Drawing on Experience: Use of Sketching to Evaluate Knowledge of Spatial Scientific Concepts

Benjamin D. Jee (b-jee@northwestern.edu)[1]
Dedre Gentner (gentner@northwestern.edu)[1]
Kenneth Forbus (forbus@northwestern.edu)[2]
Bradley Sageman (brad@earth.northwestern.edu)[3]
David H. Uttal (duttal@northwestern.edu)[1]

[1]Department of Psychology
[2]Department of Electrical Engineering and Computer Science
[3]Department of Earth and Planetary Sciences
Northwestern University
Evanston, IL 60607 USA

Abstract

How does learning affect the structure of domain knowledge? This question is difficult to address in domains such as geoscience, where spatial knowledge is paramount. We explore a new platform, called CogSketch, for collecting and analyzing participants' sketches as a means of discerning their spatial knowledge. Participants with differing levels of experience in the geosciences produced sketches of geologic structures and processes on a tablet computer running CogSketch software. This allowed for the analysis of not only the spatial-relational structure of the sketches, but also the *process* through which the sketches were constructed.

Introduction

The development of expertise is a central issue in cognitive science and education. Understanding the paths by which people acquire mastery in a domain is important both in guiding the design of training in the domain and in shaping theories of learning high-level knowledge (Bransford, Brown, & Cocking, 2000; Chi, Glaser, & Farr, 1988). To support student learning, cognitive scientists and educators need to assess the mental models that students develop at different points along the continuum of knowledge acquisition. This is a challenge in domains that are intensely spatial, such as geoscience, architecture and engineering, because evaluating students' spatial mental models requires that students produce a spatial depiction; assessing such sketches and designs is extremely laborious. In this paper we describe a new method for eliciting students' spatial knowledge through sketching.

Spatial knowledge in Geoscience

Geoscience requires learning about complex, large-scale processes that occur over large ranges of space and time. These include the structure of the Earth's interior, the deformation processes that occur within its crust (e.g., folds, faults, and fissures) and unifying both, the model of plate tectonics. All these topics are deeply spatial.

Consider the simple concept of a *fault* (see Figure 1). Faults are defined by the spatial relationship between the blocks of rock along a fracture. For the experienced geoscientist, this perceptual information is connected to causal knowledge about the formation of the structure. For the fault depicted in Figure 1, the geoscientist would notice that the block of rock above the fault plane has moved downward relative to the block of rock below the plane. This indicates that extensional forces deformed the rock until it fractured along the fault plane.

Figure 1. Example of a fault

Thus, simply to identify a basic geoscience concept, students must understand relevant spatial structures, which are intimately connected to causal knowledge. To take a more complex example, consider *subduction*, depicted in Figure 2. Understanding this process involves identifying spatial structures – e.g., that one plate is moving underneath another. Yet, noticing that structure is just the beginning. The student must also understand the driving forces: that the subduction process is causally related to both the process of seafloor spreading, which creates new ocean crust at divergent spreading zones, and differences in crustal density between converging plates, which determines which plate sinks below the other. In addition, subduction drives other relevant geologic processes, such as volcanic activity.

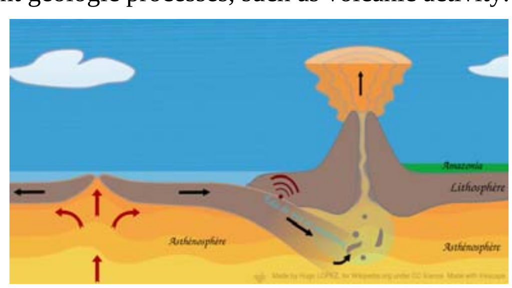

Figure 2. Diagram of Subduction

Because understanding the spatial structure of geological phenomena is crucial, students are constantly exposed to sketches of geoscience phenomena in their classes, in their texts (e.g., Marshak, 2005), and in the field, where their instructor is likely to ask them to sketch an outcrop and then explain their sketch to deepen their understanding. Thus, as part of their coursework, geoscience students are asked to produce sketches showing key spatial configurations.

What are the effects of the highly spatial tasks that geoscience students regularly engage in? Generally, as people gain expertise in a domain, their mental models emphasize key relational and goal-relevant structure, and deemphasize object-level appearances (e.g., Chi, et al., 1988; Gentner & Ratterman, 1991; Jee & Wiley, 2007). Thus, geoscience students may interpret geoscience diagrams, photographs, and outcrops in terms of their relational structure, especially key spatial relations that reflect geologic activity. Novices, however, may focus on objects and their attributes.

Sketching as an assessment of spatial knowledge

How can students' mental models of spatial phenomena be assessed – from relatively simple models for concepts such as fault, to complex models of phenomena such as plate tectonics? Clearly, verbal assessments will fall short because of the difficulty of conveying complex spatial relations through speech. The ideal solution would be to ask learners to draw the relevant structures. Sketches can reflect spatial and causal knowledge of the domain (Goel, 1995; Suwa & Tversky, 1997). Vosniadou & Brewer (1992), for example, used children's sketches to help discern their model of the Earth and solar system. With respect to geoscience expertise, people with more experience would be expected to draw geoscience-related images in terms of the spatial relations that reflect geologic activity. Moreover, if they engage a spatial model while interpreting such images, people with geoscience experience may be more likely to begin their sketches with relational information.

In addition to probing spatial and causal knowledge, another advantage of using student's own constructions is that sketching and handwriting can reveal implicit aspects of people's representations that are lost in speech. For example, Landy and Goldstone (2007) found that when people copy equations, the distance between elements of their handwritten expressions reflect their conceptual understanding. Specifically, students who understand the calculations tend to group pairs of operations that must be done earlier in the equation more closely in space than pairs that are done later in the calculation. Similarly, Cheng and Rojas-Anaya (2007) asked people at different levels of expertise to copy a set of equations using a computer stylus and tablet. The number of long pauses that participants took while writing each equation was used as an index of the size of the chunks in memory (the fewer long pauses, the larger the presumed size of the chunks). As predicted, participants with greater expertise took fewer long pauses, suggesting that they analyzed the equations into larger chunks.

These findings offer new possibilities for the assessment of learning and knowledge in spatial domains. However, the temporal and spatial qualities of students' drawings have been difficult to study because of technological limitations. One method is to score videotapes of participants sketching, but this of course is highly time-intensive. Alternatively, one can manipulate the sketching task so that the ordering of the parts of the sketch to be later analyzed. This method is used with the Rey-Osterreith Complex Figure test (a widely used assessment of neurological disorder), in which the participant uses a different colored pencil for different parts of their sketch. Yet, this would clearly be prohibitive in a classroom, and does not record finer-grained ordering.

This paper describes a new program for collecting and scoring sketches called *CogSketch*, and show how it can be used to assess student's knowledge. CogSketch accepts and spatially analyzes people's drawings. It takes advantage of recent advances in computer technology that make it possible to record ink strokes on a computer screen, and to record their spatial and temporal properties with far less effort than previous methods.

Our goal in this initial study was to determine whether participants' sketches would provide information diagnostic of expertise, and, if so, to characterize this information. A further goal was to determine whether CogSketch could help automate the process of obtaining and scoring student sketches. To do this, we asked geoscience students and novices (psychology students) to sketch using a Tablet PC running CogSketch software. The materials were of two kinds: geological formations and causal diagrams. The specific task was also varied, from tracing to copying to copying from memory.

CogSketch as a research tool

CogSketch (Forbus et al., 2008) is a program that constructs representations from human-drawn sketches and other line drawings. CogSketch is an open-domain, general-purpose sketch understanding system. Most sketch understanding systems focus on recognizing the entity that the user is sketching. CogSketch is based on the insight that when humans communicate through sketches, they describe what they are sketching verbally, rather than relying on visual recognition. Similarly, CogSketch allows users to tell it what they are sketching via a specialized interface. Thus it focuses on capturing and interpreting the spatial relations among (and within) entities, including perceptual and spatial organization.

Every sketch drawn in CogSketch is made up of *glyphs*. A glyph is a shape or object that has been drawn by the user. Each glyph is represented internally as *ink*—essentially the lines drawn by the user when creating the glyph—and *content*—a conceptual entity representing what the glyph depicts. The user presses a button to indicate that they are about to start drawing a glyph, then they draw as many ink strokes as they want, and end the glyph by pressing the button again. As a separate step, users can describe what they intend a glyph to mean by labeling it with either a concept from its knowledge base or with a word or phrase.

Importantly, CogSketch records a time stamp along with every ink point, and exports it in a format easily digested by analysis tools. This allows experimenters to examine the time sequence of events in sketching in great detail, without having to rely on video data capture and scoring. Because the user has manually segmented their ink into glyphs, there is no ambiguity about which parts belong to which depicted entity. Because users have used the interface to label their glyphs, there is no need to analyze video and audio transcripts after the fact to ascertain what labels should be used for the parts of a sketch. Thus the interface design greatly facilitates data gathering.

Examining geoscience knowledge through sketches

In this research, we explored differences in the structure of people's geoscience knowledge by analyzing their sketching. Participants with differing levels of experience in geoscience produced sketches of geologic structures and processes on a tablet computer running CogSketch software. Because CogSketch requires people to label their sketches, the finished sketches provide a rich source of data about their spatial and causal models. Further, because CogSketch records the order in which elements are drawn, the data also speak to the *process* through which the sketches were constructed.

The specific aims of this study were (a) to determine whether automatic scoring of sketches could be used to detect differences in geoscience knowledge; (b) to explore differences in the structure of geoscience knowledge between geoscience students and novices by analyzing their sketches; (c) to explore differences in the kind of task – tracing, copying, and reproducing from memory; and, most speculatively, (d) to determine whether differences in dynamic temporal construction of sketches are diagnostic of expertise in geoscience.

There were two kinds of images: causal diagrams and photographs of geological formations. These involved different specific questions and predictions. With the causal diagrams, the questions were whether geoscience students (relative to novices) include more causal information, whether they focus more on depicting relational information and less on depicting the objects present, and whether they begin their sketches with this causal/relational information[1]. For the geologic formations, the questions were whether geoscience students included more geoscience-relevant structures, and if they represent these structures differently.

Participants were asked to sketch geoscience-related images under various task conditions: tracing, copying, and drawing from memory. In addition, to distinguish whether the observed differences were related to differences in spatial reasoning vs. drawing ability, participants also sketched control images. If geoscience students use spatial and causal models to sketch geoscience images, they should differ from novices in their geoscience sketches but not on the control sketches.

Method

Participants

20 participants from Northwestern University took part: 10 geoscience students (6 graduate students and 4 undergrads) and 10 undergraduates with no prior course experience in geoscience. All were run individually.

Materials

A set of 12 images were used (Table 1). Nine of these were related to basic geoscience content; three depicted a causal process, and six were photos of geological formations. The remaining three images were controls unrelated to geoscience, but comparable to the geoscience images in that they contained parts arranged in intricate patterns. One was a causal diagram (pressure cooker), and the others were photos of moderately complex, layered objects (fruit cross-section and lasagna slice).

Table 1. Images used in the present study.

Image	Description
1) Carbon cycle	**The cycling of carbon through the atmosphere, hydrosphere, and geosphere**
2) Subduction	**A slab of ocean crust subducts beneath a slab of continental crust**
3) Mountain formation	**Volcanism associated with subduction builds a mountain belt**
4) Fault	A fracture showing evidence of movement
5) Faulted fold	A stack of rock layers that has been deformed plastically and later faulted (brittle failure)
6) Rock strata	Distinct layers of rock stacked vertically
7) Multiple folds	Folds in several locations
8) Syncline	A concave downward fold
9) Unconformity	A surface separating strata of different ages,
10) Pressure cooker	**A sealed vessel creating high pressure**
11) Fruit section	Cross-section of glade mallow fruit
12) Lasagna	Italian dish with alternating layers

Note. Images 1-9 are geoscience-related, and images 10-12 are control; **bold font** = diagram; regular font = photo.

Tasks

Each participant completed three sketching tasks. In each task, the participant was given up to 3 minutes to complete their sketch. In the tracing task, an image was displayed on the computer screen, and the participant traced it. In the copying task, a printout of an image was presented next to the computer screen, and the participant drew the image on

[1] A pilot study by Louis Gomez (personal communication) found that biology students copying a diagram would start sketching at the beginning of a process, while novices would start with visually salient parts of the diagram.

the screen. In the memory task, a printout of an image was presented for 30 seconds and then removed. The participant then drew the image from memory on the computer screen.

Procedure

All sketches were made on a tablet laptop computer running CogSketch software. The participant used a stylus to draw on the screen. The participant was first given a brief tutorial and practice with using the software. The participant was then given an overview of the tasks and procedure. For each of these tasks—tracing, copying, and drawing from memory—they were asked to try to draw the image to the best of their ability, to use the glyph function to divide the sketch into the parts that seem natural to them, and to label the parts as they saw fit. They were given up to 3 minutes per sketch, with a warning 1 minute before the time limit.

All participants completed the tasks and sketched the images in the same order: tracing task, copying task, and then memory task. The images were assigned such that each task included 1 causal diagram, 2 geoscience-related photos, and 1 control image. The images for the tracing task were 1, 4, 5, and 10; for the copying task, 2, 6, 11, and 7; and for the memory task, 8, 3, 12, and 9.

Results

Sketches of causal diagrams

For the causal diagrams (Images 1, 2, 3, and 10) participants' sketches were analyzed in terms of the number of relations and objects represented: specifically, the overall proportion of relations in the sketches, and the number of relations in the first two parts sketched (an indication of whether the initial focus was relational). Figure 3 displays a representative sample of geoscience-related sketches from experienced and inexperienced participants. Each part/glyph in each sketch was coded as either an *object* –a part that depicted some tangible entity – or a *relation* – a part that depicted the behavior or connectivity of either a tangible or intangible entity, e.g., a force. The information in both the glyphs and their labels was used for this classification. A part was scored as a relation if either the glyph itself contained relational information, such as an arrow indicating movement, or its label described the part in terms of its relation to other entities. Thus, even if the participant did not label a glyph in a relational way, it would still be coded as a relation if the glyph contained at least one symbol that related it to other parts of the sketch. Here and throughout the study, scoring was done with all identifying information removed from each sketch.

The results for the causal diagrams are shown in Table 2. As predicted, the geoscience students' sketches had a higher proportion of relations (M = .42, SD = .32) than the novices' (M = .08, SD = .11), $F(1, 69) = 48.25$, $MSE = 0.04$, p <.01. This difference was specific to the geoscience images; no such difference appeared for the control image, as confirmed by an Experience Level x Image interaction, $F(3, 69) = 5.19$, $MSE = 0.04$, p <.01.

Table 2. Relational Content of Sketches of Causal Diagrams

Image	Mean proportion of relations		Mean relations in 1st two parts	
	Geo	Novice	Geo	Novice
	Geoscience-related			
1) Carbon	.63 (.41)[a]	.03 (.06)	.60 (.46)[a]	.10 (.21)
2) Subduction	.53 (.30)[a]	.11 (.17)	.44 (.46)[a]	.00 (.00)
3) Mountain	.27 (.19)[b]	.09 (.10)	.30 (.42)	.13 (.23)
	Control			
10) Cooker	.21 (.13)	.08 (.08)	.00 (.00)	.00 (.00)

Note. Geo = Geoscience students; *SD*s in parentheses; a = p <.01; b = p <.05.

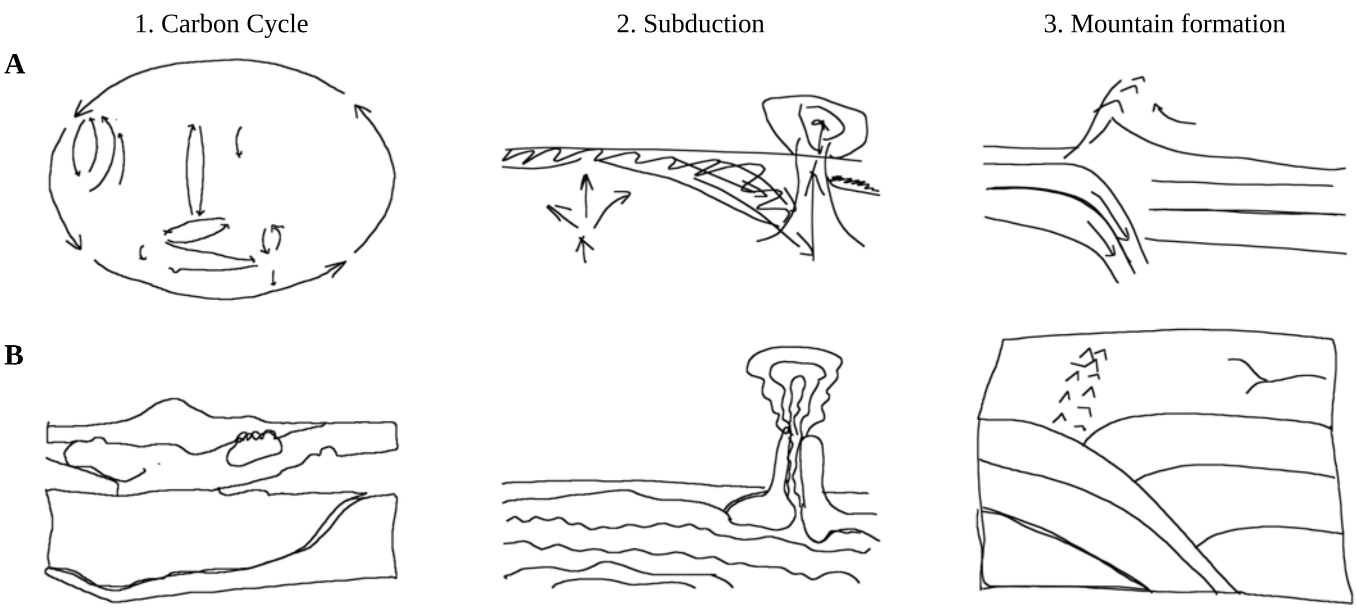

Figure 3. Sample of causal diagram sketches by geoscience students (row A) and novices (row B)

A second set of analyses examined the mean number of relations in the first two parts that participants sketched. Again, the geoscience students drew a higher proportion of relations (M = .22, SD = .06) than the novices (M = .04, SD = .06), $F(1, 71) = 5.23$, $MSE = 0.09$, $p < .05$. This difference was specific to the geoscience images, as confirmed by an Experience Level x Image interaction, $F(3, 71) = 5.23$, $MSE = 0.09$, $p < .05$. For two of the three geoscience images, the geoscience students drew a significantly higher proportion of relations, Fs > 10.17, ps <.01, but for the third (image 3) and the control image, there was no significant difference Fs < 1.47, *ns*.

Overall, participants with geoscience experience had a higher proportion of relations in their sketches of geoscience-related images. This is clear from the sample sketches in Figure 3. It is worth noting that these effects were most striking for the two images (Carbon cycle and Subduction) for which the diagram was present during sketching. For the Carbon cycle, the diagram was actually traced. Thus, these effects are not due to differences in memory for the images.

Sketches of geological formations

For the photographs of geologic formations (Images 4-9), and control photos (11 and 12), we analyzed the sketches in terms of the number of key structures depicted. Figure 4 displays a representative sample of sketches from experienced and inexperienced participants. A key geologic structure was defined as a visible feature that reflected a geologic event (e.g., the folding of rock). These key structures were identified a priori by one of the authors of the present paper (B. Sageman), who listed 2-4 key structures for each geoscience photo. Key structures for the control images were assessed by an independent rater.

The results for the photos of geologic formations are shown in Table 3. Geoscience students drew a higher proportion of key structures (M = .83, SD = .26) than the novices (M = .53, SD = .33), $F(1, 147) = 63.68$, $MSE = 0.06$, $p < .01$. This difference was specific to the geoscience images; no such difference appeared for the control images, as confirmed by an Experience Level x Image interaction, $F(7, 147) = 2.55$, $MSE = 0.06$, $p < .05$.

Table 3. Key Structures in Participants' Sketches

Image	Geo	Novice
	Geoscience-related	
4) Fault	.73 (.39)	.21 (.22)
5) Faulted fold	.56 (.12)	.30 (.20)
6) Rock strata	.90 (.21)	.68 (.34)
7) Multiple folds	.67 (.38)	.30 (.23)
8) Syncline	1.0 (.00)	.59 (.20)
9) Unconformity	.92 (.15)	.45 (.27)
Total	**.80 (.29)**	**.43 (.29)**
	Control	
11) Fruit	.86 (.17)	.77 (.24)
12) Lasagna	.97 (.13)	.91 (.08)
Total	**.92 (.15)**	**.84 (.23)**

Note. Geo = Geoscience students; *SD*s in parentheses.

Overall, the analyses suggest that geoscience experience affected whether participants' sketches of photos included key geological structures. Geoscience students were more likely to draw these structures. This is clearly seen in Figure 4. No such difference was found for the control images, which were sketched similarly by geoscience students and novices (arguing against drawing ability as a factor in the results).

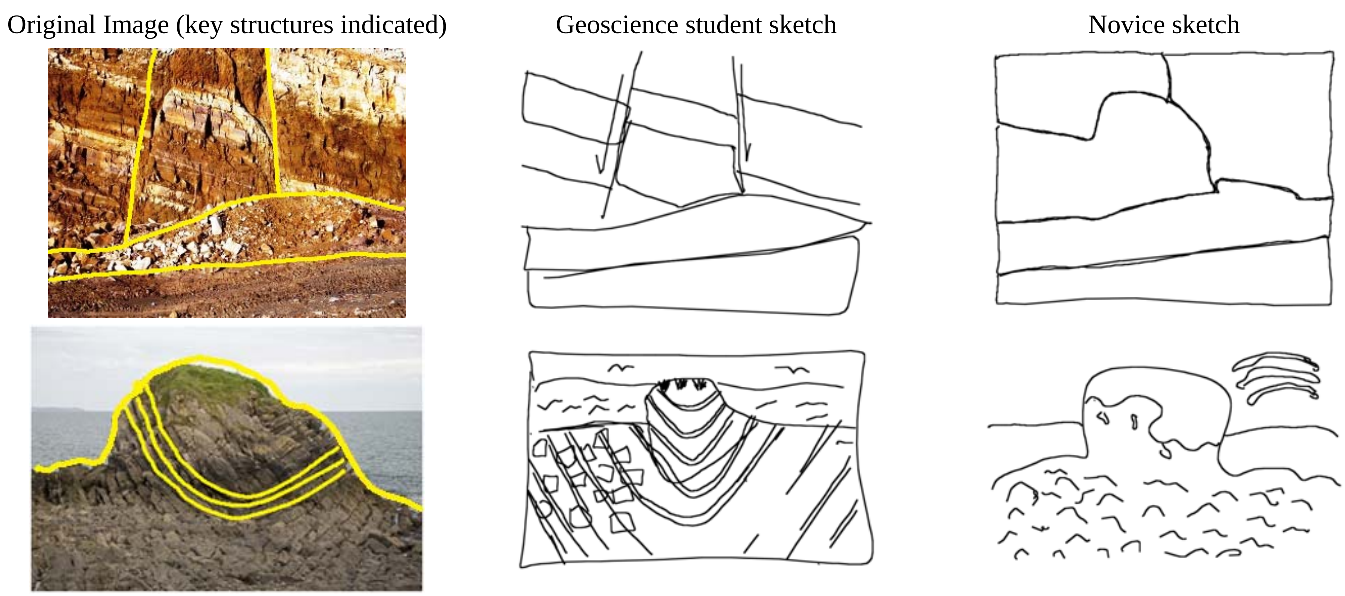

Figure 4. Sample of geological formation sketches by geoscience students and novices participants

Discussion

Our main goals were to see if sketches would reveal expertise-related differences in spatial knowledge within geoscience, and if automatic recording of online sketching by CogSketch could be used to detect differences in geoscience knowledge. The answer to both questions is yes. When sketching causal diagrams, geoscience students showed greater focus on causal relations and processes in their sketches, both overall and in their initial steps. For geological formations, geoscience students included a higher proportion of key geologic structures. No such differences appeared with non-geoscience images. Overall, this application of CogSketch was successful in distinguishing participants with different levels of domain knowledge.

In addition to validating CogSketch as a research instrument, and a potentially useful diagnostic instrument for expertise, the present study revealed interesting differences in the structure of geoscience knowledge. Experienced participants tended to include more relational information, such as arrows and relational labels, in their sketches of causal diagrams. This difference probably reflects participants' mental models of the systems (Gentner & Stevens, 1983). Heiser and Tversky (2002) found that people were more likely to produce sketches with arrows when they were provided with a functional description of a system compared to when they were presented with an object-level description. The present findings suggest that geoscience students represented the pictured systems functionally, based on their spatial/causal mental models, whereas novices, lacking such knowledge, represent them pictorially. Indeed, the sample sketches in Figure 3 support such an explanation.

We also found, as predicted, that geoscience students included more key structures in their sketches of geological formations. Surprisingly, this difference showed up strongly even in the tracing task, where we might have expected no difference. Again, this result suggests that domain-specific differences in spatial knowledge influenced how participants analyzed and represented the photos that they sketched. However, because in this study the specific images were confounded with the tasks, further research will be necessary to decouple the image and task variables. Future research will also explore using CogSketch to automatically score for the key structures, by carrying out an analogy between an expert's sketch and a participant's or student's sketch.

In conclusion, using the CogSketch platform and a set of varied sketching tasks appears to be a rich source of evidence for diagnosing differing levels of domain knowledge. The approach of the current work could be applied broadly to other domains that involve spatial learning, to educational settings, and to tests that use sketching as a diagnostic of knowledge or mental impairment. The analysis of people's sketching behavior promises to reveal much about the course of learning in spatial domains.

Acknowledgments

This research was supported by NSF grant SBE-0541957, the Spatial Intelligence and Learning Center (SILC). The authors thank Thomas Shipley, Cathy Manduca, Basil Tikoff, and Carol Ormand for their contributions to this project, Nathaniel Goldin-Meadow for assistance with data collection, and Jeffrey Usher for technical assistance with CogSketch.

References

Bransford, J. D., Brown, A., & Cocking, R. (Eds.). (2000). *How people learn: Mind, brain, experience and school.* Washington, DC: National Academy Press.

Cheng, P. C.-H., & Rojas-Anaya, H. (2008). Measuring mathematic formula writing competence: An application of graphical protocol analysis. *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society* (pp. 869-874). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H., Glaser, R., & Farr, M. (Eds.). (1988). *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Forbus, K., Usher, J., Lovett, A., Lockwood, K., & Wetzel, J. (2008). CogSketch: Open-domain sketch understanding for cognitive science research and for education. *Proceedings of the Fifth Eurographics Workshop on Sketch-Based Interfaces and Modeling.* Annecy, France.

Gentner, D., & Ratterman, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), Perspectives on thought and language: Interrelations in development (pp. 257-277). London: Cambridge University Press.

Gentner, D., & Stevens, A. L. (Eds.). (1983). Mental models. Hillsdale, NJ: Erlbaum.

Goel, V. (1995). *Sketches of thought.* MIT Press, Cambridge: MA.

Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science, 30,* 581-592.

Jee, B. D., & Wiley, J. (2007). How goals affect the organization and use of domain knowledge. *Memory & Cognition, 35,* 837-851.

Landy, D., & Goldstone, R. L. (2007). Formal notations are diagrams: Evidence from a production task. *Memory & Cognition, 35,* 2033-2040.

Marshak, S. (2005). *Earth: portrait of a planet.* W. W. Norton & Company, New York: NY.

Suwa, M., & Tversky, B. (1997). What architects and students perceive in their sketches: A protocol analysis, *Design Studies, 18,* 385-403.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24,* 535-585.