JOHN JR-HUNG LIN and SUNNY S. J. LIN

COGNITIVE LOAD FOR CONFIGURATION COMPREHENSION IN COMPUTER-SUPPORTED GEOMETRY PROBLEM SOLVING: AN EYE MOVEMENT PERSPECTIVE

Received: 17 July 2013; Accepted: 2 November 2013

ABSTRACT. The present study investigated (a) whether the perceived cognitive load was different when geometry problems with various levels of configuration comprehension were solved and (b) whether eye movements in comprehending geometry problems showed sources of cognitive loads. In the first investigation, three characteristics of geometry configurations involving the number of informational elements, the number of element interactivities and the level of mental operations were assumed to account for the increasing difficulty. A sample of 311 9th grade students solved five geometry problems that required knowledge of similar triangles in a computer-supported environment. In the second experiment, 63 participants solved the same problems and eye movements were recorded. The results indicated that (1) the five problems differed in pass rate and in self-reported cognitive load; (2) because the successful solvers were very swift in pattern recognition and visual integration, their fixation did not clearly show valuable information; (3) more attention and more time (shown by the heat maps, dwell time and fixation counts) were given to read the more difficult configurations than to the intermediate or easier configurations; and (4) in addition to number of elements and element interactivities, the level of mental operations accounts for the major cognitive load sources of configuration comprehension. The results derived some implications for design principles of geometry diagrams in secondary school mathematics textbooks.

KEY WORDS: cognitive load, configuration comprehension, eye movement, geometry diagram, problem solving

INTRODUCTION

Background

Geometry problems in secondary textbooks usually provide a simple instructional text posed with a diagram. To solve a problem, a student has to understand the text and read the diagram for familiar configurations that he/she has previously learned, as the familiar patterns may help the individual connect to relevant geometry concepts from memory that support making effective inferences and numerical computations, thus leading to a solution. Studies have found that configuration comprehension is a source of difficulty for students (Bobis, Sweller & Cooper, 1994). Unfortunately, this crucial issue for math teachers, textbook

International Journal of Science and Mathematics Education (2014) 12: 605–627

designers, students and parents has yet to be fully and carefully investigated due to the lack of proper techniques (Epelboim & Suppes, 2001). For example, the recent development of eye movement protocols made the survey of the online process of configuration comprehension possible. Accordingly, it appears that eye movement protocols have some evident advantages over traditional written or spoken protocols, as eye movement protocols do not require extra effort by the participant nor is it disruptive. Furthermore, eye movements are a genuine, integral part of the geometry configuration comprehension process.

However, using the eye movement technique to study configuration comprehension is not without its challenges. Among the many common problems inherent in this type of research, Ratwani, Trafton, & Boehm-Davis (2008) suggest that the only possibly useful approach to interpret eye movement data is to work within a specific theoretical framework.

In this study, we attempt to identify the locus of these difficulties based on the cognitive load theory (Sweller, 1988). Accordingly, we monitored students' eye movement patterns as they read and prepared to solve geometry problems. We constructed serial problems of similar triangles holding most sources of extraneous cognitive loads constant to reduce the complexity of cognitive load sources. The comparison of the pass rate and the subjective cognitive load rating (Paas, 1992) made it possible to evaluate the extent to which problem solvers comprehended the configuration, while self-report confirmed the sources of the difficulties. The use of eye movements made it possible to compare on-line configuration comprehension patterns for successful versus unsuccessful problem solvers during a realistic, complex cognitive task (Grant & Spivey, 2003; Just & Carpenter, 1985). From the perspective of eye movement, researchers can more objectively observe learners' cognitive processes during geometry problem solving. Specifically, researchers can gain insights into the difficulties that learners perceived while problem solving.

The Present Study

In general, the present study investigated (a) the perceived cognitive load and pass rate of several carefully designed geometry problems, (b) what cognitive load sources did students explicitly notice and (c) whether eye movement patterns show sources of cognitive load in configuration comprehension for various difficulty levels of problems and (un)successful levels of performance.

REVIEW OF LITERATURE

Cognitive Load

Cognitive load theory has been widely used to describe the load of a human's cognitive system when a task is performed (Chandler & Sweller, 1991; Sweller & Chandler, 1991). In the recent decades, the theory has been greatly adapted as a framework for researchers to improve instructional outcomes (Paas, Renkl & Sweller, 2003). The cognitive resource of the human cognitive system, such as attention or working memory, is limited. Therefore, a complex task would impose heavier loads on the system than a simpler task. Cognitive load consists of three components (Paas, Tuovinen, Tabbers & Van Gerven, 2003). First, the intrinsic cognitive load is primarily caused by the number of elements that must be processed simultaneously and the inter-relationship between these elements (Sweller, Ayres & Kalyuga, 2011). Second, extraneous cognitive load results from unnecessary instructional designs. Such a design imposes a higher extraneous cognitive load on learners than designs with no irrelevant materials (Park, Moreno, Seufert & Brunken, 2010). The last type of cognitive load, germane cognitive load, is defined as the load of the human cognitive system when a schema corresponding to a specific learning task is acquired or integrated with the previous schema (Holm et al. 2009; Paas & van Merriënboer, 2006). Although all three loads consume cognitive resources, they have distinct impacts on learning. For example, the intrinsic and extraneous loads tend to hinder learning, whereas germane loads facilitate learning (Plass, Moreno & Brünken, 2010).

Methods for Measuring Cognitive Load

Two distinct approaches, subjective and objective measures, were adopted to evaluate cognitive load in performing a task (Paas et al. 2003b). In subjective measures, learners are required to report the cognitive load they perceive as they implement a task. Due to the nature of the multidimensional construct of cognitive load, subjective techniques usually adopt a self-reporting questionnaire with several subscales. For instance, the National Aeronautics and Space Administration-Task Load Index (NASA-TLX) consists of six subscales, including mental demand, physical demand, temporal demand, frustration, effort and performance, which together represent mental workload in performing a task (Hart & Staveland, 1988). In contrast, Paas (1992) designed one item with a 9-point Likert scale to measure the amount of effort participants invest to complete a task.

On the other hand, objective measures often use physiological techniques to detect changes in physiological variations of cognitive load. These methods include heart activities (e.g. changes in heart rate), eye activities (e.g. ratio of pupillary dilations) and brain activities (e.g. electroencephalography) (Holm et al., 2009). Researchers use specific instruments to collect quantitative data while participants are performing a task, thus allowing the level of "online" cognitive loads to be estimated. An objective measure paradigm provides a precise and spontaneous way of measuring cognitive load. Mayer (2010) suggested that eye-tracking methodology was beneficial for gaining an understanding of the perceptual processes while learning with graphics. In addition, eye tracking can detect where viewers fixated and when they fixated (Rayner, 1998). Thus, eye tracking facilitates the understanding of not only what the attention procedure is but also how the attention procedure progresses (Mayer, 2010).

Geometry Problem Solving

Geometry problems consistently require the presentation of a diagram that is central to the problem (Laborde, 2005; Larkin & Simon, 1987; Sweller, Mawer & Ward, 1983). The diagram provides a useful, readily visualized representation of the problem (Sweller et al., 1983). Cognitive load researchers have yet to thoroughly analyze the sources of difficulty associated with configuration comprehension in reading the diagrams that accompany the geometry problems as well as when preparing to solve the problems (Just & Carpenter, 1985). Among the limited evidence, Sweller et al. (1983) found that geometry experts had learned a variety of geometric configurations to which they readily apply the appropriate theorems. Most often, cognitive load theorists use geometry problems (such as major–minor arc, interior angle, exterior angle, right-angle triangles in a two-dimensional space with the application of trigonometric ratios and Pythagorean principles with these triangles) to examine the effects of the tasks, such as worked examples (Paas & van Merriënboer, 1994; Schwonke, Renkl, Salden & Aleven, 2010) and the development of expertise from a mean-ends strategy toward a forward strategy (Sweller et al., 1983). For example, in an experiment conducted by Sweller et al. (1983), a geometric diagram was presented on a visual screen. Solvers had to find a specific angle using the following theorem: Each exterior angle of a triangle is equal to the sum of the interior opposite angles.

Eye Movement in Configuration Comprehension

The eye-tracking methodology has been applied to study online cognitive processing in reading (Rayner, 1998) and problem solving (Hegarty & Just, 1993; Hegarty, Mayer & Green, 1992; Hegarty, Mayer & Monk, 1995), as this methodology provides researchers a promising way to study what people think when they see something, such as a text or an object (Hyona, 2010). However, this method has only recently been applied to study the design principles of instructional material. Mayer (2010) commented that eye-tracking methodology contributed to understanding how a particular instructional design influences learning. Examples include adopting various eye movement indicators based on eye fixations to investigate a learner's perceptual processing during learning (Boucheix & Lowe, 2010; Meyer, Rasch & Schnotz, 2010; Ozcelik, Arslan-Ari & Cagiltay, 2010).

With respect to configuration comprehension, Carpenter and Shah (1998) suggest two cognitive stages. These include (1) the pattern recognition stage, which leads to the encoding of a visual pattern by forming a visual chunk, and (2) the interpretive stages, which translate the pattern into its quantitative and qualitative meanings and relate this information to the referents in the diagram. These processes are repeated in a cyclical manner for each visual chunk in the diagram, with each cycle interpreting a single chunk. A model proposed by Gillan (1995) suggests that there are five functional component stages when people interact with graphs during the process of completing tasks, which include (1) searching for the spatial location of an indicator, (2) encoding the value of the indicator, (3) performing arithmetic operations on the encoded values, (4) comparing the spatial relations of the indicators and (5) responding. The fourth stage in which spatial relations in a diagram must be identified for information integration is critical for geometric problem solving. Furthermore, when solvers used the graphs to make comparisons, they focused on the spatial metaphor of the graph (Gillan, 1995). Therefore, different graphs influence solving processes differently, even though the problems applied the same theorems or properties. To demonstrate this cyclical process, Carpenter and Shah (1998) examined graph readers' transitions between regions of the diagram. Readers' fixations cycled between different regions for each of the visual chunks represented in the diagram, suggesting that readers cycled between different stages of processing. As diagram complexity increased (i.e. the number of unique visual chunks increased), the number of transitions between regions of the diagram increased, suggesting that a single processing cycle was required for each chunk in the diagram. Similarly, Ratwani et al. (2008) introduce two components that are needed to form a coherent representation of the graph: visual and cognitive integration. Verbal protocols and eye movement data provide strong support for both of these components. Visual integration involved the explicit formation of visual clusters of information, while cognitive integration involved the explicit comparison of these visual clusters to the referents and were critical to other visual clusters to form a coherent representation. Thus, the visual clusters formed during visual integration served as objects similar to units that could then be used to reason about the graph during cognitive integration. In this study, we expect that when reading a geometry diagram, readers need to identify the critical chunks of the graphic for the fixation to be placed in the informational areas. The readers must then scan each visual chunk to form a proper spatial relation for a coherent representation to connect/activate familiar configurations that were previously learned to make an effective inference for problem solving.

Little research has been conducted to investigate participants' online configuration comprehension processing when solving geometry problems from a cognitive load perspective. The present study used eye tracking incorporated with a subsequent writing comprehension test to investigate both online and offline cognitive load while reading diagrams that accompany geometry problems. Experiment 1 explored the relationships between difficulty and self-reporting cognitive load, while the goals of experiment 2 were to examine what cognitive load sources students explicitly noticed/used and whether eye movement patterns show sources of cognitive load in comprehending geometry configurations.

EXPERIMENT 1

The properties of similarity are fundamental in learning geometry. According to Competence Benchmarks S-4-15 for Grades 1–9 Curriculum Guidelines in Taiwan, students should be able to comprehend the similarity properties of triangles and apply the corresponding properties to solve problems and make inferences (Ministry of Education, ROC, 2011). Similarly, the Common Core Georgia Performance Standards (CCGPS) state that students should use congruence and similarity criteria for triangles to solve problems and to prove relationships between geometric figures (Georgia Department of Education, USA, 2013). Considering the importance of the similarity properties and their ecological validity, we designed five pairs of similar triangles and asked the participants to apply similarity properties to calculate the length of a specific side of a triangle. To find an accurate solution, the individual must identify possible similar triangles and recognize the corresponding congruent angles and the corresponding sides that are of the same proportion. The diagrams were designed to have various levels of cognitive load. After the completion of each task, the individual was asked to report the level of cognitive load. The authors were interested in determining whether the self-reported cognitive loads and the pass rate changes would align with designated levels of cognitive load in configuration comprehension.

METHOD

Participants

Given that the similarity properties of triangles were taught in grade 9 in Taiwan, we extensively conducted this research using grade 9 students who had just learned the similarity properties. Using the convenience sampling method (O'Leary, 2004), 311 participants (146 males and 165 females) were selected from five junior high schools in Taiwan. All participants were 14 to 15 years of age. Consents were obtained from participants' parents and from the administrators of the schools, and all participants received a gift equivalent to approximately US $2.

Materials and Procedure

The task in this study involved five similar triangle problems. Each problem consisted of a brief statement that preceded the problem section, and a diagram with paired similar triangles was represented (mathematically denoted as ⊿ABC~⊿DEF, see Fig. 1). To control the extraneous cognitive load, the layout (relative distance and size) of the diagrams were designed to be the same. Additionally, the areas of the figures were approximately equal. Two types of cues, namely the lengths of the sides or the degrees of the angles, were conditionally provided and closely written to the corresponding sides or angles

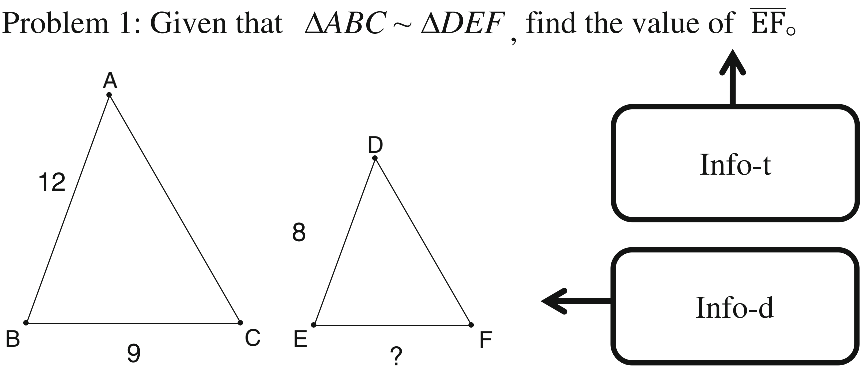

Figure 1. Problem #1 with a translational pair of similar triangles. $\overline{EF}$ refers to the distance between point *E* and *F*. *Info-t area* is the area of instruction text while *Info-d area* as the diagram area

on the appropriate diagram. To test whether the order of the five problems would influence participants' responses (denoted as the period effect) and whether the preceding problem will affect the answers to the current problem (denoted as the carry-over effect), a William's square design was used to generate five distinct sequences of the problems (Xu, 2006). That is, in different time periods, each class answered the ten problems in varying sequences. Participants were required to solve for some unknown angles or sides according to the similarity property of triangles. After performing the tasks, participants were required to answer the perceived cognitive load while solving each problem (Paas et al. 2003b). The procedures were as follows: (1) The homeroom teacher first announced the purposes of the study and the participants completed the consent form. (2) The problem sheets, with the sequence (period) of five problems randomly placed, were distributed to all students. No cues were given in the test phase, and students were not allowed access to any references. (3) Students were informed that the maximal time to solve a problem and complete the NASA-TLX questionnaire was 180 s. They were also told that they should write down the problem-solving processes as clearly as possible. (4) After completing the experiment, the results were scored by awarding one point for each correct answer and zero points if there was no answer or there was an incorrect answer. No partial scores were awarded.

With respect to the cognitive load, three characteristics, namely the number of informational elements, level of element interactions and level of mental operations were assumed to account for the increasing difficulty. Figure 1 and the diagram comprehension stages, as proposed by Ratwani et al. (2008), were used to describe the above three characteristics (see Table 1).

1. The number of informational elements: An individual would process visual integration that involves identifying a visual cluster of critical information. Reading Fig. 1, one would find the four most salient information elements (e.g. lengths of sides of two triangles = >12, 9, 8 and $\overline{EF}$). The five problems were designed to have the same (four) salient information elements (see Fig. 2).
2. Level of mental operation: In Fig. 1, the spatial relation of the pairs of similar triangles is in translation such that a simple mental operation must be performed when the individual reads one of the triangles and then moves to the right (or left) to find the similar second triangle for the corresponding angles. In the other four problems, mental operations such as flip-over or mental rotation to compare one angle in the first triangle to the corresponding angle in the second triangle would sometimes be required (see Table 1).

TABLE 1

Pass rates, means (SDs) and diagram characteristics ($n = 311$)

Problem #	*Pass rate*	*Cl1*	*Cl2*	*Cl3*	*Cl4*	*Cl5*	*Cl6*	*Number elements*	*Mental operations*	*Inter-actions A:B = C:D*
1	.78	4.28 (4.71)	3.24 (3.62)	3.23 (3.54)	3.94 (4.15)	14.80 (6.61)	3.34 (4.12)	4	Straightforward	1 step simple mapping
2	.82	4.57 (5.14)	3.54 (4.22)	3.44 (3.91)	4.22 (4.48)	14.69 (6.54)	3.11 (3.91)	4	Check correspondence angels	1 step overlap mapping
3	.83	4.33 (4.84)	3.32 (3.67)	3.47 (3.97)	4.05 (4.18)	14.88 (6.43)	3.27 (3.96)	4	Check correspondence angles	1 step flip-over mapping
4	.71	4.26 (4.86)	3.33 (3.84)	3.44 (3.90)	3.87 (4.05)	14.42 (6.89)	3.10 (3.69)	4	Check correspondence angles	1 step vertical mapping
5	.58	7.36 (6.38)	5.10 (5.21)	5.42 (5.46)	6.74 (5.86)	12.24 (7.22)	4.53 (5.00)	4[a]	Need mental rotation to check correspondence angles	2 step rotation mapping
total	–	4.96 (5.35)	3.71 (4.21)	3.80 (4.28)	4.56 (4.72)	14.21 (6.81)	3.47 (4.19)	–		

Mental operations refer to the crucial cognition to find similar-triangle pattern. For the configuration pattern of each problem, please see Fig. 2

Cl1–Cl6 cognitive load items of NASA-TLX, *Cl1* mental demand, *Cl2* physical demand, *Cl3* temporal demand, Cl4 performance, *Cl5* effort), *Cl6* frustration

[a] One being hidden

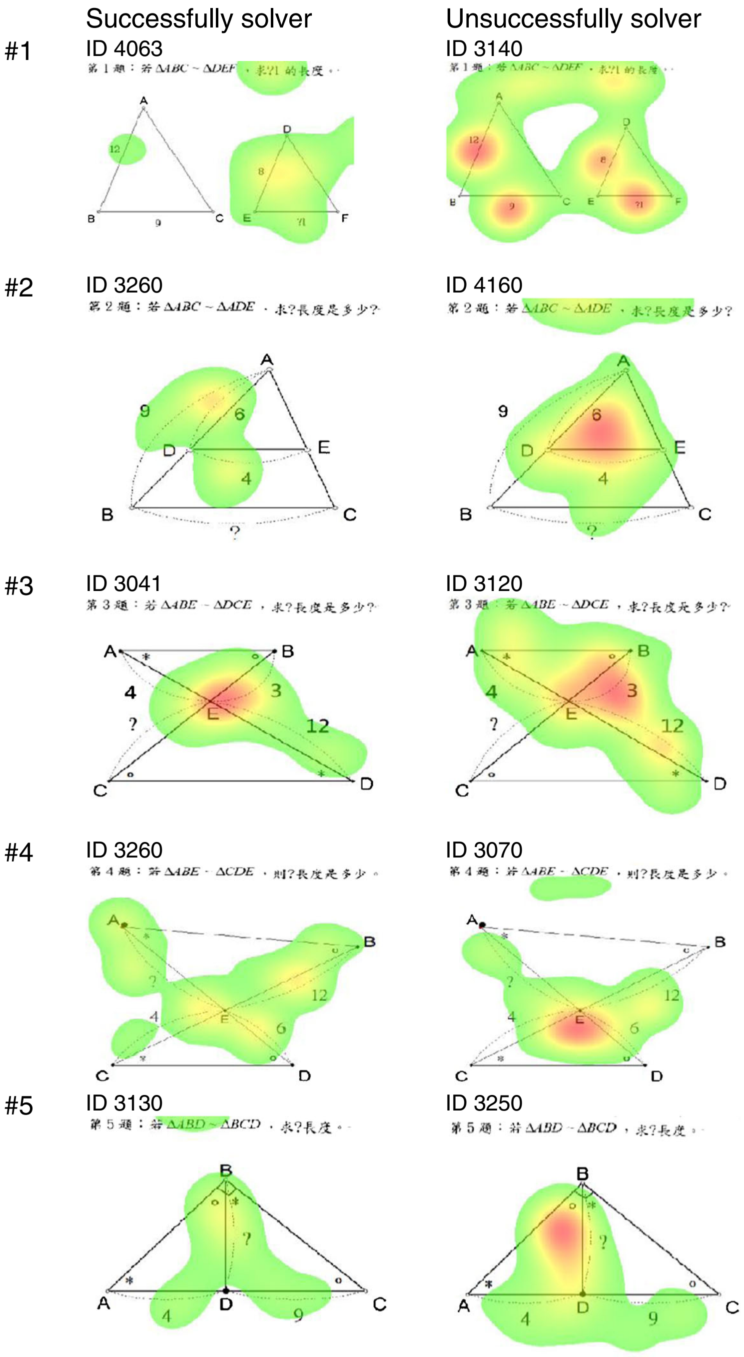

Figure 2. Heat maps of successful and unsuccessful problem solvers while solving problem #1 to problem #5

3. Level of element interactions: After stage 1, cognitive integration would occur involving the explicit comparison of the visual clusters to prior knowledge and the critical comparison to other visual clusters to form a coherent representation. The concept of similar triangles would be retrieved, the problem solver would recall that corresponding sides are in proportion, and they would then derive the appropriate equations (e.g. $12:8 = 9:\overline{EF}$ for Fig. 1). This involved one interaction among four elements. The other four problems (see Fig. 2) were designed to have only one equation to show the interaction of informational elements; however, the interaction would be a simple mapping, a flip-over mapping or a rotational mapping of the corresponding sides.

Problem-Solving Sheet

The problem-solving sheet includes two sections, the instructional section and the problem section. The instructional section explains the procedures of the experiment and illustrates how to complete the NASA-TLX questionnaire using a 20-point Likert scale survey.

Background Information and Prior Knowledge Measures

All participants were required to identify their gender, their final mathematics score for the last semester, their first monthly mathematics exam score for the current semester and their first monthly Chinese exam score. These three scores were used to measure participants' prior knowledge.

Cognitive Load Measure

A NASA-TLX questionnaire consisting of six items (mental demand, physical demand, temporal demand, performance, effort and frustration) was presented after each problem (Hart & Staveland, 1988).

RESULTS

Model 1: Effects of Problem, Prior Knowledge Order and Carry-Over on Pass Rate

Table 1 presents pass rates, means (standard deviations), numbers of informational elements, levels of mental operations and levels of information interactions. To test whether the problems had different pass rates, we controlled two error sources: the sequencing of the problems (period) and the effect of correctly solving the previous problem on the pass rate of the next problem (carry over). A multiple logistic regression was also conducted (Allison, 1999). The problem, period, carry-over and prior knowledge were the predictors, and the pass rates were the response. The result revealed that the pass rate can be significantly predicted by the aforementioned four predictors, $\chi^2 = 705.24, df = 30, N = 3,110$, $p < .001$. Upon closer examination, the pass rates were significantly different among the problems #1 to #5, $\chi^2 = 432.07, \mathrm{df} = 9, p < .001$. With regard to prior knowledge, Math had significant effects on the pass rate. For Math 1: $\chi^2 = 26.85, \mathrm{df} = 1, p < .001$; for Math 2: $\chi^2 = 90.02$, $\mathrm{df}=2, p < .001$. The scores related to exams for Chinese had no significant effect on the pass rate, $\chi^2 = 1.60, \mathrm{df} = 1, p = .201$. Furthermore, the period and the carry-over effects were not significant, $\chi^2 = 7.86, \mathrm{df} = 9, p = .55$ and $\chi^2 = 12.24, \mathrm{df} = 9, p = .200$, respectively.

Model 2: The Predictive Effect of Various Cognitive Load Indicators on the Pass Rate

A multiple logistic regression was conducted with the six NASA-TLX items as the predictors and the pass rate as the response to test whether the NASA-TLX could predict the pass rates. The results revealed that the pass rate can be significantly predicted by the NASA-TLX items, $\chi^2 = 668.83, df = 6, N = 3,110, p < .001$. Furthermore, among all six items, the pass rate was significantly predicted by mental demand, physical demand, effort and frustration. The odds ratio of mental demand was .863, implying that every unit increase in the mental demand yielded a 14.8 % decrease in the odds that a participant would correctly solve the problem.

CONCLUSION

The results of study 1 suggest that the pass rates for the five similar triangle problems were not identical when controlling for prior knowledge, problem order and the carry-over effect. Though prior math knowledge was a significant predictor of the pass rate for a problem, the pass rate for problem #5 was significantly lower than the pass rates for the other four problems, which had simpler and clearer configuration patterns of similar triangles. The low pass rate for problem #5 may suggest that the participants had difficulties in configuration comprehension, such as difficulty (1) identifying a hidden informational element for a successful solution and (2) mentally rotating a triangle to form a familiar configuration pattern of similar triangles. In addition, self-reported cognitive load supported the hypothesis that five problems had different levels of perceived cognitive load. The perceived cognitive load ratings of problem #5 were the highest among the five problems. The pass rate would be lower if mental demand or frustration ratings were higher

EXPERIMENT 2

Based on the five similar triangle problems examined in study 1, we used the eye-tracking method to compare online configuration comprehension patterns for successful versus unsuccessful solvers during a realistic, complex cognitive task. Specifically, we wanted to know whether eye movement patterns show sources of cognitive load in configuration comprehension when the participants successfully versus unsuccessfully solve the problem.

METHOD

Participants

Experiment 2 is used to investigate whether eye movement is sensitive to the perceived difficulty while problem solving. To reduce other possibilities (e.g. prior knowledge) that could result in increasing the difficulty in performing the tasks, 63 participants (23 males and 40 females), who were randomly selected from a senior high school in northern Taiwan, participated in this study. All participants were 17 to 19 years of age.

MATERIALS

All materials were controlled by a host computer that presented seven pages of slides, namely slide1 to slide7 for convenience. Slide1 was the practice page, slide2 was an introduction to the procedure of the ongoing task and Slide3 to slide7 were the five problems used in experiment 1.

Apparatus

An EyeLink 1000 desktop remote eye-tracker system (SR Research Ltd., Canada) with a sampling rate of 500 Hz and an accuracy of 0.5 degrees was used to record the eye movement of participants. The experimental materials were implemented on an Intel duo core computer running at 3.0 GHz in a windows XP service pack 2 environment and displayed on a 22-inch monitor (resolution, 1,024×768; refreshing rate, 85 Hz). Participants sat in front of the monitor at a distance of 600 mm. Before conducting the experiment, each participant had to calibrate until gaze durations could be validated. A digital pen and a touch tablet device (Wacom Corporation, Saitama Japan) were used to input participants' written answers. The screenshots of solutions for each problem were saved for subsequent analyses.

Measure

Eye Movement Data (On-line). To describe an entire picture of a participant's visual attention and configuration comprehension of the geometry diagrams, we analyzed the heat maps (shown in Fig. 2). In addition, we designed two areas of interest (AOI), the informational text (Info-t) and the informational diagram (Info-d), to monitor eye movement behaviors of solvers. Based on previous research (Mayer, 2010; Rayner, 1998), we monitored three eye movement indicators, dwell time (DT), fixation count (FC) and run count (RC). Dwell time is the summation of the duration across all fixations on the current AOI, fixation count is defined as the number of fixations falling within the AOI, and run count is defined in terms of the number of times the AOI was entered and left. The DT and FC are used to measure the time that participants spent on the relevant areas. A longer DT or more FC indicated that learners may require more cognitive processes for a specific area (Carpenter & Just, 1978). Both DT and FC are pervasively used in eye movement research on problem-solving (Carpenter & Just, 1978; Grant & Spivey, 2003). The RC is a promising indicator that represents the perceived difficulty associated with a specific area. For example, Hegarty et al. (1992) proposed that some students required more re-readings of previously fixated words for difficult problems. We expected problem solvers to have higher FCs and RCs and longer DTs on those areas they felt were especially difficult.

Perceived Difficulty Measurement (Offline). After solving each problem, the participants were required to complete the self-reporting cognitive load questionnaire. The single item constructed by Paas (1992) adopted a nine-point Likert scale that ranges from 1 (low) to 9 (high). Furthermore, responses to an open-ended question asking for the source of difficulty in solving problems were collected.

Procedure

First, participants were required to sit in front of a computer equipped with a desktop eye tracker and to adjust the seat until the distance between the participant and the monitor was approximately 600 mm. Second, prior to the experiment, each participant was calibrated to ensure the validity of eye movement data. Third, a slide was presented on the screen demonstrating how to use the handwriting device, and participants were allowed to practice on the same screen. Fourth, the definition of similarity and the corresponding formula were presented on a slide and participants were told that these concepts should be used on subsequent tasks. Fifth, the participants began solving the problems using the tablet device while the eye tracker simultaneously recorded eye movements.

RESULTS

Pass Rate and Cognitive Load

A series of one-way repeated-measures ANOVAs was conducted to test whether different problems had different pass rates. The results indicate that the pass rates varied across problems, $F(4, 53) = 29.50$, $p < 0.001$, $\eta^2 = .69$. Multiple comparisons revealed that pass rates were not significantly different among problems #1 and #2 or #3 and #4. In sum, problems #1 and #2 were the easiest, problems #3 and #4 were moderately difficult and problem #5 was of the most difficult.

A multiple logistic regression was conducted with prior knowledge and the level of cognitive load as the predictors and pass rates as the responses to test whether problems with various cognitive loads would have different pass rates after controlling for prior knowledge. The results revealed that the pass rates could be significantly predicted by the cognitive loads, $\chi^2 = 39.18$, $df = 1$, $N=285$, $p < .001$, while no significant effect of prior knowledge was observed, p's > .05. The odds ratio of the cognitive load was .70, indicating that for every unit of increase in the cognitive load rating, there was a 30 % decrease in the odds that a participant would correctly solve the problem.

Eye-Tracking Data

A series of one-way repeated-measures ANOVAs was conducted to test whether eye movement indicators would behave differently across problems. The results showed that dwell time with $F(4, 53) = 19.01$,

$p < .001, \eta^2 = .59$, fixation count with $F(4, 53) = 19.15, p < 0.001, \eta^2 = .59$ and run count with $F(4, 53) = 7.82, p < .001$ were significantly different across problems. A post hoc comparison showed that the participants had longer dwell time and greater fixation count for the more difficult problems (problems #5 > #4, #3 > #2 and #1), and run count showed a similar effect where participants scanned back and forth more frequently on problems with higher difficulty.

A series of Mann–Whitney tests was used to test whether there were differences between successful and unsuccessful problem solvers with respect to eye movement indicators. The heat maps of successful solvers are listed in the left column of Fig. 2, while the heat maps of the unsuccessful solvers are in the right column. We specifically selected the heat maps from various participants to demonstrate the general patterns. The red or orange spots (i.e. heat areas) represent locations where the information was processed for a longer time and processed more deeply by participants, while yellow and green colors represent locations where the information was only minimally processed. The eye movement patterns of unsuccessful problem solvers and the comparison between patterns of successful and unsuccessful problem solvers yields interesting evidence.

We found three characteristics of unsuccessful solvers. (1) They scanned the whole diagram extensively. For example, in problem #1, the whole heat areas are larger for an unsuccessful participant, ID 3140, compared with those of a successful participant, ID 4036. (2) They fixated lengthily at the informational areas of the triangles, such as the length of each side (in problem #1: 12, 9, 8, $\overline{\mathrm{EF}}$), which are the critically known and unknown conditions. For example, in problem #1, the unsuccessful participant, ID 3140, fixed longer, with the heat map showing four red spots compared with the green color area of the successful participant, ID 4036. (3) They looked attentively at the intersection of the triangle pairs showing red-orange colors on the interaction area (i.e. ID 4160 in problem #2, ID 3120 in problem #3 and ID 3070 in problem #4) compared with those of their successful counterparts.

The Relations Between Cognitive Load and Eye Movement Indicators

To investigate the extent to which cognitive loads are associated with eye movement indicators (i.e. FC, DT and RC) on the problem area, the linear relationships between the cognitive loads and the eye movement indicators were measured using Pearson's correlation (Barrett, Morgan,

& George, 2005). Table 2 presents the results of statistical analyses including degrees of freedom, correlation coefficients and *p* values for each problem. Correlation coefficients between cognitive loads and FCs ranged between .312 and .604. Similarly, correlation coefficients between cognitive loads and DTs ranged between .379 and .625. The results suggest that there is a consistent increase in FC and DT on the problem area as the perceived cognitive load increases.

Source of Perceived Difficulty

With regard to the source of the perceived cognitive load, 37 (66.1 %) participants reported that the need to rotate the graph was their primary source of difficulty, and 29 (51.8 %) the inactive/blur of the corresponding concepts was the primary source of difficulty. Other sources of perceived difficulty were relatively low in proportion (less than 20 % each), such as unfamiliarity with the equation, inability to find known conditions or too many unknown variables.

TABLE 2

Correlation coefficients (significant levels) between cognition and eye-movement indicators when problems were solved ($n = 311$)

Problem	*Mental operations*	*Interactions A:B = C:D*	*df*	*FC*	*DT*	*RC*
#1	Straightforward	1 step simple mapping	55	.604***	.625***	.368**
#2	Check correspondence angles	1 step overlap mapping	56	.483**	.467**	.142
#3	Check correspondence angles	1 step flip-over mapping	56	.483**	.471**	.246
#4	Check correspondence angles	1 step vertical mapping	57	.312*	.379**	.110
#5	Need mental rotation to check correspondence angles	2 step rotation mapping	56	.460**	.480**	.130

Mental operations refer to the crucial cognition to find similar-triangle pattern. For the configuration pattern of each problem, please see Fig. 2

*$p < .05$; **$p < .01$; ***$p < .001$

CONCLUSION

The results indicate that more attention and more time were given to reading the more difficult configurations (problem #5) than to reading the intermediate and easier configurations (problems #3 and #4 and problems #1 and #2).

The heat maps, dwell time and fixation counts indicate that the successful participants did not fixate as long as the unsuccessful ones on the areas with crucial information (lengths of corresponding sides) for effectively inferring and solving the problems. This comparison reveals that our participants encoded the configuration and formed visual chunks (clusters of information), a process known as “the pattern recognition stage”, as suggested by Carpenter and Shah (1998), or “the visual integration stage”, as suggested by Ratwani et al. (2008). The illustration of the successful solvers’ fixations (heat maps) did not clearly show any evidence of this processing stage.

In addition, the eye data (run count and heat maps) also showed that they experienced “an interpretive stage” (Carpenter and Shah (1998) in which the pattern was translated into its quantitative and qualitative meanings and this information was related to the referents found in the participants’ memory (prior knowledge of similar triangles). This stage is in accordance with “the cognitive integration stage”, as suggested by Ratwani et al. (2008), whereby the readers formed visual clusters during a previous visual integration stage that served as object-like units that could then be used to reason about the graph during cognitive integration. Both run counts and heat maps indicated that our participants scanned back and forth around several visual chunks (informational areas). The heat maps of the unsuccessful problem solvers were more evident than those of the successful problem solvers in demonstrating the cognitive interpretation stage in that they scanned more extensively the whole diagram and attended to the intersection of overlapping, flip-over, upside-down and rotated triangles. We suggest that unsuccessful solvers have more difficulties in forming spatial relations of visual chunks (corresponding angles and sides), translating the pattern into its quantitative and qualitative meanings and relating this information to the referents in their memories.

General Discussion

The present study designed five geometry problems and controlled the extraneous cognitive loads (Paas et al. 2003a). For example, the layout (relative distance and size) of the diagrams were designed to be the same, and the areas of the figures were approximately equal. Two types of cues, lengths of the sides or degrees of the angles, were conditionally provided and closely written to the corresponding sides or angles on the diagrams. The participants were required to solve some unknown sides according to the principles of similar triangles. Regarding the cognitive load (Sweller et al., 1983), three characteristics, including the number of informational elements, the levels of element interactions and the levels of mental operations, were assumed to account for the increasing difficulty. In study 1, we proved that problem #5 was the most difficult (lowest pass rate) and required the greatest amount of effort (largest cognitive load) when prior knowledge, problem order and carry-over effects were controlled. The major cognitive load sources in reading problem #5 were assumed to be the configuration comprehension difficulties in which one of four known conditions (lengths of corresponding sides in a pair of similar triangles) was hidden and the problem solvers had to rearrange the spatial relation of two similar triangles (i.e. to mentally rotate one of the triangles).

Using these five problems in study 2, we investigated what cognitive load sources students explicitly noticed and whether eye movement patterns revealed sources of cognitive load in configuration comprehension for various difficulty levels of problems and successful levels of performance. The self-reported cognitive load sources included the difficulties in viewing the diagrams to connect with the relevant concepts (fundamental properties of similar triangles and the length ratio of corresponding sides) and the limitations of cognitive functions to mentally manipulate the paired triangles. This result was supported by the evidence that the participants spent more time reading the most difficult configurations of problem #5 than they did reading the configurations of the intermediate and easier problems.

With respect to configuration comprehension, the participants experienced three stages: (1) visual integration, (2) spatial relation identification and (3) cognitive interpretation. Because the successful solvers were very swift in pattern recognition and visual integration for a coherent representation of similar triangles, the illustration of their fixation (heat maps) did not clearly show evidence of the processing stages. First, the comparison of the eye data between successful and unsuccessful problem solvers demonstrated that in the visual integration stage, unsuccessful problem solvers look intensively at the areas with critical information (lengths of corresponding sides), as suggested by Carpenter and Shah (1998) and Ratwani et al. (2008). Second, because of the particular design, to understand some configurations (especially problem #5), the problem solvers must scan to organize meaningful spatial relations among visual chunks or information areas, which is in accordance with the spatial processing described by Gillan (1995). We assume that if the spatial relation of visual chunks is not a difficult source, the eye data may not clearly reveal the processes involved in this stage. Finally, the visual clusters formed during the previous visual integration stage served as object-like units that could then be used to reason about the graph during cognitive integration. Both the run count and heat maps indicated that the participants scanned back and forth around several visual chunks (informational areas). The heat maps of the unsuccessful problem solvers were more evident than the successful ones in demonstrating the cognitive interpretation stage in that they scanned the whole diagram more extensively and attended to the intersection of overlapping, flip-over, upside-down and rotated triangles.

In designing geometry diagrams, designers and teachers must be aware of the sources of the extraneous load as well as the intrinsic loads. An effective instructional approach is to reveal as many sources of the intrinsic load as possible. In this study, we proved that the number of informational elements, the level of element interactions, and the level of mental operations are three major sources of cognitive loads in configuration comprehension, especially with respect to geometry problem solving. Another cognitive load comes from the fact that readers have to experience all three stages, that is, visual integration, spatial relation identification and cognitive interpretation. Thus, an effective instructional approach could include the provision of a variety of worked examples (Sweller et al., 1983) for beginning students. The diagrams in the worked examples must be well designed and vary in the number of informational elements, in the levels of element interactions and in the levels of mental operations. Teachers could then guide students through the three stages of configuration comprehension (i.e. visual integration, spatial relation identification and cognitive interpretation) using the worked examples.

Limitation

Because the aim of the present study was to investigate the relationship between perceived difficulty and eye movement, to reduce other possibilities (e.g. low prior knowledge) that could result in increasing the perceived difficulty in performing the tasks, students at grade 11 were chosen to participate in the second experiment. Although the statistical results indicate prior knowledge had no significant impact on eye movement, there is the possibility that it may be due to the effect of sample size. Further research can improve this by employing a larger sample size.

ACKNOWLEDGMENTS

This research was funded by the ROC National Science Council through the project "Reading math and science from static and motion visual displays in multimedia virtual worlds: The associations of mental model, cognitive load and eye-tracking" (NSC 100-2511-S-009-007-MY3) to the corresponding author.

REFERENCES

Allison, P. D. (1999). *Logistic regression using the SAS system: Theory and application*: SAS Publishing.

Barrett, K. C., Morgan, J. & George, A. (2005). *SPSS for intermediate statistics: Use and interpretation*: Psychology Press.

Bobis, J., Sweller, J. & Cooper, M. (1994). Demands imposed on primary-school students by geometric-models. *Contemporary Educational Psychology, 19*(1), 108–117.

Boucheix, J. M. & Lowe, R. K. (2010). An eye tracking comparison of external pointing cues and internal continuous cues in learning with complex animations. *Learning and Instruction, 20*(2), 123–135. doi:10.1016/j.learninstruc.2009.02.015.

Carpenter, P. A. & Just, M. A. (1978). Eye fixations during mental rotation. *Eye movements and the higher psychological functions, 115–133*.

Carpenter, P. A. & Shah, P. (1998). A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology-Applied, 4*(2), 75–100. doi:10.1037/1076-898x.4.2.75.

Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293–332.

Epelboim, J. & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research, 41*(12), 1561–1574. doi:10.1016/S0042-6989(00)00256-X.

Gillan, D. J. (1995). Visual arithmetic, computational graphics, and the spatial metaphor. *Human Factors, 37*(4), 766–780. doi:10.1518/001872095778995571.

Grant, E. R. & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*(5), 462–466. doi:10.1111/1467-9280.02454.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human Mental Workload, 1*, 139–183.

Hegarty, M. & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language, 32*(6), 717–742. doi:10.1006/jmla.1993.1036.

Hegarty, M., Mayer, R. E. & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixation. *Journal of Educational Psychology, 84*(1), 76–84. doi:10.1037/0022-0663.84.1.76.

Hegarty, M., Mayer, R. E. & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87*(1), 18–32. doi:10.1037/0022-0663.87.1.18.

Holm, A., Lukander, K., Korpela, J., Sallinen, M. & Muller, K. M. (2009). Estimating brain load from the EEG. *The Scientific World Journal, 9*, 639.

Hyona, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction, 20*(2), 172–176. doi:10.1016/j.learninstruc.2009.02.013.

Just, M. A. & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review, 92*(2), 137. doi:10.1037/0033-295X.92.2.137.

Laborde, C. (2005). The hidden role of diagrams in students: Construction of meaning in geometry. *Meaning in Mathematics Education*, 159–179.

Larkin, J. H. & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*(1), 65–100. doi:10.1016/S0364-0213(87)80026-5.

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction, 20*(2), 167–171. doi:10.1016/j.learninstruc.2009.02.012.

Meyer, K., Rasch, T. & Schnotz, W. (2010). Effects of animation's speed of presentation on perceptual processing and learning. *Learning and Instruction, 20*(2), 136–145. doi:10.1016/j.learninstruc.2009.02.016.

O'Leary, Z. (2004). *The essential guide to doing research*: Sage.

Ozcelik, E., Arslan-Ari, I. & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26*(1), 110–117. doi:10.1016/j.chb.2009.09.001.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429–434. doi:10.1037/0022-0663.84.4.429.

Paas, F., Renkl, A. & Sweller, J. (2003a). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1–4.

Paas, F., Tuovinen, J. E., Tabbers, H. & Van Gerven, P. W. M. (2003b). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*(1), 63–71.

Paas, F. & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*(1), 122–133. doi:10.1037/0022-0663.86.1.122.

Paas, F. & van Merriënboer, J. J. G. (2006). Optimising worked example instruction: Different ways to increase germane cognitive load. *Learning and Instruction, 16*(2), 87–91.

Park, B., Moreno, R., Seufert, T. & Brunken, R. (2010). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*.

Plass, J. L., Moreno, R. & Brünken, R. (2010). *Cognitive load theory*: Cambridge University Press.

Ratwani, R. M., Trafton, J. G. & Boehm-Davis, D. A. (2008). Thinking graphically: Connecting vision and cognition during graph comprehension. *Journal of Experimental Psychology-Applied, 14*(1), 36–49. doi:10.1037/1076-898X.14.1.36.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. doi:10.1037/0033-2909.124.3.372.

Schwonke, R., Renkl, A., Salden, R. & Aleven, V. (2010). Effects of different ratios of worked solution steps and problem solving opportunities on cognitive load and learning outcomes. *Computers in Human Behavior*.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. doi:10.1016/0364-0213(88)90023-7.

Sweller, J. Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*: Springer.

Sweller, J. & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*(4), 351–362.

Sweller, J. Mawer, R. F. & Ward, M. R. (1983). Development of expertise in mathematical problem solving. *Journal of Experimental Psychology-General, 112*(4), 639–661. doi:10.1037/0096-3445.112.4.639.

Xu, H. (2006). Design and analysis of experiments, volume 2: Advanced experimental design. *Technometrics, 48*(4), 571–571. doi:10.1198/tech.2006.s433.

National Chiao Tung University
Hsin-Chu, Taiwan, Republic of China
E-mail: sunnylin.nctu@gmail.com