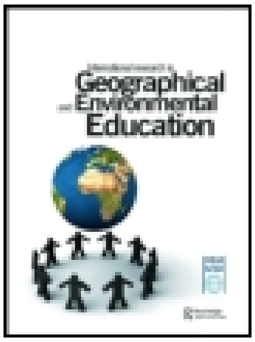

International Research in Geographical and Environmental Education

ISSN: 1038-2046 (Print) 1747-7611 (Online) Journal homepage: https://www.tandfonline.com/loi/rgee20

What improves spatial thinking? Evidence from the Spatial Thinking Abilities Test

Robert Bednarz & Jongwon Lee

To cite this article: Robert Bednarz & Jongwon Lee (2019): What improves spatial thinking? Evidence from the Spatial Thinking Abilities Test, International Research in Geographical and Environmental Education, DOI: 10.1080/10382046.2019.1626124

To link to this article: https://doi.org/10.1080/10382046.2019.1626124

Published online: 12 Jun 2019.

INTERNATIONAL RESEARCH IN GEOGRAPHICAL AND ENVIRONMENTAL EDUCATION
https://doi.org/10.1080/10382046.2019.1626124

What improves spatial thinking? Evidence from the Spatial Thinking Abilities Test

Robert Bednarz[a] and Jongwon Lee[b]

[a]Department of Geography, Texas A&M University, College Station, TX, USA; [b]Department of Social Studies Education, Ewha Womans University, Seoul, South Korea

ABSTRACT

Geography educators have been, and continue to be, interested in spatial thinking, especially since they have been convinced of its importance in their students' ability to learn and do geography. As they developed strategies to improve students' spatial thinking, they searched for assessment instruments to evaluate their interventions' effects. Because they and we found existing psychometric tests of spatial ability lacking, we developed the Spatial Thinking Ability Test. This article reviews 22 research studies that have used the test to assess spatial thinking. Both instructional strategies or methods and individuals' characteristics that lead to, or are associated with, improved spatial thinking are identified. Suggestions for further research are offered.

KEYWORDS
Spatial thinking; assessment; STAT; STEM

Introduction

Geographers have been interested in spatial thinking for a long time. This is not surprising since geography is a discipline that is often inherently spatial. In the past, cartographers were concerned about how they could effectively communicate with the users of their products. They were interested in the ability of map-readers to differentiate among different line weights, patterns, or grey tones. Geographers who studied environmental perception were also involved in how humans understood the environments in which they lived and worked. Much of the current interest in spatial thinking can be attributed to the publication of *Learning to Think Spatially* (Bednarz & Lee, 2011; National Academy of Sciences, 2006). This publication argued strongly for consideration of spatial thinking as another type of human cognition similar to verbal and numerical cognition.

Recent research provides empirical support for the importance of spatial thinking ability in the science, technology, engineering, and mathematics (STEM) domains (Moorman & Crichton, 2018; Wai, Lubinski, & Benbow, 2009). Several correlational studies suggest the relation between STEM achievement and spatial skills—high STEM achievers tend to have high spatial skill levels. Spatial ability emerged as a

CONTACT Jongwon Lee jongwonlee@ewha.ac.kr Department of Social Studies Education, Ewha Womans University, 52 Ewhayeodae-gil Seodaemun-gu, Seoul, 03760, South Korea

consistent and statistically independent predictor of selecting STEM related courses, graduate study, and other measures of STEM attainment (Sorby, Veurink, & Streiner, 2018). Moreover, identification and development of STEM talent has become a national priority (National Science Board, 2010), and research supports the contention that spatial ability is a predictor of success in these fields (Lubinski, 2010). Nevertheless, this ability is rarely measured and is relatively neglected in the general practice of teaching and learning in the K–12 setting.

Spatial thinking assessment

Once educators realized the important role of spatial thinking in their disciplines, they quickly began to explore educational interventions that might improve the spatial thinking of their students. As they developed new pedagogic approaches, they also searched for methods to assess their students' spatial thinking ability and the changes that their interventions produced. Geography educators initially turned to the wide variety of psychometric, spatial-ability tests developed, for the most part, by cognitive psychologists. In fairly short order, however, researchers began to express dissatisfaction with these assessment instruments.

Geographers and other geo-science educators offered a number of reasons for their dissatisfaction. First, these paper-and-pencil tests (e.g. paper folding, imbedded figure, or card rotation tests) assess spatial ability at a small or table-top scale. Geographers, however, were most often concerned with much larger areas than table-top-scale environments. Researchers were not convinced that demonstrated spatial ability at this smaller scale would transfer to an understanding of processes, phenomena, and problems at a typically larger geographic scale (Charcharos, Kokla, & Tomai 2015; Golledge, 1993; Mark & Freundschuh, 1995).

Second, psychometric tests were designed to measure spatial ability, usually focusing on two, spatial visualization and spatial orientation (e.g. Goldstein, Haldane, & Mitchell, 1990; Kali, Orion, & Mazor, 1997; McGlone, 1981; Miller & Santoni, 1986; Newcombe & Dubas, 1992). Spatial thinking, however, is understood to be a broader concept that encompasses spatial ability, that is, spatial ability can be understood as a subset of spatial thinking (Lee & Bednarz, 2009). In fact, Ishakawa (2013) asserts that various terms such as spatial ability, spatial cognition, or spatial intelligence are used interchangeably without clear definitions. In most instances, performance in geography is associated with the broader ability or skill, spatial thinking. To address this concern, Golledge and Stimson (1997) argued that spatial relations should be included in the definition of spatial thinking. In defining spatial relations, they included the following.

> … to recognize spatial distributions and spatial patterns, to connect locations, to associate and correlate spatially distributed phenomena, to comprehend and use spatial hierarchies, to regionalize, to orientate to real-world frames of reference, to imagine maps from verbal descriptions, to sketch map, to compare maps, and to overlay and dissolve maps (Golledge & Stimson, 1997, p. 158).

At the present time, no consensus exists about whether spatial relations is as fundamental a component of spatial ability as is visualization and orientation (Gilmartin

& Patton, 1984; Golledge & Stimson, 1997; Lohman, 1979; Montello, Lovelace, Golledge, & Self, 1999). Nevertheless, Montello et al. (1999) express agreement with Golledge and Stimson when they say, "… the restricted definition of spatial ability, as incorporated into many psychometric tests, contrasts with the richness of the general literature on spatial activities and spatial behavior…" (Montello et al., 1999, p. 517).

Other researchers have expressed similar concerns when they worry that spatial thinking in a particular discipline goes beyond what is typically measured by psychometric tests. For example Libarkin and Brick observed

> Visualization in a specific topic requires a unique set of skills; visualization of earth processes requires spatial and temporal projections not encountered in available assessment tools. Certainly, the field would benefit from instruments specifically designed for studying learning in the earth system (Libarkin & Brick, 2002, p. 453).

Charcharos, Kokla, and Tomai (2015, p. 160) express a similar sentiment when they argue that, if geospatial thinking is to be assessed well, "[t]he role of the GIscientist is crucial for the development of the spatial thinking ability test…"

Finally, the dissatisfaction with psychometric test persists as demonstrated by Ishikawa's (2016, p. 78) assertion that, "Although there is a long tradition of psychometric testing of spatial ability, formal methods for assessing spatial thinking have not been developed…" and Sharpe and Huynh's (2015, p. 178) recommendation "that new tools be developed to assess the use of geospatial thinking practices."

The concerns expressed by these researchers have been born out in many studies. For example, in a study whose purpose was to "1) investigate the pre-existing cognitive skills and abilities of students enrolled in an introductory university course of GIS and 2) examine the relationship these cognitive factors have on the students' success in the class" (Vincent, 2004, pp. 5–6), students' spatial abilities proved to have little or no relationship to their performance in the course. Spatial abilities were measured by the Hidden Pattern Test (Ekstrom, French, Harman, & Dermen, 1976), Card Rotation Test (Ekstrom et al., 1976), and the Vandenberg Mental Rotations Test (Vandenberg & Kuse, 1978) while class performance was measured by three scores: final grade, examination grades, and project grade. As might be expected, the first two scores were significantly correlated with students' grade point average. In the study's conclusions, Vincent (2004, pp. 110–111) states that "[t]he results … showed that these cognitive factors [spatial abilities] had virtually no relationship with the success of students taking the GIS class. Only two of the factors, geography attitude and learning style, were significantly correlated with the project grade."

Similarly, when Ishakawa (2013) explored relation between the spatial visualization ability of students (card rotation, surface development, cube comparison, and paper folding) and their geospatial thinking performance, he found highly significant correlations among the visualization scores. The visualization scores, however, were not strongly correlated with the scores for geospatial thinking tasks. Performance on only one of the tasks, understanding map projection, was significantly correlated with visualization ability.

These concerns, how well existing psychometric tests can assess the broader spatial thinking required in geography and the weak, if any, relationship between psychometric test scores and performance on geospatial-thinking tasks, led researchers to express a desire for a more appropriate assessment instrument.

Among the first to call for, and then develop, a test that measured factors beyond those assessed by psychometric tests were Kali et al (1997). Their instrument required students to draw cross-sections of, and imagine, block diagrams of geologic structures, an activity that they felt more accurately reflected performance in earth science.

Ishakawa (2013) noted that although assessments of spatial ability were well developed, an instrument to evaluate spatial thinking did not exist. This led him to develop a geo-spatial thinking test to assess abilities with respect spatial distributions, frames of reference, map projection, map scale, and earth's movement.

Others also created their own test instruments designed for their particular research projects (Audet & Abegg, 1996; Kerski, 2000; Lim, 2005; Meyer, Butterick, Olkin, & Zack, 1999; Olsen, 2000). Unfortunately, few of the tests were developed using rigorous, assessment-development processes (Lee & Bednarz, 2012).

It is understandable why researchers designed special-purpose tests to measure the impact of their interventions, especially given the general dissatisfaction with existing assessment instruments. It is also important to note that in at least one recent effort best-practice test-development methods were employed to create "a tool to assess learners' spatial thinking about enhanced greenhouse effect" (Skaza, 2016, p. 6).

To address the need for an instrument to assess spatial thinking, especially in a geography and earth science context, the authors developed the Spatial Thinking Ability Test (STAT). The STAT is a revised version of the Spatial Skills Test (SST). The SST was created to examine the effect of GIS learning on the spatial thinking ability of college students in 2005. The components of spatial relations as defined by Golledge and Stimson (1997) provided the structure for developing the test. Each test item was designed to measure a component or a trait of spatial relations. The spatial-skills test consists of a set of multiple-choice questions and performance tasks. It was designed to evaluate students' spatial thinking ability, including overlaying and dissolving a map, reading a topographic map, evaluating several factors to find the best location, recognizing spatially correlated phenomena, constructing isolines based on point data, and differentiating among spatial data types. The final version of the spatial skills test includes seven types of question items (Lee & Bednarz, 2009).

Development of the spatial thinking ability test

The initial step in the construction of the STAT was the delineation of the assessment objective and the description of the content to be measured. The STAT incorporated more recent work concerning key spatial thinking concepts from several studies including *Learning to Think Spatially* (National Academy of Sciences, 2006) and Gersmehl and Gersmehl's (2007) spatial thinking taxonomy. Golledge's list of geographic thinking elements presented in 2002 also served to guide the development of the STAT. This was especially useful because the elements were detailed enough to guide in the development of test items, potentially leading to an improvement of the test's content validity. During the development of STAT, a number of other factors were also considered. These factors included (1) cognitive process (i.e., maximizing spatial processes and minimizing verbal processes); (2) psychometric rationale; (3) mode of representation (text, picture, graph, map, color versus black and white, etc.); and (4) practical constraints (e.g., amount of time required to complete the test).

Using past work to develop test question categories (Gersmehl and Gersmehl 2007; Golledge, 2002; Janelle and Goodchild 2009), the STAT was designed to test eight spatial thinking ability components using sixteen test questions. The following list outlines the eight spatial thinking ability components (indicated with roman numerals) with the corresponding STAT questions: Questions 1 and 2 assess spatial thinking abilities that are related to comprehending orientation and direction (Type I). The tasks are to undertake way finding and route planning by visually navigating road maps using verbal information including participant's current location, directions to destination, and street information. Question 3 evaluates spatial thinking abilities of recognizing patterns in a map and representing the patterns in a graphic form (Type II). Question 4 assesses the student's ability to select an ideal location based on the given spatial features (such as land use, elevation, and population density) and based on comprehension of the concept of overlay (Type III). Question 5 asks students to visualize a slope profile based on a topographic map (Type IV). Questions 6 and 7 measure abilities to identify spatial correlations, either positive or negative, by comparing patterns featured in a set of maps and to represent such spatial relationships in a graphic form (Type V). In order to answer Question 8, students had to transform representations and images from one dimension to another (Type VI). Students were asked to mentally visualize 3D terrains based on a two-dimensional topographic map. Questions 9–12 concern the understanding of various types of overlaying and dissolving processes and the ability to apply them to select images by mentally executing such processes (Type VII). The last four questions (Questions 13–16) evaluate comprehension of geographic features represented as points, lines, or polygons (Type VIII). In order to solve these problems, students must understand different types of spatial data and be able to apply that knowledge to identify an appropriate data type (i.e. points, lines, or polygons) to represent various real-world geographic features (e.g. weather stations, rivers, and bus routes) (Figure 1).

The current version of the STAT is fourteen pages long and has two equivalent forms (one that can be used for a pre-test and one for a post-test) allowing for the evaluation of changes in spatial thinking skills over a period of time. Before the final version of the assessment was established, a pilot study was conducted with a relatively small number of participants (higher education students) in order to assess the validity and reliability of the test and to correct as many errors and omissions as possible (questions that have more than one correct answer, questions that are quite difficult, adjusting the required time, clarifing difficult-to-understand questions). Following the pilot, test results from 352 university students from four different U.S. states who took STAT were used to examine the reliability and validity of STAT. Analysis indicated that STAT has moderate reliability and construct validity.

Since the STAT has been developed and disseminated via dissertation work and via journal publications, many researchers have used it to measure spatial thinking. It has been used in a wide variety of situations, and over a wide age range spatial thinkers.

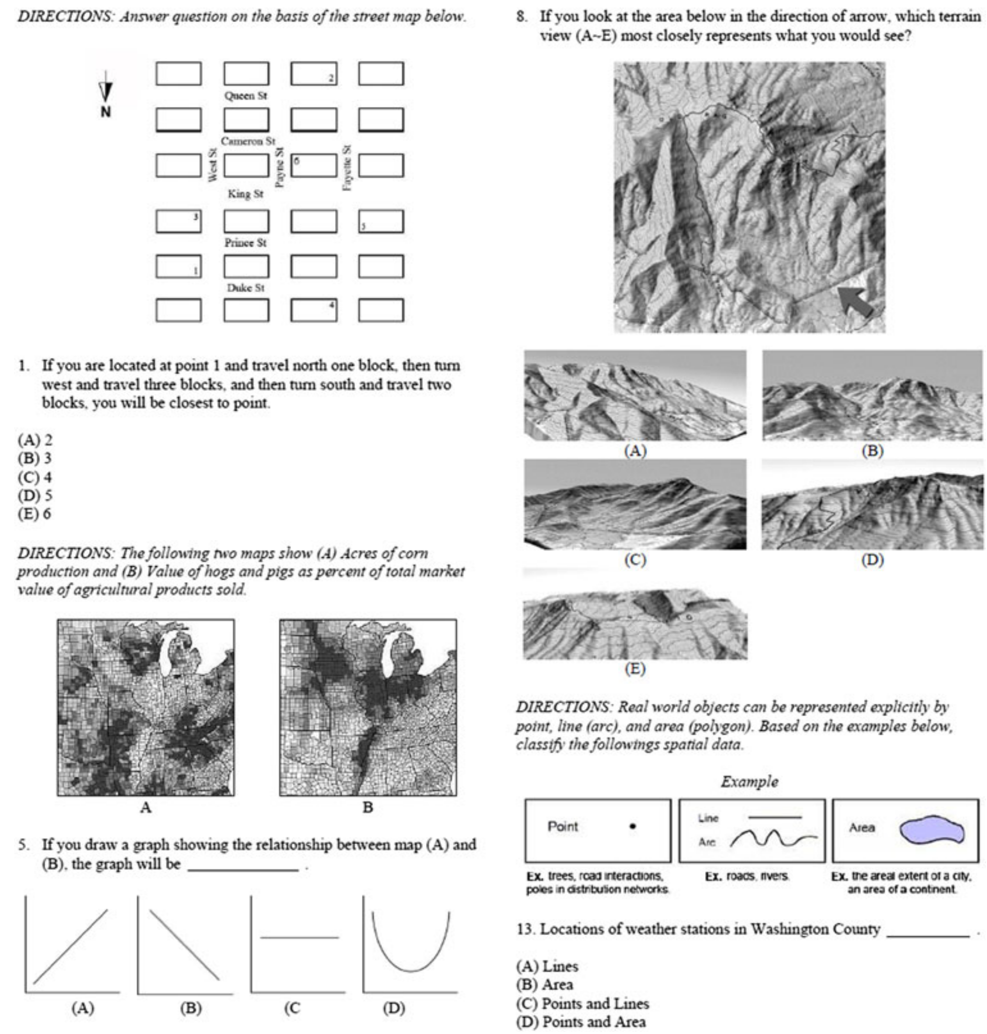

Figure 1. Selected items from the STAT.

Research results

In order to explore how the test has been used and what factors produced changes in students' spatial thinking scores, we conducted a systematic review of published research projects that adopted either SST or STAT into their research design. The review is based on works contained in relevant databases, including Web of Science, Taylor & Francis, and Google Scholar. The search protocol was complemented with a "manual" search based on key words. The initial search identified 272 studies. After eliminating those that only cited but did not administer SST or STAT, 22 studies remained, and these were analyzed in depth. The studies included in the analysis were conducted in eight countries: Japan, Indonesia, Singapore, China, Netherlands, United States, Brazil, and Rwanda (see Table 1).

In order to provide structure to the analysis of the 22 studies, we evaluated them using a set of criteria: (1) Setting (e.g. countries, age levels, study size); (2) type of

Table 1. Research studies reviewed.

Study	Setting	Test	Intervention (Research Questions)	Outcomes
Low et al. (2014)	USA; Students from three colleges	SST	ArcGIS Online course focusing on a place-based quantitative exploration of the impacts of environmental changes	Participants in the course demonstrated statistically significant improvements in spatial skills.
Low et al. (2014)	Singapore; 20 A-level geography students (16–17 year old)	STAT	Three GIS lessons concerning tectonic hazards	The GIS lessons had a positive impact on students' spatial thinking.
Kim and Bednarz (2013)	USA; 32 university students from either GIS or economic geography or education course	STAT (21 items used)	To determine effect of GIS learning on exercising spatial reasoning	GIS students improved their spatial reasoning more than education course students but not economic geography students.
Li and Liu (2018)	USA; 552 8–12th grade	STAT (8items used)	To determine the status of middle and high school students' spatial thinking ability and impact of integration of GIS (i.e. Google Earth, ESRI StoryMap) in science classes on students' spatial thinking ability.	12th grade students' spatial thinking ability was better than others. Learning with Google Earth or StoryMap enabled students to see different geographic views, such as 2D maps, 3D maps and satellite maps and the relationship between different locations, which is associated with several spatial components such as map overlay.
Jo et al. (2016)	USA; 307 university students from Texas and Georgia	STAT	Investigates the effectiveness of a Web-based GIS application as an instructional tool for world geography in improving students' spatial thinking skills	Only the experimental classes showed a significant increase in STAT scores.
Collins (2018)	USA; 327 8th grade students (digital instruction 106, traditional 111, control 110)	STAT	To determine whether spatial learning outcomes differ with respect to different instructional media	Both intervention groups' scores increased significantly, whereas the control group did not. The two intervention groups' scores did not improve in the same areas.
Hedley et al. (2013)	USA; 207 high and 96 junior high students in grades 7–12	SST	To determine the extent to which the use of geospatial technologies within a STS (student/teacher/ scientist) partnership improved the geospatial skills and atmospheric science concept knowledge of high school and junior high school students.	The mean scores of both SST and science knowledge tests increased from the pre-test to the evaluation-phase test. Both male and female students experienced approximately the same level of increase in SST scores.

(*continued*)

Table 1. Continued.

Study	Setting	Test	Intervention (Research Questions)	Outcomes
Crews (2008)	USA; 52 students (grades 9-12)	SST	To understand how a geospatial technologies professional development project for teachers and the subsequent implementation of a teacher-developed geospatial curriculum influenced student spatial literacy skills and students' interests in science and technology	Students' scores on the SST insignificantly decreased from a mean score of 10.23 to 10.12 on the post-test. Males' scores increased by 3%; females' decreased by 5.4%.
Powell and Kong (2017)	USA; 20 African studies university students	STAT	Explores how librarians' participation as instructors in week-long GIS-enhanced Digital Humanities (DH) classes can advance a variety of library objectives. The STAT was administered as a part of a background survey.	About 27 percent of the participants felt that they lacked the spatial thinking skills to further their research in the spatial humanities.
Flynn (2018)	106 American university students and 28 Ethiopian university students	STAT	Aims to determine whether a low-cost experiential-based learning method incorporating a geocaching activity can strengthen spatial thinking skills.	The results suggest that the geocaching activity significantly improved the spatial thinking of the grouping of all students. These students showed significant improvements in orientation and directional abilities, spatial overlay and dissolve, and points, networks, regions/spatial shapes and patterns.
Weakley (2010)	USA; 290 university students	SST (+ two items)	To determine the effects of a pre-service teacher education earth-science-content course, conceptually designed and inquiry-based, on the spatial thinking of university students compared to the course that follows a lecture format	The experimental group and comparison group both demonstrated statistically significant higher post-test SST scores.
Wakabayashi (2015)	Japan; 90 university students	STAT (Rev.)	To examine the factors affecting spatial thinking test scores	The scores for three question items (spatial patterns, spatial correlation, and landscape visualization) were closely related to the students' interest in geography and maps, as well as their sense of direction.

(*continued*)

Table 1. Continued.

Study	Setting	Test	Intervention (Research Questions)	Outcomes
Verma, (2015)	USA; 77 university students	STAT (only 6 items)	To assess the geospatial thinking differences of undergraduate students.	Academic major (i.e., geography major) and number of college geography courses taken influenced students' geospatial thinking.
Verma (2014)	USA; 1,479 students in 61 public universities	STAT (10 items)	To assess group variances in geospatial thinking abilities of undergraduate students	Ethnicity, along with socioeconomic status, and completed geography courses were the most important variables in understanding, influencing, and predicting undergraduate students' geospatial thinking.
Wan et al. (2017)	China; 126 middle school students	STAT	To identify factors affect spatial thinking ability.	General intelligence level, the habit of using maps, geography knowledge, and interest in geography were the factors most correlated with students' high spatial thinking ability
Shin et al. (2016)	USA; 28 elementary education majors, 38 secondary social studies education majors, and 37 geography majors university students	SST	To explore the effect of gender, age, geography courses, geographic information systems (GIS) training, or travel experiences on the spatial thinking ability	The respondents' SST scores showed statistically significant mean differences by major, gender, GIS courses, and international travel experience.
Tomaszewski et al. (2015)	Rwanda; Approximately equal numbers of students (74, 73, 75) from three schools, an urban experimental and control school and a rural experimental school	STAT	To explore relationship between gender, school grade (i.e., age), and STAT question group	Logistical regression results indicated that main effects of school, gender, question group, and interaction for school and question group were significant. The urban experimental school performed best on average and on each question group except visualizing 3-D images from 2-D information. Males outperformed females across all other variables.
Indhawati (2015)	Indonesia; 52 GIS major and 43 non-GIS major (Informatics Engineering) student	Partial STAT	To find the relationship between GIS learning outcome levels and spatial thinking levels	GIS majors scored significantly higher than non-GIS major. Results yielded no significant difference between male and female subjects on score improvement.
Bednarz and Lee (2011)	USA; 446 individuals: 52 secondary, 149 tertiary, and 245 university students.	STAT	To determine whether components of spatial thinking can be identified	The analysis suggests that spatial thinking is, in fact, not a single skill or ability but a collection, including map visualization and overlay, discerning symbology, and recognizing patterns.

(*continued*)

Table 1. Continued.

Study	Setting	Test	Intervention (Research Questions)	Outcomes
Duarte (2016)	Brazil; 268 students of six different Brazilian schools (9th grade)	STAT	To determine the level of spatial thinking ability of Brazilian students	Brazilian results were compared with those of two international studies (one from the US and three from Rwanda). There were no significant differences among groups.
Ishikawa (2013)	Japan; 32 university students	STAT	To explore the similarities and differences between spatial ability and spatial thinking	The STAT had larger correlations with the spatial ability tests than did the geospatial thinking questions (developed by Ishikawa).
Firdiansyah (2012)	The Netherlands 35 university students; 27 geography major 8 non-geography major	Partial STAT	To develop test material for evaluation of spatial thinking in geographical space.	The test instrument was developed and tested for 35 individuals for evaluation.

intervention (e.g. GIS learning); and (3) outcome(s). It is necessary to note that there are several issues that complicate the comparability of test scores across contexts, interventions, and grades. This is especially true when researchers used only a portion of the SST/STAT questions, added their own items or modified SST/STAT or modified the test in some other way. Applying standard measures to all of the studies can nevertheless help us compare the scores across the variety of research projects.

The test has been used to explore a wide range of effects and relationships. The first group of studies we identified explored the relationship between GIS learning and the development of spatial thinking skills. Thus, nine of 22 studies utilized SST/STAT to measure changes in students' spatial thinking resulting from GIS teaching and learning. Six studies examined the effect of students' GIS learning on spatial thinking skills, while the remaining three explored how teachers' professional development, focused on teaching with GIS, could influence their students' spatial thinking.

For the most part, outcomes were positive, that is, GIS learning was positively correlated with gains in students' spatial thinking scores. The effect of students' GIS learning on spatial thinking abilities seemed to be consistent regardless of target subjects' age and gender (Collins, 2018; Jo, Hong, & Verma, 2016; Kim & Bednarz, 2013; Low, Tan, & Huat, 2014; Low, Boger, & Mandry, 2014). Through learning with GIS, students are exposed to various concepts related to spatial thinking ability. For example, Collins (2018), who studied the effect of digital map-based instruction (i.e., Google Earth) on spatial thinking ability of eighth grade students, concluded that students in the experimental group were more directly exposed to instruction about utilizing points, lines, and polygons unlike students in the control group who were not exposed to similar instruction. Similarly, Li & Liu (2018, p. 33) who investigated the effect of the integration of GIS (i.e. Google Earth, ESRI StoryMap) in science classes on students' spatial thinking ability explained:

> The study results also indicated Google Earth allowed students to see more features of map, such as different scales, and various information of different locations, which associated with spatial component of "comprehending integration of geographic features" (Golledge, 2002). For the Storymap, students were allowed to see the relationship between different locations, which is associated with several spatial components such as map overlay (Golledge, 2002).

The impact of teachers' GIS training and their use of GIS in the classroom was not as uniform. Two studies found a positive impact on spatial thinking, for the other the results were mixed. For example, Hedley, Templin, Czajkowski, & Czerniak (2013) reported that use of geospatial technologies within a STS (student/teacher/scientist) partnership had a positive and significant effect on test scores of secondary school students. On the other hand, Crews (2008) who studied participants in a teacher geospatial technologies professional development project did not find similar increases in students' spatial thinking scores.

A second group of studies attempted to ascertain how students' characteristics or experiences influenced their spatial thinking abilities. A series of statistical analyses were conducted to identify what factors affected spatial thinking. The impact of subjects' personal attributes—gender, age, residential location (i.e. urban or rural), and socioeconomic status—and their experiences—number of geography courses completed, academic major, and domestic and international travel—on spatial thinking were evaluated. Geography as academic major and number of geography courses completed were the most consistently positive factors influencing subjects' spatial thinking abilities (Shin, Milson, & Smith, 2016; Wakabayashi, 2015; Wan, Lu, Du, Wang, & Ju, 2017). Verma (2014) found that both US geography and geology students scored higher than nursing, health psychology, criminal justice, education, communication, business, other social science majors, as well as students with an undeclared major. When the analysis was extended to the level of the type of test item, geography majors scored higher than students majoring in other fields, especially with respect to spatial visualization and the ability to identify spatial correlation and visualization (Wakabayashi, 2015). In addition to the number of geography courses at the college level, the number of high school geography courses completed also showed a strong and positive correlation with test scores (Verma, 2014). This result is in accordance with the findings of Lee and Bednarz (2009). They also found that students who completed two or more geo-spatial technology courses (GIS and cartography) scored significantly higher on a post-test STAT than students who took only one GIS or cartography course.

Researchers also found that improvements in spatial thinking scores were positively correlated with students' travel experiences although the relationship was relatively weak (Collins, 2018). Shin et al. (2016) found a significantly higher mean STAT score for subjects who had travelled internationally at least once than the mean score of participants who had no international travel experience. Domestic travel did not have a significant impact, however. Along the same lines, Wakabayashi (2015) found a significant relationship between outdoor activities and test scores on items measuring spatial correlation and visualization.

Other factors were also found to positively influence spatial thinking abilities in some cases. For example, Verma (2014), in her study of US students, reported that ethnicity, along with socioeconomic status were the most important variables in understanding, influencing, and predicting undergraduates' geospatial thinking ability. Scores of students with higher socioeconomic status (parents with higher annual income and higher education degree) were significantly higher than those for students with lower socioeconomic status (parents with lower annual income and no college degree). A similar socioeconomic effect was not found in a study of US secondary school students (Hedley et al., 2013) nor in a Chinese study (Wan et al., 2017).

Gender is a factor that has been explored frequently with regard to its effect on spatial thinking. Our review of the research, however, found that the gender effect was variable, insignificant, and inconclusive. Shin et al. (2016), Crews (2008), and Tomaszewski, Vodacek, Parody, & Holt (2015) found males scored higher, but Collins (2018), Hedley et al. (2013), Li and Liu (2018), Verma (2014), and Wakabayashi (2015) reported that there was no clear gender difference reflected in the test scores, leaving room for further research.

The several studies that analyzed the differences between the spatial thinking abilities of students from rural environments with those from urban environments provided inconsistent results. In Rwanda, Tomaszewski et al. (2015) found that urban students performed better than rural students, whereas in the US, Verma (2014) reported that both suburban and rural students outscored urban students. Considering that, unlike in the US, the Rwandan urban students often had better educational backgrounds and opportunities than those in rural area, the contradictory results are not surprising. Other factors that were found to correlate with participants' spatial thinking skills included the habit of regularly using maps (Wakabayashi, 2015; Wan et al., 2017), age (Li & Liu, 2018; Verma, 2014), and general intelligence level (Wan et al., 2017).

A majority, 13 of the 22 studies analyzed, used SST/STAT without modification; the remaining 9 studies omitted items, added items or further modified the tests in some fashion. For example, Verma (2014) omitted 6 STAT items, using only 10, whereas Weakley (2010) added two items to represent more specifically the content of the intervention (i.e. geography course). Culturally dependent factors imbedded in the test were also modified by users (e.g. converting English to metric units and employing maps of Africa instead of the United States). These modifications did not, for the most part, change the basic nature of spatial thinking ability components measured by the test (Tomaszewski et al., 2015). In a very limited number of cases, however, researchers transformed STAT rather dramatically. For instance, Wakabayashi (2015) revised the test by replacing images, changing the format, and renaming it J-STAT.

Analysis of the selection and use of the STAT/SST provides information about the strengths, limitations, and area for improvement regarding the assessment tools as well as the results of using the assessment tools. This information is useful for developing better assessment tools to measure spatial thinking. The strengths, limitations, and area for improvement identified in the process of selection, utilization, and analysis of the STAT/SST are summarized with respect to three criteria: format, content selection, and target population (Table 2).

The advantage of STAT/SST is that it is almost the only standardized, reliable assessment tool to evaluate spatial thinking ability (Collins, 2018; Crews, 2008; Jo

Table 2. Identified strengths and weaknesses of STAT/SST.

	Strength/advantages	Weakness/disadvantages/limitations
Format & Reliability	Standardized (Collins, 2018), tested for reliability and validity (Collins, 2018; Crews, 2008; Jo et al., 2016; Tomaszewski et al., 2015) Two equivalent forms allowing pre- and post-test (Crews, 2008) Suitable for replication studies with different populations and settings (Duarte, 2016; Flynn, 2018; Jo et al., 2016; Tomaszewski et al., 2015)	Unclear directions (Collins, 2018; Tomaszewski et al., 2015) A set of questions with low reliability (Verma, 2014)
Content coverage	Useful to determine and compare scores of sub-categories of spatial thinking ability (Tomaszewski et al., 2015) Integrating geographic content and skills (Collins, 2018)	Not covering concepts such as scale, frames of reference, regionalization (spatial classification), spatial diffusion, spatial hierarchy, and spatial analog (Verma, 2014) Knowledge and skills are mixed (Ishikawa, 2013; Wakabayashi, 2015)
Target population	Can be used for various ages from junior high students to adults (Jo et al., 2016)	Cannot used for young students (Crews, 2008) Some questions may not be age-appropriate (Collins, 2018)

et al., 2016; Tomaszewski et al., 2015). Because it employs multiple choice questions and consists of two equivalent forms allowing pre- and post-test comparison, it is relatively easy to compare the scores of the groups even if they are located in different regions or speak different languages (Crews, 2008). In addition, because a relatively large number of studies utilized STAT, it is possible to compare the outcomes of the studies (Duarte, 2016; Flynn, 2018; Jo et al., 2016; Tomaszewski et al., 2015). It should be noted, on the other hand, that some researchers stated that the directions for some items were not clear (Collins, 2018; Tomaszewski et al., 2015) and others had low reliabilities (Verma, 2014).

The validity of test's content is enhanced because it was based on an analysis of existing research about spatial thinking ability (Collins, 2018; Jo et al., 2016). Furthermore, since the STAT items are constructed to measure individual components of spatial thinking ability, it is also possible to compare not only the overall spatial thinking scores of an individual or a group, but also the differences among these components. These are important because this information can provide educational interventions.

> The results of this study show that it is likely that different spatial thinking skills are best taught by different media. Some skills may be acquired best with paper maps and others with digital maps. Therefore, when designing curriculum to teach spatial thinking skills, it may be more beneficial to focus specifically on what skills to teach and how to teach them rather than solely what media to use (Collins, 2018, p. 13).

> They were useful for identifying specific curricular intervention areas. For example … low scores in category viii questions (comprehending geographic features represented as point, line, or polygon), could indicate … technical geographic information systems (GIS) training with vector datasets as opposed to general map-reading ability that could be a curricular focus based on low scores in category I (comprehending orientation and direction) (Tomaszewski et al., 2015, p. 41).

On the other hand, a few researchers pointed out that to successfully answer some questions geographical knowledge as well as spatial thinking is required. Although some studies acknowledged and accepted that a combination of geographic knowledge and skills is useful (Collins, 2018), others would prefer that knowledge and skills be separated in an evaluation tool (Wakabayashi, 2015). Verma (2014, pp. 62–63) pointed out that STAT did not completely cover some geographic concepts and made suggestions for the inclusion of additional concepts. "STAT … omit[s] such fundamental geospatial thinking skills as scale, frames of reference, regionalization (spatial classification), spatial diffusion, spatial hierarchy, and spatial analog."

Finally, regarding the target population, some studies appreciated that STAT could be administered to a wide range of subjects, from junior high school students to adults (Jo et al., 2016). Other researchers, particularly those working with younger age groups, saw a need for additional assessment instruments appropriate for younger age groups such as elementary school students (Crews, 2008).

Conclusion

The STAT was initially developed to assess what effect GIS learning has on spatial thinking. Since its publication and dissemination, the test has been used in a wide variety of settings. It has been administered to assess the impact of many types of interventions other than GIS learning. Fortunately, the test has proven to be flexible, allowing researchers to investigate a broad set of questions concerning spatial thinking expertise and the changes in spatial thinking skills and to adapt the test to local situations and conditions.

Research that used the test for its original purpose, to evaluate the effect of GIS learning on spatial thinking, found a positive relation between the two. The reported results are robust; they apply regardless of subjects' gender or age. Only three of the four studies that investigated the secondary effect of GIS teacher training on students' spatial thinking found a positive impact.

In addition to investigating the impact of educational interventions on spatial thinking, investigators have used the test to explore relationships between the characteristics and experiences of individuals and their spatial thinking ability. To some extent, then, researchers have used STAT to assess spatial thinking similar to the way others have used psychometric tests to evaluate spatial ability. This research found that, perhaps not surprisingly, students who "did" geography at university or in high school had better spatial thinking abilities and skills. Once again, this relationship held for females and males. Travel, especially international travel, was related to spatial thinking prowess, but the relationships for socioeconomic status, gender, and urban/rural residential location were inconsistent or insignificant.

Although the test has been used in a wide variety of research designs, it has the potential to contribute in an even larger arena. For example, the authors used test scores to attempt to identify the components of spatial thinking. Several researchers have argued that spatial thinking is not a single ability or skill but a collection, and they have proposed hierarchies or lists of these components (Gersmehl & Gersmehl, 2006; Golledge, 2002; Janelle & Goodchild, 2009). The results of our analysis did not align with or confirm the components proposed by other researchers, but a factor analysis did identify five factors. In order of their significance, they included the ability to visualize and overlay maps, to understand map symbolization, to recognize map patterns and perform Boolean operations, to navigate, and to identify spatial correlation.

This study presented outcomes of research projects that assessed participants' knowledge and skills using STAT. Although the test proved useful in a wide variety of studies, our review of these projects does not support the assertion that STAT is ideal or that improvements, additions, or revisions should not be attempted. We hope that in developing new spatial-thinking assessment instruments, researchers can build on the demonstrated strengths of STAT/SST, while addressing any limitations especially by utilizing recent research results and technological advancements.

First, our experience and review of the literature leads us to recommend that any newly developed assessment instrument should be easy to use; ensure rigorous reliability though a large-scale, pilot-test implementation (and modification, if necessary); and enable pre- and post-testing through the creation of two equivalent forms. It would also be beneficial to develop a tool in the form of a web or smartphone app, making it easy for individuals to participate and for researchers to analyze and process the results. Second, the instrument should reflect the developments in the field of spatial thinking ability since the STAT was designed. In order to ensure content validity, a list of spatial thinking ability components should be established through comprehensive research, so that questions measure each component accurately and exclusively. Ideally, participants should be able to answer questions without additional geographic knowledge. The use of video, virtual reality, and 3D technologies might also help researchers measure a broader range of abilities than was possible a decade ago. Lastly, to the extent possible, tools should be appropriate for the widest set of examinees, from elementary school students to adults.

This review demonstrates that during the last 10 years, we have acquired a much greater understanding of the nature and structure of spatial thinking. We have learned more about its relationship with GIS education, and for example how GIS learning may affect specific components of spatial thinking. We also have learned more about its relationship to personal attributes and experiences. Perhaps most importantly, we have learned something about how we can improve individuals' spatial thinking abilities. Understanding which factors correlate strongly with spatial thinking abilities provides us insight into effective ways that formal and informal training can be shaped. Some aspects of spatial thinking abilities would benefit from more study. In particular, more investigation of the relation between individuals' spatial thinking and their performance and attainment in STEM are needed. Newcombe (2017) and Uttal and Cohen (2012) argued that spatial thinking abilities serve as a gateway or barrier for entry into STEM fields. *Learning to Think Spatially* also asserted that spatial thinking skills were an important factor determining success in a wide variety of disciplines particularly geosciences (National Academy of Sciences, 2006, Uttal et al., 2013). In order to study this issue properly, suitable instruments to assess and measure spatial thinking are prerequisite. We hope that STAT can assist in this research agenda and invite colleagues to use it.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Audet, R., & Abegg, G. L. (1996). Geographic information systems: Implications for problem solving. *Journal of Research in Science Teaching*, *33*(1), 21–45.

Bednarz, R. S., & Lee, J. (2011). Spatial thinking assessment; An invitation to participate. *Procedia Social and Behavioral Sciences*, *21*, 103–107.

Charcharos, C., Kokla, M., & Tomai, E. (2015, November). Assessing spatial thinking ability in geospatial thinking: Educating the future spatial citizens. In M. Kavouras & S. Sotiriou (Eds.), *Proceedings of the GEOTHNK International Closing Conference* (pp. 151–166). Pallini, Greece.

Collins, L. (2018). The impact of paper versus digital map technology on students' spatial thinking skill acquisition. *Journal of Geography*, *117*(4), 137–152.

Crews, J. W. (2008). *Impacts of a teacher geospatial technologies professional development project on student spatial literacy skills and interests in science and technology in grade 5–12 classroom across Montana* (PhD Dissertation). University of Montana.

Duarte, R. G. (2016). School cartography and spatial thinking of Brazilian students at the end of junior high school. *Boletim Paulista de Geografia*, *99*, 185–199.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for the kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Services.

Firdiansyah, D. R. (2012). *Analysis of spatial thinking tests in geographical space* (master's thesis). The Netherland: Wageningen University.

Flynn, K. C. (2018). Improving spatial thinking through experiential-based learning across international higher education setting. *International Journal of Geospatial and Environmental Research*, *5*(3). Article 4.

Gersmehl, P. J., & Gersmehl, C. A. (2006). Wanted: A concise list of neurologically defensible and assessable spatial-thinking skills. *Research in Geographic Education*, *8*, 5–38.

Gersmehl, P. J., & Gersmehl, C. A. (2007). Spatial thinking by young children: neurologic evidence for early development and "educability". *Journal of Geography*, *106*(1), 81–91.

Gilmartin, P. P., & Patton, J. C. (1984). Comparing the sexes on spatial abilities: Map-use skills. *Annals of the Association of American Geographers*, *74*(4), 605–619.

Goldstein, D., Haldane, D., & Mitchell, C. (1990). Sex differences in visual–spatial ability: The role of performance factors. *Memory & Cognition*, *18*(5), 546–550.

Golledge, R. G. (1993). Geographical perspectives on spatial cognition, In T. Gärling & R. G. Golledge (Eds.), *Behavior and environment: Psychological and geographical approaches* (pp. 16–46). Amsterdam: Elsevier Science.

Golledge, R. G., & Stimson, R. J. (1997). *Spatial behavior: A geographic perspective*. New York: Guilford Press.

Golledge, R. G. (2002). The nature of geographic knowledge. *Annals of the Association of American Geographers*, *92*(1), 1–14.

Hedley, M. L., Templin, M. A., Czajkowski, K., & Czerniak, C. (2013). The use of geospatial technologies instruction within a student/teacher/scientist partnership: Increasing students' geospatial skills and atmospheric concept knowledge. *Journal of Geoscience Education*, *61*(1), 161–169.

Indhawati, M. (2015). *Comparison of spatial thinking levels between GIS and non-GIS students in understanding* (master's thesis). The Netherlands: Wagenigen University.

Ishikawa, T. (2013). Geospatial thinking and spatial ability: An empirical examination of knowledge and reasoning in geographical science. *The Professional Geographer*, *65*(4), 636–646.

Ishikawa, T. (2016). Spatial thinking in geographic information science: Students' geospatial conceptions, map-based reasoning, and spatial visualization ability. *Annals of the American Association of Geographers*, *106*(1), 76–95.

Janelle, D. G., & Goodchild, M. F. (2009). Location across disciplines: reflection on the CSISS experience. In H. J. Scholten, N. van Manen & R. v.d. Velde (Eds.), *Geospatial technology and the role of location in science* (pp. 15–29). Dordrecht: Springer.

Jo, I., Hong, J. E., & Verma, K. (2016). Facilitating spatial thinking in world geography using Web-based GIS. *Journal of Geography in Higher Education*, *40*(3), 442–459.

Kali, Y., Orion, N., & Mazor, E. (1997). Software for assisting high-school students in the spatial perception of geological structures. *Journal of Geoscience Education*, *45*(1), 10–21.

Kerski, J. J. (2000). *The implementation and effectiveness of geographic information systems technology and methods in secondary education* (PhD Dissertation). University of Colorado, Boulder.

Kim, M., & Bednarz, R. (2013). Development of critical spatial thinking through GIS learning. *Journal of Geography in Higher Education*, *37*(3), 350–366.

Lee, J., & Bednarz, R. (2009). Effect of GIS learning on spatial thinking. *Journal of Geography in Higher Education*, *33*(2), 183–198.

Lee, J., & Bednarz, R. S. (2012). Components of spatial thinking: Evidence from a spatial thinking abilities test. *Journal of Geography*, *111*(1), 15–26.

Li, S., & Liu, X. (2018). *Integrating GIS into secondary school science curriculums: A mixed-method study on students' spatial thinking ability*. Paper presented at the annual meeting of the NARST – a worldwide organization for improving science teaching and learning through research, Atlanta, GA, March 10–13, 2018.

Libarkin, J. C., & Brick, C. (2002). Research methodologies in science education: visualization and the geosciences. *Journal of Geoscience Education*, *50*(4), 449–455.

Lim, K. Y. T. (2005). Augmenting spatial intelligence in the geography classroom. *International Research in Geographical & Environmental Education*, *14*(3), 187–199.

Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature, Technical report 8, Stanford University Aptitude Research Project*. Stanford, CA: School of Education, Stanford University.

Low, L., Tan, J., & Huat, S. J. (2014, April). GIS lessons and spatial thinking. Paper presented at The GA Annual Conference. Retrieved from https://www.smore.com/fpc1p-gis-lessons-and-spatial-thinking

Low, R., Boger, R. A., & Mandry, C. A. (2014, December). *Connecting spatial literacy and climate literacy using a place-based GIS approach in a collaborative online educational setting*. Paper presented at American Geophysical Union, Fall Meeting 2014. Retrieved from http://adsabs.harvard.edu/abs/2014AGUFMED33B3511L

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences*, *49*(4), 344–351.

Mark, D. M., & Freundschuh, S. M. (1995). Spatial concepts and cognitive models for geographic information use. In T. L. Nyerges, D. M. Mark, R. Laurini & M. J. Egenhofer (Eds.), *Cognitive aspects of human-computer interaction for geographic information systems* (pp. 21–28). Dordrecht: Kluwer Academic Publishers.

McGlone, J. (1981). Sexual variation in behavior during spatial and verbal tasks. *Canadian Journal of Psychology*, *35*(3), 277–282.

Meyer, J. W., Butterick, J., Olkin, M., & Zack, G. (1999). GIS in the K–12 curriculum: A cautionary note. *Professional Geographer*, *51*(4), 571–578.

Miller, L. K., & Santoni, V. (1986). Sex differences in spatial abilities. *Acta Psychologica*, *62*(3), 225–235.

Montello, D. R., Lovelace, L. L., Golledge, R. G., & Self, C. M. (1999). Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American Geographers*, *89*(3), 515–534.

Moorman, L. A., & Crichton, S. (2018). Learner requirements and geospatial literacy challenges for making meaning with Google Earth. *International Journal of Geospatial and Environmental Research*, *5*(3), Article 5. https://dc.uwm.edu/cgi/viewcontent.cgi?article=1074&context=ijger

National Academy of Sciences Committee on Support for Thinking Spatially (2006). *Learning to think spatially*. Washington, DC: The National Academies Press.

National Science Board (NSB) (2010). Preparing the next generation of STEM innovators. NSB, 10–33

Newcombe, N., & Dubas, J. S. (1992). A longitudinal study of predictors of spatial ability in adolescent females. *Child Development*, *63*(1), 37–46.

Newcombe, N. (2017). *Harnessing spatial thinking to support stem learning. OECD Education Working Papers, No. 161*. Paris: OECD Publishing,. Retrieved from http://dx.doi.org/10.1787/7d5dcae6-en.

Olsen, T. P. (2000). *Situated student learning and spatial informational analysis for environmental problem*, PhD dissertation. University of Wisconsin-Madison, Madison.

Powell, S., & Kong, N. N. (2017). Beyond the one-shot: Intensive workshops as a platform for engaging the library in digital humanities. *College & Undergraduate Libraries*, *24*(2–4), 516–531.

Sharpe, B., & Huynh, N. T. (2015). A Review of geospatial thinking assessment in high schools. In O. Muñiz Solari, A. Demirci & J. A. van der Schee (Eds.), *Geospatial technologies and geography education in a changing world* (pp. 169–180). IGU-CGE, Springer.

Shin, E., Milson, A. J., & Smith, T. J. (2016). Future teachers' spatial thinking skills and attitudes. *Journal of Geography*, *115*(4), 139–146.

Skaza, H. J. (2016). *Development and testing of an assessment to measure spatial thinking about enhanced greenhouse effect* [PhD dissertation]. Las Vegas: University of Nevada.

Sorby, S., Veurink, M., & Streiner, S. (2018). Does spatial skills instruction improve STEM outcomes? The answer is 'yes'. *Learning and Individual Differences*, *67*, 209–222.

Tomaszewski, B., Vodacek, A., Parody, R., & Holt, N. (2015). Spatial thinking ability assessment in Rwandan secondary schools: Baseline results. *Journal of Geography*, *114*(2), 39–48.

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education. When, why, and how? *Psychology of Learning and Motivation – Advances in Research and Theory*, *57*(35), 147–181.

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

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*(2), 599–604.

Verma, K. (2015). Influence of academic variables on geospatial skills of undergraduate students: An exploratory study. *The Geographical Bulletin*, *56*, 41–55.

Verma, K. (2014). *Geospatial thinking of undergraduate students in public universities in the United States* [PhD dissertation]. Texas State University.

Vincent, P. C. (2004). *Using cognitive measures to predict the achievement of students enrolled in an introductory course of Geographic Information Systems* [PhD dissertation]. College Station: Texas A&M University.

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.

Wakabayashi, Y. (2015). Measurement of geospatial thinking abilities and the factors affecting them. In Geographical *reports of* Tokyo Metropolitan University, *50*, 127–136.

Wan, J., Lu, X., Du, F., Wang, J., & Ju, B. (2017). Influencing factors of middle school students' spatial thinking ability: A case study on senior one students of Baiyin No. 1 Middle School in Gansu Province. *Progress in Geography*, *36*(7), 853–863.

Weakley, K. D. (2010). *The effects of an inquiry-based earth science course on the spatial thinking of pre-service elementary teacher education students* (PhD Dissertation). Western Michigan University