HOW CAN VISUAL INTERACTIONS SUPPORT DEEP LEARNING IN GEOMETRY?

by

Sarah Davies

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Educational Psychology

The University of Utah

August 2012

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Sarah Davies**

has been approved by the following supervisory committee members:

Kirsten R. Butcher	, Chair	**5/3/12** Date Approved
Robert Zheng	, Member	**5/3/12** Date Approved
Michael K. Gardner	, Member	**5/3/12** Date Approved

and by **Elaine Clark** , Chair of

the Department of **Educational Psychology**

and by Charles A. Wight, Dean of The Graduate School.

ABSTRACT

Self-explanation is a robust learning strategy, but automatic, scalable methods are needed to make it a practical strategy for large-scale implementation in classrooms. This study explored the effects of using visual interactions to engage students in self-explaining while they learned geometry using a computer-based intelligent tutoring system (ITS). The current study compared students who were asked to highlight diagram elements relevant to geometry principles during problem-solving against students who were not asked to highlight diagram elements. Verbal protocols generated during use of the ITS, as well as pre- and posttests targeting retention and transfer, were used to assess learning. Results showed that while the number of overall utterances did not differ across conditions, students who highlighted diagram elements produced a higher proportion of deep self-explanations that connected domain principles to problem diagrams and a lower proportion of shallow utterances that simply paraphrased diagram information (i.e., reading angles from the geometry diagrams). Shallow diagram utterances were negatively correlated with learning but deep diagram explanations were not correlated to learning. Thus, additional interactive elements may be needed to support successful self-explanation using visual interactions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my deep and sincere gratitude to:

- My committee chair, Dr. Kirsten R. Butcher, for her continual guidance, encouragement, and enduring patience. I could not have done this without her help!
- My committee members and program professors for their insight and support.
- My fiancé, David Waters, for cheering me on from start to finish!

CHAPTER 1

INTRODUCTION

Supporting Deep Learning

One of the continual challenges in instruction is facilitating deep learning. A well-known model of comprehension, Construction-Integration (CI), categorizes knowledge as occurring at three different levels: the surface level, the textbase, and the situation model (Kintsch, 1994). The surface and textbase representations each refer to levels of knowledge that can be encoded directly from learning materials and retained by students, whereas the situation model refers to a deeper level of understanding that is formed through the integration of new with prior knowledge.

The knowledge representation that is formed during learning determines the potential depth and breadth of its application. When students form a surface representation of the to-be-learned content, students can recall specific details such as the exact phrasing of a text. Students rarely try to form this exact representation, although it can be useful in cases where specific words and word order are central to the learning task (e.g., memorizing a poem). When a typical student tries to memorize the content of learning materials, the textbase representation usually is formed. The textbase representation contains the basic propositions drawn from a set of learning materials but does not go beyond the encountered information. Students remember concepts but may not recall the exact words or sentences used to explain them. Even when students successfully retain information by forming a textbase representation, they often fail to apply it correctly in new situations. Doing so requires creating a situation model, which is formed by the integration of incoming, to-be-learned content with existing background knowledge. The resulting situation model is a flexible representation that allows the learner to transfer and apply knowledge to new contexts.

Assessing Cognitive Processes During Learning

Ongoing research seeks to investigate how to support the construction of deep, transferrable understanding. Learning assessments, ranging from multiple choice questions to open-ended essays, are used to explore student knowledge (Messick, 1994). Performance on such assessments can provide information about students' grasp of content as well as the cognitive complexity of their problem-solving processes (Linn, Baker, & Dunbar, 1991). However, outcome assessments address the current state of knowledge, revealing little about the comprehension processes in which students have engaged during a specific learning task.

Learning processes are something that must be investigated as they happen. Computer interfaces provide one solution, since they can easily record and compare how individuals interact with a system. The step-by-step log data generated by such systems can provide a detailed record of students' efforts related to learning (Baker, Corbett, & Koedinger, 2004). For example, log data may include the amount of time spent on a task or the order in which a user completes subgoals necessary to solve a problem. Both variables might contribute to our understanding of a learner's ultimate mastery and knowledge.

However, even similar interactions with an interface do not guarantee similar cognitive processes. Some students who demonstrate mastery of a rule in simple situations can successfully apply it to more complex contexts, while others cannot (Corbett & Anderson, 1995). For example, consider a problem in which students are asked to solve the measure of an angle PLA (see Figure 1). One student might understand that the measure of PLA is 60 degrees because the corresponding angles theorem applies: angle PLA is formed by the intersection of the transversal PLS with LA, the segment parallel to SY. The resulting angles, PLA and LSY, are corresponding angles and have congruent measures. Another student might take the same amount of time to enter the same answer of 60 degrees, but do so with the shallow reasoning that the measure of PLA is probably the same as another measure provided in the diagram. Or, the student may reason that because PLA and LSY look similar, they are probably equal. Lacking a deep understanding of the corresponding angles theorem, the student in the latter example will be less likely to correctly apply the corresponding angles rule in more complex situations.

Understanding the processes that occur during learning requires a methodology that can make cognitive processing visible. Verbal protocols, in which students describe their active thought processes (Ericsson & Simon, 1993), can provide rich insight into the learning differences that produce different outcomes. Cognitive processes associated with deep learning can be assessed by analyzing the content of student utterances produced during a verbal protocol. For example, researchers have examined utterances that exhibit the integration of new and prior knowledge, the generation of inferences, and the development of predictions (Butcher & Kintsch, in press). These can be contrasted with cognitive processes that are associated with shallow learning, such as paraphrasing.

Verbal protocols can be produced as the result of a variety of experimental methods; the most common methodologies used to gather verbal protocols are think alouds and self-explanation (McNamara & Magliano, 2009). During a think aloud, students report their spontaneous thoughts to an experimenter as they study and work with learning materials. Think alouds are intended to provide insight into students' thoughts without interfering in their thought processes (Ericsson & Simon, 1993). Thus, the verbal protocols produced from a think aloud are intended to reflect individuals' naturally occurring thoughts and processing.

The self-explanation verbal protocol is distinct from the think aloud protocol in that verbalization typically is trained and prompted (rather than spontaneous), and the emphasis of learners' utterances is focused on explanations of the to-be-learned content (Chi, de Leeuw, Chiu, & LaVancher, 1994). As the name "self-explanation" suggests, the purpose of the self-explanation is for learners to explain the meaning, importance, and impact of the materials to themselves as they engage in a learning task. Unlike a think aloud protocol, a self-explanation protocol is acknowledged to change the typical processes of a learner: as discussed below, students who self-explain during learning tend to learn more than students who fail to engage in these explanations (Chi, Bassok, Lewis, Reimann, & Glaser, 1989).

Supporting Deep Learning Through Self-explanation

Self-explanation originally was studied in the context of individual differences during problem solving. To identify differences between good and poor problem solvers, Chi et al. (1989) invited students to spontaneously explain aloud to themselves what they were learning as they studied. Students varied in how often and how much they self-explained; that is, whereas all students were asked to articulate their thinking as they worked through the problems, not all students successfully engaged in explanation that attempted to reason through problem content and domain principles. The variation was significantly related to learning success: students who verbalized more explanations to themselves performed better in assessments of near and far transfer. In other words, students who spontaneously self-explained the content of a set of learning materials to themselves as they worked learned more deeply. The benefit of these self-generated explanations was termed the self-explanation effect (Chi et al., 1989).

Because the self-explanation effect was established by looking at the spontaneous learning processes of successful students, an open question was whether self-explanation was a by-product of successful learning or if self-explanation itself could be used as a robust learning strategy. Chi et al. (1994) showed that students who were trained and prompted to self-explain were better able to understand, answer complex questions and make inferences about the material that they had studied. These results showed that self-explanation is not simply a passive description of the cognitive processing of successful learners; it can be implemented as a successful strategy to promote the active integration of new knowledge with prior knowledge. Since integration is a hallmark process in the formation of a situation model (Kintsch, 1994), it is clear that the impact of self-explanation on cognitive processes during learning can account for its support in developing deep understanding during learning and problem solving.

Additional research has replicated and elaborated these findings, showing that the kinds of explanations that novice learners generate during self-explanation also affect the quality of the learning outcomes. In a study by Renkl (1997), students were prompted to self-explain as they studied probability calculation. Renkl categorized utterances according to their purpose: to understand the problem (e.g., elaboration of the problem situation and noticing coherence); to apply knowledge of problem-solving strategies to the problem (e.g., principle-based explanations, goal-operator combinations, and anticipative reasoning); and, to monitor understanding (e.g., negative monitoring and positive monitoring). See Table 1 for short descriptions of all seven codes.

In his study, Renkl (1997) found that those who learned less (as measured by pre- to posttest learning gains) focused more on elaboration of the problem situation.

Elaboration of the problem situation and noticing coherence are important to creating a mental model of the task at hand. When students elaborate the problem situation, they interpret given information to better understand the nature of the problem task. Similarly, utterances noticing coherence contribute to a mental model of the problem situation by comparing the problem at hand with other problems encountered previously. However, the weakness of these explanations may be that even a well-developed understanding of the problem is not enough for producing a solution.

Successful problem-solving requires not only understanding the problem situation, but applying knowledge of domain-specific principles to that situation. Accordingly, Renkl (1997) found that those who voiced more principle-based explanations and anticipative reasoning statements achieved better learning outcomes. Principle-based explanations, goal-operator combinations, and anticipative reasoning statements involve applications of different problem-solving strategies to the problem situation. When making principle-based explanations, students refer to a domain-specific rule and elaborate its implications for the task at hand. In geometry, for example, a student might make a principle-based explanation by explaining that the triangle sum rule means that all three angles in a triangle must add to 180 degrees. Goal-operator combinations explain how specific mathematical operations can be applied to achieve a named subgoal of the learning task. Utterances characterized by anticipative reasoning predict solution steps needed to solve a problem. The benefit of producing these explanation types suggests that students are better served by explaining the application of problem-solving strategies than by elaborating the problem itself. This is not surprising, considering that applying problem-solving strategies is likely to require a situation model able to transfer existing knowledge to novel situations.

Taken together, the studies on self-explanation show us that high-quality self-explanations – especially those that integrate to-be-learned material or examples with high-level principles from the target domain – can yield impressive learning outcomes. Unfortunately, while the associated outcomes are desirable, self-explanations are not easily elicited or evaluated in the majority of learning contexts. In order to prompt the student to engage in self-explanation during learning, typical self-explanation experiments have used a one-to-one ratio of human facilitator to participant (Butcher, 2010; Chi et al., 1989; Renkl, 1997). This factor alone presents a scalability problem in traditional classroom settings, in which one instructor is responsible for many students. Furthermore, researchers have assessed the quality of self-explanations through a time- and resource-intensive process of verbal protocol analysis. Conversations must be recorded, transcribed, and coded in order to evaluate the quality and accuracy of learners' self-explanations, and to investigate the relationship between self-explanations and learning outcomes. Since this is not efficient for widespread educational use, there is a need for practical alternatives.

Intelligent Tutoring Systems (ITSs)

How can we reap the learning benefits of self-explanation without requiring individual, human facilitation? Research on Intelligent Tutoring Systems (ITSs) may offer a potential solution. ITSs provide a digitized learning environment designed to support individual student needs. ITSs that have been built on Adaptive Control of Thought – Rational (ACT-R) theory, such as the Geometry Cognitive Tutor, seek to scaffold students in transforming declarative knowledge into procedural knowledge through problem-solving and practice (Anderson, Corbett, Koedinger, & Pelletier, 1995). Many studies have shown that ITSs support students in solving problems (Aleven, Koedinger, Sinclair, & Snyder, 1998; Anderson et al., 1995).

Like all ITSs, the Geometry Cognitive Tutor uses adaptive programming algorithms to compare a student's responses to a model of expert knowledge (Koedinger

& Aleven, 2007). This comparison allows the tutor to estimate a student's current level of knowledge and overall progress; it uses this information to compute an individual student's estimated knowledge level, select targeted problems that address skills or knowledge that it has determined that the student is lacking, and provide detailed feedback as the student works in the tutoring system. Within the Geometry Cognitive Tutor, the learning experience varies by individual but the tutoring system provides all students with a number of common scaffolds as it guides them toward mastery. As shown in Figure 2, the Cognitive Tutor sets up the problem situation, organizes the problem into a series of subgoals, provides students with a problem diagram, and makes on-demand help available (Koedinger & Aleven, 2007). As students solve missing angles, the tutor responds to student input with immediate feedback (e.g., correct or incorrect).

As discussed earlier, similar interactions with a computer interface do not guarantee similar cognitive processes. Thus, despite the demonstrated successes of ITSs, shallow learning and misconceptions are still a concern (Baker, Corbett, & Koedinger, 2004). Students may achieve the correct answer in an ITS while failing to achieve the correct understanding. In geometry for example, conceptually different principles sometimes utilize the same equation (see Table 2). A student may correctly solve an angle measure and still misattribute or misunderstand the underlying geometry rules.

An important step in preventing shallow learning in ITSs is identifying common shallow strategies. As described by Aleven et al. (1998), one shallow strategy sometimes used by geometry students in ITSs is the use of "guessing heuristics." Students guess the measurements of unknown angles based on perceived similarity to other angles or by using other angle measurements provided. A related shallow strategy in ITSs identified by Baker, Corbett, and Koedinger (2004) is termed "gaming the system." In that study, many students proceeded as quickly as possible through the tutor by quickly trying many different answers (e.g., 60, 90, 180) until they succeeded. Additionally, some students systematically made use of on-demand help in order to solve problems: since the final hint in a series basically provides the answer to students in order to allow them to proceed, some students skipped to the final hint without reading the preceding help. These "gaming" strategies reflect the need to embed features in ITSs that help students become more thoughtful and reflective in their work. Overall, shallow strategies and processes are a concern because they make it possible for students to successfully solve problems without building a deep situation model. Thus, researchers who study and develop ITSs seek to develop interactions and algorithms that can reduce or eliminate the use of shallow strategies during intelligent tutoring practice. An example of algorithmic intervention is the work by Baker, Corbett and Koedinger (2004) to detect systematic abuse of help features. However, the current research is focused on the development of interactions in the ITS that combat the use of shallow strategies. Specifically, this work expands on previous research that has attempted to implement efficient forms of self-explanation via interactive elements in ITSs.

Text-based Explanation in Intelligent Tutoring Systems

How can self-explanation be incorporated into a computer-based interface to prevent shallow learning? Instead of eliciting spoken explanations, ITSs can require text-based explanations. Hausmann and Chi (2002) investigated the general effectiveness of typed self-explanations. Their initial results showed that free-form typing inhibited spontaneous self-explanation and increased shallow techniques such as paraphrasing. Hausmann and Chi concluded that the nature of written language seems to delay and obscure the expression of cognitive processes, which is a fundamental component of successful self-explanation. However, a second experiment in the same study found that participant responses could be scaffolded: tutor prompting improved the quality and quantity of participants' typed explanations. The improved output, in turn, produced better learning outcomes. This suggests that appropriate prompts and scaffolds may support explanation-like thinking in students using ITSs.

ITS cues are likely to be most effective when they prompt students to engage in the self-explanation processes that best support learning. As noted earlier, Renkl (1997) found that even orally-produced self-explanations varied in how effective they were for learning, with explanations that connected problem-specific information to high-level domain ideas (e.g., principle-based explanations, goal-operator combinations, and anticipative reasoning) yielding the best outcomes. By extension, computer-based interventions are most likely to improve deep learning when they engage students in interactions that result in explanations or reasoning about the connection between specific problem features and high-level, domain-relevant concepts or principles.

Within ITSs, several functions already have been developed to encourage students to make principle-based "explanations." These functions do not elicit free-form, spontaneous verbal explanation; rather, they prompt students to reason about the meaning and application of appropriate problem-solving principles via interactive elements in the computer environment. Conati and VenLehn (2000) implemented drop down menus that required students to "explain" physics principles during problem-solving. This drop-down explanation consisted of a menu that students used to name the principle that justified their calculation of a numerical answer for the current problem. (Figure 3 shows an example drop-down menu for a geometry problem: after correctly solving the numerical value of an angle, the drop-down menu allows students to select the rule that justified the calculation of that numerical answer.) Although using a drop-down menu to name problem-solving rules or principles may seem to be an impoverished form of self-explanation, requiring students to select rule names helped them develop more successful problem-solving skills.

Similarly, Aleven and Koedinger (2002) compared students who self-explained their problem solving in a geometry tutor by either typing rule names or selecting them from a glossary. This simple, self-explanation condition was compared against students who did not have to justify their answers (and simply solved the geometry problems in the tutor). Aleven and Koedinger found that students who were required to choose rule names gained greater understanding and demonstrated improved transfer. Although selecting a rule name from a drop-down menu or a glossary does not specifically explain the relationship between the problem and the domain concept, the interaction supports students in developing a better understanding of the principles underlying their problem-solving.

Visual Explanation in Intelligent Tutoring Systems

ITS-facilitated “explanations” need not only be text-based. Considering the decreased output of typed explanations and the nongenerative nature of menu selections (e.g., drop-down menus), other student activities may have the potential to support deep learning in a computerized environment. Since ITSs often are used in domains that rely heavily on visual as well as textual information (e.g., mathematics), appropriate interactions with visual elements may support students in learning more deeply. In such domains, the visuals provided and the interactions they support can be especially important for promoting the integration of visual and verbal information for a more complete situation model.

Support for integrating visual information is important because students often struggle to relate multiple representations of information, such as text and pictures (Bodemer, Ploetzner, Bruchmuller, & Hacker, 2005). In geometry, for example, students have been found to rely most strongly on visual features of diagrams (vs. underlying structure or principles) when forming problem-solving memories. Lovett and Anderson (1994) found that when students were given a series of geometry problems, they had the most difficulty when the two problems utilized similar diagrams but the logic of the proofs was different. Students focused on similar elements across the problem diagrams to retrieve irrelevant domain principles, failing to recognize important diagram elements related to the underlying structure of the problem. These results showed that students lacked meaningful integration of domain knowledge and visual problem features. A similar finding was published by Kozma (2003), who found that novices in chemistry tended to fall back on easier-to-perceive surface features of chemical representations to build understanding.

How can we scaffold students to use visual representations in deep and meaningful ways? Recent studies of interactions with visual elements show promise. In ITS research comparing the learning of geometry students who entered answers into an answer box against students who entered answers directly into a geometry diagram, the latter group achieved better transfer (Butcher & Aleven, 2008). Additionally, think-aloud protocols showed that the interactive diagram group verbalized more deep thinking (Butcher, 2010). This suggests that simple interactions with visual information provided a learning benefit by prompting students to actively integrate visual and verbal information.

The promise of visual interactivity has prompted further exploration of how visual elements might support self-explanation-like reasoning. Butcher and Aleven (2009) examined the learning of geometry students who made “visual explanations” by highlighting diagram elements to justify geometry proofs. These visual explanations were generated by students following an error during problem solving. Following an error, students were asked to (correctly) identify a geometry principle that would be used to solve the problem step. Following the identification of this principle, the tutor prompted students to highlight the visual elements that are necessary to apply the principle. For example, alternate interior angles are formed when two parallel lines are intersected by a transversal. The angles on opposite sides of the transversal are “alternate interior angles.” If a student correctly identified this principle as relevant to the current problem step, the tutor would cue the student to highlight (in turn): the parallel lines, the transversal, and each of the alternate interior angles (angle 1 and angle 2). They found that the students who highlighted diagram elements achieved better long-term retention and transfer than students who did not. The prolonged duration and successful application of their learning indicated a better-developed situation model. However, the study did not explore the actual cognitive processes underlying the learning benefit.

We know that students who self-explain are better able to understand, answer complex questions and make inferences about material studied. The learning outcomes from visual explanations, as reported by Butcher and Aleven (2009), are similar to those achieved by verbal self-explanation. If diagram highlighting does yield deep student learning by facilitating self-explanation, it may provide a generative form of self-explanation that is more practical to implement than language-based explanation. This can lead to the development of other computer-based interactions that prompt self-explanation.

Alternatively, the learning benefit of highlighting may be the result of attentional cueing. That is, perhaps the highlighted diagram simply serves to direct student attention to relevant areas of the problem diagram. In that case, student generation of these highlights may not be necessary. Rather, students may be supported if the tutoring system simply provides these highlights in order to guide students’ visual attention during learning.

The current study was conducted to answer three key questions:

1) Does highlighting diagram elements result in improved retention and transfer over time (as shown previously)?
2) Does highlighting diagram elements in order to justify problem-solving steps in geometry lead students to generate more effective self-explanations?
3) How do the types of self-explanations that students generate correspond to learning outcomes, including retention and transfer?

The Current Study

This research seeks to explore how deep learning can be supported via visual explanations that are generated by highlighting diagram elements to justify problem-solving steps in an intelligent tutoring system for geometry. To determine how these prompted visual interactions influence cognitive processes, the experiment compared the learning of those required to highlight diagram elements following problem-solving errors against the learning of those not required to do so. Verbal protocols were used to obtain a rich understanding of students' cognitive processes while learning with a geometry intelligent tutoring system.

Since (as noted earlier) previous research has found a learning benefit of highlighting diagram elements, a secondary goal of this research was to replicate those results. We hypothesized that our students who highlighted diagram elements would demonstrate better retention and transfer over time. Our primary research goal was to better understand the impact of visual self-explanation on students' learning processing during intelligent tutoring practice. We hypothesized that the highlighting interaction would support more frequent generation of deep self-explanations than in the control condition (as evidenced in verbal utterances).

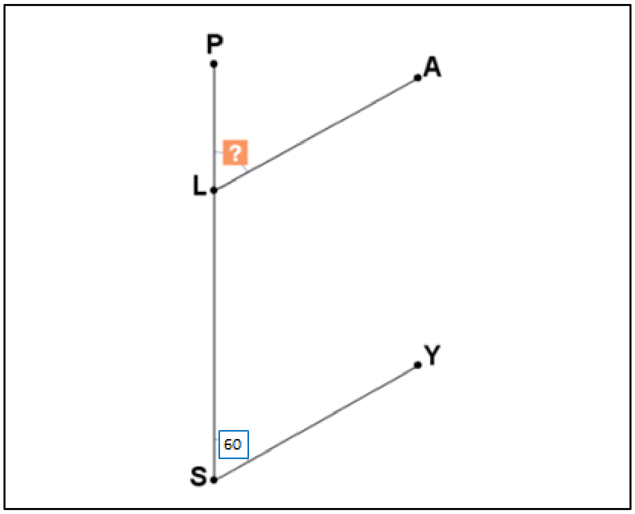


Figure 1. An example geometry problem that could be solved using geometry principles (i.e., corresponding angles) or shallow strategies.

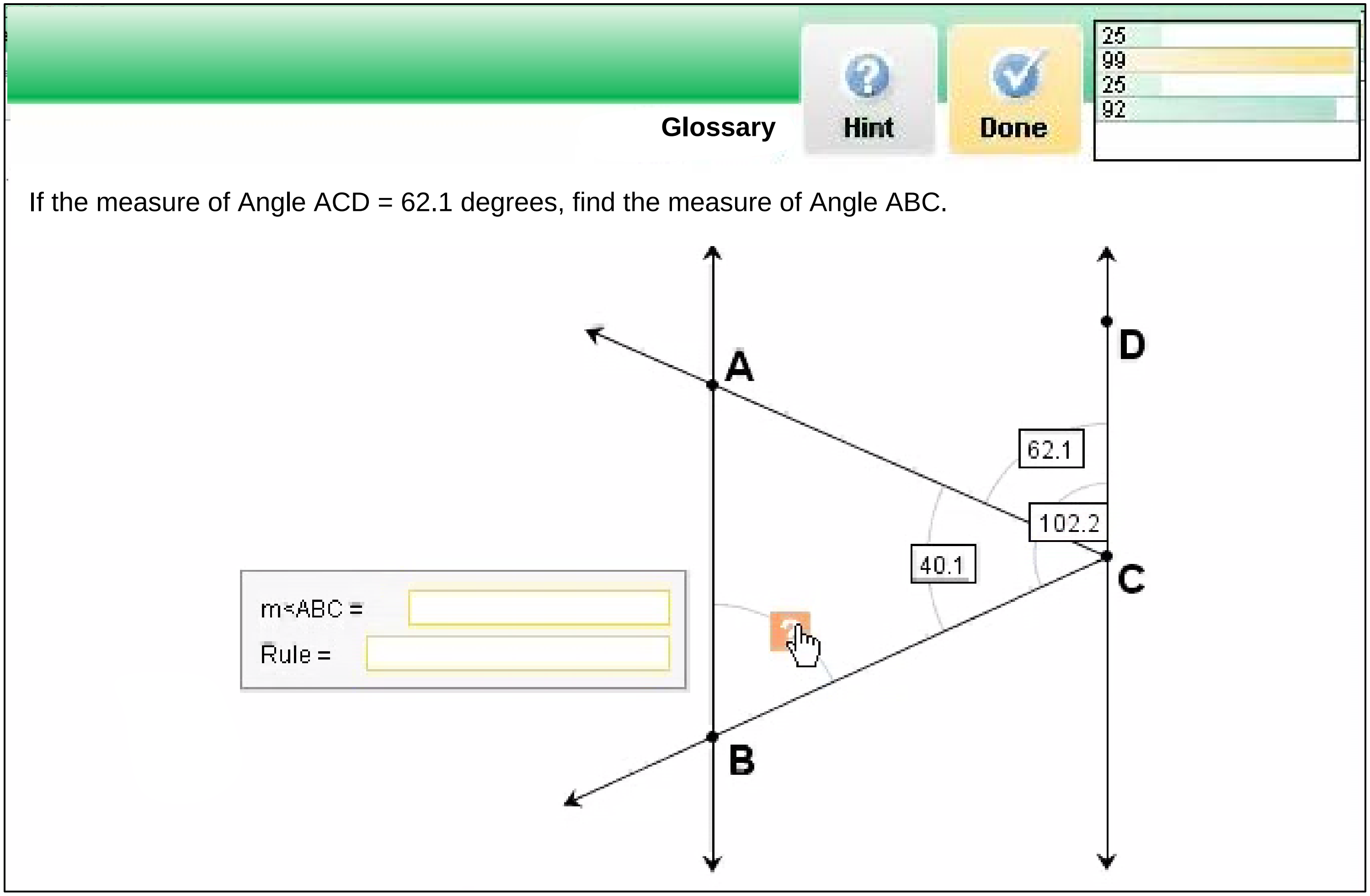


Figure 2. A screen shot of the Geometry Cognitive Tutor interface that includes: given information; an interactive diagram; and, access to a glossary and on-demand hints.

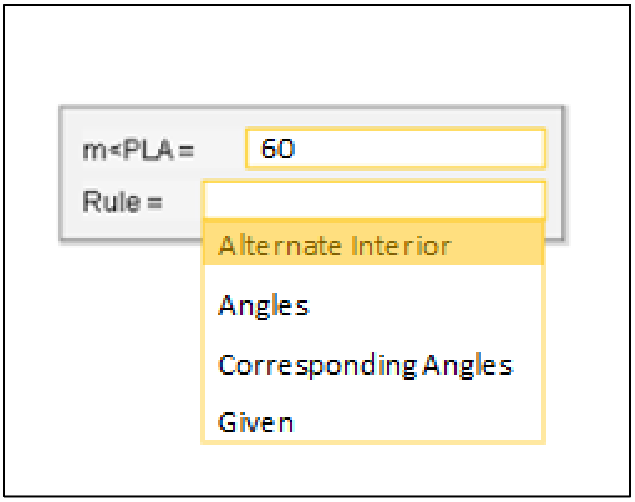


Figure 3. Example of a drop-down menu used to select a geometry rule name.

Table 1

Descriptions and examples of Renkl's (1997) self-explanation categories

Category Name	Description	Example
Principle-based Explanation	Elaborating about a principle used in problem-solving.	"Triangle sum means that all three angles in the triangle must add to 180."
Goal-operator Combination	Mentioning a goal and the operation or strategy by which it could be achieved.	"I can solve the whole angle by adding together the measures of these two adjacent angles."
Anticipative Reasoning	Predicting solutions steps needed to solve a problem.	"The next step should be to subtract the angle measure from the measure of the whole line."
Elaboration of Problem Situation	Using known information to build a mental model of the problem situation.	"The angle I just solved is this one here; it's 45 degrees."
Noticing Coherence	Referring to the similarity between current and previously-solved problems.	"This problem is the same as the barn door problem from before."
Monitoring-negative	Indicating confusion or non-understanding.	"I don't understand what I need to do."
Monitoring-positive	Indicating understanding during problem-solving.	"Oh, I see now."

Table 2

Examples of different geometry rules that utilize the same mathematical equation

Equation	Applicable Geometry Rules
$m\angle A = 180-x$	Interior Angles Same Side Linear Pair Supplementary Angles Triangle Sum
$m\angle A = m\angle B$	Alternate Interior Angles Corresponding Angles Vertical Angles

CHAPTER 2

METHODS

Participants

Participants were recruited via fliers as well as through the Educational Psychology Research Subject Pool. A total of twenty-four undergraduate students were recruited for the study (13 males and 11 females). Seventeen participants were compensated $30 for their participation, while the other 7 participants received credit in an undergraduate educational psychology course. Four participants were dropped from final analyses: two students lacked a basic understanding of mathematical principles (e.g., order of operations) which impeded their progress through the problems presented by the tutor; one student was excluded due to technical problems (the audio failed to record during the learning session); and, one participant did not complete the study.

Design

This study used a two-condition, between-subjects design to assess whether the added requirement of interactive visual explanations could improve learning beyond what is achieved with the Cognitive Tutor alone. All participants were randomly assigned to one of the two conditions.

Materials

Two versions of the Geometry Cognitive Tutor, an intelligent tutoring system, formed the experimental conditions for this study: a visual explanations version and a control version. Pre-, post-, and delayed posttest assessments, as well as a rubric for coding verbal data, also were developed for the study. Each of these materials is described below.

Geometry Cognitive Tutor

The Geometry Cognitive Tutor is an intelligent tutoring system (ITS) that supports student learning-by-doing in geometry. The Geometry Cognitive Tutor assesses students' knowledge development by analyzing their solution attempts on geometry problems and comparing students' individual knowledge to an expert model. The Tutor uses this information to select future problems to help meet identified learning needs.

Each problem presents students with textual information, including given measurements and missing angles to be solved, as well as a diagram that visualizes the problem situation described in the text (see Figure 2). Students have access to a glossary of geometry principles (see Figure 4) and a hint button that provides on-demand help. Students must solve the measures of one or more missing angles in each diagram; angles to be solved are marked with a question mark icon. Along with each numerical answer, students must type (or select) the name of the single geometry principle they used to solve the angle. When the students' answers are correct, the tutor allows them to proceed to the next problem. When the students' answers are incorrect or incomplete, the tutor outlines the incorrect answer in red to notify the student of errors (see Figure 5).

Visual Explanations Condition

In the visual explanations condition, participants clicked directly on the diagram to enter their answers (see Figure 2). When participants made an error (whether on angle measurement or geometry principle name), they could not proceed until they 1) correctly named the required geometry principle and 2) highlighted components of the diagram corresponding to the geometry principle (see Figure 6).

This condition is referred to as the visual explanations condition because students used highlighting to visually “explain” how the elements of the geometry diagram are relevant to the geometry principle being used to solve the problem step. For example, if a student entered the incorrect numerical answer for an angle measure, the tutor would mark the answer as incorrect and prompt the student to identify the name of the rule required to solve the angle. Once the student had identified the correct rule name, the student would be prompted to highlight diagram elements relevant to that rule. In the case of the corresponding angles theorem, the student would need to understand that applying this rule requires the intersection of a transversal with two parallel lines, and that this configuration yields a pair of two corresponding angles. Therefore, the student would be asked to highlight each of the relevant diagram elements: the pair of parallel lines, the transversal, and the two angles that correspond to one another (see Figure 6). Simple diagrams (encountered early during use of the tutor) may contain few additional diagram elements other than those relevant to the problem step, but complex diagrams (encountered during subsequent problems) required students to discriminate between several relevant and irrelevant diagram features when generating these explanations.

Control Condition

In the control condition, participants also clicked directly on the diagram to enter their answers (see Figure 2). They had to provide angle measurements and the names of geometry principles used, but this condition did not involve any highlighting of diagram components. Following an error, participants would retry their answers until accepted by the tutor as correct.

Problem-solving Assessments

Learning assessments were administered on three occasions during the study: at the beginning of Session 1 (pretest), at the end of Session 1 (posttest), and during Session 2, 1 week later (delayed posttest). Below are descriptions of the types of items that made up the pre- and posttests.

Pretest Assessment

The pretest consisted of two types of questions; geometry principle recognition items and problem-solving items.

Recognition Items

The recognition items were intended to assess participants' recognition of several geometry principles. Students were provided with eight simple diagrams, each indicating a pair of angles. Given the choice of several geometry principles, participants were asked to mark the principle that applied to the relationship between angles 1 and 2 as portrayed in each diagram (see Figure 7). If none of the principles applied, they were instructed to mark "None of the Above." Seven of the eight problems had the correct geometry principle available as a listed option; one of the problems merited the selection of "None of the Above." Each item was worth one point, for a total of eight points.

Problem-solving Items

The problem-solving items were intended to capture participants' pre-knowledge of both the skills and the geometry rules to be practiced and applied within the Cognitive Tutor. This assessment was made up of nine diagrams with questions asking for angle measures (numerical answers) and written explanations of how those answers were obtained (see Figure 8). All nine problems were solvable. One point was awarded for each correct numerical answer: one point for mentioning the correct name of the geometry rule used; and, one point for each correct explanation of how the problem was solved (e.g., the mathematical formula). Each problem was worth three points, one for each element of the problem, for a possible total of 27 points.

Posttest Assessment

The posttest consisted of items intended to assess retention of practiced knowledge and various transfer applications.

Practiced Items

Within the Cognitive Tutor, students in all conditions used diagrams and given information to solve for missing angles. Additionally, in all tutored items, they completed problems in which they correctly named the geometry principles used to solve those angles. To evaluate these practiced skills, the posttest assessment included seven items that again required students to solve missing angles and name the geometry principles used (see Figure 9). Each of these items was worth two points, for a total of 14 points.

Transfer Items

When students demonstrate transfer, their understanding evidences a depth and flexibility that can be applied to situations quite different from the original problem-solving. The posttest items intended to measure transfer tested students on their ability to apply and explain geometry principles. Principle-based explanations involving application and explanation evidence a deep, "situation model" level of understanding.

Visual transfer items***.*** To assess students' abilities to recognize varied configurations of geometry principles learned in the tutor, the posttest included visual items. These questions provided visual examples of diagrams and asked participants to identify instances of specific geometry principles. One type of visual transfer item included illustrations that were new and different configurations from those practiced in the tutor. Provided with a rule name, participants were instructed to circle examples of angle pairs fitting the geometry rule. They were also instructed to draw X's over nonexample angle pairs (see Figure 10). Half of the illustrations were examples and half were nonexamples. Each correct circle and each correct X received 1 point. The posttest included 30 illustrations, for a possible total of 30 points.

The second type of visual transfer item also asked participants to identify appropriate applications of geometry principles. Provided with the name of a geometry principle and a diagram with given information, participants were asked whether a specific angle could be solved by applying the named geometry principle (see Figure 11). Seven of the instances were true applications, and three were not. Participants were awarded one point for identifying true and false applications, for a total of 10 points.

Explanation transfer items***.*** To assess students' abilities to generate principle explanations in novel contexts, the posttest included formula explanation items and unsolvable explanation items. Formula explanation items were associated with solvable practiced and visual transfer items, and unsolvable explanation items were associated with unsolvable practiced and visual transfer items.

Solvable practiced and application transfer items required students to justify their answers by describing the formula of the relevant geometry principle (see Figure 9 and Figure 11). For example, if a student had solved the measure of an angle and named alternate interior angles as the rule used, the student would generate additional explanation by providing the formula of the alternate interior angle theorem. Formula generation was not a skill practiced in the tutor, and so could be used to demonstrate a deep understanding of the geometry principle. For each formula explanation associated with a practiced item or visual transfer item, participants received one point for a correct formula explanation, for a total of 14 points.

When the practiced and visual transfer items were unsolvable, participants were asked to suggest which geometry rules and missing elements could be used to find a solution (see Figure 9 and Figure 11). For example, a student might suggest that an angle adjacent to another known angle could be solved using the angle addition theorem with the measure of the whole angle encompassing the two adjacent angles. Students did not encounter any unsolvable problems in the tutor; thus, the ability to propose an accurate problem-solving solution to unsolvable items in the posttest would demonstrate knowledge transfer. For each unsolvable explanation associated with practiced and visual transfer items, participants received one point for naming a correct geometry rule and one point for correctly specifying a needed element, for a total of nine points.

Overall, the transfer items requiring explanations were worth 23 points.

Delayed posttest assessment

The delayed posttest was administered 1 week following the posttest; it contained the same numbers and types of (practiced and transfer) items as the posttest.

All items were isomorphic versions of those found on the posttest. The numerical values, illustrations, and order of items were the only significant differences.

Verbal Analysis

As participants worked with the Geometry Cognitive Tutor, they were asked to self-explain. Verbal utterances were recorded with the Morae Usability Suite and later transcribed by a professional transcription service. Verbal data are useful for distinguishing between self-explanations that are more, or less, effective.

In order to maximize the usability of the verbal data, both the first and last 15 minutes of the learning session were excluded from the analysis. During the first 15 minutes, students are still learning how to navigate the system. During the last 15 minutes, they are more likely subject to fatigue. Thus, the middle 30 minutes of the learning session were considered the most useful sample. The utterances were segmented into complex propositions (Kintsch, 1998); complex propositions are roughly equivalent to an idea unit. They focus on a single idea, concept or process – thus, one sentence may be broken into several segments (see Table 3, for an example of text segmenting). Following segmenting, each segment was assigned a code as described below.

Coding Rubric

The coding rubric was adapted from previous research using verbal protocols (Butcher, 2010; Renkl, 1997). It consisted of 24 codes belonging to seven code families. The code families categorized utterances as relating to self-explanations, diagram explanations, monitoring, planning, problem-solving techniques, reading and paraphrasing, or tutor interactions. These code families are explained in greater detail below. Table 4 provides definitions and examples of all codes that were organized under each code family.

Self-explanation Code Family

To address the question of whether highlighting diagram elements resulted in increased self-explanation, the self-explanation code family included specific codes that distinguished between different types of self-explanations. These specific codes borrowed from previous research (Butcher, 2010; Renkl, 1997). Shallow explanations categorized those utterances making use of guessing heuristics or gaming strategies. Elaboration explanations expanded on prior knowledge to clarify the problem situation. Goal-operator combinations were explanations relating mathematical operations with problem-solving goals. Utterances elaborating the meaning of geometry principles were coded as principle-based explanations. As in Renkl's (1997) study, shallow and elaboration explanations are considered less effective, while goal-operator and principle-based explanations are considered more effective for learning. Table 4 provides example utterances that would be coded in each of these categories.

Diagram Explanation Code Family

Because of the highly visual nature of geometry, additional codes were created to capture self-explanations specific to diagram interactions within the Geometry Cognitive Tutor. At the shallowest level, diagram reading consisted of just that – reading angles or information directly from the diagram (e.g., "OK, angle A - B - C"). At the deepest level, diagram explanations identified the geometrical function of diagram elements and relationships between them (e.g., "That means, Angles PLA and LSY must be the same because they are congruent angles"). Diagram interactions were utterances consisting of more than reading from the diagram but less than a complete diagram explanation – these utterances went beyond simple reading of diagram angles or features, but stopped short of fully explaining relationships or principles involved in the diagram (e.g., “So, these two should be the same”). Diagram reading and diagram interaction utterances are considered shallow or incomplete and thus, less effective; diagram explanations are considered more effective self-explanations for learning. Table 4 provides additional example utterances that would be coded in each of these categories.

Monitoring and Planning Code Families

During problem-solving, students often show different levels of metacognitive awareness. Thus, the broader metacognition family was retained to capture participant utterances indicating monitoring and planning of their own problem-solving. Positive and negative monitoring consisted of statements of comprehension or lack thereof, respectively. Statements recognizing errors made were coded as error monitoring, while statements acknowledging overall progress were coded as progress monitoring. Plans to access the glossary or on-demand help were coded as glossary planning and hint planning, respectively. Utterances about aspects of the task to work on next received the subgoal planning code. While metacognitive utterances do not support learning directly, indirectly they are important for self-regulated learning (Azevedo, 2005). Table 4 provides example utterances that would be coded in each of these categories.

Techniques Code Family

The techniques family was created to capture mentions of problem-solving actions without elaboration or explanation. When mathematical operations were described without being tied into geometry reasoning, they were coded as mathematical procedures. Numerical answer was the code for utterances stating only the measurement of an angle, without elaboration. The principle-naming code was applied to mention of specific geometry principles, also without elaboration. Principle-naming was divided into two due to the high frequency of the “given” rule in geometry problems. Because the codes in this family fail to tie in problem-solving actions with geometry reasoning, they are considered shallow and less effective for learning. Table 4 provides example utterances that would be coded in each of these categories.

Reading and Tutor Code Families

Because students frequently read textual on-screen information aloud, a reading family was created to distinguish between the reading from different sources of information. These sources included the problem statement, the glossary, and hints. A final family of codes captured utterances specific to the tutor interface. The feedback response code included all statements acknowledging tutor feedback about correct and incorrect answers. The narration code was used for participant statements about non-content interactions, such as hitting the enter key. If students commented on the functionality of the tutor interface itself, the tutor code was applied. Reading codes are considered to be more or less effective depending on their application: for example, over-reliance on hints would likely undermine a student’s learning, while prudent references to the glossary would likely support learning. Table 4 provides example utterances that would be coded in each of these categories.

Procedure

This study was run in a computer lab on the University of Utah campus. The lab was equipped with personal computers on which the participants used their randomly-assigned condition of the Geometry Cognitive Tutor. All students were run individually through the research protocol.

Session One

Upon arrival at the study, students first were guided through an informed consent procedure. After informed consent was obtained, participants were given an identification number, which was used to log in to the cognitive tutor and served as a de-identified label for all experimental materials collected during the study.

Next, each participant completed the two pretests, taking up to 10 minutes for each test. Participants were instructed to make their best guess if they were not sure of the answer or to write “I don’t know” if they were unable to attempt a solution. Tests were collected prior to moving on to the next task.

Next, participants were trained in the self-explanation procedure for about 10 minutes. The experimenter modeled a poor and a good example and then the participant practiced self-explaining. As a practice task, both the experimenter and participant used the utah.edu website to find information (e.g., the nearest location of a student parking lot) while thinking aloud. The experimenter provided feedback to the participant about their self-explanation following practice – the purpose of this feedback was to help the participant determine if she/he was explaining their actions (e.g., “I want to find out the parking costs, so I’ll choose …”) rather than narrating their activities (e.g., “I’m clicking on …”).

Following the self-explanation practice, each participant was instructed in how to work with the cognitive tutor interface. Participants were instructed to log in to the Cognitive Tutor using the identification number they had been randomly assigned. The experimenter used a static example problem provided upon logging in to the tutor to explain how to submit answers, get help using the glossary and hints, recognize errors made, navigate to the next problem, and (for the visual explanations condition only) highlight diagram components.

Next, participants spent 60 minutes using the Cognitive Tutor while verbally self-explaining. When participants paused for more than 5-10 seconds, the experimenter would remind them to keep talking aloud. When they neglected to give an explanation for their actions, the experimenter prompted participants by asking a question. For example, if a participant solved an angle without explanation, the experimenter would prompt the student with, “What made you decide to do that?”

Immediately following the tutor task, the experimenter administered the posttest. Thirty minutes were allocated for this test. At the end of the first session, participants were paid $20 and were scheduled to return 1 week later for a follow-up session.

Session Two

During the follow-up session, participants were asked to complete the delayed posttest, again with a time limit of 30 minutes. Five minutes at the end of the follow-up session were set aside for payment and debriefing. Participants received an additional $10 for completing the follow-up visit. Participants were given the opportunity to ask questions about the study; they were also provided with a debriefing form that included information about the study, as well as the principal investigator's contact information for any future questions or to find out the results of the study.

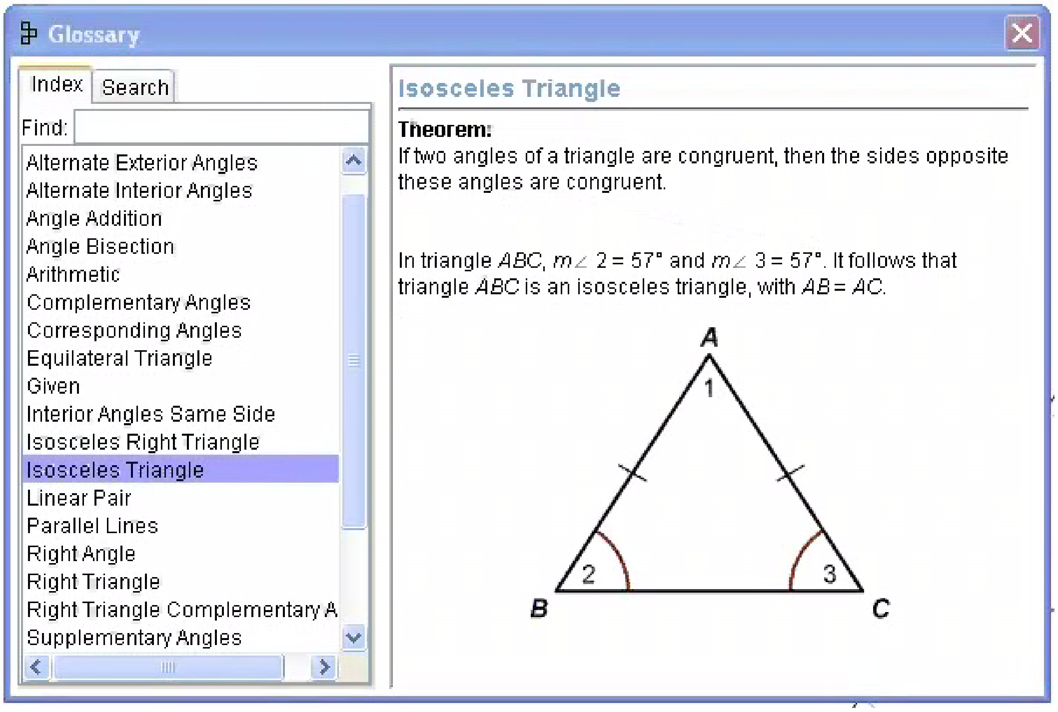


Figure 4. A screenshot of the glossary from the Geometry Cognitive Tutor, showing a list of geometry principle names (left-hand side) as well as a definition and diagram illustrating the selected principle (right-hand side).

Figure 5. The Geometry Cognitive Tutor outlines incorrect answers in red.

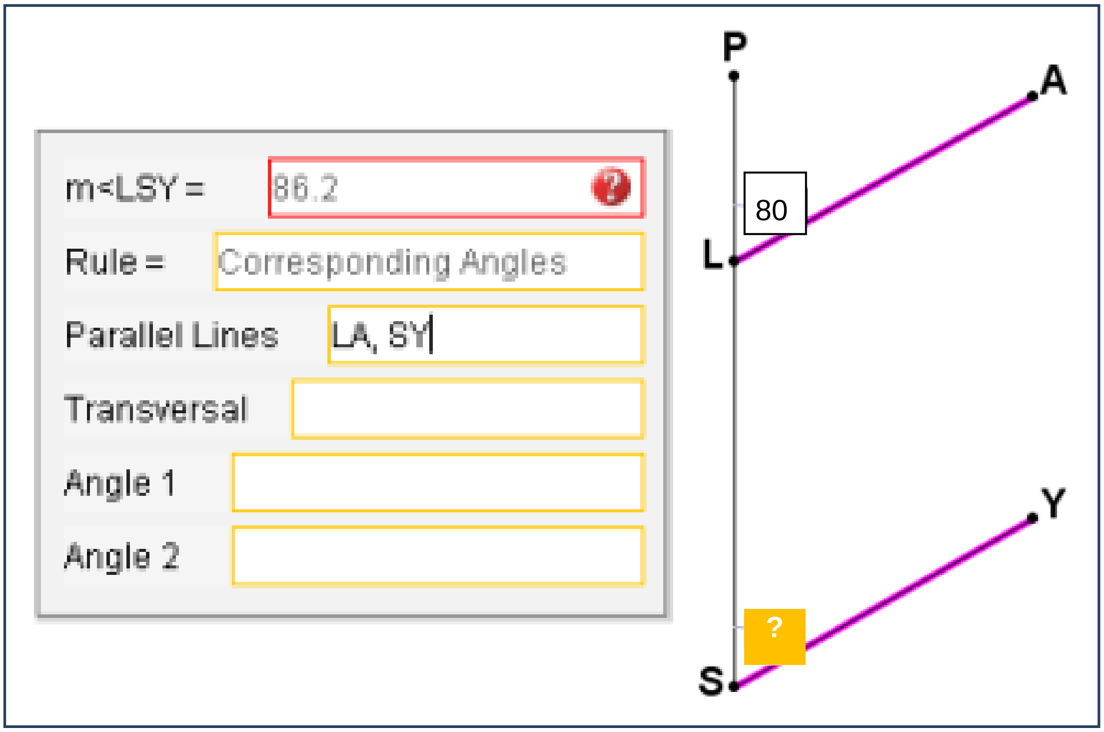


Figure 6. A screenshot from the visual explanation condition of the Geometry Cognitive Tutor. In the visual explanations condition, students who make an error must a) name the geometry rule needed and b) highlight diagram elements relevant to the rule.

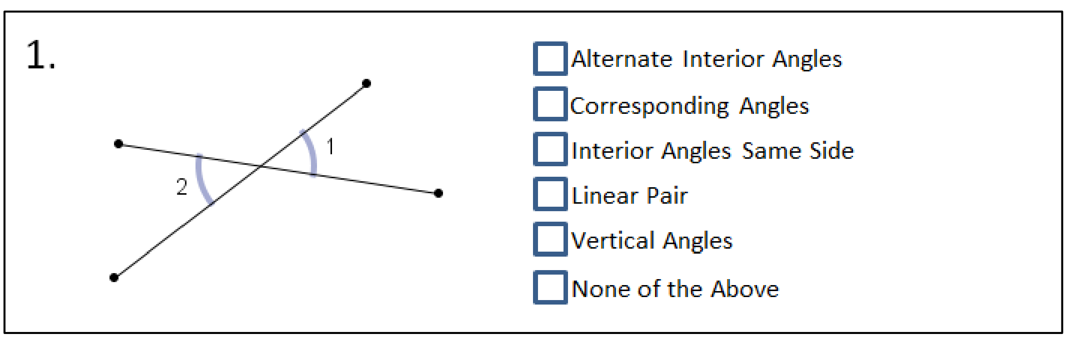
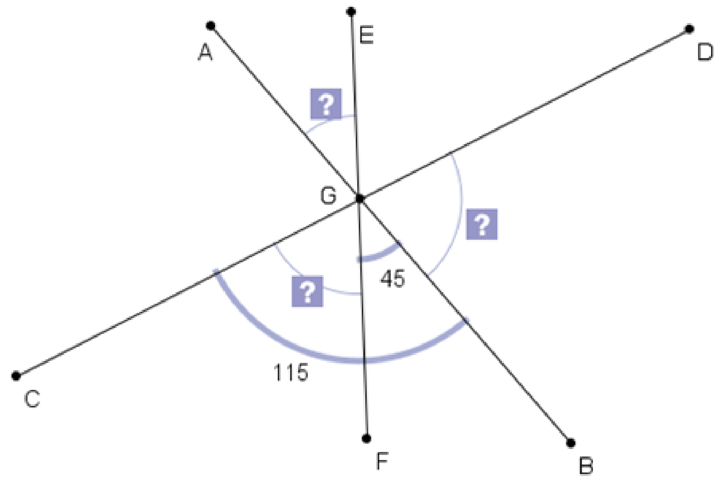


Figure 7. A sample pretest item testing recognition of geometry principles.



Given: Angle CGB is 115°, Angle FGB is 45°

1a. ∠AGE =

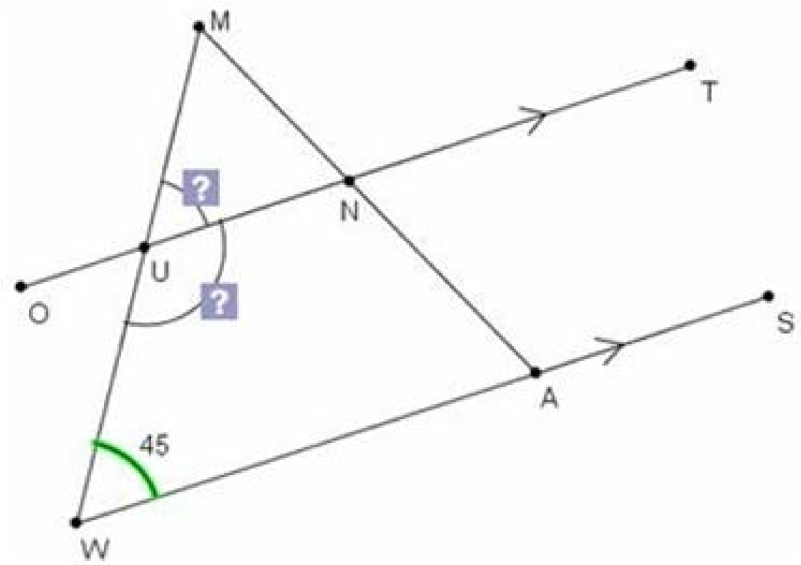
1b. Explain how you got your answer: ______________________________

____

____

____

Figure 8. A sample problem-solving item from the pretest. These items ask students to provide a) angle measures and b) written justification.



Given: Angle UWA = 45°, Segment OUNT is parallel to Segment WAS

1a. Do you have enough information to find the measure of ∠WUN?

_______ YES. The measure of ∠WUN = ______________________________

Rule: ______________________________

Formula Used: ______________________________

_______ NO. You can't with the given information, but you could find ∠WUN using:

Rule: ______________________________

If you also knew: ______________________________

Figure 9. A practiced/transfer posttest item. When solvable, these problems were comprised of a) practiced items, including angle measures and geometry rule names, and b) explanation transfer items (formula used). When unsolvable, these items required only c) explanation transfer items (the geometry rule and missing element that could be used to find a solution).

Recognition: Alternate Interior Angles

Please examine the 10 diagrams seen below. For each diagram, decide if angles 1 and 2 form a pair of ***Alternate Interior Angles***.

- If angles 1 and 2 are alternate interior angles, **circle** the diagram.
- If the angles are NOT alternate interior angles, **draw an "X"** over the diagram.
- If you are **not sure**, draw a **question mark** on the diagram.

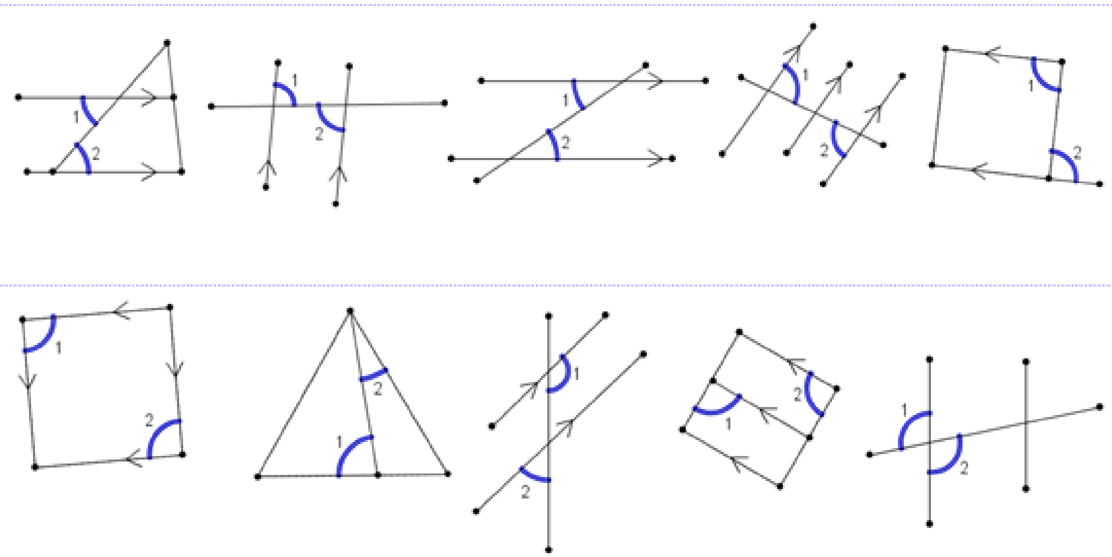
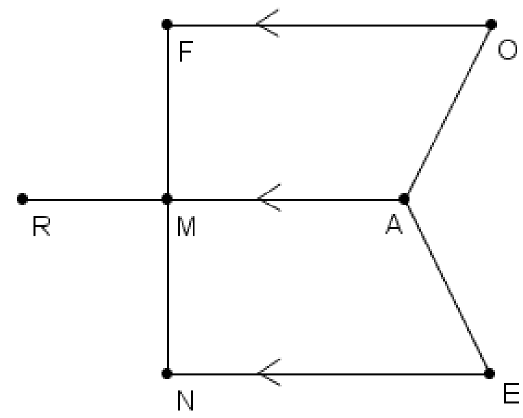


Figure 10. A visual transfer posttest item asking students to circle examples of the specified geometry rule and cross out nonexamples.



Given: Segments FO, RMA, and NE are parallel to each other

2. For **each** of the following statements about the diagram, decide if the statement is true or false (put a chekmark in the appropriate box). If the statement is false, then tell us what information you would need to use that geometry rule.

A) You can use the **corresponding angles** rule to find Angle FMA in one step if you know only the measure of Angle MNE.

- ☐ True: You could find the answer using that rule. Formula: _______________
- ☐ False: To use that rule you'd need to know ____________________________

Figure 11. A transfer posttest item. This required a) visual transfer to identify geometry principle applications as true or false. In the case of a true application, students were asked to generate b) the formula used. In the case of a false application, students were asked to suggest c) the missing element needed to solve the problem using the specified rule.

Table 3

Verbal utterances segmented into complex propositions

Transcribed Utterances	Segmented Utterances
Because I was thinking it’s the same as this one – oh, maybe not, though. So this is 41.8 and this is 63, and DAC, CB are – ABC, are - so this is 104.8, right there, so that means that, that means that B is equal to DCB. And so that means that this angle, CB whatever, is 104.8 because they’re corresponding, I think, and then – so you have 180 minus, 180 minus 104.8.	Because I was thinking it’s the same as this one – //
	oh, maybe not, though. //
	So this is 41.8 and this is 63, //
	and DAC, CB are – ABC, are- //
	so this is 104.8, right there, //
	so that means that, that means that B is equal to DCB.//
	And so that means that this angle, CB whatever, is 104.8 because they’re corresponding, I think, //
	and then – so you have 180 minus, 180 minus 104.8. //

Table 4

Verbal code families, codes, descriptions, and examples

Family	Code	Description	Example
Self-explanation	Shallow	The student explains reasoning using only a shallow basis, such as: 1) using the numbers 180 or 90 without geometry rationale; 2) using the same measure of other known or solved angles; or 3) using the same rules as the previous angles/problems solved.	And I'm just going to guess 41, since BAR is 41.
	Elaboration	The student evidences use of prior knowledge without focusing on specific actions, operators, or geometry terms.	So this angle we know is 66.5. *[indicating previously solved angle]*
	Goal Operator	The student names a (sub)goal and explicitly names an operator which will lead to this (sub)goal.	so I'm going to take 41 degrees minus 35.4 to find the new angle
	Principle-based	The student focuses on one specific geometry principle, what it means, or how it does/does not apply.	So this has to add 180, since it's a triangle.
Diagram Explanation	Diagram Reading	The student reads off information from the diagram, usually as an angle or line or other geometry figure.	XCT-
	Diagram Interaction	The student uses mouse to indicate interaction with diagram or diagram features but does not make a complete explanation.	Uh, that side, that side, and that side.
	Diagram Explanation	The student identifies diagram components as being specific geometry elements or diagram features as having specific relationships.	This line here is the transversal because it connects two parallel lines to create these angles.
Monitoring	Positive Monitoring	The student indicates knowledge or understanding.	Okay, that makes sense.
	Negative Monitoring	The student indicates confusion or a lack of knowledge about the content or about what to do.	So I'm sort of stuck.
	Error Monitoring	The student acknowledges s/he has made a mistake and/or describes the mistake made.	and I was reading it wrong, um, when I tried to put the answer in.
	Progress Monitoring	The student reflects on progress made (e.g., summarizes completed work); includes moving on to the next problem.	So that problem is done.

Table 4 (cont.)

Family	Code	Description	Example
Planning	Glossary Planning	The student voices plans to use the glossary.	so I'm gonna check the glossary.
	Hint Planning	The student plans to seek a hint.	So I need a hint and I just click here.
	Subgoal Planning	The student mentions which aspect of the task s/he is working on/will work on next.	So now to find the measure of angle IRL.
Techniques	Mathematics Procedures	The student talks about the mathematical procedures being used without explaining geometry components or referring to a goal.	so I'm just gonna go ahead and put 180 minus 72.6 minus 72.6.
	Numerical Answer	The student states the numerical measurement only, without explanation, analysis, or elaboration.	Um, so this has to be 35.5.
	Principle-Naming (Given)	The principle "given" is attributed, whether as a single word or in a longer narrative.	It's given.
	Principle-Naming (NOT Given)	One of the geometry principles is being named but no actual self-explanation occurs.	Interior angle same side.
Reading / Paraphrasing	Glossary	The student reads or paraphrases (part of) the clarifications in the Glossary text.	"Corresponding, equilateral, linear pair."
	Hint	The student reads or paraphrases (part of) the hints. Includes hints that pop up to offer help to student.	"You can find the measure of NIE using the linear pair postulate."
	Problem Statement	The student reads or paraphrases the problem or given information presented.	"Find the measure of angle IRL."
Tutor Interaction	Feedback Response	The student acknowledges system feedback.	Yes. And that is right.
	Narration	The student narrates his or her noncontent-related interactions with the tutor.	Now I'll hit enter.
	Tutor	The student comments on the tutor interface, the appearance of the problems on screen (e.g., question mark placement or highlighting), or anything related to superficial aspects of the tutor.	It's waiting.

CHAPTER 3

RESULTS

Analyses

Learning Assessments

Previous research found a learning benefit for students who highlighted diagram elements to justify geometry problem-solving. Thus, we used pre- and posttest assessments to determine if the effect would be replicated. Analyses for the practiced problem-solving component of pre-, post-, and delayed posttests were conducted using a repeated measures multivariate analysis of variance (RM-MANOVA). A mixed design was used where the between-subjects factor was Cognitive Tutor version (visual explanations condition vs. control condition). The within-subjects factor was test time (pretest, posttest, and delayed posttest). Dependent measures included performance on the skills practiced in the tutor: angle measures and geometry principle names. These scores were calculated as percent correct.

Because posttests were administered immediately after the intervention as well as 1 week later, analyses for transfer items were conducted using a repeated measures multivariate analysis of variance (RM-MANOVA). The between-subjects factor was Cognitive Tutor version, and the within-subjects factor was test time (posttest vs. delayed posttest). Dependent measures included performance on visual and explanation transfer items during the posttest and delayed posttest.

Self-explanations

An important research question was how highlighting relevant diagram elements might influence students' processes of self-explanation. To examine potential differences in overall verbal output as well as potential differences in self-explanation patterns, a multivariate analysis of variance (MANOVA) was performed. The independent variable was condition, and the dependent variables included the total number of utterances, the percent of total utterances for each code family, and the percent of family utterances for each individual code. To control for overall verbosity, the numbers of coded utterances in each code families were examined as percent of total utterances. To examine patterns within each code family, individual codes were examined as a percentage of their respective code families. Five codes were excluded from the individual and family code analyses because of low relevance or low incidence. Principle-naming of given rules was excluded from the analysis because it meant that students simply were noting that they had used "given" information to fill in angle measures provided by the tutor at the beginning of each problem. Similarly, problem statement reading was excluded since all students generally read the problem statements to begin each problem and seldom referred back to them. The entire tutor code family was excluded because its codes tended to be very low in quantity across all participants, and tutor-specific utterances were not theoretically meaningful to the purpose of the study (i.e., these comments refer to design features rather than domain-based thinking about tutored content).

Self-explanations and Learning Assessments

How do specific self-explanations contribute to retention and transfer? To answer this question, we examined the relationship between significant diagram codes and learning measures. Bivariate correlations were calculated to examine the relationship between individual explanation codes and practiced and transfer scores from both the posttest and delayed posttest assessments. To control for overall verbosity, correlations used the percent of total utterances for individual codes.

For all analyses, main effects and interactions of the independent variables were examined and conclusions were based on a standard alpha level of .05.

Findings

Learning Assessments

Practiced Items

The Geometry Cognitive Tutor has been shown to produce significant learning gains (Aleven & Koedinger, 2002; Butcher & Aleven, 2008); and, as expected, the RM-MANOVA showed a significant main effect for test time ($F_{(4, 15)} = 11.97$, $p < .001$). There was no significant effect for condition ($F < 1$), and no interaction between condition and test time ($F < 1$). Univariate tests showed that the test time effect was significant for both angle measures ($F_{(2, 36)} = 22.13$, $p < .001$) and for naming geometry rules ($F_{(2, 36)} = 23.62$, $p < .001$). A pairwise comparison (with Bonferroni's adjustment) found that overall learning gains on practiced items were significant between the pretest and posttest ($M_{\text{diff}} = .35$, $p < .001$) and between the pretest and delayed posttest ($M_{\text{diff}} = .30$, $p < .001$); the changes between posttest and delayed posttest were not significant ($M_{\text{diff}} = .04$, $p > .99$).). Means and standard deviations are shown in Table 5.

Transfer Items

On items of transfer (visual transfer and explanation transfer), the RM-MANOVA showed there was no significant main effect for condition ($F_{(2, 17)} = 1.02$, $p = .38$) or test time ($F_{(2, 17)} = 1.36$, $p = .28$), and no interaction between condition and test time ($F < 1$). See Table 6 for means and standard deviations.

Self-explanations

In terms of overall verbal output, the two conditions were similar. Means and standard deviations for family codes as percent of total utterances are listed in Table 7. Overall, the MANOVA showed no significant effect for condition ($F_{(6, 13)}$ = 2.12, *p* = .12). Within the univariate analyses, some dependent variables did not differ significantly between conditions while others did. Total utterances (the total number of coded utterances) showed no main effect of condition ($F_{(1, 18)}$ = 2.80, *p* = .11). The code families that did not differ significantly by condition included: self-explanations ($F < 1$); monitoring statements ($F < 1$); and, planning statements ($F < 1$). The reading code family showed a nonsignificant trend ($F_{(1, 18)}$ = 4.16, *p* = .06) with participants in the visual explanations condition reading slightly less (*M* = 17.1, *SD* = 4.2) than the control condition (*M* = 21.3, *SD* = 5.0). Reading utterances for this analysis included the reading of glossary information or hints. Two code families showed significant differences in the percentage of total utterances by condition: diagram explanations, and technique utterances. The percent of diagram explanations produced by participants was higher in the visual explanations condition (*M* = 24.9, *SD* = 6.9) than in the control condition (*M* = 18.3, *SD* = 4.9; $F_{(1, 18)}$ = 6.12, *p* = .02). Diagram explanation statements included diagram reading, diagram interactions, and diagram explanations. A significant effect was also seen for the percent of techniques described in utterances ($F_{(1, 18)}$ = 4.9, *p* = .04): participants in the control condition talked about techniques more frequently (*M* = 11.6, *SD* = 3.1) than students in the visual explanations condition (*M* = 8.7, *SD* = 2.5). Techniques consisted of problem-solving strategies that were not explained or tied into geometry principles (e.g., stating unexplained mathematical operations or single numerical answers, and mentioning principle names without elaboration).

Overall, condition did not have a significant effect on the distribution of individual codes within code families ($F_{(1,18)}$ = 70.9, *p* = .09). However, four codes differed significantly between conditions (see Table 8 for means and standard deviations for individual codes). For the visual explanations condition, subgoal planning utterances represented a higher proportion of the planning code family (*M* = .67, *SD* = .14) than seen for the control condition (*M* = .48, *SD* = .13; $F_{(1, 18)}$ = 9.45, *p* < .01). Subgoal planning consisted of students referring to the part of the problem they planned to work on next. The visual explanations condition also produced a higher proportion of numerical answers (*M* = .46, *SD* = .16) in the technique family than the control condition (*M* = .30, *SD* = .10; $F_{(1, 18)}$ = 6.84, *p* = .02); the numerical answer technique was used to code utterances stating angle measurements without elaboration. For a visualization of the distribution of individual codes within each family by condition, see Figures 12-17.

The two other individual codes found to differ between conditions belonged to the diagram explanations code family. As noted earlier, the diagram explanations code family included specific codes that indicate how deeply the learner was utilizing the visual information: diagram reading, diagram interactions, and diagram explanations. Diagram reading consisted only of reading from the diagram and was considered a shallow utterance; diagram explanations were considered deep utterances, consisting of more complete explanations of how geometry principles applied to the diagram. Diagram interactions consisted of more content than diagram reading but less than a complete explanation. Students prompted to highlight diagram elements in the tutor did not differ from the control condition in their proportion of diagram interactions (*F* < 1). However, as seen in Figure 13, students who were prompted to highlight diagram elements produced a higher proportion of deep diagram explanations (*M* = .48, *SD* = .16) than students in the control condition (*M* = .28, *SD* = .15; $F_{(1,18)}$ = 8.78, *p* < .01). Conversely, students who were prompted to highlight diagram elements produced a smaller proportion of shallow diagram reading (*M* = .14, *SD* = .08) than students in the control condition (*M* = .34, *SD* = .22; $F_{(1,18)}$ = 6.63, *p* = .02).

Self-explanations and Learning Assessments

Regarding the relationship between diagram explanation family codes and learning outcomes, the percent of total utterances coded as diagram reading was inversely related to performance on nearly all practiced and transfer items on the posttest and delayed posttest (see Table 9). No correlations between deep diagram explanations and learning outcomes reached significance at an alpha of .05. However, there was a slight – though nonsignificant – pattern to the correlations across the assessments. All learning outcome measures on the posttest and delayed posttest correlated negatively with the percent of diagram reading codes and all but one correlated positively with the percent of diagram explanation codes.

Because of the trend differentiating the overall quantity of reading utterances between conditions, the relationship between reading family codes and learning outcomes was also explored. The reading category was comprised of two different sources of information (reading hints vs. reading glossary items). Whereas students who focus on reading and considering glossary items are working more generatively in the system, an overreliance on hints has been shown to be a shallow strategy. Thus, correlations between posttests and delayed posttest items and the percent of total utterances made up by reading from these specific types of information sources were investigated. Overall, reading is negatively correlated to performance on the posttest and delayed posttest; however, these correlations are not significant. Although both conditions were similar in the proportions of their reading belonging to these two sources, hint reading made up a majority of reading for both the control condition (M = 62.3, SD = 29.8) as well as the visual explanations condition (M = 68.7, SD = 30.4). The percent of total utterances coded as glossary reading was positively correlated with every learning outcome; however, only its correlation with performance on posttest practiced items reached significance at an alpha of .05 (r = .589, p < .01). Meanwhile, the percent of total utterances coded as hint reading was significantly, negatively correlated with every learning outcome. See Table 10 for the magnitude and significance level of these correlations.

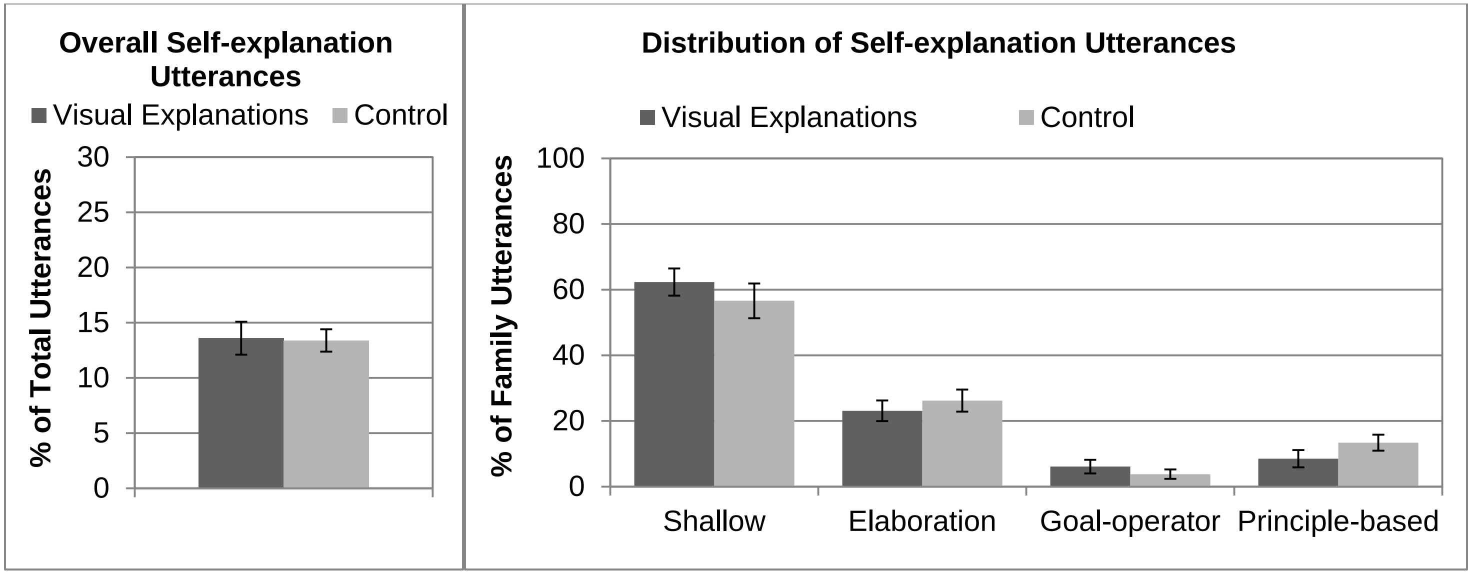


Figure 12. The percent of total utterances coded in the self-explanation code family by condition (left); and, the distribution of individual codes within the self-explanation code family by condition (right).

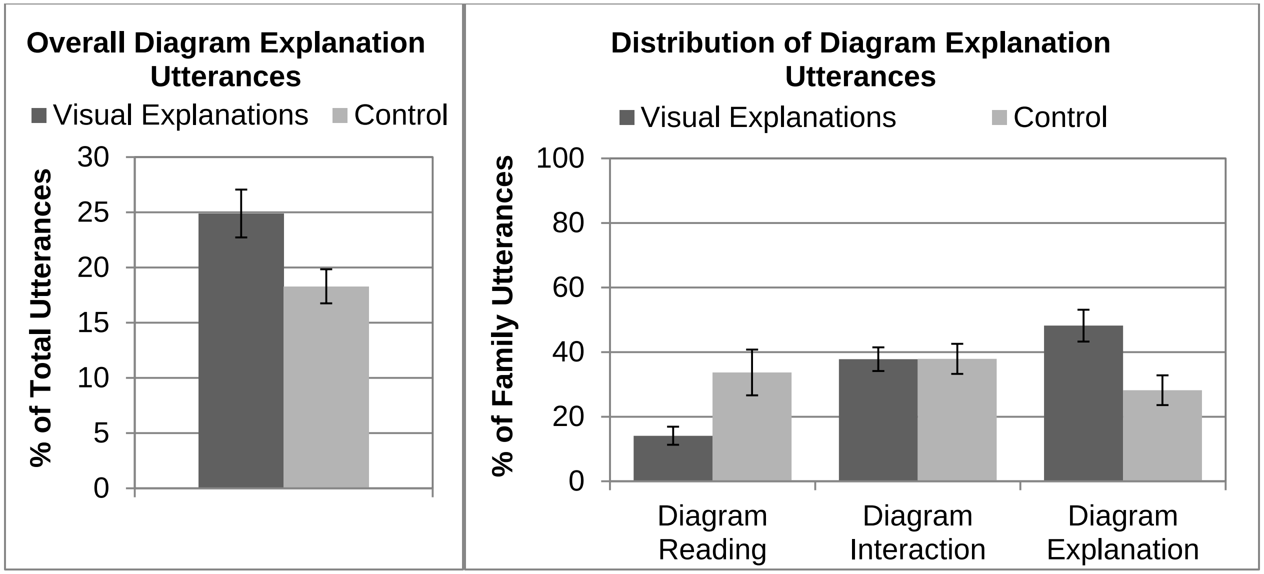


Figure 13. The percent of total utterances coded in the diagram explanation code family by condition (left); and, the distribution of individual codes within the diagram explanation code family by condition (right).

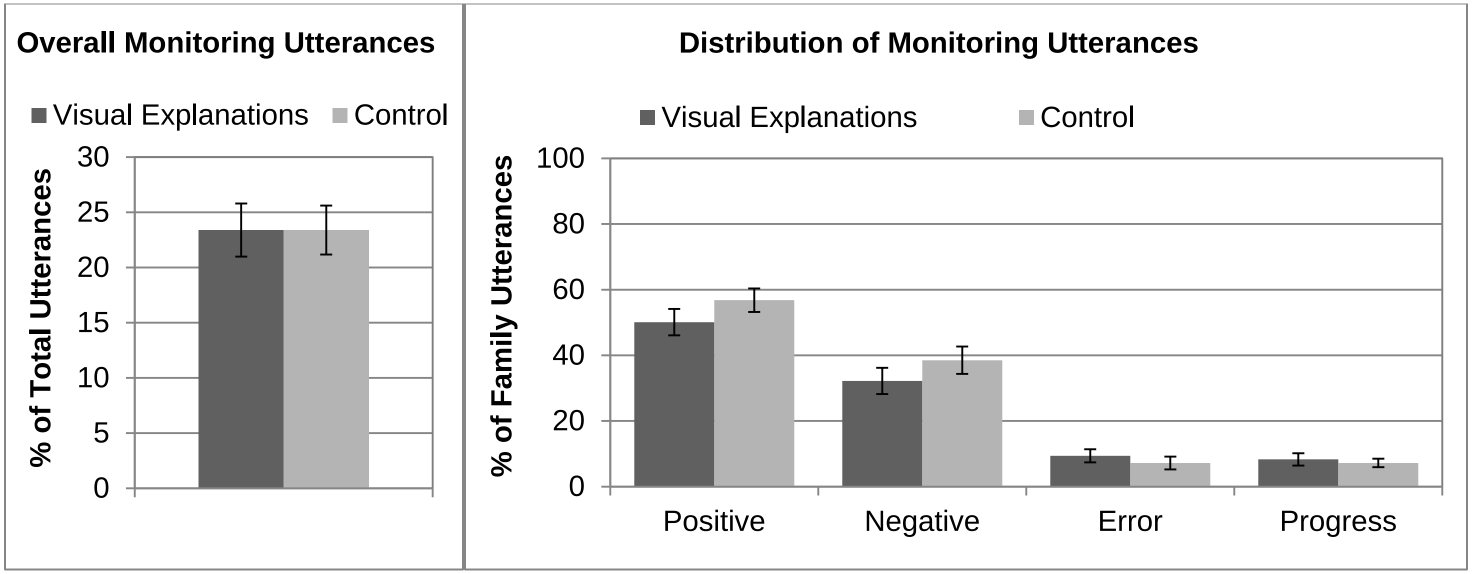


Figure 14. The percent of total utterances coded in the monitoring code family by condition (left); and, the distribution of individual codes within the monitoring code family by condition (right).

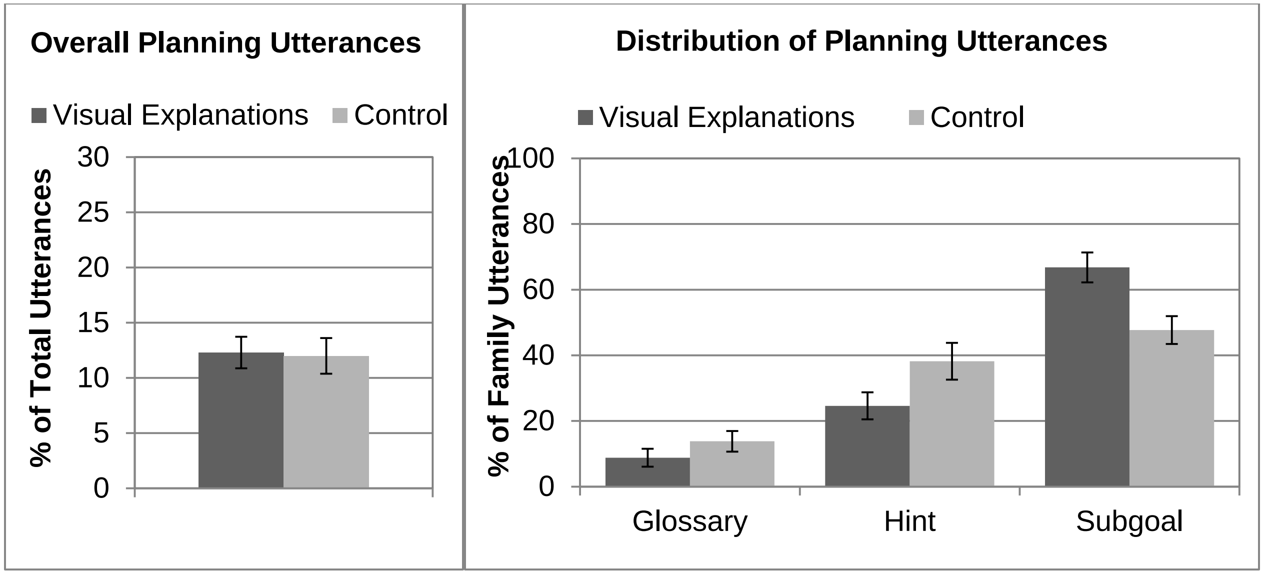


Figure 15. The percent of total utterances coded in the planning code family by condition (left); and, the distribution of individual codes within the planning code family by condition (right).

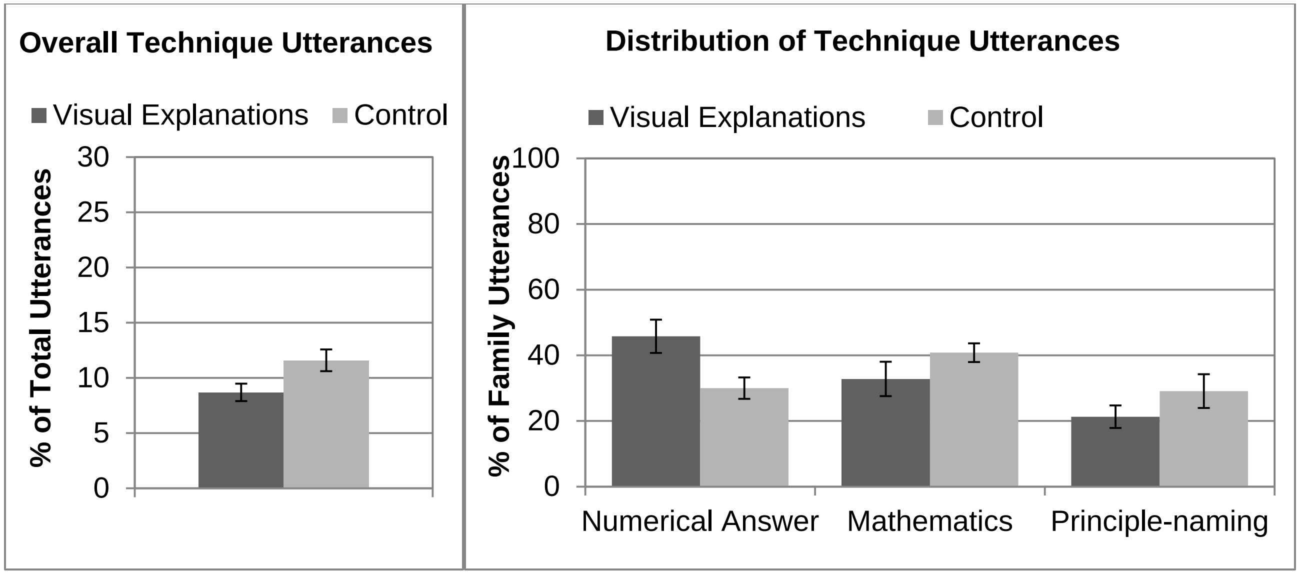


Figure 16. The percent of total utterances coded in the techniques code family by condition (left); and, the distribution of individual codes within the techniques code family by condition (right).

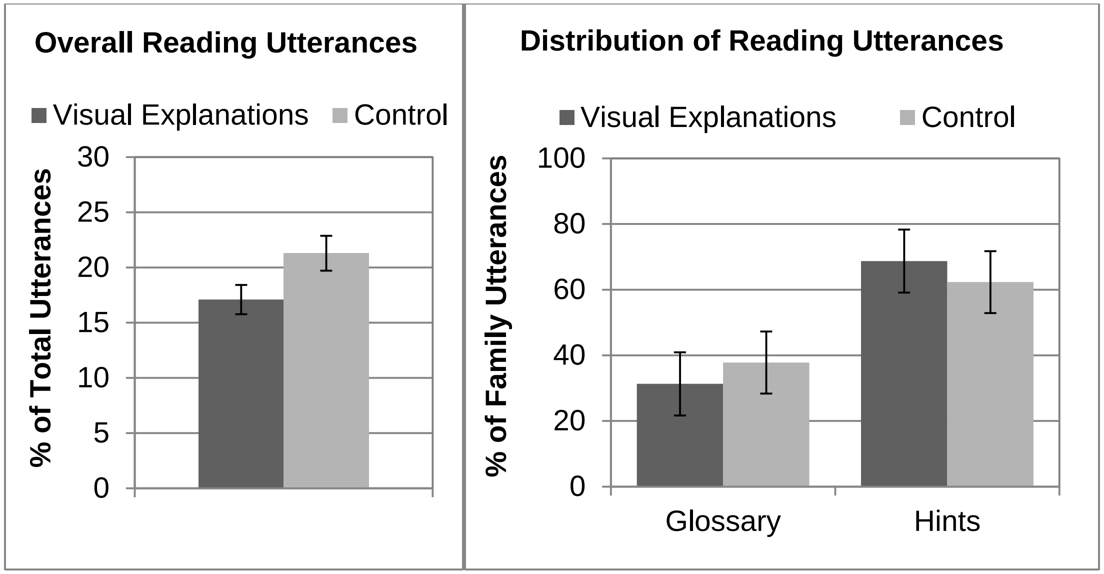


Figure 17. The percent of total utterances coded in the reading code family by condition (left); and, the distribution of individual codes within the reading code family by condition (right).

Table 5

Percent Correct Means (and Standard Deviations) for Practiced Items

	Visual Explanations (*n*=10)			Control (*n*=10)		
DV	Pretest	Posttest	Delayed Posttest	Pretest	Posttest	Delayed Posttest
Angle Measures	43.3 (26.9)	74.3 (29.2)	72.9 (24.7)	48.9 (23.5)	87.1 (21.8)	80.0 (23.5)
Geometry Rule Names	2.2 (4.7)	31.4 (28.4)	34.3 (28.7)	0 (0)	47.1 (41.5)	45.7 (39.2)

Table 6

Percent Correct Means (and Standard Deviations) for Transfer Items

	Visual Explanations (*n*=10)		Control (*n*=10)	
DV	Posttest	Delayed Posttest	Posttest	Delayed Posttest
Visual Items	70.0 (14.0)	69.5 (15.3)	72.5 (11.2)	71.5 (14.0)
Explanation Items	51.4 (36.9)	61.4 (28.8)	67.9 (30.7)	73.6 (29.4)

Table 7

Percent of Total Utterances (and Standard Deviations) for Code Families

Code Family	Visual Explanations (*n*=10)	Control (*n*=10)
Self-explanations	13.6 (4.7)	13.4 (3.2)
Diagram Explanations	24.9 (6.9)	18.3 (4.9)
Monitoring	23.4 (7.6)	23.4 (7.0)
Planning	12.3 (4.5)	12.0 (5.1)
Techniques	8.7 (2.5)	11.6 (3.1)
Reading	17.1 (4.2)	21.3 (5.0)

Table 8

Percent (and Standard Deviations) of Code Family for Individual Codes

Code Family	Code	Visual Explanations (*n*=10)	Control (*n*=10)
Self-explanation	Shallow	62.3 (13.1)	56.6 (16.7)
	Elaboration	23.1 (10.0)	26.2 (10.7)
	Goal-operator	6.1 (6.5)	3.8 (4.6)
	Principle-based	8.5 (8.3)	13.4 (7.7)
Diagram Explanation	Diagram Reading	14.1 (8.8)	33.7 (22.4)
	Diagram Interaction	37.8 (11.6)	37.9 (14.7)
	Diagram Explanation	48.2 (15.6)	28.2 (14.6)
Monitoring	Positive Monitoring	32.2 (12.7)	38.5 (11.3)
	Negative Monitoring	50.1 (12.6)	46.8 (13.2)
	Error Monitoring	9.4 (6.3)	7.2 (6.1)
	Progress Monitoring	8.3 (5.9)	7.2 (4.1)
Planning	Glossary Planning	8.8 (8.7)	13.8 (10.0)
	Hint Planning	24.6 (13.0)	38.2 (17.7)
	Subgoal Planning	66.8 (14.4)	47.7 (13.3)
Techniques	Mathematics Procedures	32.8 (16.6)	40.8 (9.0)
	Numerical Answer	45.8 (16.0)	30.0 (10.4)
	Principle-naming (NOT Given)	21.3 (10.8)	29.1 (16.2)
Reading / Paraphrasing	Glossary	31.3 (30.4)	37.8 (29.9)
	Hint	68.7 (30.4)	62.3 (29.8)

Table 9

Correlations between percent of diagram reading and diagram explanation and the posttest and delayed posttest learning assessments

	Posttest			Delayed Posttest		
	Practiced	Visual Transfer	Explanation Transfer	Practiced	Visual Transfer	Explanation Transfer
Diagram Reading	-.526*	-.482*	-.504*	-.573**	-.506*	-.397
Diagram Explanations	.118	.302	-.051	.062	.151	.116

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 10

Correlations between overall reading, proportion of glossary reading, and proportion of hint reading and the posttest and delayed posttest learning assessments

	Posttest			Delayed Posttest		
	Practiced	Visual Transfer	Explanation Transfer	Practiced	Visual Transfer	Explanation Transfer
Overall Reading	-.107	-.166	-.083	-.239	-.283	-.228
Glossary Reading	.589**	.432	.384	.423	.401	.377
Hint Reading	-.686**	-.574**	-.457*	-.624**	-.637**	-.567**

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

CHAPTER 4

DISCUSSION

Does Highlighting Diagram Elements Result in Improved Retention and Transfer over Time?

In previous research investigating visual explanations, students who highlighted geometry diagrams to justify geometry proofs achieved better long-term retention and application than students who did not (Butcher & Aleven, 2009). The purpose of our study was to replicate that finding and to use verbal protocols to investigate the impact of visual interaction on cognitive processes during learning with an intelligent tutoring system. We hypothesized that visual explanations would produce better, longer-lasting learning outcomes by supporting the generation of deep self-explanations, particularly those explanations that linked visual features of problem diagrams to relevant domain concepts.

On practiced and transfer items, the current study did not find a learning advantage for the visual explanations condition. Instead, students in both conditions made significant learning gains from pre- to posttest and sustained those gains on the delayed posttest (see Tables 5 and 6). Differences in methodology between the current study and previous research with visual explanations may explain the conflicting results in learning assessment outcomes. For example, one possibility for the lack of condition differences could be the use of verbal self-explanation. As noted in the introduction, self-explanation is a robust learning strategy that repeatedly has been found to support deeper learning (Chi et al., 1989; Chi et al., 1994; Renkl, 1997). Because students in both conditions were trained to self-explain as they worked with the tutor, the self-explanation effect may have prompted the control condition students to think more deeply about their problem-solving. Thus, the visual highlighting prompts may not have spurred explanation that was significantly greater than what can be achieved by prompting verbal explanations.

Alternatively, the brevity and laboratory setting of this study may have resulted in more shallow learning strategies among participants, thereby minimizing potential learning benefits. The previous study took place in a high school classroom, while the current study examined university undergraduate students in a laboratory setting. Away from an authentic setting, the stakes may not have been high enough to motivate participants to interact thoughtfully with the Cognitive Tutor. For example, students in the classroom tend to avoid asking for hints until after errors are made. In this laboratory study, students frequently asked for hints even before attempting problems and relied heavily on hints and glossary entries to guide them through the problems. As noted earlier, systematic abuse of hints can lower student learning (Baker, Corbett, & Koedinger, 2004; Baker, Corbett, Koedinger, & Wagner, 2004). In this study as well, hint reading was significantly negatively correlated with every learning outcome. Longer-term studies in the laboratory may be necessary for students to form sufficient levels of knowledge that will allow them to engage with tutor interventions in more authentic ways.

Does Highlighting Diagram Elements In Order to Justify Problem-solving Steps in Geometry Lead Students to Generate More Effective Self-explanations?

Though learning assessments showed no effect for condition, the verbal analysis provided insight into the role of highlighting in altering the pattern of students' self-explanations. We found that students who highlighted diagram elements to justify geometry proofs produced proportionally more deep diagram explanations and fewer shallow diagram reading utterances than the control condition. This pattern is important because students in the visual explanations condition could have responded to the highlighting prompts by continuing to simply read from the diagram, or engaging in lower level diagram interaction (in which students fail to engage in actual explanation of the visual elements). The fact that highlighting prompts served to increase students' deep diagram explanations suggests that the interaction helped them to engage in deeper thinking about the visual diagram elements as related to domain principles and concepts.

The visual explanations condition also produced fewer utterances of unexplained techniques. Technique utterances are considered shallow because they name steps in the problem-solving process without adequate justification. Visual interactions with the diagram may have prompted students in the visual explanations condition to incorporate their mathematical calculations and principle references with diagram information, transforming them into deep diagram explanations.

Although the deep process changes observed in this research show the potential for these types of visual explanations to improve the depth with which students engage with the intelligent tutoring system during problem-solving practice, the lack of differences in learning gains across conditions suggest that these changes may not have been significant enough to impact learning outcomes. As discussed earlier, longer-term usage in more authentic contexts may amplify these benefits. However, another possibility is that stronger interventions may be needed in order to influence outcomes.

How Does Explanation Type Correspond to Learning Outcomes, Including Retention and Transfer?

The significant negative correlations between diagram reading and performance on practiced and transfer items on the posttest and delayed posttest indicates that students who produced a higher proportion of diagram reading learned less overall. This result suggests that shallow diagram interactions are detrimental to learning. It is somewhat troubling that deep diagram explanations were not significantly correlated to learning outcomes. The absence of significant correlation with practiced, visual transfer, or explanation transfer items suggests that the visual explanation prompts may not have been sufficient for facilitating deeper thinking than in the control condition.

While increasing the depth of processing of visual content during intelligent tutoring, the addition of visual interaction in the form of diagram highlighting may also serve to reduce students' dependence on external sources of geometry information. Students in the visual explanations condition relied less on reading overall than did the control condition. This could mean that deeper reasoning about the diagram, evoked by highlighting cues, reduces the perceived need for guidance; or, it may suggest that the highlighting cues serve as a form of guidance in and of themselves.

For both conditions, the majority of reading utterances were made up of hint reading (as noted above). This is important to note because the source of reading material influenced learning outcomes in very different ways. Overreliance on ITS hints has been shown before to contribute to shallow learning (Baker, Corbett, Koedinger, & Wagner, 2004). Here, again, hints were strongly negatively correlated with performance on all learning measures. This indicates that students used the hint function to arrive at solutions without engaging in deep cognitive processes. In contrast, glossary reading seemed to provide a learning benefit: glossary reading was positively correlated with performance on all learning measures.

What could drive the difference in learning outcome between reading from the hints versus reading from the glossary? For the most part, the language of the hints is similar to the language found in the glossary; both the hints and glossary provide students with definitions of geometry principles. One difference between the two is that the hints are specifically catered to the problem at hand while the glossary entries are not. It may be that the tailored help in hints actually hurts novice learners by reducing their need to reason and self-explain. Another difference is that the glossary includes diagrams along with text to illustrate each principle. Consequently, the glossary may support learning by facilitating integration of geometry principles with visual applications. Alternatively, the benefit of glossary reading may lie in why and how students approach supplementary reading material. Students may access the two resources with different purposes: using hints to simply find the answer, or using the glossary to learn about geometry principles. Students may choose a different resource depending on their prior knowledge: students are likely to access hints when they have no idea how to proceed, while students who access the glossary may do so with some idea of what they are looking for and have more pre-existing knowledge with which to work. Future interventions may consider exploring methods that would encourage students to make use of the glossary and discourage overreliance on hints.

Conclusions

Participants think differently about geometry diagrams when asked to provide visual explanations by highlighting relevant diagram elements. The differences in verbalized diagram explanations suggest that prompting visual explanations can increase deep self-explanations about the diagram. The simultaneous reduction in shallow diagram utterances, shallow techniques, and reliance on supplemental information (hints and glossary) influences learning. By decreasing students' tendencies to engage simply in diagram reading, visual explanations can improve students' performance on items of retention and transfer over time. They may also serve to reduce students' dependence on external help.

However, learning outcomes in this study were not entirely consistent with earlier research of visual explanations (Butcher & Aleven, 2009). The evidence is insufficient to support the claim that increased self-explanations about the diagram improve deep understanding. The visual interaction implemented in this research may not have helped students carry their self-explanations far enough. Additional interactive applications may be needed to support the successful integration of visual and verbal information. To confirm mastery of a principle after successful problem-solving, perhaps students could be prompted to highlight angle pairs demonstrating the specified principle in new diagram configurations. Or, students could be asked to resolve unsolvable angles by highlighting diagram elements representing information that could be used with a specific geometry principle to obtain an angle measure. Or, students could be asked to select an explanation that provides a verbal rationale for why the selected principle is represented in the highlighted diagram configuration. Ultimately, the interactions involved in ITS problem-solving practice should help students build a situation model that can be applied to new, complex contexts. Further exploration should identify appropriate additions.

REFERENCES

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26*(2), 147-179.

Aleven, V., Koedinger, K. R., Sinclair, H. C., & Snyder, J. (1998). Combatting shallow learning in a tutor for geometry problem solving. In B.P. Goettl, H.M. Halff, C.L. Redfield & V.J. Shute (Eds.), *Intelligent Tutoring Systems, Fourth International Conference, ITS '98. Lecture Notes in Computer Science 1452* (pp. 364-373). Berlin: Springer Verlag.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *Journal of the Learning Sciences, 4*(2), 167-207.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist, 40*(4), 199-209.

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2004). Detecting student misuse of intelligent tutoring systems. In J. C. Lester, R. M. Vicari, & F. Paraguaçu (Eds.), *Intelligent tutoring systems* (Vol. 3220, pp. 54-76): Springer Berlin / Heidelberg.

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the Cognitive Tutor classroom: When students “game the system”. *Proceedings of ACM CHI 2004: Computer-Human Interaction*, 383-390.

Butcher, K. R. (2010). *How diagram interaction supports learning: Evidence from think alouds during intelligent tutoring*. In A. Goel, M. Jamnik & N.H. Narayanan (Eds.), *Diagrammatic Representation and Inference: Lecture notes in computer science, volume 6170* (pp. 295-297). Berlin/Heidelberg: Springer.

Butcher, K. R., & Aleven, V. (2008). Diagram interaction during intelligent tutoring in geometry: Support for knowledge retention and deep understanding. In B.C. Love, K. McRae & V.M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1736-1741). Austin, TX: Cognitive Science Society.

Butcher, K. R., & Aleven, V. (2009). *Visual self-explanation during intelligent tutoring: More than attentional focus?* Paper presented at the 13th biennial conference of the European Association for Research on Learning and Instruction (EARLI), Amsterdam, Netherlands.

Butcher, K. R., & Kintsch, W. (in press). Text comprehension and discourse processing. In A.F. Healy & R.W. Proctor (Eds.), *Handbook of psychology, volume four: Experimental psychology* (*2nd Ed.)*. Hoboken, NJ: Wiley.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145-182.

Chi, M. T. H., de Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.

Conati, C., & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education, 11*, 389-415.

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*, 253-278.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev ed.). Cambridge, MA: MIT Press.

Hausmann, R. G., & Chi, M.T. (2002). Can a computer interface support self-explaining? *Cognitive Technology, 7*(1), 4-14.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294-303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*(3), 239-264.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*(2), 205-226.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

Lovett, M. C., & Anderson, J. R. (1994). Effects of solving related proofs on memory and transfer in geometry problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(2), 366-378.

McNamara, D. S., & Magliano, J. P. (2009). Self-explanation and metacognition: The dynamics of reading. . In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Handbook of Metacognition in Education* (p. 449). New York, NY: Routledge/Taylor & Francis Group.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13-23.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*(1), 1-29.