Chapter 15
Visualizing Student Navigation of Geologic Block Diagrams

Karen S. McNeal, Rachel Atkins, and Elijah T. Johnson

Introduction

Spatial thinking is defined by Uttal et al. (2013) as, "the mental processes of representing, analyzing, and drawing inferences from spatial relations…between objects…or…within objects" (p. 367). There are many spatial skills, and different disciplines and fields have varying names and definitions for them. However, the consensus is that spatial reasoning is an essential skill necessary for success in the STEM (science, technology, engineering, and mathematics) domains as these disciplines require students to visualize past or theoretical phenomena based on spatial relationships between elements in nature or a provided diagram. The importance of spatial reasoning ability in the STEM domains is solidified through a study by Wai et al. (2009) that completed an eleven-year longitudinal study and aligned their results to 50 years of preexisting psychological data to conclude that spatial ability is correlated with STEM achievement and career paths. Forty-thousand random participants were tracked for over 11 years to assess their self-organization into careers based on mathematical, verbal, and spatial ability. The data showed that students with high spatial ability excelled in physical science, math/computer science, and engineering in terminal bachelor, master, and doctorate degrees. Students with lower spatial reasoning ability tended to self-organize into education, law, and business. The occupations they pursued after college strongly resemble those same trends. Similarly, Kell and Lubinski (2013) suggest that students may self-organize into their majors and careers based on their spatial thinking ability, whereas students may

K. S. McNeal (✉) · E. T. Johnson
Department of Geosciences, Auburn University, Auburn, AL 36849, USA
e-mail: ksm0041@auburn.edu

R. Atkins
North Carolina State University, Raleigh, NC 27659, USA

self-select out of STEM domains due to the amount of spatial reasoning ability they possess, especially when underserved by academic institutions.

Visualizing three-dimensional (3D) structures is a challenge in science, technology, engineering, and mathematics (STEM) learning (Milner-Bolotin & Nashon, 2012). "Students in these fields are required to reason about objects or features that occur at spatial scales too large or small to be directly observed. Consequently, 3D phenomena are often illustrated using visual representations such as diagrams" (Gagnier et al., 2017, p. 884). 3D thinking is particularly important in chemistry, biology, engineering, and the geosciences. In chemistry education, students' performances on a variety of chemistry problems including problem-solving (e.g., stoichiometry) and multistep calculations, as well as balancing chemical equations were linked to their spatial thinking skill development (Staver & Jacks, 1988; Carter et al., 1987). Instructional interventions that worked to build students' spatial skills through practice lead to significant increases in student chemistry test performance (Small & Morton, 1983; Tuckey et al., 1991).

In modern biology and engineering education, 3D and 4D visualization also plays an important role in students learning both the concepts and skills needed to accurately reason about specific phenomenon (e.g., technical drawing, graphical representations of biochemical structures, or embryo development; Milner-Bolotin & Nashon, 2012). Biology and engineering classroom interventions have shown that providing students graphical software to scaffold their development of 3D and 4D skills supported student success in biochemistry classes compared to those that did not receive the software intervention (Richardson & Richardson, 2002), and in engineering, students that were exposed to such technology had gains in spatial rotation and space relations, assessed using pre-post test scores (Sorby, 2009).

The geosciences also have a high requirement for building spatial thinking skills (Sanchez & Wiley, 2014), many of which include 3D visualization. One of the spatial skills a geoscientist must employ is mental brittle transformation, which is the ability to mentally break and reconstruct objects (Resnick & Shipley, 2013). Another skill is mental rotation, which involves a person's ability to turn a 2D or 3D object about an axis (Shepard & Metzler, 1971) and may be activated by a stratigrapher examining the position of overturned strata. Yet another skill is spatial orientation, which requires an understanding of perspective and the relation of an object to a frame of reference (Ramful et al., 2017). For example, geoscience students in field camp, a capstone course required by many geology programs, may employ spatial orientation in navigating the field environment and marking the relative positions of outcrop features on a map. A more complex skill that geoscientists employ is spatial visualization which represents multiple associated tasks. Linn and Petersen (1985) define spatial visualization as "spatial ability tasks that involve complicated, multistep manipulations of spatially presented information" (p. 1484). A unique skill to the geosciences includes penetrative thinking or visualizing the subsurface or interior of an object using clues from the visible parts of the object (Alles & Riggs, 2011), like how a structural geologist generates a 2D cross-section that represents a 3D phenomenon in the real world in a drawing that represents the bedforms present at the Earth's surface and in geologic outcrops. Often these phenomena are represented in geologic block

diagrams that illustrate geologic structures at scales ranging from centimeters to tens of kilometers. However, students struggle to interpret the 3D spatial relations conveyed in these diagrams (Gagnier et al., 2017). As such, we have focused on penetrative thinking skills in this study as they are key to interpreting geologic block diagrams, which are the spatial representations employed in this work.

Relevant Work in Geoscience Education Research

Many students may have natural spatial thinking ability, while others may lack this skill, which could make learning geological concepts more challenging (Ishikawa & Kastens, 2005).

Since spatial reasoning ability has been shown to be malleable (Uttal et al., 2013), many researchers have explored interventions designed to train spatial thinking skills in the geosciences. For example, Titus and Horsman (2009) performed a semester-long study training students from two different populations in spatial visualization. This study assessed the effect of spatial training with Spatial Intelligence and Learning Center (SILC)-verified assessment strategies and improved undergraduate student performance, as well as students' overall course grades. In a study conducted by Ormand et al. (2014), students in a variety of geoscience courses, from introductory to senior level major courses, and with various degrees of incoming spatial ability experienced spatial skill gains from simply being exposed to spatial concepts during the course. Gold et al. (2018) found that regular, short interventions throughout an academic semester improve students' spatial thinking skills significantly with a moderate to large effect size when compared to an instruction-as-usual control group. They also found that about 15% of the students improved their spatial skills to the point that they would be considered high enough for those commonly entering/continuing in STEM.

Spatial training using new technologies has been a growing area in the geosciences. McNeal et al. (2020) conducted a study aimed to understand the impact of using an augmented reality (AR) sandbox on students' topographic map performance. They found that students with higher spatial ability tended to perform better on the task with the AR sandbox than those with lower spatial ability, but that this performance gap was mitigated with more structured activities in the sandbox. This finding led to the idea that perhaps the AR sandbox could support students' spatial thinking skills. Johnson and McNeal (in review) have since shown that the AR sandbox has the potential to support student spatial skill development. They implemented activities with students in a lab environment to aid their development of spatial orientation, spatial rotation, and spatial visualization skills. Results indicate that the AR sandbox may have the greatest potential to assist students in developing their spatial visualization skills. Spatial visualization was the area in which students identified the most challenges and the least strategies during their problem-solving.

Spatial training with new technologies in the geosciences has also been used with geographical information systems (GIS) to explore whether GIS could impact

students' spatial thinking (Lee & Bednarz, 2009; Kim & Bednarz, 2013). Lee and Bednarz (2009) grouped multiple GIS activities into spatial skills categories and administered a spatial skills test before and after the GIS course. The test showed gains in the students' spatial reasoning ability, showing that technology that has a high spatial component, such as GIS, can train spatial thinking ability.

Eye-Tracking in the Geosciences

Geoscience education researchers in classroom, field, and lab environments employ eye-tracking approaches using stationary devices and portable headsets in their studies to gain insights into student cognition. For example, a field study by Maltese et al. (2013) used eye-tracking headsets on students to both investigate its viability for observing and detailing students' experiences in the field and evaluate the variety of information that can come from eye-tracking in the field. Although there were operational and technical challenges related to using the eye-tracker in the field, the scene video they acquired elucidated how students were engaging with the geology and each other in the field.

Eye-tracking was also used by McNeal et al. (2014) to evaluate and revise an online curriculum, *EarthLabs*. College undergraduates interacted with the online modules while being eye-tracked to determine how they were engaging with the material to improve the online *EarthLabs* curriculum and user experience. Evidence from eye-tracking revealed that although students were engaging with the text portions of the modules more than the images, that engagement declined over time as the students worked through the activities in the module. Additionally, students generally found charts, graphs, and questions embedded in the text to be most useful, however, they experienced difficulty engaging with graphs depicting change over time. Learning what students are paying attention to as well as how their engagement varies over time from their eye movements in this study speaks to the usefulness of eye-tracking for user-testing.

The effectiveness of eye-tracking for user-testing is also expounded upon in a study by Maudlin et al. (2020) where male and female decision-makers and students were eye-tracked to explore gender-differences in visual attention on a decision-support website, *PINEMAP DSS*. *PINEMAP DSS* is a website service that communicates climate impacts to loblolly pine forests in the southeastern United States. Since this information is primarily for forest service professionals and decision-makers, testing the usability of the medium and getting insight into how different users interacted with the content was of high importance. The researchers found that males paid more attention to the data and map features of the website, while females paid less attention to the data itself and more time evaluating other features of the website including tabs, map legends, and text. Males also outperformed females on the questions they were given about the information on the websites. These results can be used to revise website-based tools so that content creators can effectively communicate with their intended audiences.

Atkins and McNeal (2018) also explored how users interacted with climate information by eye-tracking students looking at climate change graphs. This study compared the eye movement and attention patterns of undergraduate and graduate students to identify how knowledge, skill, and expertise affect performance on fact-extraction and extrapolation tasks. They found that undergraduates spent more time on graphical elements not pertinent to the content (i.e., axes, title, legend), while graduate students spent more time interpreting the provided data. They also found that undergraduate students with high graphical skills performed similarly to graduate students. By exploring the cognitive limitations of novice students on climate change graph understanding, scientists can improve those graphs to help communicate their findings more effectively and educators can focus their instruction on scaffolding student to be able to interpret scientific graphs.

In this study, we use eye-tracking in a similar way: to understand the visual attention of students while problem-solving. Spatial reasoning ability as a suite of cognitive skills can be improved with training (Uttal et al. 2013), so understanding the challenges students have with these skills is a first step in developing interventions designed to improve those skills. A spatial thinking skill that has been shown to be challenging for students in the geosciences is penetrative thinking, particularly when studying geologic block diagrams. Currently, eye-tracking has not been applied to understanding how students navigate block diagrams in the geosciences. We aim to investigate how students visually navigate geologic block diagrams. More specifically, we look to identify emergent patterns between students who do well (high performers) and students who do poorly (low performers) in solving geologic block diagrams. Finally, we highlight common errors made by students as identified by their visual navigation patterns while solving geologic block diagrams and provide future directions for investigating the cause of these errors.

Methods

Participants

The 58 participants in this study consisted of 45 undergraduates enrolled in an Earth Systems Science course at a large land-grant university in the southeastern United States and 13 graduate students in a graduate program at the same university. Our participants ranged in class rank from freshman to graduate students with at least 1 year of experience ages 17 to >23 with a median age of 20 years and a male to female ratio of 31:27. Participants were recruited from two sections of an introductory Earth Systems Science course taught by different instructors and received a $20 Amazon gift card as compensation after completion of the pre-test and eye-tracking study outside of class. Graduate students participated on a volunteer basis and were recruited from a departmental listserv. Human subject's research approval

from the Institutional Review Board (IRB) was obtained before recruitment and the commencement of any research activities.

Experimental Design and Instrumentation

This study used two versions of the Geologic Block Cross-sectioning Test (GBCT; Orman et al., 2014). The first 16-question version was used as a pre-test to establish participants' visual penetrative thinking skills and the ability to recognize the correct vertical cross-section through a geology block diagram prior to eye-tracking. For undergraduates, the test was administered at the end of an Earth Systems Science class period, while graduate students were asked to take the pre-test immediately before the eye-tracking study. After pre-testing, participants were asked to solve five selected problems of varying difficulty from the second version of the GBCT assessment. These problems addressed geologic concepts such as dipping beds, faulted horizontal strata, dipping transverse beds, and plunging folds (Fig. 15.1). During this second assessment, participant eye movements were tracked while solving each question using a Tobii TX300 eye-tracker and participants were also asked to state their answer to the presented problem aloud.

The eye-tracker was attached to a 23-inch computer monitor, collecting at 300 Hz, and did not come in physical contact with participants. Calibration was completed
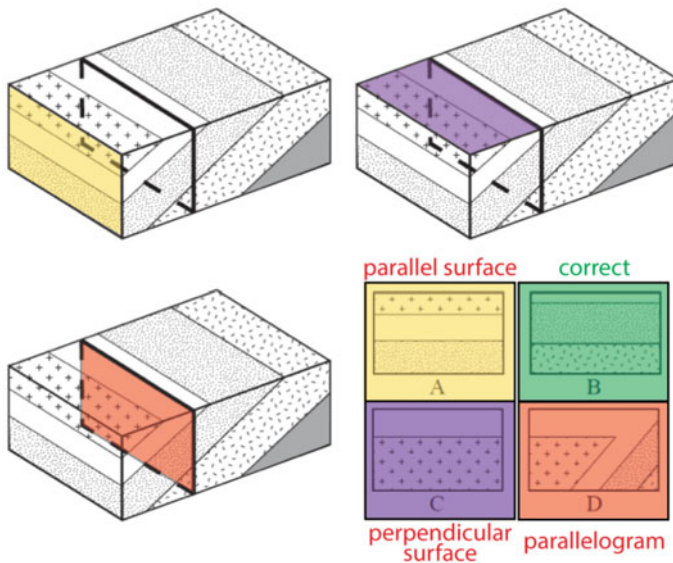

Fig. 15.1 (above) Highlights specific errors depicted in the response choices and where they may be found in the 3D block diagram. Choice B is the correct response for this question

for each participant to ensure accuracy and precision among participant trials. Participants sat ~65 cm from the monitor and gazed at the computer screen to view the provided graphs with an unobstructed view. The noncontact nature of this technology allows for the capture of natural eye movements, compared with instrumentation that is worn by the participant. The system allowed for corrective lenses to be worn without affecting results.

Ormand et al. (2014) created the 16-question Geologic Block Cross-sectioning Test which has since undergone multiple assessments to ensure validity and reliability internally and between the two versions. It was also developed to specifically address common misconceptions people have about geologic block diagrams (Kali & Orion, 1996) (see Fig. 15.1 for examples). Participants are given the same three instructions for all diagrams: "1. Study the geologic structure that is displayed in the 3D block diagram, 2. Visualize what the cross-section of that geologic structure would look like on the surface of the vertical plane intersecting the block and 3. Choose the multiple choice answer that illustrates the structure along that plane. Where more than one answer appears to be possible, choose the MOST LIKELY answer." For both the pre-test and eye-tracking test, an example with its correct answer was provided to allow participants to practice the format before being assessed.

Each question consists of a 3D box with an illustrated geologic problem inside, and a dark box highlighting the horizontal transect where the participant is asked to mentally slice through the diagram, with four possible answers to choose from (Fig. 15.1). The types of errors included in the four possible answers for each question fall under the two broad categories of penetrative and non-penetrative answers (Ormand et al., 2014). Penetrative errors reflect an attempt to visualize the inside of the structure, but do so incorrectly, whereas non-penetrative errors indicate an individual's inability to mentally penetrate the block, subsequently their answers reflect one of the visible sides of the diagram (Kali & Orion, 1996).

Data Analysis

Two aspects of eye movements that are most often studied include saccades and fixations. Saccades are the short periods of rapid eye movement between fixations that redirect participant gaze from one fixation to another (Ramat et al., 2008). These can occur up to four times in a second and participants are effectively blind while they occur (Land, 2012). Fixations are the points between saccades where the eye is nearly stationary for relatively longer periods of time (~70–100 ms). These eye movements are of particular interest as it is during these viewing times that mental processing takes place (Bojko, 2013).

A correlation analysis was done to determine the relationship between pre-test score and score on the five-question eye-tracking test. All 58 participants yielded a moderately strong correlation coefficient of 0.6581, indicating that the five selected questions appropriately represent participants' overall performance on the assessment. Therefore, we use performance on the more thorough 16-question pre-test to
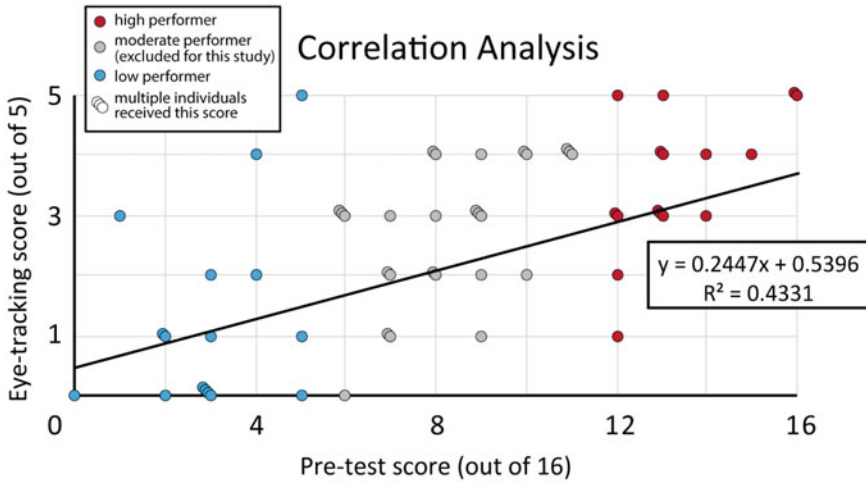
Fig. 15.2 (above) Correlation analysis between questions asked during eye-tracking (eye-tracking score) and pre-test scores reveals a moderately strong correlation, therefore we are confident that the five questions selected for the eye-tracking study yield valid results. Additionally, the pre-test was used to group participants into high performers (red points) and low performers (dark blue points) using quartiles. These groupings are used for the remainder of the analysis. Points have been jittered to indicate multiple points at each location

bin participants into **low performers** ($n = 16$, pre-test scores: 0–5 out of 16) and **high performers** ($n = 16$, pre-test scores: 12–16) by quartiles (Fig. 15.2).

Results

From the GBCT assessment, we were able to confirm that high performers on the pre-test answered the eye-tracking questions correctly more often than low performers. Further, when high performers answered questions incorrectly on the eye-tracking assessment, they all made the same error, whereas low performers made multiple different errors. Additionally, we were able to determine that the most common error made by all participants was a parallelogram error, indicating the possibility that they were solely visualizing within the parallelogram itself, and not taking into account the behavior of the layers in the rest of the diagram. This was confirmed in our eye-tracking results when comparing the visual patterns of high and low performers. For example, Fig. 15.3a shows high performers, who also answered questions correctly more often, focusing their attention on the face that indicates that the layers are dipping to inform their answer selection (see the top right corner of the bolded box for question #2). Conversely, low performers do not focus their attention in this same place, but rather distribute their gaze throughout the parallelogram, congruent with their most often incorrectly selected answer (parallelogram).
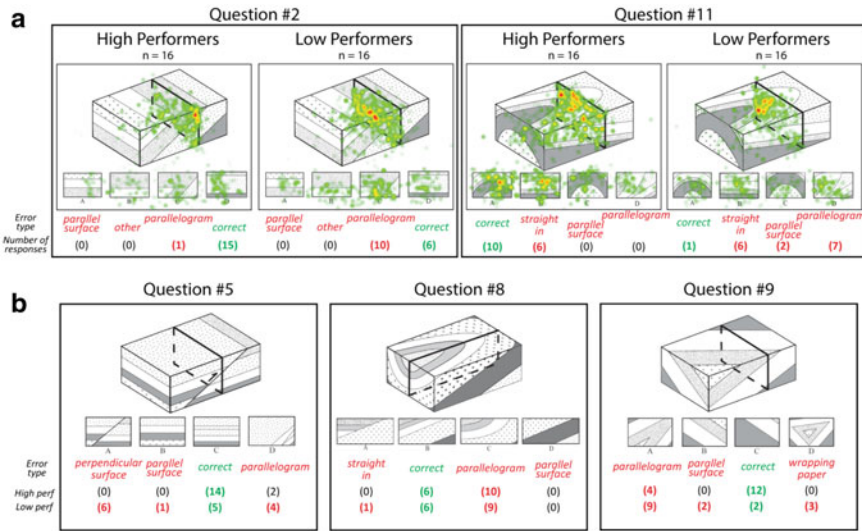
Fig. 15.3 (above) Part a shows examples of eye-tracking results from two of the five questions asked during the eye-tracking assessment. Areas of red indicate high concentrations of visual gaze and areas in green correspond to low concentrations of gaze. The type of error that each response choice depicts is indicated in RED LETTERING below each choice, GREEN LETTERING indicates correct answer, and the number of responses is in parentheses. Part b is a summary of the number of high and low performer responses for the remaining three questions without the eye-tracking heat maps for simplicity

Out of the five eye-tracking questions, only #8 and #11 had penetrative errors (labeled "straight in") as selectable options. These types of errors are unique because they indicate the participants attempt to see through the diagram and visualize the internal structure. The difference between high and low performers is highlighted in Fig. 15.3a for question #11. The only error made by high performers was a penetrative error, indicating that those who answered incorrectly were still attempting to mentally penetrate the diagram and visualize in 3D. Conversely, not only were the number of correct responses fewer for low performers, but their incorrect answers spanned all options, indicating an inability to visualize in 3D. The eye-tracking results from both high and low performers reflect these differences, showing the gaze concentrations of high performers distributed throughout the diagram, whereas low performers remain mostly in the parallelogram (Fig. 15.4).

The distribution of gaze throughout the diagram, particularly for a complicated problem such as the ones depicted in questions #9 & 11, indicates the viewers attempt at understanding the overall behavior of the geologic layers and interpreting how they may look at different angles.
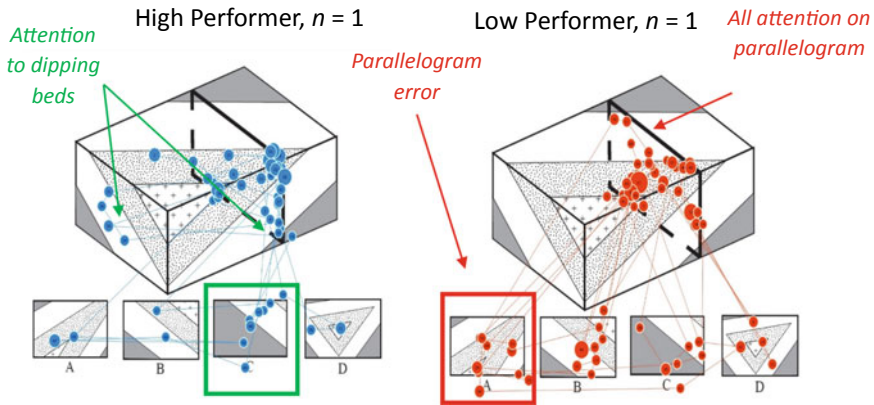
Fig. 15.4 (above) Examples of eye-tracking gaze plots from one high and one low performer from question #9. The green arrows show what the high performer is paying attention to and the selection of the correct answer indicated by the green box. The red arrows indicate the parallelogram error of the low performer and the resulting incorrect response option that was chosen as indicated by the red box

Discussion and Future Directions

Overall, eye-tracking results from all five questions indicate that high performers allocate proportionately more of their visual gaze outside of the bolded parallelogram in the provided block diagrams. We interpret this to indicate that these participants are likely attempting to visualize the whole structure (i.e., on all sides), and using this to determine their answer. By using observations from all sides of the diagram, they are able to develop a complete story, as opposed to selecting pieces of the diagram (i.e., what's depicted inside the parallelogram) to inform their answer. This exploratory work has helped to show eye-tracking as a useful tool to understand how individuals navigate spatial diagrams. This study was able to highlight the differences between high and low performers as revealed in pre-test scores and in eye-tracking results.

Some of the most informative next steps to build on findings from this work would be the addition of concurrent think alouds and/or a post-assessment interview, along with the addition of multiple spatial skill pre-assessments. The addition of a qualitative metric would provide insight into why individuals selected their answers and combined with correlations between spatial skill competencies acquired by additional assessments, could help identify specific areas to target spatial training to facilitate skill improvement. Recommendations for more rigorous spatial skill assessments include the Purdue Visualization of Rotations Test (PVRT) (Guay, 1976) for mental rotation, the Educational Testing Service (ETS) Hidden Figures test (Ekstrom et al., 1976) for disembedding skills, and Planes of Reference test (Titus & Horsman, 2009) for penetrative thinking. Furthermore, while our initial study used five questions for eye-tracking analysis, we recommend increasing the number of questions to increase statistical robustness.

Another question worthy of investigation that could be informed by eye-tracking is "do patterns exist among eye movements between diagram and answer options?" and the connection that those patterns may have with the types of cognitive processes being used to solve the problem (i.e., inductive vs. deductive reasoning). For example, if a participant investigated the answer options before viewing the diagram, they may have solved the problem by eliminating incorrect options. Conversely, first fixations at the main diagram may indicate the development of an answer before looking at the choices and may indicate a higher confidence in 3D visualization. The investigation of these spatiotemporal patterns combined with additional metrics (i.e., gender, pre-test performance) and/or assessments could answer unique questions that traditional paper testing is not able to.

We caution future researchers interested in expert/novice comparisons to perform a thorough assessment of a range of spatial skills before binning participants into expert and novice categories. Research has shown that spatial abilities are a combination of one's prior experiences that can be traced all the way back to childhood, their innate ability, and exposure to formal training (Gold et al., 2018). Not only do spatial abilities vary extensively across student populations, it has also been shown that with appropriate training, spatial skills can be improved (Lee & Bednarz, 2009; Uttal et al. 2013; Ormand et al., 2014; Gold et al., 2018). For these reasons, it is important to confirm that perceived experts (i.e., domain experts) actually exhibit high levels of spatial abilities.

Despite the needs for continuing research, this study provides insights about how to better support students' 3D problem-solving skills, especially among high and low performers. Potential classroom activities that could help mediate this performance gap is to pair high and low performers together to solve spatial problems. This would require some pre-testing of students at the beginning of a course, but such distributed expertise pairing could be helpful to students. Additionally, the replaying of eye-tracking scan patterns of high performers before completing a geologic block diagram problem may help to improve the performance of all learners. Both of these suggestions require further testing to document any potential student learning gains through intervention studies in the geosciences. However, these are activities that our research eludes to as potential actions that instructors could take to help support the learners in their classrooms develop 3D spatial problem-solving skills.

Conclusions

This exploratory eye-tracking study provided unique insights not yet obtained by the geoscience education research community about how high and low spatial performing students navigate geologic block diagrams, a 2D visualization tool used in the geological sciences to represent conditions within a 3D geologic formation. The results showed that there were differences in the visual attention that high and low performers made on the diagrams. These differences aligned with the correct/incorrect selection on the geologic block diagram assessment used in this study where high performers

tended to make more fixations on all faces of the diagram where low performers tended to fixate on more specific areas of the diagram. This trend indicated that high performers seemed to be able to "see the big picture" whereas low performers could not when solving 3D visualization problems in the geosciences. Additional research is recommended to expand on and verify our exploratory results, however, the use of eye-tracking within the context of understanding how students solve spatial problems in the geosciences has provided new insights that with continued research can be used to inform how best to scaffold students as they build their spatial skills in the geosciences, a STEM field that has a high requirement for multiple spatial thinking skills to be developed among learners.

References

Alles, M., & Riggs, E. M. (2011). Developing a process model for visual penetrative ability. *Geological Society of America Special Papers, 474,* 63–80.

Atkins, R. M., & McNeal, K. S. (2018). Exploring differences among student populations during climate graph reading tasks: An eye tracking study. *Journal of Astronomy and Earth Sciences Education (JAESE), 5*(2), 85–114.

Bojko, A. (2013). *Eye tracking the user experience: A practical guide to research.* Rosenfeld Media.

Carter, C. S., Larussa, M. A., & Bodner, G. M. (1987). A study of two measures of spatial ability as predictors of success in different levels of general chemistry. *Journal of Research in Science Teaching, 24*(7), 645–657.

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

Gagnier, K. M., Atit, K., Ormand, C. J., & Shipley, T. F. (2017). Comprehending 3D diagrams: Sketching to support spatial reasoning. *Topics in Cognitive Science, 9*(4), 883–901.

Gold, A. U., Pendergast, P. M., Ormand, C. J., Budd, D. A., & Mueller, K. J. (2018). Improving spatial thinking skills among undergraduate geology students through short online training exercises. *Journal International Journal of Science Education, 40*(18).

Guay, R. B. (1976). *Purdue spatial visualization test.* West Lafayette, IN: Purdue Research Foundation.

Ishikawa, T., & Kastens, K. A. (2005). Why some students have trouble with maps and other spatial representations. *Journal of Geoscience Education, 53*(2), 184–197.

Kali, Y., & Orion, N. (1996). Spatial abilities of high-school students in the perception of geologic structures. *Journal of Research in Science Teaching, 33*(4), 369–391.

Kell, H. J., & Lubinski, D. (2013). Spatial ability: A neglected talent in educational and occupational settings. *Roeper Review, 35*(4), 219–230.

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

Land, M. F. (2012). The operation of the visual system in relation to action. *Current Biology, 22*(18), R811–R817.

Lee, J., & Bednarz, R. (2009). Effect of GIS learning on spatial thinking. *Journal of Geography in Higher Education, 33*(2), 183–198. https://doi.org/10.1080/03098260802276714.

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.

Maltese, A. V., Balliet, R. N., & Riggs, E. M. (2013). Through their eyes: Tracking the gaze of students in a geology field course. *Journal of Geoscience Education, 61*(1), 81–88.

Maudlin, L. C., McNeal, K. S., Dinon-Aldridge, H., Davis, C., Boyles, R., & Atkins, R. M. (2020). Website usability differences between males and females: An eye-tracking evaluation of a climate

decision support system. *Weather, Climate, and Society, 12,* 183–192. https://doi.org/10.1175/WCAS-D-18-0127.1.

McNeal, K. S., Libarkin, J. C., Ledley, T. S., Bardar, E., Haddad, N., Ellins, K., & Dutta, S. (2014). The role of research in online curriculum development: The case of EarthLabs climate change and Earth system modules. *Journal of Geoscience Education*, *62*(4), 560–577.

McNeal, K. S., Ryker, K., Whitmeyer, S., Giorgis, S., Atkins, R., LaDue, N., & Pingel, T. (2020). A multi-institutional study of inquiry-based lab activities using the Augmented Reality Sandbox: impacts on undergraduate student learning. *Journal of Geography in Higher Education, 44*(1), 85–107.

Milner-Bolotin, M., & Nashon, S. M. (2012). The essence of student visual-spatial literacy and higher order thinking skills in undergraduate biology. *Protoplasma, 249*(1), 25–30.

Ormand, C. J., Manduca, C., Shipley, T. F., Tikoff, B., Harwood, C. L., Atit, K., & Boone, A.P. (2014). Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study. *Journal of Geoscience Education*, *62,* 146–154.

Ramat, S., Leigh, R. J., Zee, D. S., Shaikh, A. G., & Optican, L. M. (2008). Applying saccade models to account for oscillations. In *Progress in brain research* (Vol. 171, pp. 123–130). Elsevier.

Ramful, A., Lowrie, T., & Logan, T. (2017). Measurement of spatial ability: Construction and validation of the spatial reasoning instrument for middle school students. *Journal of Psychoeducational Assessment, 35*(7), 709–727.

Resnick, I., & Shipley, T. F. (2013). Breaking new ground in the mind: An initial study of mental brittle transformation and mental rigid rotation in science experts. *Cognitive Processing, 14*(2), 143–152.

Richardson, D. C., & Richardson, J. S. (2002). Teaching molecular three-dimensional literacy. *Biochemistry and Molecular Biology Education, 30*(1), 21–26.

Sanchez, C. A., & Wiley, J. (2014). The role of dynamic spatial ability in geoscience text comprehension. *Learning and Instruction, 31,* 33–45.

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

Small, M. Y., & Morton, M. E. (1983). Research in college science teaching: Spatial visualization training improves performance in organic chemistry. *Journal of College Science Teaching, 13*(1), 41–43.

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

Staver, J. R., & Jacks, T. (1988). The influence of cognitive reasoning level, cognitive restructuring ability, disembedding ability, working memory capacity, and prior knowledge on students' performance on balancing equations by inspection. *Journal of Research in Science Teaching, 25*(9), 763–775.

Titus, S., & Horsman, E. (2009). Characterizing and improving spatial visualization skills. *Journal of Geoscience Education, 57,* 242–254. https://doi.org/10.5408/1.3559671.

Tuckey, H., Selvaratnam, M., & Bradley, J. (1991). Identification and rectification of student difficulties concerning three-dimensional structures, rotation, and reflection. *Journal of Chemical Education, 68*(6), 460–464.

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.

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 Psycholog*y, *101*, 817–835, https://doi.org/10.1037/a0016127.

Karen S. McNeal, Ph.D. is the College of Science and Mathematics Molette Endowed Professor in the Department of Geosciences at Auburn University, Auburn, Alabama, USA. Her research is in the field of geoscience education and geocognition where she applies mixed methods research

(e.g., qualitative and quantitative approaches) to investigate geoscience teaching and learning challenges in formal and informal settings. Her research group is interested in addressing questions that examine how to engage individuals in geoscience related materials, the misconceptions and mental models people hold about complex Earth systems, the psychomotor responses (e.g., eye movements, skin conductance) that occur when people are exposed to geoscience related materials and experiences, and the best practices in teaching and learning about the Earth. Her team conducts research in the broad areas of: (i) Earth systems science understanding, (ii) climate change understanding and communication, (iii) spatial thinking, (iv) active learning, (v) co-production of science, and (vi) diversity and inclusion in the geosciences. She has published over 60 peer-reviewed papers and received ~$25 in external funding from a variety of federal and state funding sources. She has served as Editor for Research for the Journal of Geoscience Education and has been the President of the Geoscience Education Research Division of the National Association of Geoscience Teachers (NAGT).

Rachel Atkins is a Ph.D. candidate at North Carolina State University in Raleigh, North Carolina, USA. Her research spans multiple disciplines including tectonic and fluvial geomorphology, planetary geology and geoscience education. Her solid Earth geology research investigates how planetary surfaces evolve over human and geologic timescales as a result of anthropogenic, global contraction and river processes. Her discipline-based education research experience has centered on investigating how students learn geoscience concepts in an effort to understand how to improve teaching practices. Some of these projects include employing eye-tracking to determine how novice students visually navigate climate change graphs and using the eye-tracking technology to determine how students explore and solve 3-Dimensional problems (e.g., block diagrams) in 2-Dimensional space (e.g., a piece of paper or computer screen). These skills have been shown to be potential barriers for retention of geoscience students. Additionally, she has been involved with projects investigating the effectiveness of using Augmented Reality Sandboxes (ARS) on student learning. Some of the eye tracking projects that Rachel has been involved with have been published in the Journal of Astronomy and Earth Science Education, Journal of Geography in Higher Education and Weather, Climate and Society. She has served as a reviewer for the Journal of Geoscience Education since 2017 and received a two-year NASA Future Investigator of Earth and Space Science and Technology grant in 2019 and a Southeast Climate Adaptation Science Center Global Change Fellowship in 2015. Rachel is also an active member of several professional organizations including the National Association of Geoscience Teachers (NAGT), Geological Society of America (GSA) and the Association for Women Geoscientists (AWG).

Elijah T. Johnson is a Ph.D. student at Auburn University in Auburn, Alabama USA, a National Science Foundation Graduate Research Fellow, and a geoscience education researcher. He investigates the role and development of spatial thinking in STEM and how the geosciences can promote and improve spatial learning across disciplines. His research employs augmented-reality technology to enhance the learning environment of introductory geoscience classrooms through training topographic map skills. One of these technologies is the augmented-reality sandbox, which he is using to create student-informed curricula that effectively teach topography and improve spatial thinking skills concurrently. He is also involved in ongoing research in the field of health geography where he is using GIS to map and analyze the spatial distribution of geology, groundwater quality, demography, and chronic illness in the Southern United States. Eli is an active member of several professional organizations including the National Association of Geoscience Teachers (NAGT) and the Geological Society of America (GSA).