilities. However, after several years of use and repeated copying, the image quality of the MRT declined and it was deemed necessary by Peters et al. (1995) to redraw the test. Peters et al. (1995) used computer software to redraw the original images from the MRT, creating the MRT-A. The MRT-A is a paper-and-pencil test comprised of 24 items. Each item contains five three-dimensional drawings of cubical figures which includes one target figure on the left and four answer choices on the right. An example of the items and instructions for the MRT-A and a practice item can be seen in Figure 3.1.

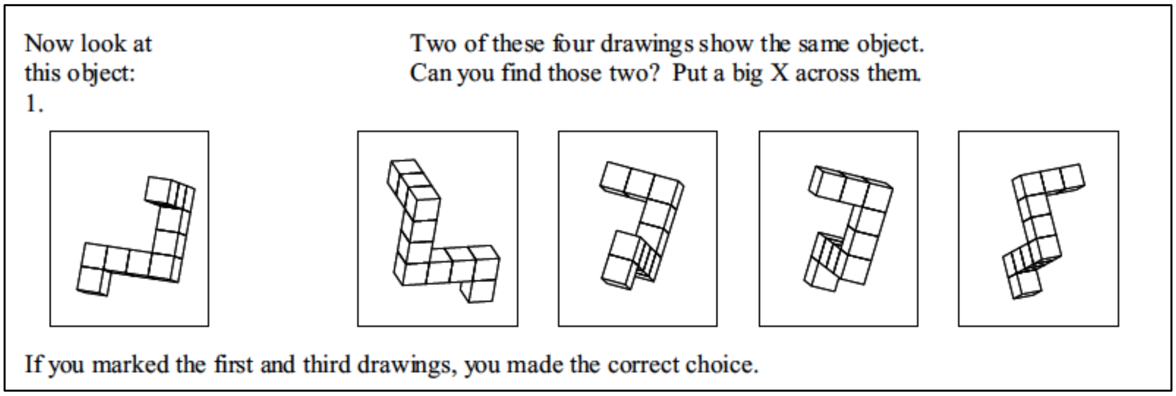

Figure 3.1 Practice Item Included on MRT-A

The task presented in the example involves identifying which 2 of the answer choices are identical to the target but rotated along the y-axis. The 2 remaining answers are mirror-images of the target and cannot become identical to the target by rotation. The MRT-A is designed to be administered in 2 parts with 3 minutes to complete the first 12 items, followed by a 2-3-minute break, and then another 3 minutes to complete the remaining 12 questions. The test is traditionally scored by assigning one point to an item if, and only if, both correct answer choices are selected (maximum score of 24 points). Alternatively, scores can be calculated by assigning one point to each correct figure chosen and subtracting one point for each incorrect figure selected (maximum score of 48).

Previous research has shown that sex has been an influential factor on MRT scores (Voyer & Saunders, 2004; Voyer & Doyle, 2010) and sex differences have remained robust across several published studies (Peter et al., 1995). Sex differences in spatial performance has been shown to increase with age (Linn & Petersen, 1985; Voyer et al., 1995) and therefor age could also influence MRT performance. Although the MRT was designed for participants of age 14 and older (Vandenberg & Kuse, 1978) it has been shown that it can reliably measure mental rotation performance in primary school-aged children (Titze et al., 2010). However, the instrument items are typically modified, replacing original images, when used with younger children (e.g., Hawes et al., 2015). Additionally, it has been shown that the conventional scoring method (one point per item if and only if both are correct) has also contributed to sex differences in performance (Voyer et al., 1995). This suggest that sex, age, and scoring procedures could potentially impact MRT score reliability.

Although the MRT by Vandenburg and Kuse (1978) is well-established, validated, and has shown relatively high reliability scores across settings, there exist the need to empirically assess the reliability and validity of the MRT-A for several reasons. First, Peters et al. (1995) simply reports the reliability estimates from Vandenburg and Kuse (1978), failing to report score reliability for the newly redrawn version of the instrument in various settings. Second, though it was stated that the MRT-A was exactly redrawn from the original MRT, inconsistencies were found in the literature concerning the number of items on the instruments. The MRT-A consists of 24 items while the original MRT consists of only 20 items. An earlier researcher (Bryant, 1982) used a 24-item MRT, however, they too merely cited Vandenburg and Kuse (1978) and the source of the four additional items has not been found. Finally, the MRT-A is now the most commonly used measure for mental rotation abilities, and as such it is important that the reliability estimates for this specific version of the test be reported. Synthesizing previously reported MRT-A reliability estimates across studies provided a weighted average of score reliability estimates for the instrument as well as allow for the examination of characteristics that may impact reliability estimates across studies.

3.1.2. Reliability

The misuse of the term reliability in educational and psychological research still exists. Far too often authors inaccurately state that an instrument is reliable, which has led to the misconception that reliability is a characteristic of an instrument, when really it is a property of the scores produced by that instrument (Thompson & Vacha-Haase, 2000). Reliability illustrates to what extent scores yielded by an instrument administered to a target population, at a particular time, and under certain conditions are consistent and reproducible (Crocker & Algina, 1986; Onwuegbuzie & Larry, 2000). Therefore, the reliability of scores produced by an instrument can be influenced by study and sample characteristics.

As a consequence, researchers have underestimated the importance of reporting reliability (Cousin & Henson, 2000) and fail to provide reliability estimates for data collected (Holland, 2015), or they simply cite previously reported reliability estimates, a practice known as reliability induction (Vacha-Haase et al., 2000). Reliability estimates fluctuate across administrations (Crocker & Algina, 1986; Vacha-Haase et al., 2002; Vacha-Haase & Thompson, 2011) because anything that could potentially affect scores could also affect reliability (Barnes et al., 2002). Given the diversity across studies (Vacha-Haase, 1998) and importance of score reliability in all quantitative analyses

(Vacha-Haase & Thompson, 2001) it is pertinent that authors report reliability coefficients for their data. The absence of these data can create several challenges such as limiting the generalizability and reducing the likelihood of study replication (Williams & Young, 2021).

In attempts to improve reliability reporting practices and highlight the importance of score reliability, a methodological approach emerged to explore and examine reliability estimates across studies. This approach is most commonly known as a reliability generalization (Vacha-Hasse, 1998) and provides a means for researchers to illustrate the integrity of scores produced by an instrument as well as characterize study features that may predict variations in score quality.

3.1.3. Reliability Generalization

Reliability generalization (RG) studies are employed to generalize the reliability of scores produced by an instrument across various administrations, identify study or sample characteristics which may cause variability across score reliabilities, and determine circumstances in which score reliability may become unacceptable (Caruso, 2000; Vacha-Hasse, 1998; Vacha-Haase et al., 2002; Vacha-Haase & Thompson, 2011). Given there is not yet a generally accepted set of procedures for conducting RG studies (Botella et al., 2010), a variety of approaches and statistical procedures have been applied (e.g., descriptive statistics analysis, correlational analysis, analysis of variance (ANOVA), and multiple regression). However, the application of meta-analytic techniques in RG studies have been most commonly used to estimate an average score reliability across studies, determine variation in reliability, and identify possible moderator variables that may be influencing score reliability (Caruso, 2000; Kieffer et al., 2010). Therefore, a meta-analytic approach was utilized in the present RG study. In RG, with the goal of identifying sources of measurement error across studies using the same instrument, a study becomes the component of analysis, the reliability coefficient becomes the dependent variable, and scale, study, or sample characteristics become possible predictors (Cousin & Henson, 2000). Results from RG studies can provide the researcher with insights into overall score reliability in prior applications as well as sources of error which can assist researchers in future planning and study design decisions (Henson & Thompson, 2002). Prior to conducting RG, several methodological and statistical decisions have to be made; many of those decisions first rely on which reliability coefficient type will be collected and reported. An assortment of statistical tests exists to measure score reliability (e.g. test-retest, split-half, Kuder-Richardson), however, the most common reliability coefficient used is Cronbach's (1951) coefficient alpha (α) — the average of intercorrelations between items (Kaplan & Saccuzzo, 2012). In the present RG study, the analysis of previous applications of the MRT-A was guided by the following questions: (1) What is the overall reliability, using Cronbach's alpha, of the redrawn mental rotation test scores? (2) What is the risk of publication bias? (3) How heterogenous are reliability estimates across various applications of the redrawn mental rotation test? (4) To what extent do primary study characteristics including region, mean age, scoring procedures, instrument modifications, and percent female moderate the average weighted Cronbach's alpha reliability estimate?

3.2. Methods

The Peters et al. (1995) MRT-A has been heavily utilized by researchers in various fields and with various age groups to measure participant's mental rotation abilities. Several researchers (i.e., Erkoc et al., 2013; Quaiser-Pohl et al., 2006; Turgut, 2015) have reported that the selection of the MRT-A was due to the high reliability yet do not report reliability based on respective researchers' own scores from their primary study and rely on previously published reliability estimates. The goal of the present RGM is to synthesize previously reported MRT-A reliability estimates across studies in order to provide an expected range of score reliability for the instrument as well as identify characteristics which may impact reliability estimates across studies.

3.2.1. Literature Search

The first step in the process was to identify articles that utilized the MRT-A instruments to measure participants' mental rotation abilities. First, all articles that cited Peters et al. (1995) were identified (*n*=1122) through Google Scholar. Then, a search within citing articles was conducted using the following keywords and Boolean operators: *"mrt"* AND *"reliability"* OR *"internal consistency"* OR *"Cronbach's alpha"* OR *"coefficient alpha."* This resulted in 314 hits which were included in the full-text analysis.

3.2.2. Data Collection

Following the literature search a full-text analysis was conducted to identify studies for inclusion. The final selection phase examined full-text of articles for inclusion based on the following criteria: (1) reported in English, (2) reported the use of

MRT-A, and (3) reported coefficient alpha score reliability based on data collected. For inclusion criteria (2) studies were included only if the standard 24 item MRT-A test form and/or a shortened 12 item version was used. Any other modifications made to the test form (e.g., changes to the pictures) disqualified the study from inclusion with the exception of 1 study that had the test instructions translated to German (Glück et al., 2007). Figure 3.2 provides a detailed flowchart of the inclusion and exclusion decisions that lead to the final dataset.

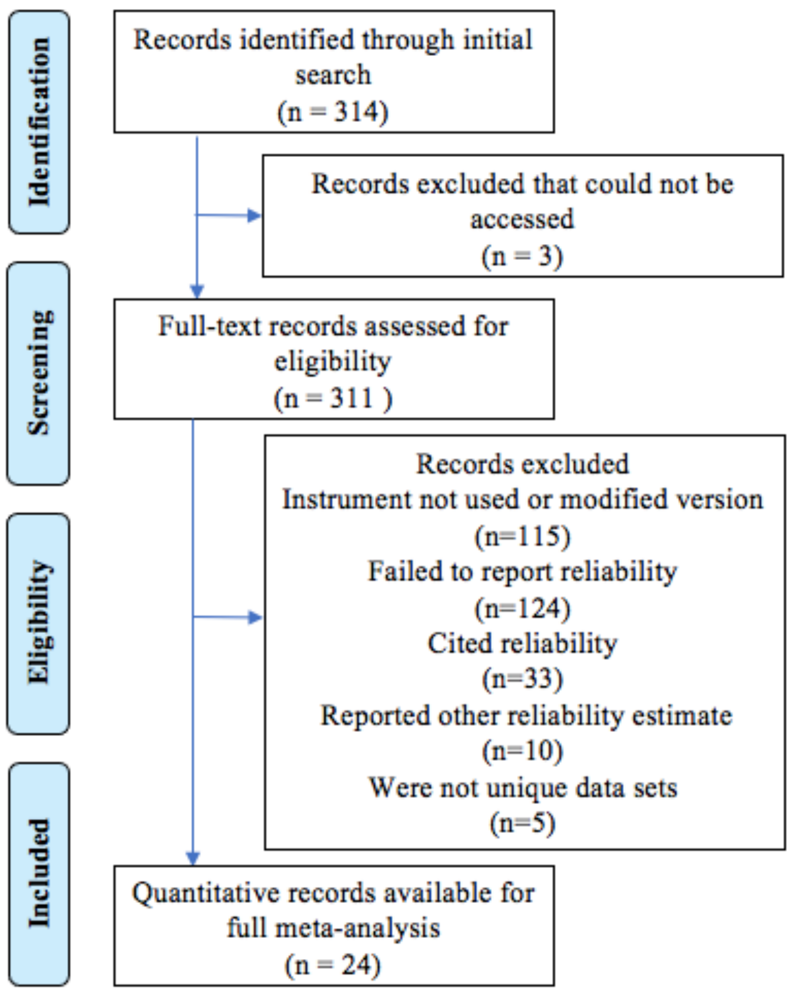

Figure 3.2 Flowchart of inclusion and exclusion decisions

3.2.3. Coding Procedures

After identifying the 24 primary studies to be included a 20-item coding form was used by 2 coders to gather information on important variables. The form was used to collect the following information: author(s), year, article type, study design, region (North America vs other), setting (educational vs clinical), sample size, mean age, sample type (adolescents, college students, adults), alpha for pre-test, alpha for post-test, scoring procedures, administration modifications, percent female, sample size, group means (pre & post), group standard deviations (pre & post), and scale variance. The initial overall agreement rate between the two coders was 88%. All discrepancies were examined, and complete consensus was reached.

3.2.4. Moderator Variables

The study characteristics assessed as moderators included region (dichotomous variable indicating whether the study was conducted in North America or another region), items (dichotomous variable indicating whether the test administered was the full 24 item version or a shortened 12 item version), scoring (dichotomous variable indicating whether the test were scored using 24 or 48 point maximum), sex (continuous variable indicating the percentage of participants that identified themselves as female), and age (continuous variable indicating the average age of each sample).

3.2.5. Transformation of Cronbach's Alpha

Cronbach's alpha is both bounded and not normally distributed; accordingly, Cronbach's alpha estimates violate the assumption that effect sizes are normally distributed in a meta-analysis (Rodriguez & Maeda, 2006). Therefore, prior to modeling, a transformation of the alpha coefficients is necessary. The Hakstian-Whalen (1976), Fisher's *r* to *Z*, and Bonett transformation (Bonett, 2002) are all transformation methods that are used in the RG literature (Semma et al., 2019). Markedly, Fisher's *r*-to-*Z* transformation is not recommended for the transformation of alpha (Henson & Thompson, 2002; López-López et al., 2013), so both the Bonett and Hakstian-Whalen transformations was conducted. The formulas for transformation and variances (see Table 3.1) are suggested by Sánchez-Meca and associates (2013) among other sources. The D'Agostino (1970) test of skewness & Anscombe-Glynn test for kurtosis (Anscombe & Glynn, 1983) were employed to determine which transformation was more effective in normalizing the effect-size distribution (skewness close to 0 and kurtosis close to 3).

Table 3.1 Formulas for Transformation of Cronbach Alpha.

Transformation	Coefficient	Back-transformation	Sampling Variance
Hakstian-Whalen	$T_i = \sqrt[3]{1 - \hat{\alpha}_i}$	$\hat{\alpha}_i = 1 - T_i^3$	$V(T_i) = \frac{18J_i(n_i - 1)(1 - \hat{\alpha}_i)^{2/3}}{(J_i - 1)(9n_i - 11)^2}$
Bonett	$L_i = \ln(1 - \lvert\hat{\alpha}_i\rvert)$	$\hat{\alpha}_i = 1 - e^{L_i}$	$V(L_i) = \frac{2J_i}{(J_i - 1)(n_i - 2)}$

Note. $\hat{\alpha}_i$: coefficient alpha for the *i*th study. n_i: sample size for the *i*th study. J_i: number of items of the test used in the *i*th study.

3.2.6. Statistical Analysis

After transformation of Cronbach alpha coefficients, a random-effects model was utilized to calculate a weighted average of the transformed alpha coefficients and assess initial heterogeneity across studies. The estimated overall weighted mean and confidence intervals was then back transformed to provide a comparable overall reliability coefficient. Homogeneity of the transformed alpha coefficients was assessed by visually examining forest plots (see Fig. 3) as well as examining the Q statistic, I^2, and $\hat{\tau}^2$ values provided. After assessing homogeneity, we elected to use a random-effects model to compute overall transformed alpha coefficients to account for within-study sampling error and between-studies heterogeneity. Categorical and continuous moderator analyses were conducted to examine the potential relationships of study-level characteristics on the heterogeneity of effect-sizes using two types of mixed effects models. When analyzing categorical moderators, ANOVA-like models were used, while meta-regression was used when assessing continuous moderators. Publication bias was visually assessed using funnel plots which included imputations for potentially missing effects with the trim-and-fill method (Duval & Tweedie, 2000). Egger's regression test was also used to assess funnel plot asymmetry (Egger et al., 1997). All statistical analyses were conducted using the statistical software R (R Core Team, 2017) using the packages *metafor* (Viechtbauer, 2010), *moments* (Komsta & Novometsky, 2015), *ggpubr* (Kassambara, 2018), and *Hmisc* (Harrell, 2020).

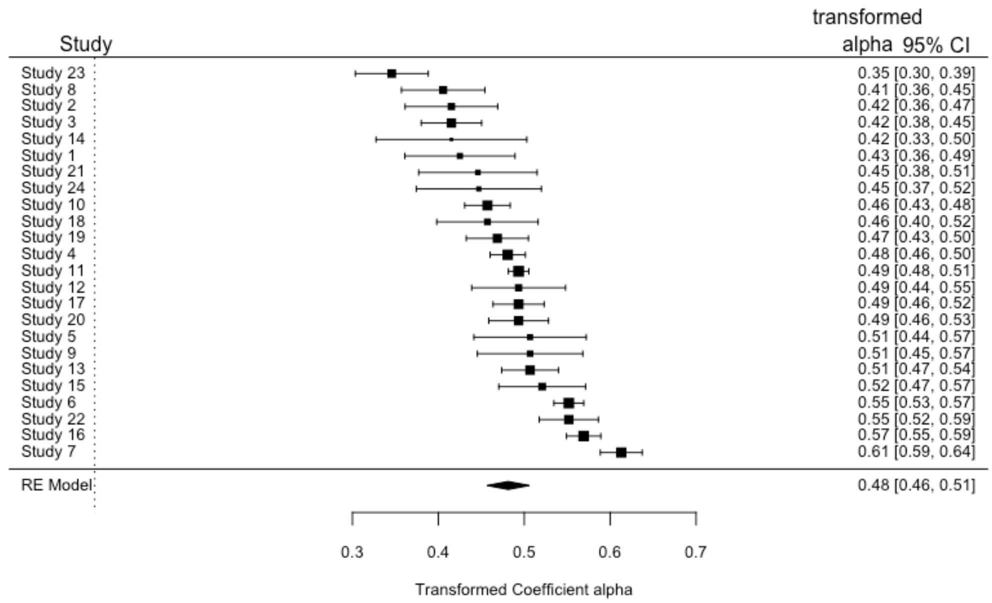

Figure 3.3 Forest plot for transformed alpha coefficients

3.3. Results

Alpha coefficients were reported for 24 unique samples across the included 24 articles ranged from 0.72 to 0.94. Study and sample characteristics, including moderator variables and reported alpha coefficients are shown in Table 3.2.

Table 3.2 Primary Study Characteristics

Study #	Authors	Year	Region	Mean Age	Sex	n	Scoring	Mean	SD	Alpha
1	Atit & Rocha	2020	0	-	74.39	82	0	6.78	4.32	0.81
2	Atit et al.	2013	0	20.6	67.24	116	0	8.16	-	0.8
3	Berkowitz	2017	1	19.5	14	254	0	16.1	3.96	0.8
4	Berkowitz et al.	2021	1	21.25	49.7	593	0	13.4	4.92	0.86
5	Bogomolova et al.	2020	1	-	-	58	0	7.12[a]	-	0.88

Notes. Author denoted with * = used 12 item version of the test; Region 0 = North America, 1= Other region; scoring 0 = 24 point max, 1 = 48 point max; mean denoted with [a] are weighted group means

Table 3.3 Continued

Study #	Authors	Year	Region	Mean Age	Sex	n	Scoring	Mean	SD	Alpha
6	Bors & Vigneau	2011	0	21	65.22	624	0	11.46[a]	-	0.91
7	Brase & Hill	2017	0	19.9	47.26	237	0	11.77	-	0.942
8	Casey et al.*	2017	0	6.72	100	138	1	15.96	4.19	0.79
9	Casselman et al.	2021	0	-	75	64	0	8.39[a]	-	0.88
10	Doyle et al.	2012	0	19.63	67	403	0	9.6[a]	-	0.84
11	Geiser et al.	2006	1	16.8	50.21	1695	0	11.1[a]	-	0.87
12	Georges et al.	2017	1	23.38	49.38	81	0	13.23	5.35	0.87
13	Gignac et al.	2016	1	19.8	68	211	0	10.77	5.01	0.88
14	Glück et al.	2007	1	-	7.14	42	0	15.5[a]	-	0.8
15	Hart et al.	2017	0	19.75	43	85	0	12.66	5.12	0.89
16	Hone et al. *	2020	0	18.94	55.61	419	0	14.4	-	0.92
17	Leon*	2014	0	31.52	100	256	1	36.74[a]	-	0.87
18	Miller & Halpern	2013	0	18.17	36.36	77	0	23.35	-	0.84
19	Newbold	2020	0	18.56	71.1	211	0	8.27	4.93	0.85
20	Osapczuk et al.	2014	1	19.7	66	201	0	11.15	5.07	0.87
21	Sanandaji et al.*	2021	0	-	41.67	60	1	17.5[a]	-	0.83
22	Voyer et al.	2006	0	19.66	52.87	157	0	14.31[a]	-	0.91
23	Yurt & Tünkler	2016	1	-	54.7	234	0	7.85	4.27	0.72
24	Zahner et al.*	2017	0	16.9	60	64	-	4.7	2.9	0.831

Notes. Author denoted with * = used 12 item version of the test; Region 0 = North America, 1= Other region; scoring 0 = 24 point max, 1 = 48 point max; mean denoted with [a] are weighted group means

The raw alpha values were transformed using the Hakstian-Whalen transformation. The distribution of transformed Cronbach's alphas appeared to be normal distributed of transformed Cronbach alpha with skewness of 0.09 and a kurtosis of 2.98. The RE model provided and overall estimate of the weighted average Cronbach's alpha of $\hat{\mu}$ = .86 with a 95% confidence interval (CI) of [.84, .88]. There was a sizeable amount of effect-size heterogeneity, given Q (23) = 300.86 p < .0001, $\hat{\tau}^2$ = .0033, and I^2 = 93.05%. In our assessment for publication bias, the funnel plot and Trim-and-Fill results indicated potential plot asymmetry (see Figure 3.4). However, the results from Egger's regression test (t = -1.8924, p = 0.0717) are not statistically significant which would suggest that publication bias is not of concern. For our moderator analysis we examined three categorical moderators using ANOVA-like methods and two continuous moderators using meta-regression; results are summarized in Table 3.3.

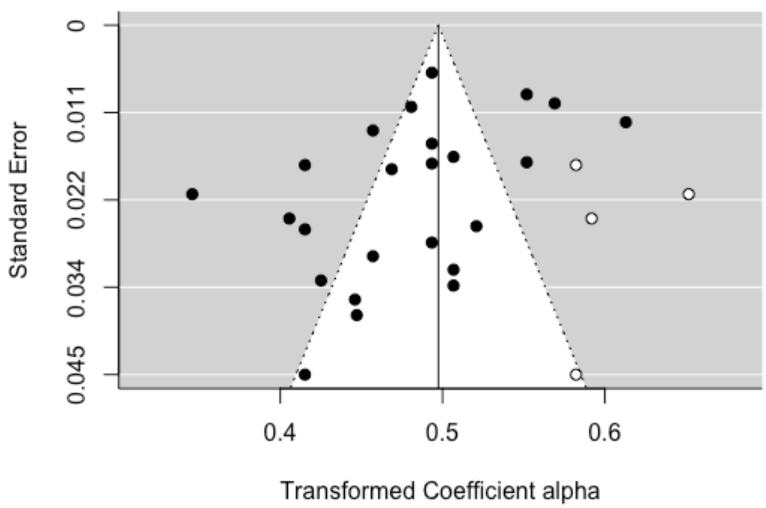

Figure 3.4 Funnel plot of transformed alpha coefficients. Filled circles denote observed alpha coefficients. Open circles denote Trim-and-Fill imputed coefficients

Table 3.4 Moderator Analysis Results

Moderator [Q_b]	K_j	Mean(SE)	95% CI	Q_{wj}
Overall Model	**24**	**0.86(0.038)**	**[0.84, 0.88]**	
Region [$Q_b(1)$ = 55.8864 p < 0.0001]				
North America	15	0.89(0.013)	[0.89, 0.90]	182.53**
Other Regions	9	0.86(0.013)	[0.85, 0.87]	62.44**
Items [$Q_b(1)$ = 7.49, p = 0.0062]				
24-item	19	0.88(0.010)	[0.87, 0.88]	237.63**
12-item	5	0.89(0.023)	[0.88, 0.90]	55.74**

Table 3.5 Continued

Moderator [Q_b]	K_j	Mean(SE)	95% CI	Q_{wj}
Scoring [$Q_b(1) = 10.5187$, $p = 0.0012$]				
24 point max	20	0.88(0.010)	[0.875, 0.88]	278.42**
48 point max	3	0.85(0.036)	[0.83, 0.87]	9.53*
	K	Coefficient(SE)	95% CI	Q_e
Sex [$Q_m(1) = 0.0204$, $p = 0.8865$]	23	0.0003(0.0018)	[-0.0036, 0.0042]	303.32**
Age [$Q_m(1) = 1.367$, $p = 0.2424$]	18	0.011(0.009)	[-0.007, 0.029]	227.74**

Note. *p < 0.01 and **p < 0.001. Kj is the number of studies in the jth group. CI = Confidence Interval.

The categorical moderators, region ($Q_b(1) = 55.88$ $p < 0.001$), items ($Q_b(1) = 7.49$ $p = 0.0062$), and scoring ($Q_b(1) = 10.52$, $p = 0.0012$) all explained a statistically significant amount of heterogeneity. Thus, whether or not the study took place in the North America, the number of items on the test (12 or 24), and the score procedures (24 or 48 point max) all had an impact on reliability estimates. However, the weighted mean reliability estimates across the groups were close in value ranging from .85-.89. In addition, there still remained a statistically significant portion of heterogeneity left unexplained; all of the groups demonstrated significant within-group variability. The moderators sex (percent female) and age (mean age) were analyzed as continuous moderators. Sex ($Q_m(1) = 0.020$, $p = 0.89$) and age ($Q_m(1) = 1.37$, $p = 0.24$) did not explain a statistically significant portion of the variability and there remained a significant portion of heterogeneity left unexplained.

3.4. Discussion

The primary objectives of this RG study were to determine the reliability of the redrawn Mental Rotation Test (MRT-A; Peters et al., 1995), assess heterogeneity in reliability estimates, and look at specific sample characteristics as moderators. Overall, the MRT-A was found to be reliable across samples, with a weighted average Cronbach's alpha of .86 and a small 95% confidence interval. A statistically significant amount of heterogeneity was observed. A moderator analysis led to conclusions that the region from which the study was conducted, the number of test items used, and the scoring procedure was a statistically significant predictor of variability. However, sex and age were not statistically significant predictors of variability. Additionally, there remained a significant portion of heterogeneity left unexplained which would imply there exist other factors that are impacting the variability of reliability estimates across samples. Investigating additional models and moderators in attempts to locate the source(s) of variability across samples is difficult due to poor reporting practices across studies (e.g., not reporting study type, sample demographics, or standard deviation of observed scores).

Our findings highlight the alarming issue that too many researchers continue with poor practices, failing to report reliability estimates for their samples. Of the 191 unique samples examined that administered the MRT-A (meeting inclusion criteria 1 and 2), less than 18% reported a reliability coefficient for their sample. About 65% of the studies failed to report any form of reliability for their samples, and about 17% cited previously reported reliability. Citing and using prior reliability estimates for current data has been termed reliability induction (Vacha-Haase et al., 2000). Our findings provide further evidence of the importance of reporting score reliabilities rather than inducting reliability from other publications.

4. FOSTERING ELEMENTARY STUDENTS' SPATIAL ABILITY DEVELOPMENT WITH COMPUTER-AIDED DESIGN SOFTWARE

4.1. Introduction

The rich discussion of how to increase science, technology, engineering and mathematics (STEM) interest and success is ubiquitous in educational research and institutions. There exists great potential to enhance STEM education through the development of students' spatial abilities (Khine, 2017). Spatial ability is a fundamental cognitive resource for STEM learning and has been identified as a marker for success in STEM fields (Berkowitz & Stern, 2018). Although spatial ability has been revealed to contribute validity to mathematical and verbal reasoning skills, there has been limited focus placed on the development of spatial abilities in schools (Basham & Kotrlik, 2008). Spatial training has been identified as a significant gap in K–12 curriculum by the National Research Council's report (2006). The National Council of Teachers of Mathematics' (2000) Geometry Standards recommend that the development of spatial skills be included in early childhood education and that these should be facilitated through hands-on experiences with a variety of geometric objects making use of technology to dynamically transform simulations of two- and three-dimensional (3D) objects.

Spatial ability was once thought of as a static element of intelligence, but it has been confirmed, by various researchers, that development and improvement of spatial ability is possible given appropriate experiences, opportunities, and practice (Cakmak et al., 2014; De Lisi & Wolford, 2002; Ogunkola & Knight, 2019; Uttal & Cohen, 2012).

Research has shown that spatial abilities can be developed and enhanced through rich experiences with 3D modeling programs (Matthews & Geist, 2002) such as computer-aided design (CAD) software. Activities utilizing CAD software naturally situate learners in an environment centered around design and the engineering process. Therefore, purposefully placing CAD activities within a design-based learning (DBL) framework, emphasizing the engineering design process, has the potential to increase students' spatial abilities as well as students' STEM interest and success.

4.2. Conceptual Framework

The theoretical underpinnings of this study are situated in constructivism. Constructivism, as an approach to teaching and learning, is strongly influenced by the works of John Dewey (1933/1997), Jean Piaget (1950/2003), Lev Vygotsky (1978), and Jerome Bruner (1960/1977), with the common thread being that developing meaning through constructing one's own knowledge is worth more than the acquisition of large sets of knowledge that can easily be forgotten (Lutz & Huitt, 2004). Constructivism is a hands-on, learner-centered model where learning is viewed as the result of mental construction, taking place when new knowledge is built upon prior experiences through the active process of construction (Clark, 2018; Hadjerrouit, 2005; Pritchard, 2014). In constructivist learning models, learners are active participants while instructors are facilitators, and knowledge is constructed as learners create meaning as opposed to acquiring it (Bodner, 1986; Clark, 2018). Through this study, we propose the utilization of DBL practices, rooted in constructivism, to implement CAD interventions to increase elementary students' spatial abilities.

DBL combines the conceptual learning aspect of knowledge seeking and idea formation with the material aspect involving the construction of prototypes, models, and end products (Seitamaa-Hakkarainen, 2011). Aligning with constructivist learning models, DBL is a problem-based, self-directed approach that engages students in interdisciplinary design projects with authentic outcomes (Puente et al., 2013). The engineering design process is a key element of DBL and provides opportunities for students to appreciate multiple ideas, approaches, and solutions to real-world problems while also learning the importance of failure and perseverance (Lachapelle & Cunningham, 2014). Incorporating the engineering design process with interdisciplinary design-based learning activities facilitates the transferring and use of technological skills and knowledge (Demirel & Coşkun, 2010). Additionally, the use of CAD tools, which is a staple in engineering education, has increasingly become more popular in primary and secondary education (McConnell, 2015). Researchers have linked students' ability to design 3D objects in CAD software to their spatial abilities (Company et al., 2004; Contero et al., 2006). Explorations and design-based projects that require the use of CAD software can be incorporated in the classroom to improve spatial ability (Magana et al., 2019; Onyancha et al., 2009; Smith, 2018). The conceptual framework, as shown in Figure 4.1, theoretically guide the proposed research study.

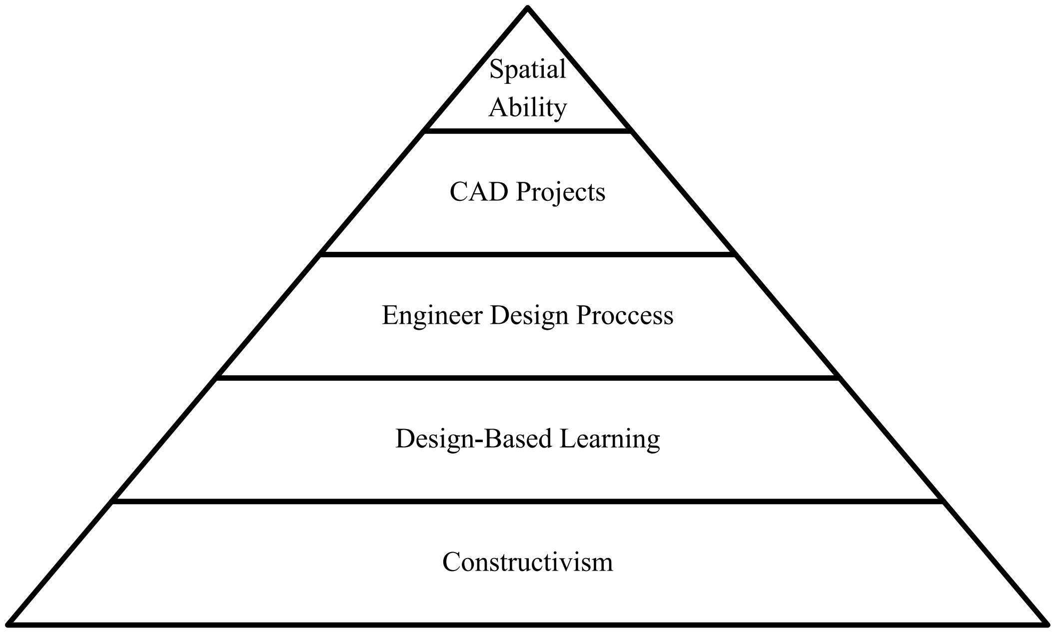

Figure 4.1 Conceptual Framework for Improving Spatial Abilities Through CAD Projects

At the base of the pyramid, the theoretical underpinnings of the proposed research are grounded in constructivism. Moving up the pyramid, DBL describes the instructional strategies based on constructivism that was utilized. From within a DBL environment, the engineering design process was emphasized and used to provide a structure in which students approach the CAD projects. The CAD projects engage students in using TinkerCAD software to work through various interdisciplinary design tasks. Through the CAD projects, at the top of the pyramid, spatial ability development was targeted and assessed. The literature review that follows defines and highlights results from previous research on DBL, the engineering design process, CAD software, and spatial abilities.

4.3. Literature Review

4.3.1. Design-Based Learning

Rooted in constructivism, DBL is a problem-based, hands-on, self-directed approach in which students engage in interdisciplinary design projects with authentic outcomes (Puente et al., 2013). Such design projects have been used to motivate and engage students in a variety of elementary, middle, and high school classrooms (Chandrasekaran et al., 2013). It is suggested by constructivists that working on relevant and meaningful activities and projects intellectually engages students and they become personally invested in their learning, which occurs both during the engineering design process as well as through sharing products (Lee & Hannafin, 2016). Providing opportunities for students to challenge their assumptions through creative design and experimentation towards the creation of a meaningful artifact embodies the idea of constructivist learning (Jun et al., 2017). Students can experience this through DBL and design-focused environments, which promote conceptual understanding (Blikstein, 2013) and where activities such as programming, making, teaching, and collaborating provide rich contexts for learning (Papert & Harel, 1991).

DBL incorporates many approaches, such as Learning by Design (Kolodner et al., 1998), project-based learning (Capraro et al., 2021; Thomas, 2000), and problem-based learning (Barrows, 1985; Gijselaers, 1996). DBL engages students in solving real-life design tasks resembling those of engineering professionals while providing opportunities for students to reflect on the learning process (Mehalik & Schunn, 2006; Puente et al., 2013). In DBL environments, students build, design, and create their own authentic products, solutions, and prototypes (Chandrasekaran et al., 2013; Chen & Chiu, 2016), which engages them to take an active role in their learning, use creative problem-solving strategies, make personal connections to knowledge, and reflect on their outcomes (Bekker et al., 2015; Puente et al., 2013; Fortus et al., 2004; Kim et al., 2015; Smith, 2018). Studies have found that DBL has the potential to promote students' spatial abilities (McConnell, 2015; Smith, 2018; Yildiz & Ozdemir, 2018), academic achievement (Ellefson et al., 2008), computational thinking (Jun et al., 2017), interest, engagement, motivation (Chandrasekaran et al., 2013; Doppelt et al., 2008), and creativity (Chen & Chiu, 2016). DBL activities, as previously established, utilize the engineering design process and interdisciplinary activities to engage students in meaningful learning experiences.

DBL has in the past involved the construction of models in a hands-on fashion using materials available in the classroom; however, technology available for design has made huge strides in becoming more user-friendly and available (Ratto & Ree, 2012). The use of CAD or 3D modeling tools is a staple in engineering education but is becoming more popular in primary and secondary education (McConnell, 2015), and research on the implications of educational CAD tools within design learning has grown (Magana et al., 2019; Smith, 2018). McConnell (2015) found within a DBL activity, students who used CAD software to draw their designs were more engaged, scored higher on their spatial ability test, and were more confident in their results than those who created paper drawings. Similarly, Smith (2018) observed in the CAD phase of the DBL activity that students used several spatial abilities, were highly engaged, and increased their use of geometry vocabulary. Integrating technology tools, such as CAD, into DBL activities or projects provides students with dynamic interdisciplinary learning opportunities to actively investigate and construct authentic innovative design solutions (Chandrasekaran et al., 2013; McConnell, 2015) and has the potential of improving students' spatial abilities and STEM achievement.

4.3.2. Engineering Design Process

The emphasis on teaching and utilizing the engineering design process is foundational in DBL environments or activities. The engineering design process can be seen in several forms, but Cunningham (2009) adapted the engineering design process for the elementary classroom. She defined this process in five cyclic stages: (1) ask, (2) imagine, (3) plan, (4) create, and (5) improve (Cunningham, 2017). The engineering design process motivates students to learn mathematics, science, reading, writing, and communication by granting students ownership of a project or process (Rogers & Portsmore, 2004). Engaging students in the engineering design process to solve real-life problems has been shown to have positive learning outcomes (English & King, 2015; Rogers & Portsmore, 2004; Williams et al., 2016).

Engineering design activities are naturally interdisciplinary because the innovation of design centered on human experiences is not limited to disciplinary boundaries (Luccarelli et al., 2019; Rogers & Portsmore, 2004). Activities or projects can be considered interdisciplinary if the application integrated knowledge, language, and concepts from two or more disciplines (Jacobs, 1989; Meeth, 1978; Newell & Green, 1982). Interdisciplinary activities have been seen to positively impact learners' motivation, problem-solving ability (Williams et al., 2016), interest in STEM (Riechert & Post, 2010), logical thinking skills (Demirel & Coşkun, 2010), and creativity (Darbellay et al., 2017). Engaging students in interdisciplinary learning tasks provides students with opportunities to learn and make meaningful connections between different disciplines (Chen, 2007). Engaging students in engineering design activities provides opportunities for students to engage in interdisciplinary real-word tasks and has the potential to increase students' success in STEM.

4.3.3. CAD Software

CAD software is commonly used for drafting, designing, and developing various machinery components (Sharma & Dumpala, 2015) and is widely used by designers because such software offers the possibility of creating complex designs (Martin & Velay, 2012) in a way that makes them more accessible to others. CAD software and manufacturing tools are ubiquitous in today's product commercialization environment, and students entering this environment need to be proficient in using these tools (Johnson & So, 2015). The use of CAD software could be an effective tool in mathematics classes because it is based on the formation and transformation of geometrical objects (Falcón, 2011). Furthermore, using CAD software can increase motivation and mathematics achievement because doing so combines problem-solution thinking, design, and production and allows for the development of reflection and critical thinking (Huleihil, 2016).

There are many CAD software platforms, such as SketchUp, SolidWorks, TinkerCAD, AutoCAD, and Maya, all of which are used to create 2D renderings of 3D designs. TinkerCAD is a browser-based CAD software that allows users to manipulate a variety of 3D figures in order to develop unique designs for 3D printing. TinkerCAD is user-friendly, for non-professionals, and ideal for young children (McConnell, 2015). Utilizing CAD software such as TinkerCAD can be incorporated into design projects and has the potential to increase students' spatial abilities.

CAD software is an essential tool in engineering education (Chester, 2007) and has advanced the creative design process by addressing past constraints and limitations of spatial expression (Chang, 2014). Utilization of 3D modeling and visualization tools, such as CAD software, has become increasingly more popular in K–12 formal and informal learning environments according to various researchers (e.g., Allan et al., 2018; Dasgupta et al., 2019; Erkoc et al., 2013; McConnell, 2015; Shavalier, 2004; Toptas et al., 2012; Vieira et al., 2016). Explorations and design-based projects requiring the use of CAD software can be incorporated in classrooms to improve spatial ability (Onyancha et al., 2009), engage students, and improve their attitudes towards mathematics (Falcón, 2011), which has been known to have positive influences on STEM achievement.

4.3.4. Spatial Ability

Spatial ability, first conceived by Galton (1879) as a visualizing faculty, has been defined as the capability to represent, transform, generate, and recall symbolic information (Linn & Petersen, 1985) or well-structured abstract visual images (Lohman, 1979). It encompasses understanding, manipulating, and reorganizing or interpreting relationships visually (Tartre, 1990). Spatial abilities are highly essential to problem solving in daily life (Erkoc et al., 2013), especially when manipulating and processing visual information (Rafi et al., 2005). Spatial abilities allow individuals to use concepts of shape in concrete and abstract ways to make and use things in the world, to navigate, and to communicate (Basham & Kotrlik, 2008). Tasks, such as mentally transforming objects to pack the trunk of a car, knowing left from right when viewing objects from different perspectives, and using a map to find your way, are all examples of activities that involve spatial abilities (Shavalier, 2004). Spatial abilities are used regularly in everyday life, especially in areas of logical reasoning and problem solving.

The complexity of spatial ability is more generally understood through the description of three commonly used subcategorizations: spatial perception, spatial visualization, and mental rotation (Linn & Petersen, 1985). Spatial perception is the ability to establish spatial relationships from visual data with respect to personal orientation (McConnell, 2015; Seabra & Santos, 2008). Spatial perception is important for physical movement, orientation, and representation of real-world physical conditions (Matthews & Geist, 2002). Spatial visualization refers to the mental manipulation of spatial information to imagine the relative movement of an image or object and determine how its spatial configuration would appear if repositioned, folded, rotated, or otherwise transformed (Linn & Petersen, 1985; Salthouse et al., 1990; Seabra & Santos, 2008). Spatial visualization tasks should engage multistep analytic procedures (Toptas et al., 2012). Mental rotation is the ability to rotate 2D or 3D figures mentally (Linn & Petersen, 1985) to verify how they would look from a different angle or perspective (Moè, 2018). Mental rotation encompasses the visual inspection and mental simulation of an object's rotation in space (Hegarty & Waller, 2005). All three of these components are important for learning and understanding mathematics and can improve students' problem-solving and reasoning skills.

Since the mid-1900s, the link between mathematics achievement and spatial ability has been of interest to educational researchers and psychologists (Bishop, 1980), and the connection between them has been shown in decades of prior research (e.g., Carr et al., 2020; Casey et al., 1997; Ganley & Vasilyeva, 2011; Rabab'h & Veloo, 2015; Verdine et al., 2014). In recent years, spatial ability has increasingly gained researchers' interest due to correlational evidence linking it with educational performance in science and mathematics (Buckly et al., 2018; Höffler, 2010). In multiple longitudinal studies, researchers followed a large population of participants from adolescence to adulthood and found that spatial ability was a predictor of long-term achievement in STEM (Shea et al., 2001; Wai et al., 2009). Researchers have also noted that individual differences in spatial ability noticeably influenced students' decisions to avoid or approach STEM domains (Lubinski & Benbow, 2006; Robertson et al., 2010). In a study of more than 500 participants, researchers found that typically those who had higher spatial ability than verbal ability at age 13 identified mathematics or science as their favorite subject, went on to secure STEM degrees, and ultimately ended up in a STEM career (Shea et al., 2001). Spatial training can indeed improve STEM learning, retention, and interest (Uttal & Cohen, 2012). Therefore, starting spatial training at a younger age could make an important difference in students' success in STEM majors and fields. Interventions and activities such as experiences with CAD and 3D printing are an excellent way to utilize advanced technologies to develop and improve students' spatial abilities. The current research study suggests the use of TinkerCAD projects in a constructivist learning environment can improve elementary students' spatial abilities.

4.4. Methods

In the present exploratory study, a one-group pretest/posttest design was conducted to explore the relationship between the utilization of CAD software in engineering design projects and students' spatial abilities. To determine the implications of using CAD software in an informal educational experience in order to influence students' spatial abilities, as measured by mental rotation skills, data were collected before and after a week-long CAD intervention. The current study was guided by the following research question: How does utilizing TinkerCAD software as the primary tool for engineering design projects foster growth and development of elementary students' spatial abilities as measured by mental rotation?

4.4.1. Participants

Participation in the study is based on a convenience sample of students (n = 31) who attended a one-week elementary STEM summer camp at a research-intensive university. The sample included eight female (27%) and 22 male (73%) students. Participants' ages ranged from seven to 11 (M_{age} = 9.24 SD_{age} = 0.99).

4.4.2. Instrument

The redrawn Vandenberg and Kuse Mental Rotation Test (MRT-A) was used to assess participants' spatial abilities as measured by mental rotation skills. The MRT-A was recreated by Peters et al. (1995) from the original paper and pencil mental rotation test by Vandenberg and Kuse (1978). The instrument contains two parts, with 12 items on each part, and there are five 3D drawings of cubical figures per item. Each item contains one target figure on the left and four answer choices on the right. The participants identify which two of the answer choices are identical to the target but rotated along the y-axis. The two other answers are mirror images of the target and thus could not become identical to the target by rotation (see Figure 4.2).

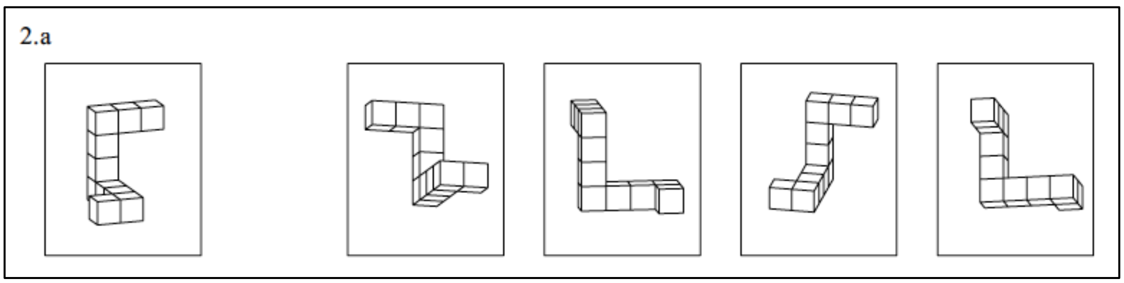

Figure 4.2 Example Item from MRT-A

For the present study, only the first part (first 12 items) of the test was administered, and participants were given four minutes to complete those 12 items. Score procedures consisted of one point being awarded for each correct answer (two correct answers per item) and one point being subtracted for each incorrect answer, yielding a maximum of 24 points. The MRT-A was selected because it is widely used, validated, and has been shown to be highly reliable in various settings. The internal consistency for this sample was measured using Cronbach's (1951) alpha coefficient; score reliability was acceptable with a value of .81 for the pretest and .72 for the posttest. All participants were tested both before and after the intervention at the same time of day by the same test administrator and with the same instructions.

4.4.3. CAD Software Used for Instruction

TinkerCAD was the CAD software introduced and utilized in the present study. TinkerCAD, owned by Autodesk, is a free online 3D modeling software that runs in a web browser and is known for its simple interface and ease of use. TinkerCAD allows users to start their designs with 3D geometric primitives (Avila & Bailey, 2016) that can be combined and manipulated. Geometric primitives are basic geometric shapes (e.g., sphere, cube, cylinder, pyramid) that can be assembled with others to construct more complex shapes (Boubekeur et al., 2019). Starting with these basic 3D shapes provides a much simpler mode to create complex shapes and objects. TinkerCAD was selected because it provides an easy early training ground to introduce solid modeling and 3D printing to younger or less experienced students.

4.4.4. Intervention

During a one-week STEM summer camp, participants were assigned and placed in a 3D printing class. The class met for two hours each day, Monday through Friday. On the first day, participants took the pretest and listened to an introduction to 3D printing, the engineering design process, and were provided guidelines for the first introduction project: “Make Your Own Trophy” (see Table 4.1). They created TinkerCAD accounts, explored the software, and created their trophy. On the second day, they received three open-ended final projects to individually choose from (see Table 4.1). After students picked the project that were of interest to them they were guided to start their project using the stages of the engineer design process and checked in with the instructor after completing each stage. Days three and four were project workdays where the participants worked independently on their final projects. On the fifth day, participants finished their final project and shared their designs.

Table 4.1 Intervention Project Descriptions

Project Title	Description
Make Your Own Trophy	Create an award for yourself to receive at the end of the camp that is less than 3"x3"x3", has your name, and consists of at least 6 shapes.
Option 1: Historical Structure	Choose and replicate any historical structure of your choice. Research and select your structure, plan & sketch your 2D design, create your 3D design, then present a brief history of the structure, selection reason, design processes, and your final product.
Option 2: Create a Character	Create a character from a TV show, movie, or game OR make a cartoon avatar of yourself. Sketch a 2D drawing of the character, create your 3D design, then present a history/description of your character with all of the design aspects.
Option 3: Design your Dream Room	Create your dream room. Brainstorm some objects or attributes for your dream room and search online for reference images, plan & sketch your 2D room blueprint with measurements to scale, create your 3D design, then present the details of your dream room, design process & steps (including measurements), and your final product.

4.4.5. Data Analysis

Quantitative data were collected pre and postintervention to assess participants' spatial abilities as measured by mental rotation skills. First, data were analyzed using a paired sample *t* test to investigate the statistical differences between mental rotation skills before and after the intervention. Prior to conducting the paired sample *t* test, Q plots and box plots of the pretest and posttest scores were analyzed to assess score distribution and check for outliers. In addition, the Shapiro-Wilk test (Shapiro & Wilk, 1965) for normality was conducted. Second, to provide a visual representation of the results, 95% confidence intervals (CIs) for the pretest and posttest scores were examined. Finally, Cohen's *d* effect size estimates were computed to quantify the magnitude of the difference between pretest and posttest scores. Statistical package Stata 17 was used in the aforementioned data analyses.

4.5. Results

The results of the paired sample *t* test [t (29) = 4.524, p < .001] indicated there was a statistically significant difference in the pretest scores (M = 12.30, SD = 6.53) and posttest scores (M = 16.17, SD = 5.91). The Q plot of the difference scores (see Figure 4.3) show that data were moderately skewed but considering the small sample size (n = 31) the spread was relatively close to a normal distribution. The box plots of the difference scores (see Figure 4.4) indicated that no outliers were present. In addition, the Shapiro-Wilk test demonstrated the score distribution did not depart significantly from normality for difference scores (W = .98, p = 0.71). The 95% CIs for the pretest and posttest scores (see Figure 4.5) confirm there existed, on average, a considerable gain in mental rotation skills following the intervention.

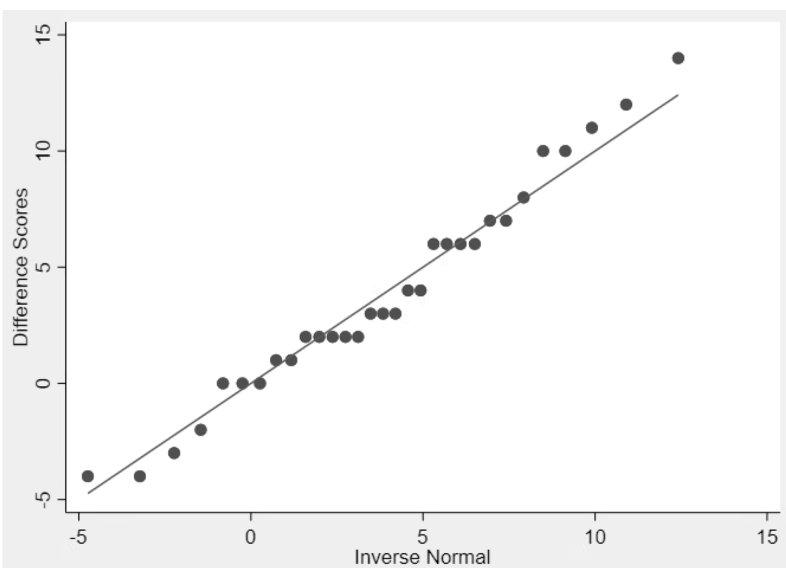

Figure 4.3 Q plots of pre-post MRT-A observed difference scores

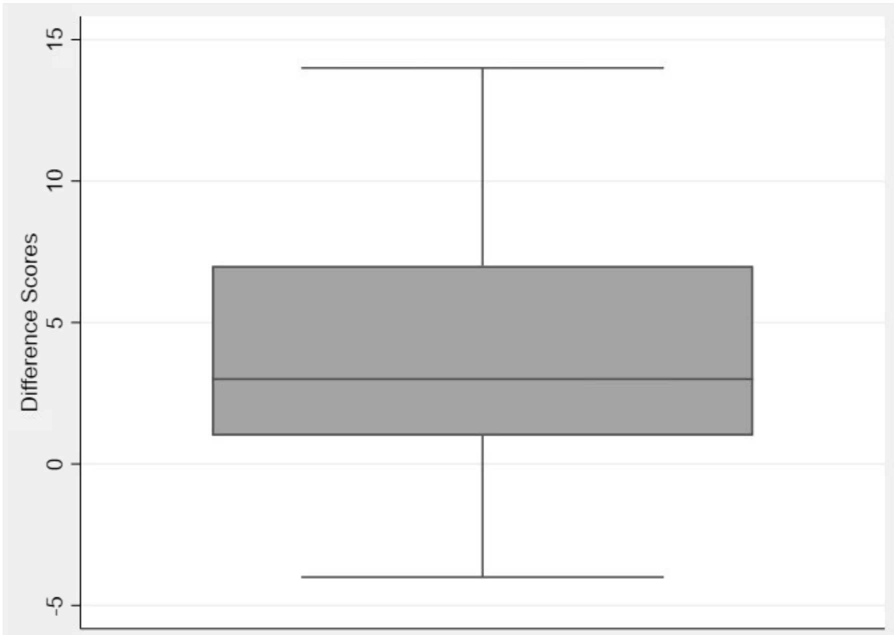

Figure 4.4 Box plots of pre-post MRT-A performance

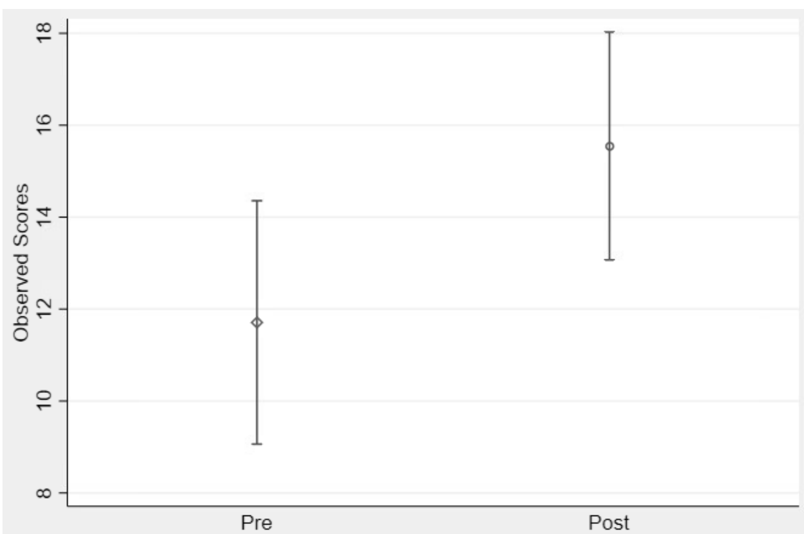

Figure 4.5 95% CI plots by pre-post MRT-A observed scores

Given the sample size, the Cohen's *d* value in this study (*d* = 0.55, 95% CI [.04,1.05]) provides a more robust and comparable characterization of the effectiveness of the intervention than significance testing. The Cohen's *d* value indicates that the standard deviation of the posttest scores increased by 0.55 from the pretest scores. The effect size calculated is lower than, but comparable to, those found in the literature. Erkoc et al. (2013) found the implementation of a CAD software with 8th grade students had an effect size of 2.7 on participants' spatial abilities. Additionally, based on means and standard deviations reported by Onyancha et al. (2009) and Martín-Dorta et al. (2008), the implementation of CAD software had effect sizes of 1.94 and 0.68, respectively, with university-level engineering students. It can be concluded based on the results that exposure to CAD software had a positive impact on elementary students' mental rotation abilities.

4.6. Discussion

Overall, the results indicated that spatial abilities, as measured by mental rotation, were improved by the CAD intervention. Framing CAD software use and spatial training in DBL and the engineer design process provided us the opportunity to show that self-directed learning is authentic and allows knowledge production without rote memorization. In the present study there was no time spent reviewing or practicing mental rotation confirming that these skills can be developed through application without direct instruction.

Our results align with previous research showing that through the use of 3D tools spatial abilities can be developed (Kösa & Karakuş, 2018; Medina Herrera et al., 2019). This work is important because supporting the development of spatial abilities has the potential to increase students' engagement, STEM achievement, and interest in various STEM careers (Lubinski & Benbow, 2006; Roberstson et al., 2010; Shea et al., 2001;

Wai et al., 2009). In addition, we observed that students who interacted with and utilized CAD software to create designs were more engaged and were more confident in their results, which aligns with the findings of previous research (McConnell 2015; Smith, 2018).

Incorporating CAD software design projects and learning opportunities into the classroom can increase spatial ability development in more students which could lead to increased inclusivity in STEM. The CAD software used in the present study is a free web-based platform and therefore can be used by any school with internet connection. The act of 3D printing students designs is not a requirement to the spatial training thereby similar design projects, utilizing free CAD software, could be incorporated to be more equitable to low-income schools or underserved student populations. In addition, Google Sketchup now also offers a free web-based platform, which could be used with more advanced learners, allowing for differentiation with the projects. Ultimately, expanding access to STEM education could create a more inclusive and equitable STEM workforce (O'Rourke, 2021).

The present study does come with limitations to be considered. A one-group pre-post design was used, and as such it is possible that the changes observed may not be accredited to the CAD software experience but instead may reflect other threats to internal validity, such as history and testing. Additionally, given the data were collected at a one-week STEM summer camp, the intervention length is shorter than may be desired. While considering the limitations, results corroborate a positive relationship between spatial abilities and CAD interventions with elementary students, similar to results seen in other studies with university students (e.g., Kinsey et al., 2008; Martín-Dorta et al., 2008; McConnell, 2015; Onyancha et al., 2009). Experiences with CAD software, such as those described in the present study, have the potential to increase engagement in STEM domains and prepare learners for STEM-related careers.

4.7. Conclusion

In this article, the relationship between the use of CAD software and the development of spatial ability was investigated. The following research question was examined: How does utilizing TinkerCAD software as the primary tool for engineering design projects foster growth and development of elementary students' spatial abilities as measured by mental rotation? There was a statistically significant difference in students' mental rotations score from pre to post intervention, and the calculated Cohen's *d* effect size of 0.62 indicates that the CAD intervention had a positive impact on students' mental rotation skills. Therefore, we concluded that employing technologies such as CAD and 3D printing into design projects is an excellent way to utilize advanced technologies to develop and improve students' spatial abilities, which can lead to improved STEM interest, achievement, and success.

For future research, expanding this research within a formal elementary classroom with a larger sample, an extended intervention window, and a control group would allow for a more robust investigation of the impact interactions with CAD software in design-based projects have on students' spatial abilities. In addition, adding a measure to evaluate participants' STEM achievement and perceptions pre and post intervention could provide valued insights into the relationship between perceptions, CAD, spatial abilities, and STEM achievement.

5. CONCLUSIONS

5.1. Intellectual Merit

Improving science, technology, engineering, and mathematics (STEM) interest and achievement is a pervasive theme in educational research and institutions. There exists potential to enhance STEM education through the development of students' spatial abilities (Khine, 2017), however, there is little attention placed on promoting and training spatial abilities in elementary and secondary schools (Chapter 2; DeSutter & Stieff, 2017). Spatial abilities permit individuals to use concepts of shape in concrete and abstract ways to generate and utilize objects in the world, to navigate, and to communicate (Basham & Kotrlik, 2008). Spatial ability is a fundamental cognitive resource for STEM learning and has been identified as a marker for interest and success in STEM (Berkowitz & Stern, 2018; Shea et al., 2001; Uttal & Cohen, 2012; Wai et al., 2009). Spatial abilities have been acknowledged as a significant gap in K-12 curriculum (NRC, 2006) despite being identified as a critical skill that should be included in early childhood education (NCTM, 2000). Various researchers have confirmed the development and improvement of spatial abilities is possible given appropriate experiences, opportunities, and practice (Guo et al., 2022; Montag et al., 2021; Uttal & Choen, 2012).

Spatial ability development can be promoted through experience with three-dimensional modeling programs such as computer-aided design (CAD) software (Matthews & Geist, 2002; Onyancha et al., 2009). Experiences in 3D virtual environments has the benefit of simulating the authenticity and visibility of the objective world, can make abstract knowledge concrete, and varies from other approaches to spatial training (e.g., engineering drawing and sketch training) because it provides students with clearer object visualization (Bliksteing et al., 2017; Guo et al., 2022). Although the relationship between 3D design software and spatial abilities have become more commonly observed in educational settings such as engineering (i.e., Contero et al., 2006), architecture (i.e., Falcón, 2011), and interior design (i.e., Zavotka, 1986) they have yet to establish a strong presence in early childhood education (Chapter 2). I have provided evidence that spatial ability development can be fostered through experience with three-dimensional modeling programs such as computer-aided design (CAD) software in various educational settings (Chapter 2) and have shown that this holds true with elementary students (Chapter 4).

After exploring and synthesize previous research on the implementation of 3D modeling software in various educational settings and identifying a gap in the research (Chapter 2) I set off to explore the relationship between elementary students' spatial abilities and experiences with CAD software (Chapter 4). In order to investigate this relationship, I needed to gauge the students' spatial abilities before and after their experiences with the CAD software and therefore had to select how to measure their spatial abilities. Given the characteristics of CAD software and 3D design projects I selected to measure students' mental rotation skills to assess their spatial abilities.

Mental rotation is a necessary spatial ability that arises in several aspects of life and often involves tasks that require individuals to dynamically represent and spatially transform figures mentally (Montag et al., 2021). Researchers from several fields have studied mental rotation skills for decades, resulting in the formation of numerous instruments that measure mental rotation skills, however, the widely used Mental Rotation Test (MRT) by Vandenburg and Kuse (1978) was among the first paper-and-pencil group test created and validated to measure mental rotation abilities. After many years of use, the MRT was redrawn using CAD software (Peters et al., 1995) to restore and enhance the images on the instrument creating the MRT-A. I selected to use the MRT-A because it has been widely utilized in various fields with various age groups to measure participants' mental rotation abilities. Although the original MRT had been previously validated I noticed that often those that used the redrawn version reported that the selection of the MRT-A was due to the high reliability yet do not report reliability based on respective researchers' own scores from their primary study and rely on previously published reliability estimates often of the original instrument. Therefore, prior to utilizing the MRT-A in our investigations I wanted to establish its reliability in previous settings.

After an exhaustive search and synthesis of the literature I was able to establish that the MRT-A possessed good reliability in various settings (Chapter 3) which held true in our investigations (Chapter 4). In addition to establishing instrument reliability, our reliability generalization (Chapter 3) exposed some alarming issues in concerns with poor reporting practices. Of the 191 unique samples that utilized the MRT-A to assess spatial ability, less than 18% reported a reliability coefficient for their sample, about 65% of the studies failed to report any form of reliability for their samples, and about 17% cited previously reported reliability. This is concerning because reliability estimates can and often will fluctuate across administrations (Crocker & Algina, 1986; Vacha-Haase et al., 2002; Vacha-Haase & Thompson, 2011) because anything that could possibly affect scores could also affect reliability (Barnes et al., 2002). Given the diversity across studies (Vacha-Haase, 1998) and importance of score reliability in all quantitative analyses (Vacha-Haase & Thompson, 2001) it is important that authors report reliability coefficients for their data. The absence of these data can limit the generalizability and reduces the likelihood of study replication (Williams & Young, 2021). Therefore, our reliability generalization not only achieved our goal of establishing a generalized reliability for the MRT-A but also hopefully encourage better reporting practice and inform researchers of meta-analytic techniques that can be used to evaluate instrument reliability across samples (Chapter 3).

In the fourth chapter of this dissertation, I established that experiencing and utilizing CAD software in design projects has led to positive effects on students' spatial ability outcomes. Our results align with previous research showing that through the use of 3D tools spatial abilities can be developed (Kösa & Karakuş, 2018; Medina Herrera et al., 2019). Integrating technologies such as CAD into design projects in the classroom is an excellent way to develop spatial abilities, expand access to STEM education, and employ advanced technologies which can lead to improved STEM interest, achievement, and success. While an initial positive relationship was explored and important insights into the possibilities of promoting spatial ability through CAD software were gained there were study limitations (i.e., sample size; intervention length; one-group design) that could have skewed results and changes observed may not be attributed to the CAD software exposure but instead may reflect other threats to internal validity. With the results of this dissertation a large-scale study, with a longer intervention window, could be conducted to investigate the effectiveness of the use of CAD in the development of spatial ability.

5.2. Broader Impacts

There is usefulness in this work to society and for various populations of students and teachers. Through this dissertation study, I have shown a link between students' spatial abilities and CAD design projects (Chapter 3 & 4). Supporting the development of spatial abilities has the potential to increase students' confidence and engagement (McConnol, 2015; Smith, 2018), STEM achievement (Buckly et al., 2018; Höffler, 2010), and interest in various STEM careers (Shea et al., 2001; Wai et al., 2009). Additionally, it has been found that spatial abilities have improved after training regardless of how students' previous skills, experiences, grades, etc. are (Šafhalter et al., 2020). Therefore, it could be expected that fostering the development of spatial abilities in the classroom would be benefit all learners which could then in turn could help towards bridging the equity gap in STEM education.

As I have described throughout this dissertation study, spatial abilities can be improved through the implementation of 3D design projects using CAD software as the primary design tool. Although I focused on the improvement of students' spatial abilities, these types of interventions can also impact other factors influencing students in a beneficial manner. For starters, "3D design through in-depth integration with traditional education, builds a personalized, interesting, and open comprehensive innovative practical teaching mode" (Guo et al., 2022, p. 181). Engaging students in design activities can stimulate STEM knowledge and problem-solving skills promoting STEM career selection (English et al., 2017; Lamb et al., 2018). Experiences with CAD and 3D printing has been shown to positively influence specific mathematics and real-life skills as well as student interest and motivation (Kwon, 2017). Design and 3D printing projects promote self-engagement and enthusiasm within classroom activities (Medina Herrera, 2019), supports the development of students' problem solving abilities (Bliksteing et al., 2017), and nurtures students' creativity (Eisenberg, 2013).

In recent years, 3D printing technologies have become more accessible and available for use in schools. This is due to the increased application of 3D printing technologies in industry and educational settings as well as the decreased cost in hardware and software (Huang et al., 2019). Teaching 3D design and modeling is believed to offer innovative new learning experiences for students that develop and improve several 21st century skills such as inventive and critical thinking (Benzer & Yildiz, 2019; Fujiwara & Jones, 2019; Trust & Maloy, 2017). Therefore, with greater access to CAD and 3D printers, a rising number of educators have the opportunity to expose students to interventions and activities that develop and improve spatial abilities, increase student interest and success in STEM, as well as develop valuable 21st century skills.

In this dissertation, I presented three articles in which I proposed, developed, and evaluated early versions and implementations of CAD software design projects. I trust this work can serve as a foundation for future explorations into spatial ability training through 3D design and printing projects using CAD software.

REFERENCES

Allan, D., Vettese, S., & Thompson, P. (2018). A study of children's relationship with making and use of CAD in collaborative, informal environments and the implications for institutional learning environments. *Design and Technology Education, 23*(3), 10–26.

Anscombe, F. J., & Glynn, W. J. (1983). Distribution of kurtosis statistic b_2 for normal statistics. *Biometrika*, *70*(1), 227–234. https://doi.org/10.1093/biomet/70.1.227

Atit, K., & Rocha, K. (2021). Examining the relations between spatial skills, spatial anxiety, and k-12 teacher practice. *Mind, Brain, and Education, 15*(1), 139–148. https://doi.org/10.1111/mbe.12274

Atit, K., Shipley, T. F., & Tikoff, B. (2013). Twisting space: Are rigid and non-rigid mental transformations separate spatial skills?. *Cognitive Processing, 14*(2), 163–173. https://doi.org/10.1007/s10339-013-0550-8

Augustyniak, K., Murphy, J., & Phillips, K. D. (2004). Psychological perspectives in assessing mathematics learning needs. *Journal of Instructional Psychology, 32*(4), 277–286.

Avila, L., & Bailey, M. (2016). A computer graphics back-to-school special. *IEEE Computer Graphics and Applications*, *36*(5), 95–96. https://doi.org/10.1109/MCG.2016.104

Bairaktarova, D. (2017). Coordinating mind and hand: The importance of manual drawing and descriptive geometry instruction in a CAD-oriented engineering design graphics class. *Engineering Design Graphics Journal, 81*(3), 1–16.

Barnes, L. L., Harp, D., & Jung, W. S. (2002). Reliability generalization of scores on the Spielberger state-trait anxiety inventory. *Educational and Psychological Measurement*, *62*(4), 603–618. https://doi.org/10.1177%2F0013164402062004005

Barrows, H. S. (1985). *How to design a problem-based curriculum for the preclinical years.* Springer.

Basham, K. L., & Kotrlik, J. W. (2008). The effects of 3-dimensional CADD modeling on the development of the spatial ability of technology education students. *Journal of Technology Education, 20*(1), 32–47.

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, *41*(2), 257–278. https://doi.org/10.1111/j.2044-8317.1988.tb00901.x

Bekker, T., Bakker, S., Douma, I., van der Poel, J., & Scheltenaar, K. (2015). Teaching children digital literacy through design-based learning with digital toolkits in schools. *International Journal of Child-Computer Interaction, 5*(1), 29–38. https://doi.org/10.1016/j.ijcci.2015.12.001

Benzer, A. I., & Yildiz, B. (2019). The Effect of Computeraided 3D Modeling Activities on pre-service teachers' spatial abilities and attitudes towards 3D modeling. *Journal of Baltic Science Education*, *18*(3), 335–348.

Berkowitz, M. (2017). *In the minds of STEM beginners: On cognitive abilities, working memory, and first year mathematics* (Doctoral dissertation, ETH Zurich).

Berkowitz, M., & Stern, E. (2018). Which cognitive abilities make the difference? Predicting academic achievements in advanced STEM studies. *Journal of Intelligence*, *6*(4), 48–71. https://doi.org/10.3390/jintelligence6040048

Berkowitz, M., Gerber, A., Thurn, C. M., Emo, B., Hoelscher, C., & Stern, E. (2021). Spatial abilities for architecture: Cross sectional and longitudinal assessment with novel and existing spatial ability tests. *Frontiers in Psychology, 11*(1), 4096.

Bishop, A. J. (1980). Spatial abilities and mathematics education—A review. *Educational Studies in Mathematics, 11*(3), 257–269. https://doi.org/10.1007/BF00697739

Blikstein, P. (2013). Digital fabrication and 'making'in education: The democratization of invention. *FabLabs: Of Machines, Makers and Inventors*, *4*(1), 1–21.

Blikstein, P., Kabayadondo, Z., Martin, A., & Fields, D. (2017). An Assessment instrument of technological literacies in makerspaces and FabLabs. *Journal of Engineering Education, 106*(1), 149–175.

Bodner, G. M. (1986). Constructivism: A theory of knowledge. *Journal of Chemical Education*, *63*(10), 873–878. https://doi.org/10.1021/ed063p873

Bogomolova, K., van der Ham, I. J., Dankbaar, M. E., van den Broek, W. W., Hovius, S. E., van der Hage, J. A., & Hierck, B. P. (2020). The effect of stereoscopic Augmented Reality visualization on learning anatomy and the modifying effect of visual-spatial abilities: a double-center randomized controlled trial. *Anatomical Sciences Education, 13*(5), 558–567.

Bonett, D. G. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics*, *27*(1), 335–340. https://doi.org/10.3102/10769986027004335

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009*). Introduction to Meta-Analysis*. John Wiley.

Bors, D. A., & Vigneau, F. (2011). Sex differences on the mental rotation test: An analysis of item types. *Learning and Individual Differences, 21*(1), 129–132.

Botella, J., Suero, M., & Gambara, H. (2010). Psychometric inferences from a meta-analysis of reliability and internal consistency coefficients. *Psychological Methods*, *15*(4), 386–397. https://doi.org/10.1037/a0019626

Boubekeur, T., Kaiser, A., & Ybanez Zepeda, J. A. (2019). A survey of simple geometric primitives detection methods for captured 3D data. *Computer Graphics Forum*, *38*(1), 167–196. https://doi.org/10.1111/cgf.13451

Brase, G. L., & Hill, W. T. (2017). Adding up to good Bayesian reasoning: Problem format manipulations and individual skill differences. *Journal of Experimental Psychology: General, 146*(4), 577–591. https://doi.org/10.1037/xge0000280

Bruner, J. (1977). *The process of education*. Harvard University Press. (Original work published 1960)

Bryant, K. J. (1982). Personality correlates of sense of direction and geographic orientation. *Journal of Personality and Social Psychology*, *43*(6), 1318–1324. https://doi.org/10.1037/0022-3514.43.6.1318

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*(3), 947–972. https://doi.org/10.1007/s10648-018-9432-z

Budinoff, H., & McMains, S. (2018). Relationships between spatial visualization ability and student outcomes in a 3D modeling course. *Engineering Design Graphics Journal, 82*(2), 17–23.

Cakmak, S., Isiksal, M., & Koc, Y. (2014). Investigating effect of origami-based instruction on elementary students' spatial skills and perceptions. *Journal of Educational Research*, *107*(1), 59–68. https://doi.org/10.1080/00220671.2012.753861

Capraro, R. M., & Capraro, M. M. (2002). Myers-Briggs type indicator score reliability across: Studies a meta-analytic reliability generalization study. *Educational and Psychological Measurement, 62*(4), 590–602.

Capraro, R. M., Capraro, M. M., Young, J., & Barroso, L. R. (Eds.). (2021). *STEM project-based learning: Integrated engineering for a new era.* Aggie STEM.

Carr, M., Horan, E., Alexeev, N., Barned, N., Wang, L., & Otumfuor, B. (2020). A longitudinal study of spatial skills and number sense development in elementary school children. *Journal of Educational Psychology*, *112*(1), 53–69. http://doi.org/10.1037/edu0000363

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

Caruso, J. C. (2000). Reliability generalization of the NEO personality scales. *Educational and Psychological Measurement*, *60*(2), 236–254. https://doi.org/10.1177/00131640021970484

Casey, B. M., Lombardi, C. M., Pollock, A., Fineman, B., & Pezaris, E. (2017). Girls' spatial skills and arithmetic strategies in first grade as predictors of fifth-grade analytical math reasoning. *Journal of Cognition and Development, 18*(5), 530–555.

Casey, M. B., Nuttall, R. L., & Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. *Developmental Psychology, 33*(4), 669–680. https://doi.org/10.1037/0012-1649.33.4.669

Casselman, M. D., Eichler, J. F., & Atit, K. (2021). Advancing multimedia learning for science: Comparing the effect of virtual versus physical models on student learning about stereochemistry. *Science Education, 105*(6), 1285–1314.

Chandrasekaran, S., Stojcevski, A., Littlefair, G., & Joordens, M. (2013). Project-oriented design-based learning: Aligning students' views with industry needs. *International Journal of Engineering Education, 29*(5), 1109–1118.

Chang, Y. (2014). 3D-CAD effects on creative design performance of different spatial abilities students. *Journal of Computer Assisted Learning, 30*(5), 397–407. http://doi.org/10.1111/jcal.12051

Chen, C. H., & Chiu, C. H. (2016). Employing intergroup competition in multitouch design-based learning to foster student engagement, learning achievement, and creativity. *Computers & Education*, *103*(1), 99–113. https://doi.org/10.1016/j.compedu.2016.09.007

Chen, W. (2007). Interdisciplinary teaching: Integration of physical education skills and concepts with mathematical skills and concepts. In L. B. Yurichenko (Ed.), *Teaching and teacher issues* (pp. 101–120). Nova Science.

Cherney, I. D. (2008). Mom, let me play more computer games: They improve my mental rotation skills. *Sex Roles, 59*(11–12), 776–786.

Chester, I. (2007). Teaching for CAD expertise. *International Journal of Technology & Design Education*, *17*(1), 23–35. https://doi.org/10.1007/s10798-006-9015-z

Clark, K. R. (2018). Learning theories: Constructivism. *Radiologic Technology*, *90*(2), 180–182.

Company, P., Contero, M., Piquer, A., Aleixos, N., Conesa, J., & Naya, F. (2004). Educational software for teaching drawing-based conceptual design skills. *Computer Applications in Engineering Education, 12*(4), 257–268. https://doi.org/10.1002/cae.20023

Contero, M., Naya, F., Company, P., & Saorin, J. L. (2006). Learning support tools for developing spatial abilities in engineering design. *International Journal of Engineering Education, 22*(3), 470–477.

Cousin, S. L., & Henson, R. K. (2000, January 27). *What is reliability generalization, and why is it important?* Paper presented at the annual meeting of the Southwest Educational Research Association (ED445077). https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED445077

Coutrot, A., Silva, R., Manley, E., De Cothi, W., Sami, S., Bohbot, V. D., Wiener, J. M., Hölscher, C., Dalton, R. C., Hornberger, M., & Spiers, H. J. (2018). Global determinants of navigation ability. *Current Biology, 28*(17), 2861–2866.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart & Winston.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(1), 297–334. https://doi.org/10.1007/BF02310555

Cunningham, C. M. (2009). Engineering is elementary. *The Bridge*, *30*(3), 11–17.

Cunningham, C. M. (2017). *Engineering in elementary STEM education: Curriculum design, instruction, learning, and assessment*. Teachers College Press.

D'Agostino, R. B. (1970). Transformation to normality of the null distribution of G1. *Biometrika*, *57*(1), 679–681. http://doi.org/10.2307/2334794

Darbellay, F., Moody, Z., & Lubart, T. (2017). *Creativity, design thinking and interdisciplinarity.* Springer. https://doi.org/10.1007/978-981-10-7524-7

Dasgupta, C., Magana, A. J., & Vieira, C. (2019). Investigating the affordances of a CAD enabled learning environment for promoting integrated STEM learning. *Computers & Education*, *129*(1), 122–142. https://doi.org/10.1016/j.compedu.2018.10.014

De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Genetic Psychology*, *163*(3), 272–282. http://doi.org/10.1080/00221320209598683

Demirel, M., & Coşkun, Y. D. (2010). Case study on interdisciplinary teaching approach supported by project-based learning. *The International Journal of Research in Teacher Education*, *2*(3), 28–53.

DeSutter, D., & Stieff, M. (2017). Teaching students to think spatially through embodied actions: Design principles for learning environments in science, technology, engineering, and mathematics. *Cognitive Research: Principles and Implications, 2*(22). https://doi.org/10.1186/s41235-016-0039-y

Dewey, J. (1997). *How we think*. Dover Publications. (Original work published 1933)

Doppelt, Y. (2003). Implementing and assessing project-based learning in a flexible environment. *International Journal of Technology and Design Education, 13*(3), 255–272. https://doi.org/10.1023/A:1026125427344

Doppelt, Y., Mehalik, M. M., Schunn, C. D., Silk, E., & Krysinski, D. (2008). Engagement and achievements: A case study of design-based learning in a science context. *Journal of Technology Education*, *19*(2), 22–39.

Doyle, R. A., Voyer, D., & Cherney, I. D. (2012). The relation between childhood spatial activities and spatial abilities in adulthood. Journal of Applied Developmental Psychology, 33(2), 112–120.

Duesbury, R. T., & O'Neil, H. F., Jr. (1996). Effect of type of practice in a computer-aided design environment in visualizing three-dimensional objects from two-dimensional orthographic projections. *Journal of Applied Psychology, 81*(3), 249–260. https://doi.org/10.1037/0021-9010.81.3.249

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(1), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Ellefson, M. R., Brinker, R. A., Vernacchio, V. J., & Schunn, C. D. (2008). Design-based learning for biology: Genetic engineering experience improves understanding of gene expression. *Biochemistry and Molecular Biology Education*, *36*(4), 292–298. https://doi.org/10.1002/bmb.20203

English, L. D., & King, D. T. (2015). STEM learning through engineering design: Fourth-grade students' investigations in aerospace. *International Journal of STEM Education*, *2*(1), Article 14. https://doi.org/10.1186/s40594-015-0027-7

Erkoc, M. F., Gecu, Z., & Erkoc, C. (2013). The effects of using Google SketchUp on the mental rotation skills of eighth grade students. *Educational Sciences: Theory and Practice, 13*(2), 1285–1294.

Falcón, R. M. (2011). Integration of a CAS/DGS as a CAD system in the mathematics curriculum for architecture students. *International Journal of Mathematical Education in Science & Technology, 42*(6), 737–750. https://doi.org/10.1080/0020739X.2011.573871

Fortus, D., Dershimer, R. C., Krajcik, J., Marx, R. W., & Mamlok-Naaman, R. (2004). Design-based science and student learning. *Journal of Research in Science Teaching, 41*(10), 1081–1110. http://doi.org/10.1002/tea.20040

Fujiwara, Y., & Jones, L. K. (2019). Developing 21st-century skills through STEM integration and global collaboration using 3D printing and CAD. In *Integrating 3D Printing into Teaching and Learning* (pp. 55–71). Brill.

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

Ganley, C. M., & Vasilyeva, M. (2011). Sex differences in the relation between math performance, spatial skills, and attitudes. *Journal of Applied Developmental Psychology, 32*(4), 235–242. https://doi.org/10.1016/j.appdev.2011.04.001

Geiser, C., Lehmann, W., & Eid, M. (2006). Separating "rotators" from "nonrotators" in the mental rotations test: A multigroup latent class analysis. *Multivariate Behavioral Research, 41*(3), 261–293.

Georges, C., Hoffmann, D., & Schiltz, C. (2017). How and why do number-space associations co-vary in implicit and explicit magnitude processing tasks? *Journal of Numerical Cognition, 3*(2), 182–211.

Gignac, G. E., Shankaralingam, M., Walker, K., & Kilpatrick, P. (2016). Short-term memory for faces relates to general intelligence moderately. *Intelligence, 57*(1), 96–104. https://doi.org/10.1016/j.intell.2016.05.001

Gijselaers, W. H. (1996). Connecting problem-based practices with educational theory. *New Directions for Teaching and Learning*, *68*(1), 13–22. https://doi.org/10.1002/tl.37219966805

Glück, J., Dünser, A., Steinbügl, K., & Kaufmann, H. (2007). Warning: Subtle aspects of strategy assessment may affect correlations among spatial tests. *Perceptual and Motor Skills, 104*(1), 123–140.

Guo, S., Wang, X., Deng, W., Hong, J., Wang, J., & Wu, Y. (2022). Whose spatial ability benefits from learning with 3D design? From the perspective of learning analysis. *Journal of Educational Technology & Society*, *25*(1), 179–192.

Hadjerrouit, S. (2005). Constructivism as guiding philosophy for software engineering education. *ACM SIGCSE Bulletin*, *37*(4), 45–49. http://doi.org/10.1145/1113847.1113875

Hakstian, A. R., & Whalen, T. E. (1976). A k-sample significant test for independent alpha coefficients. *Psychometrika*, *41*(1), 219–231. https://doi.org/10.1007/BF02291840

Harrell Jr, F. E. (2020). Hmisc: Harrell miscellaneous. *R package version 0.2. The Comprehensive R Archive Network (CRAN).* https://cran.r-project.org/web/packages/Hmisc/Hmisc.pdf

Hart, S., Daucourt, M., & Ganley, C. (2017). Individual differences related to college students' course performance in calculus II. *Journal of Learning Analytics, 4*(2), 129–153.

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.

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), *Handbook of visuospatial thinking* (pp. 121–169).

Henson, R. K. (2000). Demystifying parametric analysis: Illustrating canonical correlation as the multivariate general linear model. *Multiple Linear Regression Viewpoints, 26*(1), 11–19.

Henson, R. K., & Thompson, B. (2002). Characterizing measurement error in scores across studies: Some recommendations for conducting "reliability generalization" studies. *Measurement and Evaluation in Counseling and Development*, *35*(2), 113–127. https://doi.org/10.1080/07481756.2002.12069054

Higgins, J. P. T., & Green, S. (Eds). (2011). *Cochrane handbook for systematic reviews of interventions* Version 5.1.0. The Cochrane Collaboration. Retrieved from www.handbook.cochrane.org.

Hilton, E. C., Paige, M. A., Williford, B., Wayne Li, Hammond, T. A., & Linsey, J. S. (2017). Engineering drawing for the next generation: Students gaining additional skills in the same timeframe. *Proceedings of the ASEE Annual Conference & Exposition*, 4470–4481.

Höffler, T. (2010). Spatial ability: Its influence on learning with visualizations—A meta-analytic review. *Educational Psychology Review, 22*(3), 245–269. https://doi.org/10.1007/s10648-010-9126-7

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability methods: A note on the frequency of use of various types. *Educational and Psychological Measurement, 60*(4), 523–531.

Hone, L. S., Scofield, J. E., Bartholow, B. D., & Geary, D. C. (2020). Frequency of recent binge drinking is associated with sex-specific cognitive deficits: Evidence for condition-dependent trait expression in humans. *Evolutionary Psychology, 18*(4), 1–14.

Howes, A., Kaneva, D., Swanson, D., & Williams, J. (2013). Re-envisioning STEM education: Curriculum, assessment and integrated, interdisciplinary studies. *A report for The Royal Society's Vision for Science and Mathematics Education.* London, UK. https://royalsociety.org/~/media/education/policy/vision/reports/ev-2-vision-research-report-20140624.pdf

Huang, T. C., Chen, M. Y., & Lin, C. Y. (2019). Exploring the behavioral patterns transformation of learners in different 3D modeling teaching strategies. *Computers in Human Behavior, 92*(1), 670–678.

Huleihil, M. (2016, November 8–11). *3D printing technology as innovative tool for math and geometry teaching applications* [Paper presentation]. 5th Global Conference on Materials Science and Engineering, Taichung City, Taiwan. http://doi.org/10.1088/1757-899X/164/1/012023

Jacobs, H. H. (1989). *Interdisciplinary curriculum: Design and implementation* (ED316506). ERIC. https://files.eric.ed.gov/fulltext/ED316506.pdf

Johnson, M., & So, Y. Y. (2015, June 14–17). *Examining the interaction of spatial visualization ability and computer-aided design and manufacturing course performance* [Paper presentation]. 122nd ASEE Annual Conference & Exposition, Seattle, WA, United States.

Jun, S., Han, S., & Kim, S. (2017). Effect of design-based learning on improving computational thinking. *Behaviour & Information Technology*, *36*(1), 43–53. https://doi.org/10.1080/0144929X.2016.1188415

Kaplan, R. M., & Saccuzzo, D. P. (2012). *Psychological testing: Principles, applications, and issues* (8th ed.). Cengage Learning.

Kassambara, A. (2018). ggpubr: 'ggplot2' based publication ready plots. *R package version 0.2. The Comprehensive R Archive Network (CRAN).* https://cran.r-project.org/web/packages/ggpubr/ggpubr.pdf

Khine M. S. (2017) Spatial cognition: Key to STEM success. In M. S. Khine (Ed.), *Visual-spatial ability in STEM education* (pp. 3–8). Springer. https://doi.org/10.1007/978-3-319-44385-0_1

Kieffer, K. M., Reese, R. J., & Vacha-Haase, T. (2010). Reliability generalization methods in the context of giftedness research. In B. Thompson & R. F. Subotnik (Eds.), *Methodologies for conducting research on giftedness* (p. 89–111). American Psychological Association. https://doi.org/10.1037/12079-005

Kim, P., Suh, E., & Song, D. (2015). Development of a design-based learning curriculum through design-based research for a technology-enabled science classroom. *Educational Technology Research and Development*, *63*(4), 575–602. https://doi.org/10.1007/s11423-015-9376-7

Kinsey, B. L., Towle, E., & Onyancha, R. M. (2008). Improvement of spatial ability using innovative tools: Alternative view screen and physical model rotator. *Engineering Design Graphics Journal, 72*(1), 1–8.

Koetsier, J. (2014, June 24). Gamer globe: The top 100 countries by 2014 game revenue. VentureBeat. Retrieved from https://venturebeat.com/2014/06/24/gamer-globe-the-top-100-countries-by-2014-game-revenue/

Kolmos, A. (2002). Facilitating change to a problem-based model. *The International Journal for Academic Development, 7*(1), 63–74. https://doi.org/10.1080/13601440210156484

Kolodner, J. L., Crismond, D., Gray, J., Holbrook, J., & Puntambekar, S. (1998). Learning by design from theory to practice. *Proceedings of the International Conference of the Learning Sciences*, *98*(1), 16–22.

Komsta, L., & Novometsky, F. (2015). *Moments: Moments, cumulants, skewness, kurtosis and related tests (R Package Version 0.14)*. https://CRAN.R-project.org/package=moments

Kösa, T., & Karakuş, F. (2018) The effects of computer-aided design software on engineering students' spatial visualisation skills. *European Journal of Engineering Education, 43*(2), 296–308. https://doi.org/10.1080/03043797.2017.1370578

Kurtuluş, A., & Uygan, C. (2010). The effects of Google Sketchup based geometry activities and projects on spatial visualization ability of student mathematics teachers. *Procedia-Social and Behavioral Sciences, 9*(2010), 384–389.

Kwon, H. (2017). Effects of 3D printing and design software on students' overall performance. *Journal of STEM Education: Innovations & Research, 18*(4), 37–42. https://www.learntechlib.org/p/181996/

Lachapelle, C. P., & Cunningham, C. M. (2014). Engineering in elementary schools. In S. Purzer, J. Strobel, & M. Cardella (Eds.), *Engineering in pre-college settings: Synthesizing research, policy, and practices* (pp. 61–88). Purdue University Press.

Lee, E., & Hannafin, M. J. (2016). A design framework for enhancing engagement in student-centered learning: Own it, learn it, and share it. *Educational Technology Research and Development*, *64*(4), 707–734. https://doi.org/10.1007/s11423-015-9422-5

Leon, A. (2014). *Stereotype threat: The effect of mindfulness on gender differences in spatial reasoning* (Doctoral Dissertation). Illinois Institute of Technology.

Linn, M. C., & 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 reanalysis of the correlational literature* (Technical Report No. 8). Stanford University. https://apps.dtic.mil/dtic/tr/fulltext/u2/a075972.pdf

López-López, J. A., Botella, J., Sánchez-Meca, J., & Marín-Martínez, F. (2013). Alternatives for mixed-effects meta-regression models in the reliability generalization approach: A simulation study. *Journal of Educational and Behavioral Statistics, 38*(5), 443–469. http://doi.org/10.3102/1076998612466142

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences, 49*(4), 344–351. https://doi.org/10.1016/j.paid.2010.03.022

Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years: Uncovering antecedents for the development of math-science expertise. *Perspectives on Psychological Science*, *1*(4), 316–345. https://doi.org/10.1111/j.1745-6916.2006.00019.x

Luccarelli, M., Tillman, S., Lay, R., Grundmeier, A.-M., & Högsdal, S. (2019). Sustainable design education for elementary schools: Interdisciplinary development of new educational models through design thinking. *International Journal of Design Education*, *13*(4), 1–25. https://doi.org/10.18848/2325-128X/CGP/v13i04/1-25

Lutz, S., & Huitt, W. (2004). Connecting cognitive development and constructivism: Implications from theory for instruction and assessment. *Constructivism in the Human Sciences*, *9*(1), 67–90.

Magana, A. J., Elluri, S., Dasgupta, C., Seah, Y. Y., Madamanchi, A., & Boutin, M. (2019). The role of simulation-enabled design learning experiences on middle school students' self-generated inherence heuristics. *Journal of Science Education and Technology*, *28*(4), 382–398. http://doi.org/10.1007/s10956-019-09775-x

Martín-Dorta, N., Saorin, J. L., & Contero, M. (2008) Developing of a fast remedial course to improve the spatial abilities of engineering students. *Journal of*

Engineering Education, 97(4), 505–513. https://doi.org/10.1002/j.2168-9830.2008.tb00996.x

Martin, P., & Velay, J. (2012). Do computers improve the drawing of a geometrical figure for 10–year-old children? *International Journal of Technology & Design Education*, *22*(1), 13–23. https://doi.org/10.1007/s10798-010-9140-6

Marulcu, I., & Barnett, M. (2013). Fifth graders' learning about simple machines through engineering design-based instruction using LEGO™ materials. *Research in Science Education, 43*(5), 1825–1850. https://doi.org/10.1007/s11165-012-9335-9

Marunić, G., & Glažar, V. (2014). Improvement and assessment of spatial ability in engineering education. *Engineering Review*, *34*(2), 139–150.

Matthews, D., & Geist, E. A. (2002). Technological applications to support children's development of spatial awareness. *Information Technology in Childhood Education Annual, 2002*(1), 321–336.

McConnell, W. J. (2015). *The impact of design-based modeling instruction on seventh graders' spatial abilities and model-based argumentation* [Doctoral dissertation, Old Dominion University]. ODU Digital Commons. https://digitalcommons.odu.edu/teachinglearning_etds/21/

Medina Herrera, L., Castro Pérez, J. & Juárez Ordóñez, S. (2019). Developing spatial mathematical skills through 3D tools: Augmented reality, virtual environments and 3D printing. *International Journal on Interactive Design and Manufacturing (IJIDeM), 13*(4), 1385–1399. https://doi.org/10.1007/s12008-019-00595-2

Meeth, R. L. (1978). Interdisciplinary studies: A matter of definition. *Change: The Magazine of Higher Learning, 10*(7), 10–11. https://doi.org/10.1080/00091383.1978.10569474

Mehalik, M. M., & Schunn, C. (2006). What constitutes good design? A review of empirical studies of design processes. *International Journal of Engineering Education, 22*(3), 519–532.

Miller, D. I., & Halpern, D. F. (2013). Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learning and Individual Differences, 26*(1), 141–152.

Moè, A. (2018). Mental rotation and mathematics: Gender-stereotyped beliefs and relationships in primary school children. *Learning and Individual Differences, 61*(1), 172–180. https://doi.org/10.1016/j.lindif.2017.12.002

Montag, M., Bertel, S., de Koning, B. B., & Zander, S. (2021). Exploration vs. limitation–An investigation of instructional design techniques for spatial ability training on mobile devices. *Computers in Human Behavior*, *118*(2021), 106678. https://doi.org/10.1016/j.chb.2020.106678

Muramatsu, K., Wangmo, S., & Wangchuk, Y. (2019, November). e-Design Education Using a 3D Printer Based on Design Thinking at Primary School. In *ECEL 2019 18th European Conference on e-Learning* (p. 412). Academic Conferences and publishing limited.

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

National Research Council. (2006). *Learning to think spatially.* The National Academies Press.

National Science Teachers Association [NSTA]. (2013). Teaching STEM in 3D. *NSTA Reports, 25*(2), 1–8.

Newbold, T. (2020*). Sensemaking and leader creative performance: The effect of understanding stakeholders' mental models* (Doctoral Dissertation). University of Oklahoma.

Newell, M. H., & Green, W. J. (1982). Defining and teaching interdisciplinary studies. *Improving College and University Teaching, 30*(1), 23–30. https://doi.org/10.1080/00193089.1982.10533747

O'Rourke, B. (2021, November 18). Growing gap in STEM supply and demand*. The Harvard Gazette*. https://news.harvard.edu/gazette/story/2021/11/increasing-access-and-opportunity-in-stem-crucial-say-experts/

Ogunkola, B., & Knight, C. (2019). Technical drawing course, video games, gender, and type of school on spatial ability. *The Journal of Educational Research*, *112*(5), 575–589. https://doi.org/10.1080/00220671.2019.1592092

Onwuegbuzie, A. J., & Daniel, L. G. (2000). *Reliability generalization: The importance of considering sample specificity, confident intervals, and subgroup differences.* Paper presented at the annual meeting of the Mid-South Educational Research Association (ED448204).

Onyancha, R. M., Derov, M., & Kinsey, B. L. (2009). Improvements in spatial ability as a result of targeted training and computer-aided design software use: Analyses of object geometries and rotation types. *Journal of Engineering Education*, *98*(2), 157–167. https://doi.org/10.1002/j.2168-9830.2009.tb01014.x

Ostapczuk, M., Wagner, M., & Musch, J. (2014). Revisiting the Rybakov Figures. Swiss *Journal of Psychology, 73*(2), 57–67.

Papert, S., & Harel, I. (1991). *Constructionism: Research reports and essays.* Ablex Publishing.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach.* Erlbaum.

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. https://doi.org/10.1006/brcg.1995.1032

Piaget, J. (2003). *The psychology of intelligence*. Routledge. (Original work published 1950)

Pritchard A. (2014). *Ways of learning – Learning theories and learning styles in the classroom.* Routledge.

Puente, S. M. G., van Eijck, M., & Jochems, W. (2013). A sampled literature review of design-based learning approaches: A search for key characteristics. *International Journal of Technology and Design Education*, *23*(3), 717–732. https://doi.org/10.1007/s10798-012-9212-x

Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and*

Individual differences, *40*(3), 609–619. https://doi.org/10.1016/j.paid.2005.07.015

Rabab'h, B., & Veloo, A. (2015). Spatial visualization as mediating between mathematics learning strategy and mathematics achievement among 8th grade students. *International Education Studies*, *8*(5), 1–11. http://doi.org/10.5539/ies.v8n5p1

Rafi, A., Anuar, K., Samad, A., Hayati, M., & Mahadzir, M. (2005). Improving spatial ability using a web-based virtual environment (WbVE). *Automation in Construction, 14*(1), 707–715. https://doi.org/10.1016/j.autcon.2004.12.003

Ratto, M., & Ree, R. (2012). Materializing information: 3D printing and social change. *First Monday*, *17*(7), Article 3968. https://doi.org/10.5210/fm.v17i7.3968

Riechert, S. E., & Post, B. K. (2010). From skeletons to bridges & other STEM enrichment exercises for high school biology. *American Biology Teacher, 72*(1), 20–22.

Robertson, K. F., Smeets, S., Lubinski, D., & Benbow, C. P. (2010). Beyond the threshold hypothesis: Even among the gifted and top math/science graduate students, cognitive abilities, vocational interests, and lifestyle preferences matter for career choice, performance, and persistence. *Current Directions in Psychological Science*, *19*(6), 346–351. https://doi.org/10.1177/0963721410391442

Rodriguez, J., & Genaro Rodriguez, L. (2016). *Comparison of spatial visualization skills in two approaches to entry-level graphic courses*. Proceedings of the 2016 ASEE annual conference, New Orleans, LA.

Rodriguez, M. C., & Maeda, Y. (2006). Meta-analysis of coefficient alpha. *Psychological Methods*, *11*, 306–322.

Rogers, C., & Portsmore, M. (2004). Bringing engineering to elementary school. *Journal of STEM Education: Innovations and Research*, *5*(3), 17–28.

RStudio Team (2019*). RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA. http://www.rstudio.com/.

Šafhalter, A., Vukman, K. B., & Glodež, S. (2016). The effect of 3D-modeling training on students' spatial reasoning relative to gender and grade. *Journal of Educational Computing Research, 54(*3), 395–406.

Salthouse, T. A., Babcock, R. L., Skovronek, E., Mitchell, D. R. D., & Palmon, R. (1990). Age and experience effects in spatial visualization. *Developmental Psychology*, *26*(1), 128–136. https://doi.org/10.1037/0012-1649.26.1.128

Sanandaji, A., Grimm, C., West, R., & Sanchez, C. A. (2021). Developing and validating a computer-based training tool for inferring 2D cross-sections of complex 3D structures. *Human Factors*. https://doi.org/10.1177/00187208211018110

Sánchez-Meca, J., López-López, J. A., & López-Pina, J. A. (2013). Some recommended statistical analytic practices when reliability generalization studies are conducted. *British Journal of Mathematical and Statistical Psychology*, *66*(3), 402–425. https://doi.org/10.1111/j.2044-8317.2012.02057.x

Seabra, R., & Santos, E. (2008). Evaluation of the spatial visualization ability of entering students in a Brazilian engineering course. *Journal for Geometry and Graphics, 12*(1), 99–108.

Seitamaa-Hakkarainen, P. (2011). Design based learning in crafts education: Authentic problems and materialization of design thinking. *Design Learning and Well-Being, 4*(1), 3–14.

Semma, B., Henri, M., Luo, W., & Thompson, C. G. (2019). Reliability generalization of the meaning in life questionnaire subscales. *Journal of Psychoeducational Assessment*, *37*(7), 837–851. https://doi.org/10.1177%2F0734282918800739

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591–611.

Sharma, G. V. S. S., & Dumpala, R. (2015). Teaching of mechanical engineering concepts through three-dimensional geometric modeling. *International Journal of Mechanical Engineering Education*, *43*(3), 180–190. https://doi.org/10.1177%2F0306419015603013

Shavalier, M. (2004). The effects of CAD-like software on the spatial ability of middle school students. *Journal of Educational Computing Research, 31*(1), 37–49. https://doi.org/10.2190%2FX8BU-WJGY-DVRU-WE2L

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. https://doi.org/10.1037/0022-0663.93.3.604

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Smith, S. (2018). Children's negotiations of visualization skills during a design-based learning experience using nondigital and digital techniques. *Interdisciplinary Journal of Problem-Based Learning*, *12*(2), Article 4. https://doi.org/10.7771/1541-5015.1747

Sorby, S. A., & Baartmans, B. J. (2000) The development and assessment of a course for enhancing the 3-D spatial visualization skills of first year engineering students. *Journal of Engineering Education,89*(3), 301–307. https://doi.org/10.1002/j.2168-9830.2000.tb00529.x

Tartre, L. A. (1990). Spatial orientation skill and mathematical problem solving. *Journal for Research in Mathematics Education, 21*(3), 216–229. https://doi.org/10.2307/749375

Thomas, W. J. (2000). *A review of research on project-based learning.* The Autodesk Foundation. https://tecfa.unige.ch/proj/eteach-net/Thomas_researchreview_PBL.pdf

Thompson, B. (1991). A primer on the logic and use of canonical correlation analysis. *Measurement and Evaluation in Counseling and Development, 24*(1)*,* 80–95.

Thompson, B. (2003). Understanding reliability and coefficient alpha, really. In B. Thompson (Ed.), *Score reliability* (pp. 3–30). SAGE.

Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement, 60*(2), 174-195. https://doi.org/10.1177%2F0013164400602002

Titze, C., Jansen, P., & Heil, M. (2010). Mental rotation performance and the effect of gender in fourth graders and adults. *European Journal of Developmental Psychology, 7*(4), 432–444.

Toptas, V., Celik, S., & Karaca, E. T. (2012). Improving 8th grades spatial thinking abilities through a 3D modeling program. *Turkish Online Journal of Educational Technology, 11*(2), 128–134.

Trust, T., & Maloy, R. W. (2017). Why 3D print? The 21st-century skills students develop while engaging in 3D printing projects. *Computers in the Schools*, *34*(4), 253–266.

Tumkor, S. (2018). Personalization of engineering education with the mixed reality mobile applications. *Computer Applications in Engineering Education*, *26*(5), 1734–1741.

Turgut, M. (2015). Individual differences in the mental rotation skills of Turkish prospective teachers. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, *5*(1) (EJ1061107). https://files.eric.ed.gov/fulltext/EJ1061107.pdf

U. S. Department of Energy. (2017, January 30). Vehicles per capita: Other regions/countries compared to the United States. Retrieved from https://www.energy.gov/eere/vehicles/fact-962-january-30-2017-vehicles-capita-other-regionscountries-compared-united-states

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how? *Psychology of learning and motivation, 57*(1), 147–181. https://doi.org/10.1016/B978-0-12-394293-7.00004-2

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement*, *58*(1), 6–20. https://doi.org/10.1177%2F0013164498058001002

Vacha-Haase, T., & Thompson, B. (2011). Score reliability: A retrospective look back at 12 years of reliability generalization studies. *Measurement and Evaluation in Counseling and Development*, *44*(3), 159–168. https://doi.org/10.1177%2F0748175611409845

Vacha-Haase, T., Henson, R. K., & Caruso, J. C. (2002). Reliability generalization: Moving toward improved understanding and use of score reliability. *Educational and Psychological Measurement, 62*(1), 562–569. https://doi.org/10.1177%2F0013164402062004002

Vacha-Haase, T., Kogan, L. R., & Thompson, B. (2000). Sample compositions and variabilities in published studies versus those in test manuals: Validity of score reliability inductions. *Educational and Psychological Measurement, 60*(4), 509–522.

Vacha-Haase, T., Kogan, L. R., Tani, C. R., & Woodall, R. A. (2001). Reliability generalization: Exploring variation of reliability coefficients of MMPI clinical scales scores. *Educational and Psychological Measurement*, *61*(1), 45–59. https://doi.org/10.1177%2F00131640121971059

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*(2), 599–604. https://doi.org/10.2466%2Fpms.1978.47.2.599

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 mathematics skills. *Child Development, 85*(3), 1062–1076. https://doi.org/10.1111/cdev.12165

Veritas Health Innovation (2021). *Covidence systematic review software*. Melbourne, Australia. Available at www.covidence.org

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Vieira, C., Goldstein, M. H., Purzer, Ş., & Magana, A. J. (2016). Using learning analytics to characterize student experimentation strategies in the context of engineering design. *Journal of Learning Analytics*, *3*(3), 291–317. https://doi.org/10.18608/jla.2016.33.14

von Károlyi, C. (2013). From Tesla to Tetris: mental rotation, vocation, and gifted education. *Roeper Review*, *35*(4), 231–240. https://doi.org/10.1080/02783193.2013.829547

Voyer, D., & Doyle, R. A. (2010). Item type and gender differences on the Mental Rotations Test. *Learning and Individual Differences, 20*(5), 469–472.

Voyer, D., & Saunders, K. A. (2004). Gender differences on the mental rotations test: a factor analysis. *Acta Psychologica, 117*(1), 79–94.

Voyer, D., Butler, T., Cordero, J., Brake, B., Silbersweig, D., Stern, E., & Imperato-McGinley, J. (2006). The relation between computerized and paper-and-pencil mental rotation tasks: A validation study. *Journal of Clinical and Experimental Neuropsychology, 28*(6), 928–939.

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,* 250–270.

Vygotsky, L. (1978). *Mind in society*. Harvard University Press.

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), 817–835.

Williams, A. M. & Young, J. (2021). Reliability generalization meta-analyses in mathematics education research: A research synthesis. *International Journal of Education in Mathematics, Science, and Technology* (IJEMST), *9*(4), 741–759. https://doi.org/10.46328/ijemst.1434

Williams, J., Roth, W. M., Swanson, D., Doig, B., Groves, S., Omuvwie, M., Ferri, R. B., & Mousoulides, N. (2016). *Interdisciplinary mathematics education*. Springer.

Workman, J. E., & Zhang, L. (1999). Relationship of general and apparel spatial visualization ability. *Clothing and Textiles Research Journal, 17*(4), 169–175.

Yildiz, S. G., & Ozdemir, A. S. (2018). The effects of engineering design processes on spatial abilities of middle school students. *International Journal of Technology and Design Education*, *30*(1), 127–148. https://doi.org/10.1007/s10798-018-9491-y

Yurt, E., & Tünkler, V. (2016). A study on the spatial abilities of prospective social studies teachers: A mixed method research. *Educational Sciences: Theory and Practice, 16*(3), 965–986.

Zahner, W., Dai, T., Cromley, J. G., Wills, T. W., Booth, J. L., Shipley, T. F., & Stepnowski, W. (2017). Coordinating multiple representations of polynomials: What do patterns in students' solution strategies reveal?. *Learning and Instruction, 49*(1), 131–141.

Zavotka, S. L. (1986). Analysis of three dimensional computer graphics animation to teach spatial skills to interior design students. *Journal of Interior Design*, *12*(1), 45–52. https://doi.org/10.1111/j.1939-1668.1986.tb00085.x