TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 9 (2017) 883–901

ISSN:1756-8757 print / 1756-8765 online
DOI: 10.1111/tops.12233

This article is part of the topic "Sketching and Cognition," Kenneth Forbus and Shaaron Ainsworth (Topic Editors). For a full listing of topic papers, see: http://onlinelibrary.wiley.com/doi/10.1111/tops.2017.9.issue-4/issuetoc

Comprehending 3D Diagrams: Sketching to Support Spatial Reasoning

Kristin M. Gagnier,[a] Kinnari Atit,[b] Carol J. Ormand,[c,d] Thomas F. Shipley[e]

[a]*Science of Learning Institute, Johns Hopkins University*
[b]*Department of Psychology, Northwestern University*
[c]*Science Education Resource Center, Carleton College*
[d]*Department of Geoscience, University of Wisconsin-Madison*
[e]*Department of Psychology, Spatial Intelligence and Learning Center, Temple University*

Received 27 April 2015; received in revised form 24 March 2016; accepted 26 September 2016

Abstract

Science, technology, engineering, and mathematics (STEM) disciplines commonly illustrate 3D relationships in diagrams, yet these are often challenging for students. Failing to understand diagrams can hinder success in STEM because scientific practice requires understanding and creating diagrammatic representations. We explore a new approach to improving student understanding of diagrams that convey 3D relations that is based on students generating their own *predictive* diagrams. Participants' comprehension of 3D spatial diagrams was measured in a pre- and post-design where students selected the correct 2D slice through 3D geologic block diagrams. Generating sketches that predicated the internal structure of a model led to greater improvement in diagram understanding than visualizing the interior of the model without sketching, or sketching the model without attempting to predict unseen spatial relations. In addition, we found a positive correlation between sketched diagram accuracy and improvement on the diagram comprehension measure. Results suggest that generating a predictive diagram facilitates students' abilities to make inferences about spatial relationships in diagrams. Implications for use of sketching in supporting STEM learning are discussed.

Keywords: Sketching; STEM education; Diagram understanding; Spatial reasoning; Analogical reasoning

Correspondence should be sent to Kristin M. Gagnier, Science of Learning Institute, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218. E-mail: kristin.gagnier@jhu.edu or Thomas F. Shipley, Temple University, Department of Psychology, 1701 North 13th Street, Philadelphia, PA 12122. E-mail: tshipley@temple.edu

1. Introduction

Visualizing three-dimensional (3D) structures is a challenge that pervades science, technology, engineering, and mathematics (STEM) learning. 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. Examples range from dash-wedge diagrams, which illustrate atomic-scale spatial configurations, to geologic block diagrams, which illustrate geologic structures at scales ranging from centimeters to tens of kilometers. While these types of representations are ubiquitous in STEM learning (Ainsworth, Prain, & Tytler, 2011; Cromley et al., 2013; Hegarty, 2005; Newcombe & Stieff, 2012), students struggle to interpret the 3D spatial relations conveyed in these diagrams (Kali & Orion, 1996; Rapp, Culpepper, Kirkby, & Morin, 2007; Stull, Hegarty, Dixon, & Stieff, 2012). A failure to understand these representations can be a barrier to success in STEM, as a key aspect of scientific practice is both understanding and self-generating these types of representations (Ainsworth et al., 2011; Nersessian, 2008). While it is common for STEM educators to give students frequent opportunities to practice interpreting diagrams of 3D features, it is less common to give them frequent opportunities to generate their own visual representations through sketching (Ainsworth et al., 2011). Recently, however, there has been a surge of research interest in the role of learner-generated drawings in science learning (Prain & Tytler, 2012; Tytler, Haslam, Prain, & Hubber, 2009; Van Meter, Aleksic, Schwartz, & Garner, 2006; Van Meter & Firetto, 2013; Van Meter & Garner, 2005; Waldrip, Prain, & Carolan, 2010). Generally speaking, this work has suggested that sketching activities are beneficial for science learning. In this study we examined whether sketching can support a critical part of science learning, the understanding of diagrams that convey 3D spatial relations.

Work on student sketching has shown that it is a promising tool both for facilitating inferential reasoning (Gobert, 2000, 2005; Gobert & Clement, 1999; Johnson & Reynolds, 2005) and assessing knowledge (e.g., Jee et al., 2014; Johnson & Reynolds, 2005; Matlen, Atit, Goksun, Rau, & Ptouchkina, 2012). For example, Gobert and Clement (1999) had students learn about plate tectonics using lessons that included (a) reading and summarizing, (b) sketching diagrams that represented main text ideas, or (c) reading the text alone. Performance on comprehension measures revealed that students in the sketching group performed better on summative assessments, which assessed inferential reasoning. Jee et al. (2014) found that the types of information conveyed in a student sketch (such as process and relational information) predicted their level of geoscience content knowledge. Johnson and Reynolds (2005) suggest the use of concept sketches to promote organization and consolidation of information in the geosciences. Concept sketches are sketches that students annotate to show process, concepts, and spatial relationships among geologic regions or concepts.

Sketching has a long history in STEM and in geology in particular (e.g., Johnson & Reynolds, 2005, and references therein; Mogk & Goodwin, 2012; Turner & Libarkin, 2012). Geology is the study of the Earth: an oblate spheroid, the vast majority of which is inaccessible to humans. Geologists, therefore, have to be able to reason about 3D phenomena that are often only directly observable on the surface of the Earth, and which are commonly obscured by vegetation or other land cover. Sketching is a tool commonly used by expert geologists to record observations, make predictions, and recognize and evaluate hypotheses. Because sketching requires the scientist to generate a coherent and internally consistent representation, it facilitates the discovery of details and relations that would have otherwise gone unnoticed (Tikoff, 2014). Sketching is also often used as a pedagogical tool to help students reason about 3D structures that are not directly visible. For example, sketching is used in the field to reason about and *predict* the likely connection between different outcrops (visible exposures of rock on the Earth's surface). A common field trip format has students visiting multiple outcrops, sketching structural features at each outcrop, and then sketching how the different outcrops connect below the Earth's surface (Shipley, Tikoff, Ormand, & Manduca, 2013).

As sketching is commonly used in geology to support 3D visualization, we saw an opportunity to use sketching to facilitate student understanding of diagrams that depict 3D relations. We reasoned that one explanation for why these diagrams are so challenging is that they require students to make spatial inferences about 3D relations not visible in the diagram. Drawing on Tikoff (2014) and on our conversations with geology professors, we hypothesized that sketching spatial inferences (i.e., making spatial predictions) involves several cognitive processes that support understanding of diagrams that convey 3D spatial relations. First, it requires the sketcher to visualize and focus on within- and between-object spatial relationships and generate a spatial prediction regarding this visualization, and the act of sketching supports this visualization as the sketch is being created. Second, the sketcher must then align his or her prediction to the diagram space (Forbus, Usher, & Tomai, 2005). Third, the sketcher has to generate a coherent representation in which the drawn lines consistently correspond to some feature of the world (Van Meter & Firetto, 2013). Fourth, a sketch provides a permanent record of the sketcher's prediction and thus allows for self-guided comparison against the correct answer.

We approached the problem of learning to visualize 3D spatial relations in the context of geologic block diagrams (see Fig. 1). These diagrams show two sides and the top of a block that represents a 3D geologic structure, and they illustrate how layers of rock extend into the earth in three dimensions. These diagrams are often used to represent 3D spatial relationships, yet students often err when making spatial inferences from these diagrams (Alles & Riggs, 2011; Atit, Gagnier, & Shipley, 2015; Kali & Orion, 1996; Ormand et al., 2014). Errors students make when interpreting these diagrams range from failing to perceive that the block conveys any 3D relations, to assuming that the interior of the block is identical to one exterior face, to failing to interpolate information from multiple sides of the block to visualize how the layers extend in (Alles & Riggs, 2011; Kali & Orion, 1996). Ongoing research in our laboratory surveyed four introductory geology textbooks[1] and found that approximately 18% of all diagrams present in these textbooks are block diagrams designed to convey volumetric information (Atit et al., 2015). As these diagrams are prevalent in introductory texts, students who

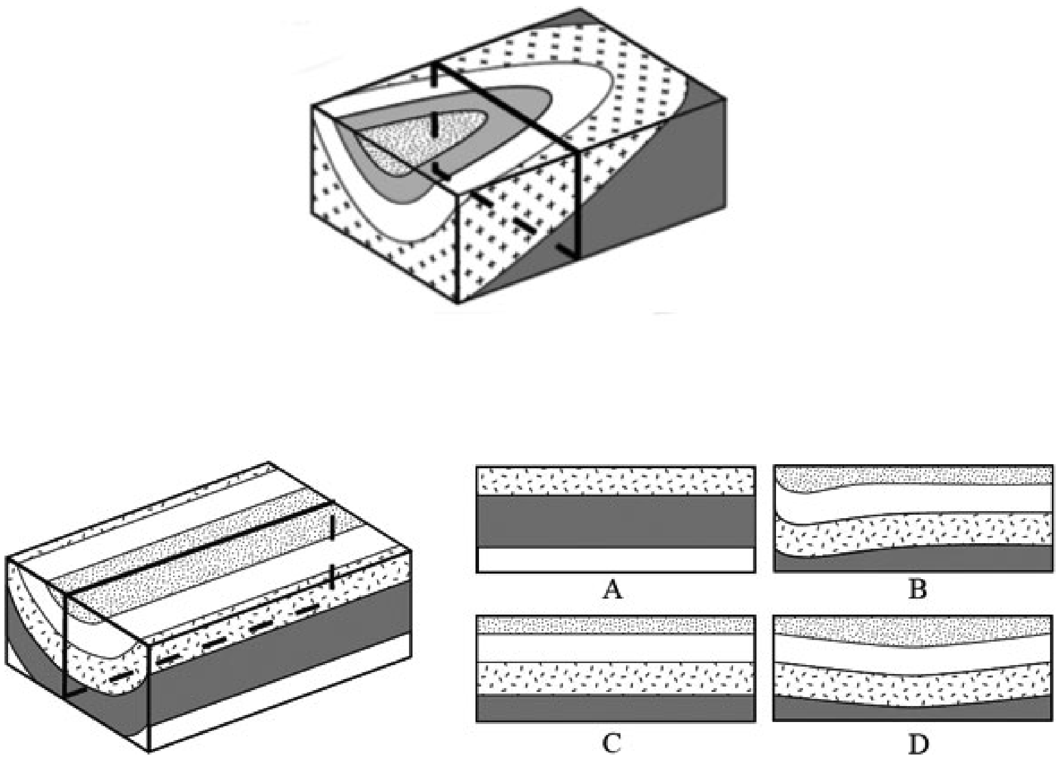

Fig. 1. Top: An example of a geologic block diagram, illustrating one possible configuration of folded rock layers. Students in a classroom might be asked to sketch the cross-section produced by the cut indicated in the middle of the block. Bottom: An example of a geologic block diagram test item (taken from the Geologic Block Cross-sectioning Test; Ormand et al., 2014). Participants are directed to visualize and select the cross-section, or 2D surface, produced by the pictured cut. The correct answer is C.

fail to understand them may experience significant challenges understanding key geological concepts.

To investigate the efficacy of predictive sketching, we gave undergraduate psychology students pre- and post-measures of geologic block diagram understanding in which they had to reason about the 3D spatial relationships presented in the diagram to select the cross-section that would be produced by the pictured cut (example item shown in Fig. 1). Although this assessment uses representations of geologic structures, no prior knowledge of geology is required to interpret them; study participants were instructed to imagine a knife cutting into the diagram and pushing off the front face, similar to cutting into a loaf of bread so you can see the inside of the loaf. We selected undergraduate psychology students because prior work in our laboratory found no difference in performance on our measure of geologic block understanding across psychology and introductory geology students (Ormand et al., 2014). Thus, our participants are a good proxy for geology students at the beginning of their training, and any benefit of our intervention would likely be effective for such students.

Between the tests, students engaged in a learning intervention in which they viewed color photographs of 3D Play-Doh models of geologic blocks and diagrams of geologic models (see examples in Fig. 2). In the *predictive sketching* condition, students sketched diagrams predicting the internal structure of the model and then immediately were shown the correct answer and asked to compare and evaluate their diagram. We hypothesized that predictive sketching requires students to *visualize* internal spatial relationships in the model and then *convey* this visualization in an external representation (a sketch). To examine whether sketching is necessary for improvement or if simply *visualizing* a prediction would also facilitate diagram understanding, we compared improvement in the sketching condition to a *visualization without sketching* condition, in which the participants predicted, but did not sketch, the interior structure of the model. Performance in both conditions was compared to a control condition, the *copying* condition, which mimicked the traditional classroom practice of copying information into notes; participants simply copied the models on paper and thus did not engage in visualizing or predicting the internal structure.

If predictive sketching facilitates diagram understanding, then we expect improvement from pre-test to post-test in the predictive sketching condition but not in either the

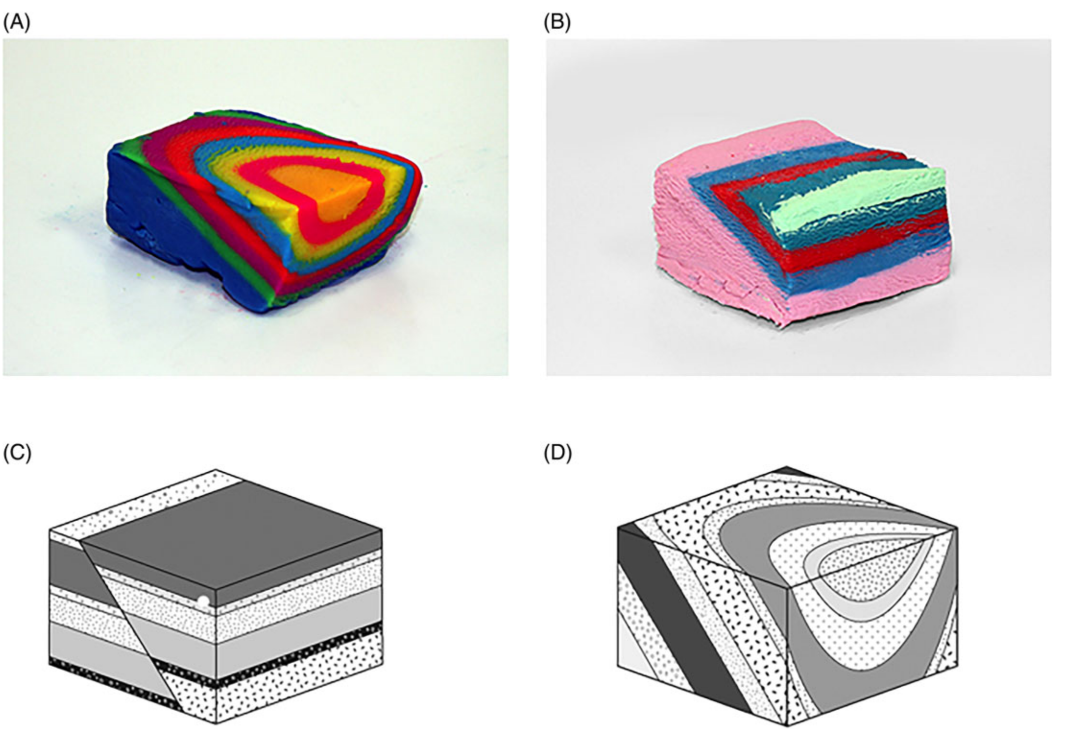

Fig. 2. The four structures used as stimuli in order of presentation during the intervention. (A) Gently plunging fold; (B) dipping layers; (C) faulted horizontal layers; (D) steeply plunging fold.

visualization without sketching or copying conditions, as previous research in our laboratory has shown no test–retest improvement by simply taking the test twice (Ormand et al., 2014).

2. Method

2.1. Participants

Participants were 105 undergraduates from a large urban university (70 women), fulfilling a requirement for an introductory psychology course.

2.2. Stimuli

The stimuli consisted of color photographs and line drawings of four types of geologic block models and slices through those block models (see Fig. 2). These four models represent common geologic structures that introductory geology students learn about: dipping layers, horizontal layers cut by a reverse fault, and gently and steeply plunging folds. The Play-Doh models show prototypical views of these structures and were created using seven different colored layers of Play-Doh. The line drawings were created using Adobe Illustrator. The images were approximately 13.5 $\times$ 9 cm.

2.3. Apparatus

Stimuli were presented in a PowerPoint presentation on a 24-inch flat-screen Apple Macintosh monitor. Participants were seated approximately 50 cm away from the monitor.

2.4. Design and procedure

2.4.1. Design

Participants were tested individually and randomly assigned to the *predictive sketching* (n = 35), *visualization without sketching* (n = 35), or *copying* (n = 35) conditions.

2.4.2. Procedure

Participants first completed a pre-test measure of items from the Geologic Block Cross-sectioning Test (GBCT; Ormand et al., 2014). This test consisted of seven multiple-choice questions similar to that shown in Fig. 1. Participants were instructed to select the answer choice that best represents the cross-section produced by the pictured cut and were given 4 min to complete the seven questions.

Following the pre-test, participants were given 50 min to complete a self-paced learning intervention. During the intervention they viewed a PowerPoint presentation in which they saw vertical cuts (i.e., cuts that were straight down into the model) into four geologic structures: a shallowly plunging fold, dipping layers, faulted horizontal layers, and a steeply plunging fold. Structures 1 and 2 (a shallowly plunging fold and dipping layers) were Play-Doh models and 3 and 4 (faulted horizontal layers and a steeply plunging fold) were line drawings, as shown in Fig. 2. In the predictive sketching condition, students viewed a model (as shown in Fig. 3A top left) and generated a sketch of what the model would look like following the indicated cut (Cut 1 in figure). After completing

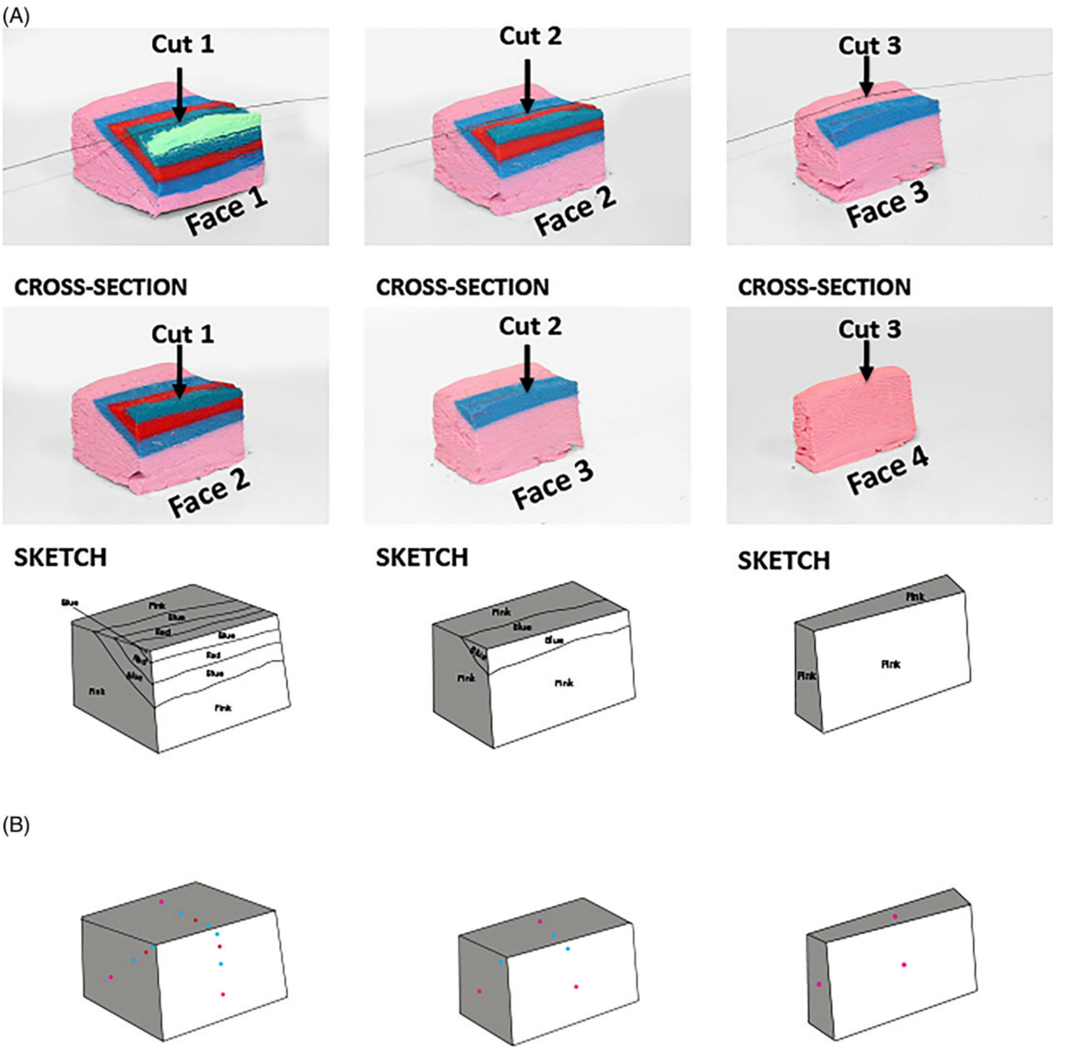

Fig. 3. (A) An illustration of the method. Students saw the top left photograph and generated a diagram of what the model would look like after Cut 1. After generating their diagram, they viewed the correct answer (middle left) and the correct diagram (bottom left) spatially aligned on the screen and compared their diagram to the correct answer. The sequence repeated for Cuts 2 and cut 3. (B) An example of the dot placed in the visualization without sketching condition.

each sketch, students wrote an explanation of how they used the layers that were visible on the front, side, and top of the block to make their prediction. They then viewed the correct answer (a photograph of the structure after it has been cut, Fig. 3A left middle) and a diagram of the sliced block spatially aligned with the original model to promote comparison (as shown in Fig. 3A left bottom) and actively compared their sketch to the correct answer by indicating what was different between their sketch and the correct answer. This sequence was then repeated for Cuts 2 and 3 into the model. While the participant sketched their prediction for the model after Cut 2, the answers for Cut 1 were visible. Similarly, while sketching their prediction for Cut 3, the answers for Cuts 1 and 2 were visible. This was done to help guide participants' spatial inferences by allowing them to perceive the changes from the front face to the inside in the previous cut. Once the correct answers to Cut 3 were shown, a new model appeared and this sequence was repeated for the next model for a total of 12 possible sketches or until 50 min had elapsed.

The visualization without sketching condition was identical in stimuli and instructions with the following exceptions. Instead of sketching, participants placed one dot to indicate the middle of both the height and the width of each layer on each face of the block, as shown in Fig. 4. To explain this, they were shown an example of a sketched diagram and the dot pattern for that diagram. In the visualization without sketching condition, participants used colored pencils for their dots to indicate the color of the layer they were marking. This allowed us to code the dot diagrams for accuracy. None of the participants reported being color anomalous or otherwise having any difficulty matching the colored pencils to the colored layers. As in the sketching condition, after generating their dot representation, students wrote an explanation of how they used the layers that were visible on the front, side, and top of the block to make their prediction. Thus, any benefit of self-explanation should occur in both conditions. Students then compared their dot representation to the correct answers. Finally, they copied the dot diagram of the correct answer. This copying task encouraged them to spend more time thinking about how to convey the correct answer by placing a dot to represent the middle location of each layer and was included because pilot work in the visualization without sketching condition revealed that participants in that condition were generating a greater number of diagrams than in the sketching condition. Including this copying task allowed us to avoid confounding the qualities of the intervention with number of sketches participants generated.

The copying condition was designed to mimic learning in a typical classroom where students copy diagrams and notes into their notebook, and to take the same amount of time as participants took in the predictive sketching condition, thus equating the time on task across conditions. Participants in the copying condition viewed the same images in the same order as the other two conditions and engaged in a task that took the same amount of time as predictive sketching, but they focused their attention on external information. They were instructed to estimate the amount of paint it would take to paint the top, front, and side of the first model (i.e., the block before it had been cut). They were given a base amount (e.g., it would take 0.5 gallons of red paint to paint all of the red layers on this block) and asked to use that amount to estimate how much paint it would

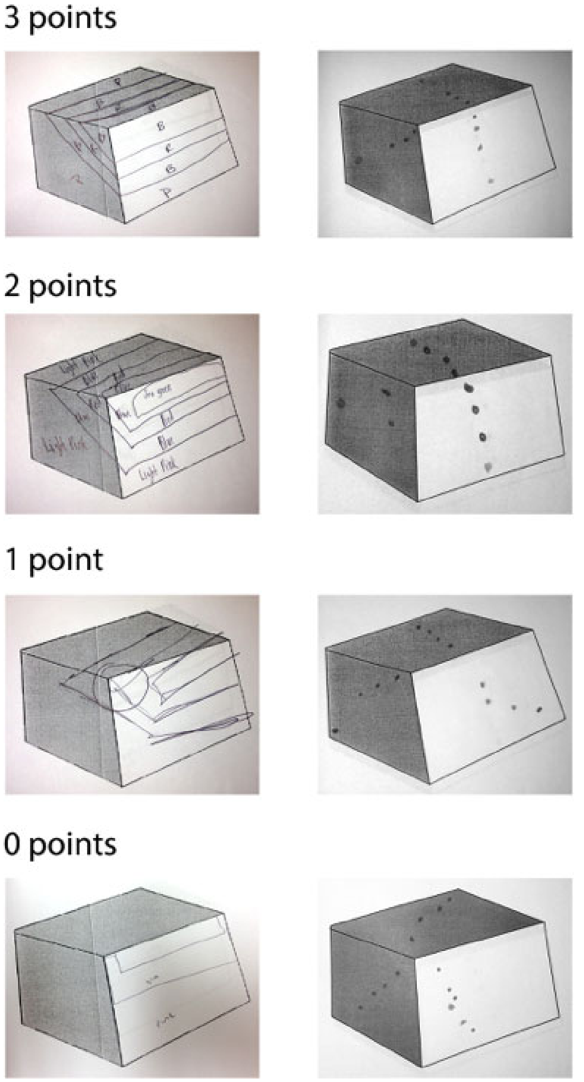

Fig. 4. Examples of diagrams in predictive sketching (left) and visualization without sketching (right) conditions, illustrating the range of accuracy. *3 points*: Sketch: Layers on all faces are correct and connected. Dot: Dots are in the correct order and position with respect to each other on all faces. *2 points*: Sketch: The shape of the layers on the cross-section face is incorrect, but the layers are shown correctly on the other two faces and are connected appropriately. Dot: Dots on the left-side face are not in correct location with respect to each other, but dots on the other two faces are in the correct positions and sequence. *1 point*: Sketch: Layers on all three faces are not the correct shapes, but they are connected appropriately. Dot: Dots in the cross-section and on the side face are not in the correct relative positions, but there are corresponding dots on each face. *0 points*: Sketch: No layers are drawn on the top and side faces and the cross-section layers are incorrect. Dot: Positions of the dots relative to each other are not correct on either the cross-section or the top. The number of dots on each side differs, indicating lack of correspondence across the faces of the block. Dot diagrams were drawn in color.

take to paint all of the other colors on each side of the block. Participants then copied the block after it was cut, thus generating a diagram of the same structure as the sketching group, but without engaging in visualization or prediction. By equating the time on task across the conditions while engaging participants in a task that did not focus them on predicting spatial relationships, we could examine whether there was any benefit to learning from simply copying the correct answer (as is common in class). As in the other conditions, they viewed what their sketch should have looked like and were asked to compare their diagram with the correct answer.

All participants saw the stimuli in the same order. The first structure was a Play-Doh model of a shallowly plunging fold, and the second structure was a Play-Doh model of dipping layers. The third and fourth structures were line diagrams of faulted horizontal layers and a steeply plunging fold, respectively. All blocks are shown in Fig. 2. We started with Play-Doh models because we hypothesized that it would be easier for participants to visually interpret 3D objects than line diagrams. Line drawings were included at the second half of the experiment to scaffold participants' transition from reasoning about structures in photographs of 3D models to reasoning about how 3D spatial properties are conveyed in 2D diagrams, which are representative of the kinds of diagrams used in STEM courses and which study participants would see on the post-test. More specifically, we used the Play-Doh model of a shallowly plunging fold to scaffold participants' understanding of plunging folds in general, and we used the Play-Doh model of dipping layers to scaffold participants' understanding of parallel, planar layers.

Participants were given up to 50 min to complete all 12 sketches. The time each participant spent on the learning intervention was recorded. If they had not completed the last sketch within the allotted time, they were instructed to stop and proceed to the next phase of the experiment. After the PowerPoint presentation, all participants were given a post-test measure of seven new items from the GBCT that had been equated for difficulty with the items used in the pre-test based on previous work in our laboratory.

2.4.2.1. Diagram coding rubric: To examine the effect of sketching-relevant spatial information on learning how to interpret diagrams, we graded each student-generated diagram based on three criteria: (a) whether the predicted shapes of the layers in the cross-section were correct, (b) whether the shapes of the layers on the other block face and top were correct, and (c) whether the layers drawn on all three faces were connected. Each sketch was given 1 point for meeting a criterion, or 0 if not, for a total possible score of 3. Examples of sketches ranging from a score of 3 to 0 are shown in Fig 4. Reliability for the sketch coding was established by having a second coder score 20% of the sketches in each condition. Inter-rater reliably was $\kappa = .82$, ($p < .01$, $n = 195$ sketches), 95% CI (0.77, 0.87). For this analysis we did not code for either the qualitative character of errors (e.g., the layers were not in the correct order) or quantitative character (e.g., boundaries were in the wrong location or shape). Future work will analyze the nature of participants' errors and provide a more detailed analysis to further understand the relationship between sketch errors and learning.

3. Results and discussion

3.1. Performance on Geologic Block Cross-sectioning Test

To examine whether sketching a diagram predicting spatial relationships led to greater improvement than either visualizing the interior but not sketching, or copying the correct answer, we compared improvement from the pre-test to the post-test across the three conditions. The mean number correct on the pre- and post-tests for each condition is shown in Fig. 5.[2] A one-way ANOVA revealed no differences in pre-test performance across conditions, $F(1, 102) = 1.2$, $p = n.s.$ As is evident in Fig. 5, only the sketching condition yielded significant improvement from pre- to post-test. A 3 (condition) × 2 (pre- and post-score) mixed measures ANOVA on the mean number correct in each condition revealed no improvement from pre-test to post-test, $F(1, 102) = 1.8$, $p = n.s$, no effect of condition, $F < 1$, and a condition × improvement interaction, $F(2, 102) = 4.0$, $p < .05$. Pairwise comparisons revealed the interaction was driven by the improvement from pre-test to post-test in *the predictive sketching conditio*n, $t(34) = 2.8$, $p < .01$, but no improvement in the visualization without sketching condition ($t(34) = .00$, $p = n.s.$) or the copying condition ($t(34) = .72$, $p = n.s.$). Planned contrasts reveal that the improvement in the predictive sketching condition was greater than in the visualization without

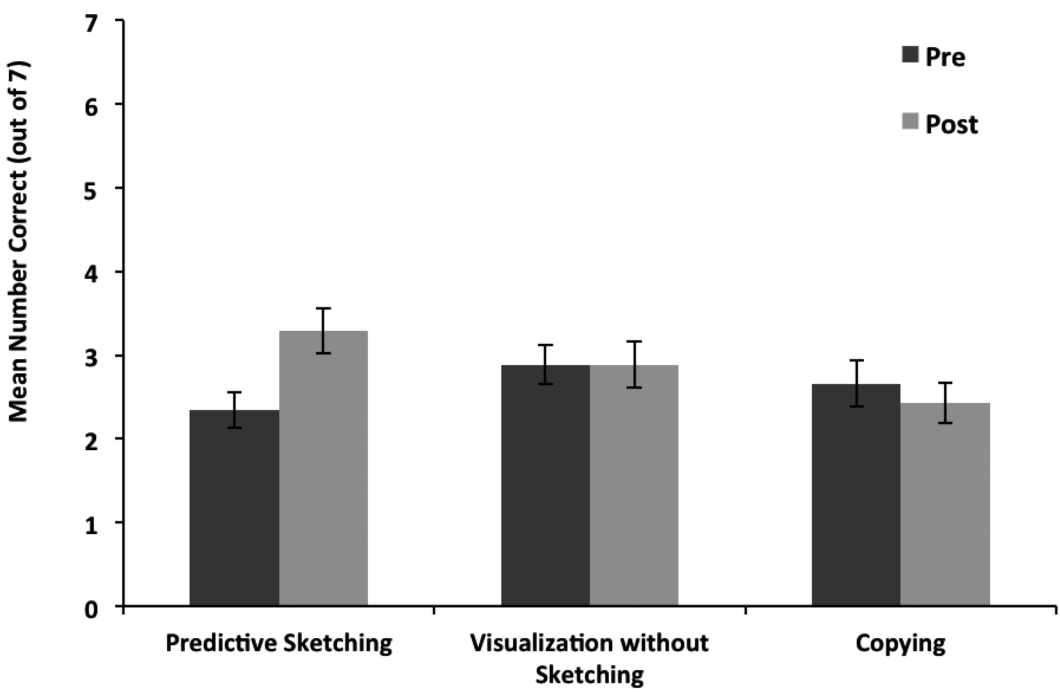

Fig. 5. The mean number of correct answers (out of 7) on the pre-test and the post-test in each condition. The error bars show the standard error of the mean in each condition.

sketching condition, $t(102) = 2.2$, $p < .05$, $d = .51$, and the copying condition, $t(101) = 2.7$, $p < .05$, $d = .58$.

The mean number of sketches or diagrams completed during the intervention was 9.0 (SD = 2.4) in the predictive sketching condition, 7.1 (*SD* = 2.4) in the visualization without sketching condition, and 10.4 (*SD* = 1.9) in the copying condition. A one-way ANOVA revealed that although we attempted to equate the number of diagrams completed during the intervention in each condition, there was still a difference, $F(2, 101) = 18.1$, $p < .01$. Thus, participants in the visualization without sketching condition completed an average of 1.9 fewer sketches than participants in the predictive sketching group. One possible interpretation of these results is that participants in the visualization without sketching condition did not improve on diagram understanding because they did not generate as many diagrams as the predictive sketching condition. However, two findings suggest that is not the case. First, an ANCOVA, comparing improvement from the pre-test to the post-test with the number of sketches completed entered as a covariate, found that the overall pattern of results still holds; there is a significant improvement by condition interaction, $F(2, 101) = 3.7$, $p < .05$. Second, there is no relationship between the number of diagrams generated and improvement, $r = -.26$, $p = .14$. Thus, although participants in the visualization without sketching condition completed fewer sketches, there is no evidence that completing more sketches results in greater learning gains. We conclude, instead, that the *process* of generating a sketch predicting the interior structure of the block facilitated understanding of the diagrams, regardless of the number of diagrams completed.

3.2. Diagram quality results

To better understand the relationship between sketching and learning, we examined whether the quality of the sketched diagram was related to improvement from pre- to post-test. In the predictive sketching condition, there was no relationship between diagram accuracy and performance on the pre-test, $r(31) = .21$, *n.s.* There was, however, a positive relationship between diagram accuracy and both post-test performance, $r(31) = .54$, $p < .01$, and improvement from pre- to post-test, $r(31) = .37$, $p < .05$.[3] In contrast, in the visualization without sketching condition, even though students were making predictions about the interiors of the blocks, we did not find a relationship between diagram accuracy and pre-test performance, post-test performance, or improvement ($r(33) = .22$, *n.s.*, $r(33) = .23$, *n.s.* and $r(32) = .05$, *n.s.*, respectively). The pattern observed for the copying condition was the same as for the visualization without sketching condition ($r(33) = .09$, *n.s.*, $r(33) = .19$, *n.s.*, $r(33) = .07$, *n.s.*, respectively).

In sum, *visualizing and sketching a diagram predicting spatial relationships that are not directly observable* led to improvements in diagram understanding. These improvements were greater than those found for both visualizing without sketching and copying the correct diagram (sketching without visualizing and predicting). Improvements when sketching a diagram that predicts unseen spatial relationships cannot be attributed to self-explanation as both the predictive sketching and visualizing without sketching conditions included a self-explanation component. Rather, the predictive sketching appears to selectively facilitate learning.

4. General discussion

Diagrams that convey 3D relations are common in STEM yet challenging for students. We examined whether predictive sketching could facilitate student understanding of diagrams that convey 3D spatial information. We focused on sketching because it is often used in the geosciences to promote 3D visualization and as a tool for making spatial predictions. However, to our knowledge, no work has examined (in a controlled laboratory setting) whether sketching could facilitate diagram comprehension. We found that students who accurately sketched diagrams predicting spatial relations inside a geologic block model and then compared their prediction against the correct answer improved on a measure of 3D diagram understanding. This was compared both to students who visualized their prediction for the internal structure but did not sketch their predictions and also to those who viewed the same images but copied the visible surfaces and thus did not make spatial predictions. In addition, we developed a simple rubric to evaluate the quality of the diagrams. Students who accurately sketched the internal spatial relationships in their diagrams demonstrated greater improvement on a measure of 3D diagram understanding than students who did not accurately convey this information in their diagrams. These results suggest that sketching diagrams predicting unseen spatial relations supports spatial reasoning about 3D relationships conveyed in diagrams.

There are a number of reasons why sketching that predicts spatial relationships may lead to improved understanding of geologic block diagrams. First, to interpret geologic block diagrams, the student has to make inferences about spatial relationships inside the model that cannot be directly observed. We suggest that asking students to sketch a spatial prediction first requires them to make 3D spatial inferences about the model that are not directly visible, and then to *map or align* those 3D inferences into the diagrammatic space. To create a diagram of the interior structure of the model, students had to establish a mapping between their prediction of the 3D relationships in the model and their diagram. Predictive sketching may support this analogical mapping between real 3D relationships present in a model and the depiction of these relationships in a diagram (see Forbus et al., 2005, for a discussion of sketching and analogical learning).

Second, sketching a prediction creates a permanent representation of the student's mental prediction that can be compared to the correct answer and evaluated to promote learning. To generate a predictive sketch, the student must consider *what* will be on the inside of the model and *how* to represent that information in the diagram. Participants in the visualization without sketching condition were asked to visualize *what* would be on the inside of the model, but without sketching they did not go through the process of envisioning *how* to portray these spatial relationships in a diagram. Conversely, participants in the copying condition engaged in the process of *how* to represent 3D spatial information in a diagram, but they did not have to figure out *what* the interior would look like.

Third, the act of sketching forces the learner to form a conceptual model, which can be used to make a prediction (see Shipley & Tikoff, 2016, for a review of reasoning from conceptual models in geology; and see Haysom & Bowen, 2010, for a discussion of the value of predictions in science learning). Predictive sketching forces the student to create an external representation of their conceptual model. This externalization is a permanent record of their prediction, which can be used by learners to evaluate the completeness and coherence of their conceptual model and to see where their prediction was correct and where it was incorrect. This process gives learners feedback and provides an opportunity to revise their conceptual model. Previous work has argued that feedback facilitates learning by encouraging students to develop the capacity to monitor the quality of their work while the work is being produced (Sadler, 1989). An important property of feedback in this paradigm is that it is spatial in nature. *Spatial feedback* may be more efficient for learning to interpret diagrams than simply correct or incorrect, as it provides additional information about the qualitative and quantitative nature of the error (e.g., location, direction, shape, etc.) in real time while the student is working. In this study we provided participants with only minimal instructions regarding how to compare and evaluate their diagram against the correct answer. Future work is needed to determine what types of comparison instructions lead to the most learning. Finally, sketching is active and can be conceived of as a constructionist task, which may have benefits beyond the facilitation of spatial thinking. Work within the drawing literature has shown that compared to various other activities, drawing increases motivation and engagement (see Ainsworth et al., 2011 for review). Indeed, effectively supporting spatial thinking may, in and of itself, have motivational benefits.

These findings, together with the work of Gobert and colleagues (Gobert, 2000, 2005; Gobert & Clement, 1999), indicate that sketching is particularly effective for supporting science learning when students have to *transform* information. In this work, simply copying information was not enough to support learning from sketching. Instead, only students who transformed visible information to generate a conceptual model, and then made spatial inferences about the model and sketched their inferences, improved on 3D diagram understanding. Importantly, predictive sketching allowed this transformation, while participants who transformed visible information into spatial predictions without sketching did not improve. Similarly, Gobert and Clement (1999) found that when students generated sketches that transformed key ideas in the text into visual representations, they were better able to understand spatial and dynamic processes related to plate tectonics compared to those who simply summarized text passages.

This invites the question of why transforming information is important. It might be that the process of having to transform information requires the student to develop a detailed and structured conceptual model of the object of study. This conceptual model has to be one that formally represents properties of the object in a way that allows a transformation to be applied (see Gallistel, 1990, for discussion of the formal relations between spatial representations and transformations). The sketch may provide a tool for organizing this model. The spatial feedback provided in this study may have allowed correction of this model if it was incorrect; broadly speaking, such learning could be characterized as *spatial accommodation*. The task of copying might not require a sufficiently detailed conceptual model to result in new learning.

4.1. Relationship between diagram accuracy and performance

The other notable finding that emerged from our studies was a relationship between the quality of the information conveyed in students' diagrams and their performance on the post-test and improvement from pre-test to post-test. We found that students who made correct inferences about the internal structure of the diagrams and then correctly mapped these inferences to the diagram space were more likely to perform better on the post-test and improve from pre- to post-test than those students who did not. Critically, this relationship was not observed for the pre-test, showing that it is not the case that students who perform better are also better at sketching from the beginning. In neither the visualization without sketching condition nor the copying conditions did we find evidence that diagram accuracy is related to performance.

Schwamborn, Mayer, Thillmann, Leopold, and Leutner (2010) found that students who generated high-accuracy drawings during a learning intervention on chemical processes scored better than students who generated low-accuracy drawings. Our findings add to this by showing that students who generate high-accuracy predictive diagrams were likely to perform well on measures of learning to use similar 3D diagrams. However, the mechanism(s) of learning through sketching remains an open question. One possibility is that accurate sketches are an index of spatial learning throughout the intervention. Alternatively, better sketching throughout the intervention might result in improvements in spatial learning. We are currently working to tease apart the contributions of various aspects of predictive sketching to learning.

4.2. Educational lmplications

Many STEM fields require students to reason about 3D relations that cannot be directly observed because of the spatial scales on which they occur. Consequently, these relations are illustrated in diagrams. Students struggle to understand and learn from diagrams (Cromley & Byrnes, 2012; Newcombe, 2013), and diagrams that convey 3D spatial information are particularly challenging (e.g., Alles & Riggs, 2011; Kali & Orion, 1996; Rapp et al., 2007; Stull et al., 2012). Our results suggest that instructors may foster students' understanding of diagrams by structuring opportunities for students to sketch predicted spatial relationships that are not directly visible in the diagram and then compare their predictive diagrams to the correct answer.

Reasoning about the 3D structure of objects is relevant for many other STEM disciplines in tasks such as thinking about conic sections in math (Boe, 1968; Davis, 1973; Piaget & Inhelder, 1956), thin sections in biology (Russell-Gebbett, 1984), recognizing 3D neuroanatomical structures in whole brain slices (Chariker, Naaz, & Pani, 2011), and the structure of roots or cavities in dentistry (Cohen & Hegarty, 2007; Khooshabeh & Hegarty, 2010). Although additional research is needed to explore the process of learning about these topics, we speculate that predictive sketching might facilitate performance on tasks that require students to visualize in three dimensions, across the STEM disciplines and beyond.

One of the critiques of using sketching in the classroom is that sketching is time consuming and sketches are laborious for instructors to grade. One of the values of our intervention is that evaluation is taken out of the hands of the instructor and put into the hands of the student. By showing students the correct answer after they have generated their own predictive diagram, instructors make it possible for students to compare their self-generated diagram to the correct answer and evaluate their own conceptual model. We do not know whether improvement would be greater if instructors were grading students; however, here we demonstrate that learning can occur without the instructor providing feedback. If the instructor were to grade sketches, the success of the diagram-scoring rubric developed here in predicting improvement in diagram comprehension suggests the rubric may be a simple and effective way to evaluate student readiness to learn from block diagrams in geology classes. This rubric is only useful for evaluating one type of 3D diagram, but it indicates the potential for developing very simple rubrics for evaluating other 3D diagrams.

A growing body of research in education has examined learning gains from student-constructed versus student-completed diagrams (Cromley et al., 2013; Van Meter & Garner, 2005). Some researchers have argued that students engage in a broader range of inferential mental processes when generating their own drawings (Van Meter et al., 2006). However, others suggest that early-stage learners may find it particularly difficult to generate drawings. Our results do not support this later concern. Our participant population was psychology undergraduates who had limited knowledge of the earth sciences and they succeeded on the task. However, the highly structured sequence of sketching, coupled with immediate visual feedback, may have been critical to students' success in self-generating diagrams in our study.

4.3. Conclusions

Diagrams that represent 3D relations, such as geologic block diagrams, are challenging for students. We show that sketching a diagram predicting a spatial relationship that students cannot see led to improved understanding of geologic block diagrams. Copying without making predictions and making predictions without sketching, in contrast, did not lead to improved understanding of geologic block diagrams. Furthermore, we found a positive correlation between sketched diagram accuracy and improvement on the diagram comprehension measure. We suggest that having students make predictive sketches about the unseen features of 3D objects, and then comparing their sketches to the correct answer, is a strong strategy for developing students' 3D spatial visualization skills and their understanding of the kinds of 3D diagrams common in STEM.

Acknowledgments

This research was supported by a grant to the Spatial Intelligence and Learning Center, funded by the National Science Foundation (SBE-0541957 and SBE-1041707). We thank Play-Doh for the contribution of materials for our research. We thank Molly Kaiser, Ashley Johnson, Katelyn Wholey, Stephany Wilson, and David Zaslav for their valuable assistance in collecting data and Kristin O'Connell for drafting the line drawings of geologic block diagrams that were used for stimuli.

Notes

1. (a) *Earth: Portrait of a Planet by Marshak*; (b) *How Does Earth Work? Physical Geology and the Process of Science* (2nd ed.) by Smith and Pun; (c) *Exploring Geology* (3rd ed.) by Reynolds, Johnson, Morin and Carter; and (d) *Earth: An Introduction to Physical Geology* (10th ed.) by Tarbuck, Lutgens and Tasa. Each book has at least 80 adoptions.
2. Note that the mean number correct scores are fairly low (2.5 of 7 = 36%). This is a challenging test for all students. Whereas introductory psychology students do not differ from introductory geology students, advanced undergraduate geology students do significantly better, with class averages in a previous study ranging from 57% to 74% (Ormand et al., 2014).
3. Two sketching participants were removed from the diagram accuracy analysis because they were bivariate outliers. Pedhazur (1997) suggests using the DFBETA criterion recommended by either Belsey, Kuh, and Welsch (1980) or Mason, Gunst, and Hess (1989) for small sample sizes, 2/sqrt (n) and 3/sqrt (n), respectively. DFBETA is a measure of the change in the regression coefficient as a consequence of deleting a participant (Pedhazur, 1997). For our data the Belsley cut-off is .39 and the Mason cut-off is .51. The DFBETA values for these two participants were 1.06 and .55; thus, by either criterion both should be excluded from any analysis.

References

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, *333*(6046), 1096–1097.

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

Atit, K., Gagnier, K., & Shipley, T. F. (2015). Students' gestures help their penetrative thinking skills. *Journal of Geoscience Education*, *63*(1), 66–72.

Belsey, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: John Wiley.

Boe, B. L. (1968). A study of the ability of secondary school pupils to perceive the plane sections of selected solid figures. *The Mathematics Teacher*, *61*(4), 415–421.

Chariker, J. H., Naaz, F., & Pani, J. R. (2011). Computer-based learning of neuroanatomy: A longitudinal study of learning, transfer, and retention. *Journal of Educational Psychology*, *103*(1), 19.

Cohen, C. A., & Hegarty, M. (2007). Individual differences in use of external visualisations to perform an internal visualisation task. *Applied Cognitive Psychology*, *21*(6), 701–711.

Cromley, J. G., Bergey, B. W., Fitzhugh, S., Newcombe, N., Wills, T. W., Shipley, T. F., & Tanaka, J. C. (2013). Effects of three diagram instruction methods on transfer of diagram comprehension skills: The critical role of inference while learning. *Learning and Instruction*, *26*, 45–58.

Cromley, J. G., & Byrnes, J. P. (2012). Instruction and cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(5), 545–553.

Davis, E. J. (1973). A study of the ability of school pupils to perceive and identify the plane sections of selected solid figures. *Journal for Research in Mathematics Education*, *4*, 132–140.

Forbus, K. D., Usher, J., & Tomai, E. (2005). Analogical learning of visual/conceptual relationships in sketches. In *Proceedings of the national conference on artificial intelligence* (Vol. 20, No. 1, p. 202). London: AAAI Press.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.

Gobert, J. D. (2000). A typology of causal models for plate tectonics: Inferential power and barriers to understanding. *International Journal of Science Education*, *22*(9), 937–977.

Gobert, J. D. (2005). The effects of different learning tasks on model-building in plate tectonics: Diagramming versus explaining. *Journal of Geoscience Education*, *53*(4), 444.

Gobert, J. D., & Clement, J. J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching*, *36*(1), 39–53.

Hegarty, M. (2005). Multimedia learning about physical systems. In *The Cambridge handbook of multimedia learning* (pp. 447–465). New York: Cambridge University Press.

Jee, B. D., Gentner, D., Uttal, D. H., Sageman, B., Forbus, K., Manduca, C. A., Ormand, C. J., Shipley, T. F., & Tikoff, B. (2014). Drawing on experience: How domain knowledge is reflected in sketches of scientific structures and processes. *Research in Science Education*, *44*(6), 859–883.

Johnson, J. K., & Reynolds, S. J. (2005). Concept sketches-using student-and instructor-generated, annotated sketches for learning, teaching, and assessment in geology courses. *Journal of Geoscience Education*, *53*(1), 85.

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.

Khooshabeh, P., & Hegarty, M. (2010). Inferring cross-sections: When internal visualizations are more important than properties of external visualizations. *Human-Computer Interaction*, *25*(2), 119–147.

Mason, R., Gunst, R., & Hess, J. (1989). Analysis of completely randomized designs. In *Statistical design and analysis of experiments*. Hoboken, NJ: Wiley.

Matlen, B. J., Atit, K., Goksun, T., Rau, M. A., & Ptouchkina, M. (2012). Representing space: Exploring the relationship between gesturing and geoscience understanding in children. In C. Stachniss, K. Schill, & D. Uttal (Eds.), *Spatial Cognition VIII: International Conference, Spatial Cognition 2012 Kloster Seeon, Germany, August/September 2012 Proceedings* (pp. 405–415). Berlin: Springer.

Mogk, D., & Goodwin, C. (2012). Learning in the field: Synthesis of research on thinking and learning in the geosciences. *Geological Society of America Special Papers*, *486*, 131–163.

Nersessian, N. (2008). Inquiry: How science works; Model-based reasoning in scientific practice. In R. Duschl & R. Grandy (Eds.), *Teaching scientific inquiry: Recommendations for research and implementation* (pp. 57–79). Rotterdam: Sense Publishers.

Newcombe, N. S. (2013). Seeing relationships: Using spatial thinking to teach science, mathematics, and social studies. *American Educator*, *37*(1), 26–31; 40.

Newcombe, N. S., & Stieff, M. (2012). Six myths about spatial thinking. *International Journal of Science Education*, *34*(6), 955–971.

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*(1), 146–154.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Orlando, FL: Harcourt Brace.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space*. London: Routledge.

Prain, V., & Tytler, R. (2012). Learning through constructing representations in science: A framework of representational construction affordances. *International Journal of Science Education*, *34*(17), 2751–2773.

Rapp, D. N., Culpepper, S. A., Kirkby, K., & Morin, P. (2007). Fostering students' comprehension of topographic maps. *Journal of Geoscience Education*, *55*(1), 5.

Russell-Gebbett, J. (1984). Pupils' perceptions of three-dimensional structures in biology lessons. *Journal of Biological Education*, *18*(3), 220–226.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18* (2), 119–144.

Schwamborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a generative activity and drawing as a prognostic activity. *Journal of Educational Psychology*, *102*(4), 872.

Shipley, T. F., & Tikoff, B. (2016). Linking cognitive science and disciplinary geoscience practice. In R. W. Krantz, C. J. Ormand, & B. Freeman (Eds.), *Earth, mind, and machine: 3D structural interpretation*. American Association of Petroleum Geologists, Hedberg Series number 6.

Shipley, T. F., Tikoff, B., Ormand, C., & Manduca, C. (2013). Structural geology practice and learning, from the perspective of cognitive science. *Journal of Structural Geology*, *54*, 72–84.

Stull, A. T., Hegarty, M., Dixon, B., & Stieff, M. (2012). Representational translation with concrete models in organic chemistry. *Cognition and Instruction*, *30*(4), 404–434.

Tikoff, B. (2014). Sketching in the geosciences. Talk presented at a special workshop on Sketching in Science Education. May, Chicago, IL.

Turner, S., & Libarkin, J. (2012). Novel applications of Tablet PCs to investigate expert cognition in the geosciences. *Computers & Geosciences*, *42*, 162–167.

Tytler, R., Haslam, F., Prain, V., & Hubber, P. (2009). An explicit representational focus for teaching and learning about animals in the environment. *Teaching Science*, *55*(4), 21.

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, *31*(2), 142–166.

Van Meter, P. (2013). Cognitive model of drawing construction: learning through the construction of drawings. In. G. Schraw, M. T. McCrudden, & D. Robinson (Eds.), *Learning through visual displays* (pp. 247–280). Charlotte, NC: Information Age Publishing.

Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review*, *17*(4), 285–325.

Waldrip, B., Prain, V., & Carolan, J. (2010). Using multi-modal representations to improve learning in junior secondary science. *Research in Science Education*, *40*(1), 65–80.