Developing and Refining Mental Models in Open-Ended Learning Environments: A Case Study

☐ Kevin Oliver
Michael Hannafin

This qualitative case study focused on the nature of science learning through open-ended problem solving. Twelve eighth graders were asked to find, frame, and resolve subproblems associated with structural failures resulting from earthquakes. Coded interviews, artifacts, and observations from the four-week study suggested students only partially derived accurate mental models about earthquake engineering problems. Recommendations for improving student problem understanding in open-ended environments include the explication of student hypotheses related to problems, and the continual testing of belief via analogical reasoning, research, communication, and tool use.

☐ Traditional instructional approaches typically organize and present information consistently with what experts judge as correct or accurate; students, in turn, are expected to adopt this standard as their own. However, recent evidence suggests that students find expert conceptions difficult to comprehend-especially in the teaching and learning of science (Snir & Smith, 1995). Students often hold naive theories about the nature of everyday as well as scientific phenomena, making the adoption of expert conceptions problematic (Land & Hannafin, 1996). Constructivist science teaching and learning approaches, featuring open, student-centered inquiry and authentic problems, may offer advantages over classrooms that emphasize rote memorization of expert conceptions. In student-centered environments, individuals interpret and synthesize information to generate original solutions. Personally relevant problem understanding emerges within an individually constructed mental framework (Hannafin & Land, 2000).

Open-ended learning environments (OLEs) typically include four components: (a) enabling contexts, (b) resources, (c) tools, and (d) scaffolds (Hannafin, Land, & Oliver, 1999). Enabling contexts provide realistic frameworks wherein problems are situated; resources allow students to frame and resolve problems; tools assist students in processing, manipulating, or discussing information; teacher and tool-based scaffolds guide learners' problem-solving strategies or processes. Open-ended learning contexts provide authentic science problems and tasks along with scaffolding questions that guide student inquiry (Hannafin, Hall, Land, & Hill, 1994). As students seek to resolve problems, they establish, revise, and advance their understanding using available resources and tools with which to manipulate ideas. Students evolve and induct "expert-like," canonical models of a system or concept as they progressively fine-tune their understanding rather than simply adopting, complying with, or adhering to external expectations (cf. McCaslin & Good, 1992).

While the four components of open-ended learning are unique in their combination (i.e., contexts, resources, tools, and scaffolds), each is associated with curricular reform movements in science education. The Science, Technology, and Society (STS) movement advocates the use of interesting problems or contexts, preferably of local interest, and student research with material and human resources to inform and solve those problems (National Science Teachers Association, 1990). Technology resources or tools are also recommended in STS modules to support processes such as communication between instructors, peers, and external human resources (Daas, 1999). The learning cycle teaching method also recommends the use of an open-ended problem during an exploration phase, followed by teacher appraisal of student understanding to scaffold the level at which basic concepts are introduced and later applied (Barman, 1996). The learning cycle is a type of discovery learning that emphasizes skills over content and processes over products (Bruner, 1966; Heywood, Heywood, & Donovan, 1992; Howe & Jones, 1993). Discovery learning emphasizes inquiry-oriented objectives over traditional content-oriented objectives, and depends heavily on the instructor to scaffold student understanding through such techniques as "Socratic" questioning (Hammer, 1997). These reforms seek to develop students who are scientifically literate on multiple dimensions, moving toward the ability to use a scientific vocabulary, apply procedural knowledge and skills, and conceptualize science within a larger social context (Bybee, 1996).

The study of mental models has yielded important perspectives as to how and why open-ended learning systems work (Calvi, 1997; Dimitroff, 1992; McGregor, 1994; Pitts, 1995; Sutton, 1994). A mental model is a representation of one's personal understanding of a system or concept. It may consist of declarative knowledge (i.e., propositions and schemas), procedural knowledge (i.e., rules and skills), or assumptions (Kyllonen & Shute, 1989). Functionally, mental models are used both to explain and predict various phenomena (Vosniadou, 1994); structurally, they are dynamic and can be spawned instantaneously to reflect emerging representations of complex and novel situations (Halford, 1993; Johnson-Laird, 1990; Wiser, 1995).

Mental models are often structurally ingrained and thus resilient to change because they are tied to "everyday experience and . . . years of confirmation" (Vosniadou, 1994, p. 49). Even children as young as 3–5 years of age develop naive theories based on their informal observations and personal sense-making of everyday phenomena such as rain, fire, and stars (Glynn & Duit, 1995; Isaacs, 1974). Children rarely enter the classroom with expert-like understanding; rather, they have informal experiences and prior knowledge, often naive and incomplete, related to concepts to be studied. When given problems and answers from an expert point of view in formal education settings, efficiency may be attained but coherence is often sacrificed (Wiser, 1995). Students often encounter difficulty deriving deep understanding via traditional didactic approaches, such as lectures and worksheets (Land & Hannafin, 1997). When provided open-ended situations and appropriate scaffolding, however, students can gradually reconcile their naive models with those of experts. Hybrid mental models emerge that reflect formative detail and assumptions, models used to formalize and test beliefs. Student-apprentices in such environments, with the tutelage of teachers and other mentors, help to refine and fine-tune these formative models as their understanding deepens (Baxter, 1995; Brown, Collins, & Duguid, 1989; Vosniadou, 1994).

In constructivist learning environments, the goal of encountering and addressing student's naive theories and models forms the basis for designing and implementing lessons and activities. Naive mental models are coherent for the learners who hold them, thus experiences that serve to illustrate their limitations are critical to promote heightened awareness and revision (Hannafin, Hannafin, Land, & Oliver, 1997). Contemporary researchers have designed numerous constructivist computer systems to improve science instruction, including many emphasizing problem solving. These systems may emphasize contextualized problems such as Jasper Woodbury (Cognition and Technology Group at Vanderbilt, 1992), visualization tools for open-ended collaborative inquiry such as CoVis (Edelson, Pea, & Gomez, 1996), or tools for socially constructing knowledge such as the computer-supported intentional learning environment (CSILE) (Scardamalia et al., 1992).

Recently, researchers have explored the potential of technology to create, test, and revise student mental models. Some tools have enabled students to externalize their mental models visually in the form of graphical relationships among the elements or nodes of a concept (see, e.g., TextVision, Kommers & DeVries, 1992, and SemNet, [Fisher, 1992]). Model-It contains a suite of tools through which students can progress from simple conceptual representations of presumed relationships to formalisms that permit the testing of the assumptions underlying a student's mental model (Stratford, Krajcik, & Soloway, 1998). Model It allows learners initially to define connections between related factors, then to analyze the model for its validity in representing those relationships, and eventually to revise and retest the models as evidence confirms or refutes their assumptions (Jackson-Metcalf, Krajcik, & Soloway, 2000).

This paper reports the findings of a qualitative case study using an on-line OLE. The case study design was "embedded" with specific units or grouped pairs of students rather than holistic in focus regarding the "global nature of a program" (Yin, 1994, p. 42). The goal of the study was to examine the influence of technology-mediated scaffolding and tools on the learning and mental model development of students who engage open-ended science problems.

METHOD

Participants and Setting

The participants were 1 teacher and 12 students in an eighth-grade science class at a rural middle school. The students included 6 females and 6 males, all White, and approximately 13 years of age. Most had limited prior experience with computers, and both the teacher and students were only beginning to use technology as a tool in everyday science teaching and learning. Participant selection was purposive rather than random: It represented a group of students being introduced to the scientific concepts under study for the first time. The research was conducted in a computer lab with 24 Macintosh computers that were not connected to the Internet or to a local area network. WebWhacker software (Blue Squirrel, 2000) was used to download more than 75 Web sites that were then stored locally on each computer. The students' regular science classroom was used to conduct interviews when not occupied by classes.

Materials

To facilitate finding, framing, and resolving open-ended science problems, students were given access to and training in the use of several tools from the "knowledge integration environment" (KIE) (Slotta & Linn, 2000). KIE is an online system designed to help K–12 students acquire the skills needed to understand complex scientific ideas. KIE was selected because of its significant development history, demonstrated applicability for science in the middle grades, and open-ended affordances: problem context, resources, scaffolds, and tools. It enables students to build on their intuitions about science and then test, revise, and reformulate their ideas. KIE is based on the scaffolded knowledge integration framework (SKIF): identifying new goals for learning, making thinking visible, encouraging lifelong learning, and providing social supports (Linn, 1995).

While engaging in scientific exploration via KIE, students begin by formalizing their positions on a scientific problem in the form of a question; that is, they externalize a mental representation of the problem and candidate solution. Students review available resources on the World Wide Web (WWW), guided by the online scaffolding, peers, and the teacher, to determine their relevance to their hypotheses. They then identify and analyze evidence regarding their position, and select and present evidence in the form of a scientific argument to support their revised position.

The instructors selected related Web pages to initiate student inquiry into an open-ended problem. We placed the same 13 Web pages on each computer to introduce students to the many complex variables involved in earthquake engineering. Based on these initial Web resources, students were exposed to and could choose to solve any of the following problems:

- How to build on soils of varying density
- Which materials to select in the design of a building
- How to reinforce or brace a building
- How to construct columns to support weight
- How best to lay out a building's floorplan.

The initial Web sites contained background information designed to catalyze student inquiry, but did not isolate correct answers to the problems posed. Rather, they contained the types of information that, given a range of solution alternatives, might be analyzed to clarify, confirm, or contradict the assumptions underlying a student's problem frame. In addition, books, magazine articles, and 65 other Web sites, providing further information on basic earthquake and engineering-related problems, were placed on each computer. The students' task was not to find the correct answer among the available resources, but to generate, test, and revise their approaches using the resources as reference material.

Two KIE tools are especially significant in this study: Sensemaker and Mildred. Students use Sensemaker to sort Web evidence into categories and integrate it into problem-related categories (see Figure 1), providing a means through which students organize and classify problem-related information per their emerging conceptual understanding. Students can create clusters, add or delete evidence, and modify the clusters. Mildred allows students to record notes for specific activities (e.g., generate research questions) and to annotate or summarize Web evidence (see Figure 2). Students also rate the relevance of the evidence gathered to their problem and solution as they evaluate various resources. As students progress through different project activities and access different Web pages, new Mildred windows appear in which to store their notes. Conceptual scaffolding questions were placed in the student note-taking window to help students focus on key concepts (e.g., "What is the problem described by this Web page?").

We provided additional scaffolding in the form of procedural prompts and metacognitive hints using the corresponding KIE affordance. For example, procedural prompts provided instructions for completing specific activities, including "What additional information do you need to understand your earthquake problem?" "Ask yourself, what scientific issues are related to your earthquake problem?" and "Generate as many scientific questions as possible." Metacognitive hints suggested appropriate strategies for working on a specific activity, such as "One way to design your Web pages is to present your solution on the main home page, then link out to evidence that supports or helps explain why your solution is needed." The teacher provided similar procedural and metacognitive scaffolding during the students' daily work, but in order to observe the emergence of mental models through tool use and research, she purposefully did not tell students during the study if their emergent ideas and thinking were incomplete. Feedback focused on how to best accomplish specific tasks and strategies for self-assessment of understanding.

Finally, the KIE toolbar contained buttons to external tools for students to create design drawings and Web pages. Although extensive in-class discussion of student solutions took place, KIE's communication tool—designed to aid in the exchange of feedback and ideas among students and teachers—was not utilized. The computers were not networked, rendering useless those KIE features designed to promote on-line collaboration between and among participants.

Instruments

Several methods and instruments may be utilized to depict student mental models, including interviewing students prior to and following an intervention (Wiser, 1995), asking students to think aloud while performing a task (Kelly, 1995), reas-

Figure 1 ☐ KIE's Sensemaker tool illustrating an organizational scheme for a select set of Web sites by one student group.

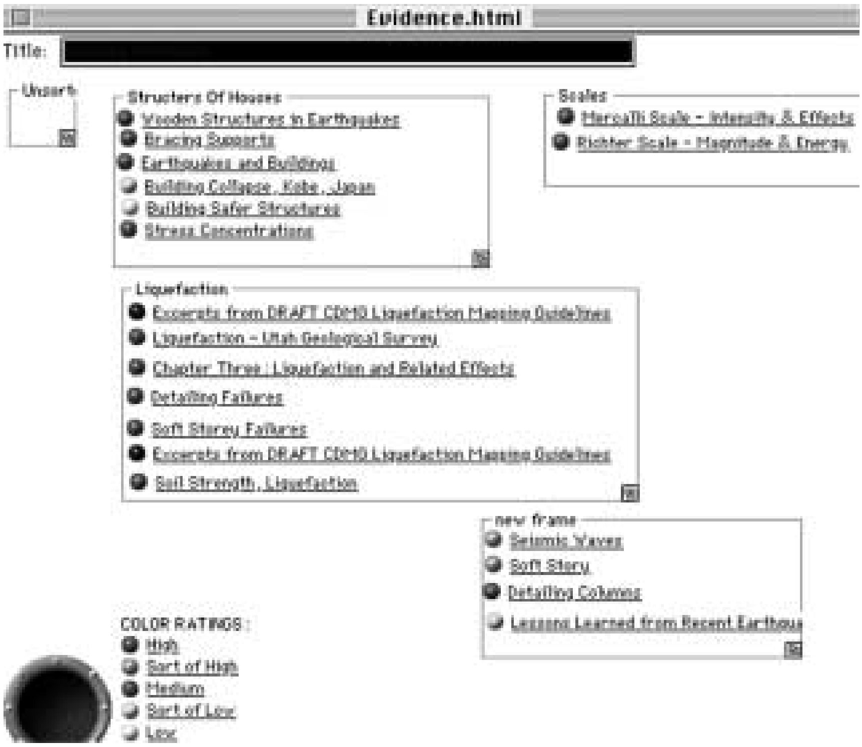
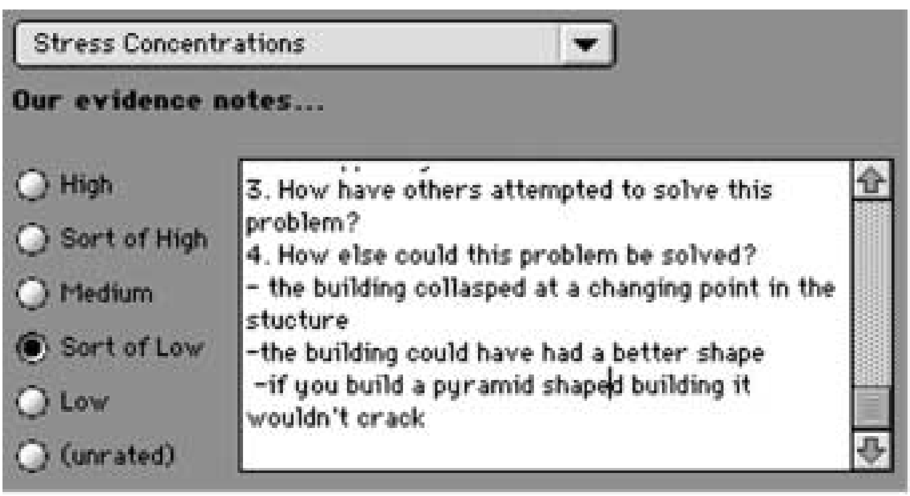


Figure 2 ☐ KIE's Mildred note-taking window for a specific Web page, Stress Concentrations, illustrating conceptual question prompts from the instructor and replies from a student group.



sembling a set of cards with concepts written on them (Rowe & Cooke, 1995), entering thoughts in a journal (Harel & Papert, 1993; Yager, 1995), filling in a blank template or structure (Naveh-Benjamine & Lin, 1991), and drawing pictures or concept maps of understanding (Glynn, 1997).

In this study, an ordered tree instrument was administered after the study to depict students' conceptual structuring of core study concepts (Naveh-Benjamine & Lin, 1991; Naveh-Benjamine, McKeachie, Lin, & Tucker, 1986). This instrument was selected because of its efficient nature and computer-based algorithm for auto-generating expressions of student understanding (Reitman & Rueter, 1980). The ordered tree technique allows student understanding to be compared to an expert model; more importantly, students' personal models or categorization schemes can be analyzed individually. To utilize the ordered tree technique, the instructor selects and groups a set of course concepts (see Table 1), then prepares a set of four pages containing the same set of concepts in scrambled order (see Figure 3 for one page or trial). Students are asked, "to arrange the concepts in a vertical order so that concepts closely related in terms of their meaning in the course appear close to each other" (Naveh-Benjamine & Lin, 1991, p. 10). Students are not asked to arrange concepts hierarchically; the computer algorithm interprets this structure across multiple term-arranging trials. At least four trials are recommended to provide the computer analysis software with enough data to generate an expression of student understanding. During the study, each student ordered four pages with the

Table 1 ☐ Sixteen study terms given to students for ordering with appropriate super- and subordinate relationships illustrated

- Seismicity (I)
    - Richter (D)
        - Magnitude (L)
    - Mercalli (N)
        - Intensity (H)
    - Body Waves (E)
        - Compressional (A)
            - Primary (M)
        - Shear (K)
            - Secondary (G)
- Liquefaction (C)
    - Soil (O)
- Earthquake Engineering (J)
    - Columns (B)
        - Detailing (F)
    - Soft Story (P)

Table 2 ☐ Interview questions used to probe students' science understanding

Activity	*Questions Asked*
Problem finding	• What scientific issues did you find to focus on as your major problem? • Was it hard to select that issue? • Did any of the tools help you reflect and make a decision? • What have you learned about topic *X* so far that you didn't know before? • What do you think you know but are not sure about? • Where did you get your information? • What questions do you still have?
Problem framing	• Have you found information on one site that is similar / different from another site? • What have you learned about your earthquake problem since the last time we spoke? • What has surprised you about the topic? • Can you make any recommendations about how to best deal with this problem? • How could you conduct an experiment to test this / these ideas?
Problem resolving	• What was your final recommendation or solution for your topic? • Did any computer tools help you make these decisions? • Did any of the tools help you defend or support your ideas? • What evidence did you find to support your position? • Have you learned anything new about topic *X* since the last time we spoke? • Do you think the problem has been adequately addressed by your solution? • After sketching your ideas and making web pages describing your ideas, would you make any changes to your solution?

Figure 3 ☐ Sample ordered tree trial using 16 concepts selected from course materials.

Put related terms in the right order on the lines. For example, if “C-rabbit” and “F-Easter” were listed in the group below, you would write those terms next to each other on the lines because they are related.

F - detailing	N - Mercalli	J - earthquake engineering	P - soft story
M - primary	O - soil	G - secondary	L - magnitude
I - seismicity	C - liquefaction	E - body waves	B - columns
K - shear	D - Richter	A - compressional	H - intensity

Example:

F Easter

C rabbit

Trial 1 of 4

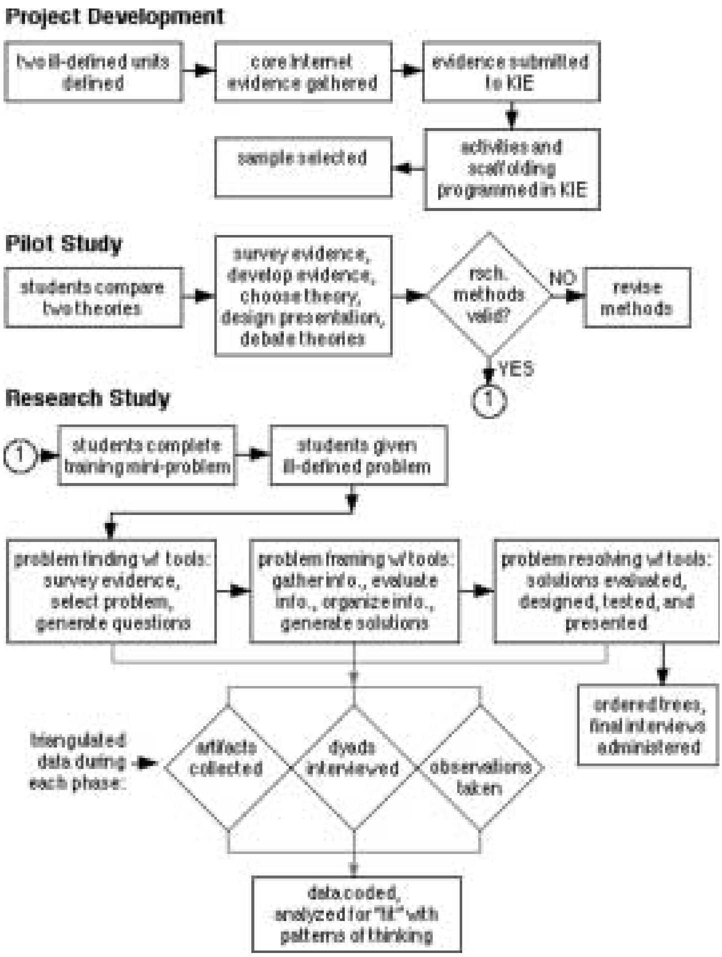
same concepts, one per day over four separate days. Each trial took approximately 10 minutes. Computer expressions were generated following, not during, the study. With limited prior exposure to the 16 listed concepts, we did not expect students to understand concept relationships prior to the study.

A semistructured interview protocol was also developed for student interviews. Three separate protocols were used for interviews following student completion of activities during problem finding, framing, and resolving. Questions were related to students’ tool use and thinking, science understanding, contextual or environmental issues, attitudes, and distribution of workload. Questions related to science understanding, the focus of this paper, are listed in Table 2.

Procedure

Figure 4 provides a visual overview of the procedures used during project development, during the pilot study, and during the formal research study. Prior to the research study, a pilot study and additional training were completed. The pilot study involved students in a debate of two conflicting earth science theories: catastrophic versus uniformitarian earth change. The pilot study was conducted over three weeks to train students on the tools and to

Figure 4 ☐ Flowchart of research procedures.



validate research instruments and methods. In addition, the pilot study was used to test and refine the observation methods and protocol and to establish the trustworthiness of the observations. While the researcher (observer) and the teacher (participant) were present each day during the pilot and formal studies, their roles were sufficiently different that it was not possible to match observations to establish interrater reliability. Instead, they met regularly during both studies to review the researcher's data and conclusions.

A two-day training phase was then held just prior to the research study to reintroduce students to the KIE tools they had used during the pilot a few weeks earlier. The training phase was intended to ensure that students knew how to access and use the available resources, tools, and scaffolds, and to acclimatize both the teacher and students to technology-enhanced teaching and learning in the computer laboratory.

Prior to the formal study, the teacher initially grouped the 12 students into heterogeneous high-low ability dyads based on seventh-grade Iowa Test of Basic Skills (ITBS) reading scores, because reading was expected to be an important skill for extracting information from Web resources. Interpersonal conflicts surfaced shortly after the study began. Therefore, individuals in two dyads were reassigned. This resulted in four heterogeneous and two homogeneous dyads (one high- and one low-reading ability).

Each dyad was provided two computers: one with the all the Web pages, the other with the Web pages and the KIE tools. This was done to increase the probability that dyads would use the tools collaboratively as they attempted to find, frame, and resolve their problems. Students in each dyad were designated as *Number 1* or *Number 2* and had alternate computer responsibilities each day. If "Number 1" was announced at the start of class, the corresponding students would operate the KIE tools; the dyad partner used the other computer containing only the Web resources.

The research study was conducted over four weeks. Students spent approximately one hour each day in the computer lab. An open-ended design problem was posed to each dyad: "How can building collapse be prevented during earthquakes?" This complex, macro problem encompasses several underlying micro problems, including liquefaction and soil quality, strength of building materials, external building layouts, internal building designs, column detailing, and bracing. Students were asked to find one relevant subproblem using the database of 13 Web sites, frame their selected problem with 65 additional Web sites and 45 print resources, and then resolve their problem with a sound solution plan. Rather than teaching students explicit earth science facts related to earthquakes (e.g., types of earthquake waves), students acquired concepts and content as they worked through the process of framing their self-selected problems. The goal was not for students to memorize a defined body of knowledge, but to learn and apply relevant information to their specific subproblems. Student presentations took place during the fifth week.

Before beginning their work, students were given a list of expected performance criteria (see Table 3). To reinforce student use of these criteria, they were also embedded in the KIE tools as both procedural instructions and metacognitive hints. The same criteria were used by the researcher as guidelines during classroom observations and by the teacher to assess student performance.

Students spent an average of six to seven days completing specific activities during the three phases: (a) problem finding, (b) framing, and (c) resolving. During problem finding, students were asked to survey the 13 core Web pages in the database, recording notes with their Mildred tool and sorting Web evidence into categories that might suggest a subproblem with their Sensemaker tool. After surveying the initial Web evidence, dyads were asked to select one problem for in-depth focus. Finally, dyads were asked to generate relevant science questions for the hazard they selected (e.g., Why do buildings on sandy soil collapse more often during earthquakes than those on bedrock?).

During problem framing, dyads were asked to frame their own questions by individually researching the additional print and Web resources provided. The Web sites were in no particular order, varied in size from 1 to 80 pages, and ranged in quality from highly valid research findings to suspect corporate advertising. To facilitate their research, students could incorporate new Web links into their existing Sensemaker diagrams and record new notes using Mildred. After framing their questions, dyads were asked to brainstorm potential solutions for their subproblem.

During problem resolving, dyads were asked to select their best solution option. A solution evaluation form provided scaffolding to help students identify quality or limiting aspects in their initial ideas. Dyads were asked to create design drawings and an experimental plan to test their selected solution. These activities were designed to help the dyads detail and elaborate their solutions. Finally, they were asked to develop a presentation in the form of Web pages to describe their solution to others. Each dyad presented its Web pages to the class; then the teacher and other students were permitted to ask questions and suggest alternative designs.

To facilitate analysis, artifacts were collected at the end of each major project phase. This method was intended to document the evolution of student products (e.g., notes, Sensemaker diagrams) that may have elaborated or changed over time through tool use. Each dyad was interviewed for approximately 25 min at the end of the three project phases (three interviews per dyad) so they could reflect on processes in each dimension. Two unstructured teacher interviews were conducted at the end of students' problem framing and resolving phases. Questions from the student interview protocol were used to prompt the teacher to reflect on evidence

Table 3 ☐ Performance criteria and observation guidelines

Students, we are about to begin a design project in which you must create a product to solve a problem associated with plate tectonics. You will select a problem to work on, gather scientific evidence to help you make decisions about how to address that problem, and finally design a product as a solution to the problem. Here is what we expect of you during this design project:

Processes	*Solutions*
• You will seek and find information related to your problem. • If you can't find the information you need, you will ask for assistance. • You will organize and keep track of your information, breaking it into categories (consider who, what, when, where, why categories, and others based on scientific concepts that relate to your problem). • You will consider, "how truthful or useful is this information" (who is the author, do others agree with it, is it based on science or research, when was it developed). • You won't use information that seems untruthful or is impractical in some way. • You will describe how other people have tried to solve this problem. • You will summarize how the information you find helps you solve the problem.	• You will generate many possible solutions. • You will describe how these solutions address the problem. • Your solutions will address different aspects of the problem. • Your solutions will be original and not something someone else has done. • Your solutions will be practical or "doable" (consider time needed, money needed, materials needed). • You will select "the best" solution to work on and describe why it is the best. • You will design one solution to address your problem (sketches or a model). • You will decide how to test your solution. • You will create a Web page that describes your solution, how it may be better than other solutions, and how it may be limited.

of student thinking, science understanding, and attitudes. All interviews were recorded on audiotape and transcribed. Further, all students completed four ordered tree trials, one per day over four consecutive days, to determine how they conceptually arranged core study concepts.

Data Sources

Because students worked in dyads, some data sources represented a pair and some the individual within the pair; both types of data were gathered. During each class period, field notes were taken on individual student tool use, attitudes, and environmental factors. Further, each dyad was interviewed and student artifacts were collected at the end of problem finding, framing, and resolving. The eight artifacts collected are listed in Table 4 along with the approximate time they were collected. The problem to be addressed and the related artifacts were generated and analyzed by dyad; individual student activities and data sources were designed to reveal and interpret within-dyad variability within the technology-enhanced problem-solving activities.

Data Analysis

The TIGER data analysis program (Hirtle, 1980) developed for the ordered tree technique was utilized to enter raw data and output a computer expression of student concept organization (see Figure 5). The analysis software "capitalizes on people's tendency to list all items in one chunk of information before moving on to the next chunk"; an algorithm is used to find the sets of information "chunked" by students and express these as an "ordered tree" or "a student's knowledge structure from a course" (Naveh-Benjamine & Lin, 1991, p. 14). For each student, the ordered tree output was examined for evidence of correctly chunked associations (e.g., student A correctly chunked or associated the term *soil* with the term *liquefaction*). Output was also examined for evidence of correctly identified hierarchical relationships (e.g., student *B* correctly chunks Mercalli scale, Richter scale, magnitude, and intensity together, *and* places them in the proper super-subordinate relationships).

To interpret the computer expression in Figure 5, visual diagrams were generated to display both the chunked terms and the hierarchical term relationships identified by each student.

Table 4 ☐ Artifacts collected and their association with the three study phases

Artifact	Description and When Collected
1. Research questions & answers	During problem finding, all students wrote science questions as activity notes with their Mildred tool to guide them strategically in searching for specifics.
2. Argument graphs	During problem finding and framing, students could use the Sensemaker tool to sort the 13 original Web sites and any additional self-selected Web evidence into problem-related categories.
3. Notes	During problem finding and framing, students could use their Mildred tool to record notes about evidence and to describe potential solution options.
4. Selected evidence	During problem framing, students could select and reference additional Internet evidence or off-line resources.
5. Evaluation forms	During problem framing, students could choose to fill out information-quality evaluation forms to help evaluate self-selected Web evidence and a solution evaluation checklist to help select their best solution option.
6. Design drawings	During problem resolving, students were asked to use ClarisWorks to create design sketches for their inventions.
7. Testing plan	During problem resolving, students could use Mildred to write testing or experiment plans for their inventions.
8. Web page presentation	During problem resolving, students were asked to use the Netscape Web-page editor to create Web pages for presenting their ideas to others.

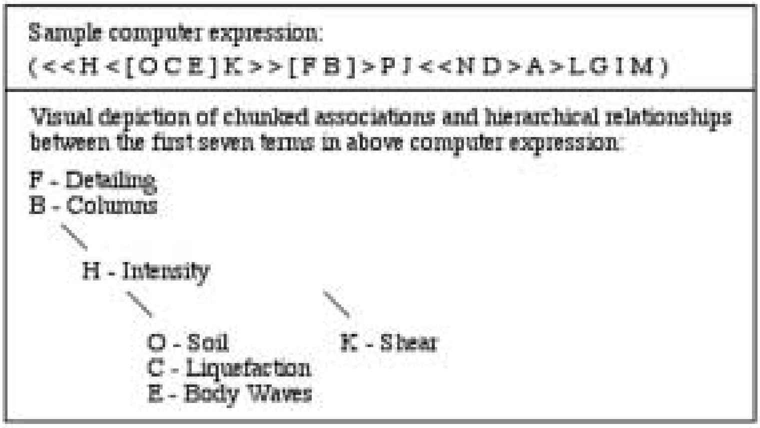
Since the first seven terms lie within a complete set of caret characters, three to the left and three to the right (< < < > > >), we determine that this set of terms represents one stand-alone tree of conceptual understanding. We further determine that terms *F* (detailing) and *B* (columns) are chunked together, because they are shown within brackets in the computer expression. We determine that terms *F* and *B* are superordinate to the single term *H* (intensity), because terms *F* and *B* lie within the bounds of one caret character or one hierarchical level, while term *H* lies within the bounds of two caret characters or two hierarchical levels. We further determine that term *H* at two carets deep, is superordinate to terms *O* (soil), *C* (liquefaction), *E* (body waves), and *K* (shear), at three carets deep. At the lowest hierarchical level, we determine that terms *O*, *C*, and *E* are chunked together because they are shown within a pair of brackets, but they are not associated with term *K* outside of the brackets.

Analysis of interview transcripts and student artifacts transpired via coding and reduction, and then display and pattern seeking (Miles & Huberman, 1994). Much of the data collected in this study underwent pattern matching or typological analysis (Lecompte & Preissle, 1993) which involved using a "start list of codes" early in data analysis (Miles & Huberman, 1994, p. 58). Pattern matching involves the use of preconceived propositions or typologies to suggest inferences in the absence of any alternative patterns (Yin, 1994). A creative and critical-thinking skills typology was developed by consolidating other thinking skills typologies (Ennis, 1991; Gitomer & Duschl, 1995; Marzano et al., 1991). The data were coded using these indicators to seek evidence of thinking exhibited by students as they used tools. When a data sample did not match existing descriptors in the typologies, additional codes were created.

Because we did not require students to use tools to generate some artifacts, but many elected to use them nonetheless, the thinking that emerged was attributed to the tool use. In contrast, we directed students to complete certain activities (e.g., "Write five questions about your selected problem." "Generate five possible solutions for your problem"). These artifacts did not inform the degree to which tools influenced thinking about problems, but rather, what students understood about their problems at different stages in the project. These artifacts were coded in an open fashion along with field notes and interview data (e.g., "Student understands problem details, makes unfounded assumptions").

Codes were inserted into a FileMaker Pro database to facilitate sorting. As mentioned, some

Figure 5 ☐ Sample computer expression generated by TIGER data analysis software, depicting the aggregation of four ordered tree trials completed by one student.



Note: Each letter represents one of the 16 study concepts.

indicators of problem understanding were associated with tools (e.g., "Does not integrate additional problem evidence into Sensemaker diagram") while others were associated with a specific activity (e.g., "Identifies misconception during presentations"). Confirming and disconfirming instances of problem understanding were described in detail to serve "categorical aggregation" as suggested for case-study analysis (Creswell, 1998; Stake, 1995). Two or more codes with shared properties suggested categories.

Six partially-ordered checklist matrices were created to display thinking and tool use by dyads (Miles & Huberman, 1994). When an instance of thinking skill (e.g., information evaluation) was present for a given dyad, a mark was placed on the matrix in a position corresponding to the study dimension in which the thinking took place (e.g., problem framing). New categories were generated by studying the matrices for commonalties in thinking (e.g., "the first dyad rarely used metacognitive scaffolding; did the second, did the third, . . .?"). The process was consistent with discriminant sampling suggested by Lincoln and Guba (1985). Theme development was accomplished by studying the semantic relationships between the categories, such as inclusions ("*X* is a kind of *Y*"), cause-effects ("*X* is a cause of *Y*"), and means-ends ("*X* is a way to do *Y*") (Spradley, 1979). Three themes were found related to student tool use: (a) the use of tools to collect basic information, (b) limitations in high-order information processing with tools, and (c) differential uses of scaffolding. This paper focuses on student problem understanding; detailed information on themes is available in Oliver (1999) and Oliver and Hannafin (2000).

Validity and Reliability

All coding and analysis were completed by the researcher as part of a dissertation study. Error and bias were addressed without multiple raters by testing the instruments and protocols in a pilot study for their ability to extract salient data and by regularly discussing emergent study findings with the teacher, students, and doctoral committee.

External validity is, to an extent, inherently limited in small-scale case studies such as this one. Findings are generalizable only to the extent that they can be compared to similar research sites, events, or populations (Borman, LeCompte, & Goetz, 1986). Internal validity is also a concern for case studies that seek to explain or interpret a set of events (Yin, 1994). To ensure that the events of the case were assessed accurately, the study employed typologies or preexisting coding schemes in a pattern-matching logic to promote internal validity (Yin, 1994). Second, participant viewpoints were solicited via interviews to describe study events from the participants' personal, emic perspective rather than the researcher's external, etic perspective. Third, in addition to positive instances of tool use and thinking, discrepant and negative cases of student tool use and thinking were cited and described. Discrepant cases can help modify or broaden an emerging framework, while negative cases may indicate the framework is incorrect altogether (Borman et al., 1986). Fourth, member checks were conducted with students and the teacher to cross-check emergent study findings in an effort to reduce errors in interpreting data sources. Artifacts were printed each week of the study, and were used to prompt students to reflect on their thinking during interviews.

As noted previously, the reliability of the researcher's observations could not be established statistically, but their veracity was confirmed with the teacher through routine discussions of field notes. Since isolating the myriad of intervening social variables was neither feasible nor the focus of this study, reliability reflects the degree to which the results cited are "consistent with the data collected" (Merriam, 1998, p. 206).

FINDINGS

Dyad Profiles

A summary of dyad demographics, performance, and interaction patterns during the study is contained in Table 5. Descriptive labels are provided to characterize the overall interactions between dyad partners. The following profiles reflect the make-up, interactions, and work habits of both the dyad as a whole and the individual students. Names used are pseudonyms.

Apathetic Dyad. Edie and Jenny engaged the open-ended problem with little intention; their interest level for project activities was generally low; and they generated minimal artifacts to represent their understanding. Both Edie and Jenny were classified *high reading ability* students as a result of their being regrouped because of a conflict between students. The pair worked well together socially, stating they liked their group, but they rarely functioned as a problem-solving team. Edie expressed a dislike of computers, while Jenny indicated that she enjoyed working with them. Jenny made critical project decisions without Edie's help: selecting their problem ("Edie wasn't here yesterday when we picked our problem"), selecting their final solution option ("I picked the best solution"), and drawing their design ("We just went in there and drew our picture, well actually, I did, then [Edie] colored it 'cause she wasn't there"). Their selected problem was broad, as they attempted to resolve two problems rather than one: building materials and shapes. Edie and Jenny resolved these problems with an unoriginal solution: a pyramid constructed of wood.

Hurried Dyad. Mark and Jill rushed through project activities, generated simple artifacts, and gathered limited evidence to inform a relevant solution. Their solution was a house on springs, which was original but somewhat unrelated to their chosen problem topic: determining the best building materials to use in earthquake-prone areas (e.g., wood). Mark, the *high ability* student, used most of the computer tools on his own without discussing the project with his partner. Although Mark worked busily throughout the project, he rushed through activities and used tools carelessly (e.g., writing short, incomplete notes). Jill, the *low ability* student, indicated a dislike for computers and a preference for more traditional learning via books. Jill participated more completely in the creative-problem-resolving phase of the study, drawing sketches and generating Web pages. Mark said he would rather work alone than in a group, while Jill indicated she preferred group work. Neither student seemed to benefit from the partnership; their performance on the ordered tree task was the poorest among the 12 participants.

Self-Directed Dyad. Nate and Keith each relied on his own abilities, used few tools, and asked few questions, yet adequately resolved their selected problem of column design with a unique multilayered disc solution. Nate was in the school's gifted program and was classified as *high ability*. Keith was in the school's Title I Program, reading at only a fourth-grade level. Nate and Keith worked well as a team; they were observed discussing their problem on several occasions. For instance, they contrasted individual design drawings, each using an individual computer while discussing these ideas. However, Nate apparently took the lead of this group in terms of thinking and tool use; Keith was typically nonresponsive during dyad interviews. However, he performed well following the study, associating three groups of study terms together through the ordered tree task. Nate's performance was the best among the 12 participants. However, while he indicated that he liked group work, he stated, "not this time; I do all the work." Keith simply indicated that he liked the group work.

Confused Dyad. Ginny and Gary struggled from the beginning of the project to find a relevant problem, and rushed through problem finding and framing, capturing less information than any other dyad. Ginny and Gary, both low-ability students, were regrouped into their homogeneous dyad after the aforementioned classroom argument. These students often seemed confused and resolved their selected bracing problem by simply restating an unoriginal solution—a steel beam. Ginny and Gary rarely worked together as a team. Ginny completed most of the note taking, sorting, and development work with the tools, while Gary withdrew and worked only on the group's design drawing. It is difficult, therefore, to explain why Gary's ordered tree score was third best among participants. Ginny applied herself to the tasks more than Gary did, but she did not appear to understand the content as well. While Ginny worked with tools to process the dyad's specific problem, Gary was observed browsing, possibly concentrating more on the Web material describing general earthquake problems.

Methodical Dyad. Ruth and Jared worked patiently through each activity, used most tools, and spent more time than most groups framing their selected problem, but were rushed to finish remaining activities at the end of the project. They framed methodically their selected bracing problem, and generated more original solutions than did any other dyad—five. Ruth and Jared worked together well and both indicated that they liked their group work. They were observed discussing the project together on several occasions, and were the only pair observed discussing and using the solution rating form to help select their best solution option. Adding together their individual scores, only one dyad associated more study terms together on the ordered tree task.

Solitary Dyad. Randy and Tisha did not operate as a dyad. Tisha was frequently absent from class during the study, leaving Randy to complete most of their work. Randy changed his initial problem of building shapes to liquefaction. He framed and resolved this second problem by recommending an original solution of floating structures weighted down with "ballast." Randy used tools more effectively than other students. He generated more new ideas for his earthquake engineering problem than did any other student during problem finding. Randy was the only student to integrate additional Web resources with his original core set of Web resources using the Sensemaker tool. On the ordered tree instrument, Randy associated five groups of study terms together and identified four hierarchical relationships between them. This conceptual arrangement represented the best effort among the 12 students.

Mental Model Refining

To clarify the emergence of student models, the findings are organized by the three study dimensions: problem finding, framing, and resolving.

Understanding during problem finding. Evidence collected during problem finding suggested some students had difficulty fractionating the macro earthquake engineering problem into manageable subproblems. All dyads were directed to write a problem statement for their selected research problem. The Hurried Dyad, the Self-Directed Dyad, and the Solitary Dyad selected specific earthquake engineering problems—materials, columns, and shapes. The Confused Dyad and the Methodical Dyad, however, selected the macro problem of bracing, into which several other problems could fit: wall bracing, ceiling bracing with beams, foundation bracing with wheels and rollers, and even appropriate materials for bracing. The final group, the Apathetic Dyad, indicated they intended to frame two problems rather than one: "Our problem is that some material is not strong enough to

Table 5 ☐ Summary of dyad attributes and performance during study

Dyads	Individuals	Ability Level	Tool Use	Solution Quality	Liked Group?
Apathetic	Edie	High	Reluctant to use tools	Unoriginal	Yes
	Jenny	High	Reluctant to use tools	but relevant	Yes
Hurried	Mark	High	Rushed use of some tools	Original	No
	Jill	Low	Used only generation tools	but Irrelevant	Yes
Self-Directed	Nate	High	Limited tool use	Original	No
	Keith	Low	Limited tool use	and relevant	Yes
Confused	Ginny	Low	Used most tools	Unoriginal	—
	Gary	Low	Used only generation tools	but relevant	Yes
Methodical	Ruth	Low	Used most tools	Original	Yes
	Jared	High	Used most tools	Partly relevant	Yes
Solitary	Randy	High	Used most tools	Original	No
	Tisha	Low	Used few to no tools	and relevant	Yes

protect a building during the earthquakes to keep it from collapsing and to find a shape that would help protect the building from collapsing." The instructor described the difficulty many students experienced with problem finding:

> narrowing down . . . focus. That was hard for them, to narrow down to get their problem. They really didn't understand that. They wanted to stick with "buildings collapse during earthquakes," well, yes they do, but why do they collapse?

Students were asked to generate research questions about their selected problems. Each question was evaluated using criteria applied previously (Slotta & Linn, 2000):

1. Is the question *specific* to a component of the selected problem?
2. Is the question *relevant* to some aspect of the selected problem so its answer would help inform and frame that problem? and
3. Is the question *productive* and likely to lead to an experiment as written?

Students were provided metacognitive scaffolding that suggested how to write appropriate scientific questions in this manner. Most dyads recorded four or five questions. Table 6 catalogs the quality of student questions by the criteria.

Three groups, the Self-Directed, Methodical, and Solitary Dyads, each generated one specific question. For example, Nate and Keith of the Self-Directed Dyad wrote: "Do thicker columns fare better than skinnier ones?" Many nonspecific questions were identified that focused on a dyad's entire problem. For example, the Hurried Dyad simply asked, "What materials are best?" Specific questions would indicate that students understood some of the variables involved in their problems (e.g., "Will a building constructed of both wood and steel material withstand a P-wave?"). Most student questions met the relevance criterion, focusing on their selected problems. However, few questions met the productive criterion. Per Slotta and Linn's (2000) guidelines, productive questions are phrased as if planning an experiment, while unproductive questions are phrased as if expecting an answer from an authority figure. For example, the Apathetic Dyad asked, "What are good shapes for buildings in earthquake areas," as if expecting an answer from an authority. Only two dyads generated a productive ques-

Table 6 ☐ Nonspecific, relevant, nonproductive research questions

Criterion	Specific	Not	Relevant	Not	Productive	Not
Apathetic Dyad		5	4	1		5
Hurried Dyad		4	3	1		4
Self-directed Dyad	1	4	5		1	4
Confused Dyad	1	2	3			3
Methodical Dyad	1	4	5			5
Solitary Dyad	1	2	2	1	1	2
Totals	4	21	22	3	2	23

tion that might suggest an experiment (e.g., "Which building is more likely to fall over in an earthquake: one shaped like a cylinder, or one shaped like a square?").

Part of the difficulty with problem finding is synthesizing evidence into categories. While students were reviewing the 13 original Web pages, they could create category headings and sort them into problem-related categories using Sensemaker. Four of six dyads organized evidence into categories, but these students did not use their organizational schemes to find and select subproblems. For example, while the Self-Directed Dyad correctly sorted Web pages under such topics as "collapsing due to materials" and "design error," they selected the columns problem to resolve. Further, Ruth and Jared of the Methodical Dyad created one problem-related topic heading (soil) but chose to resolve the bracing problem. Only Ginny of the Confused Dyad selected a problem related to a Sensemaker category, but she oversorted the evidence into 10 separate categories, indicating a struggle to identify patterns and relationships.

Understanding during problem framing. Evidence collected during problem framing suggested that students had difficulty extracting information relevant to selected problems from the resources provided. This inability to adequately identify interrelated problem variables left students with only partial understanding and misconceptions about their problems.

Sensemaker could be used during problem framing to integrate additional hyperlinks that students found to be related to their problem categories, or to modify original categories in light of new evidence. Only Randy of the Solitary Dyad integrated additional Web evidence with his original Web evidence categories. During interviews, he indicated that he had tried to identify underlying causes of the shapes problem by these actions: "I tried to put in some extra things that could've helped [cause the shapes problem], like liquefaction could have helped . . . ." The Sensemaker tool would ideally help students to reason inductively by capturing confirming instances of phenomena such as wooden buildings collapsing or steel cracking. These new links could be sorted with the original evidence to identify patterns.

Students completed problem framing by generating solution options. This activity was designed to promote inductive inferencing or idea generating (e.g., "Because of these problem variables we've identified, we suggest this solution which addresses *X*, *Y*, and *Z*."). Student ideas generally did not address and were not based on evidence gathered during problem framing. The instructor suggested that many students generated solutions at random: "some of the design solutions are quite good . . . even though they may just be ideas off the top of the student's head . . . without any evidence to support it." When asked to describe the resources they found to develop their solution ideas, many students simply referenced a single Web page rather than several. Nate of the Self-Directed Dyad indicated, "Most of the evidence we used to support this is like the isolated [columns], they use it to make it just shift a bit rather than cracking apart." In contrast, one of the Methodical Dyad partners, Jared, stated: "We were really thinking about the pyramid shape and how that helped, and so we thought maybe some other

Table 7 ☐ Quality of student solution ideas

Criteria	Original Yes	Original No	Relevant Yes	Relevant No	Elaborate Yes	Elaborate No	Partial Yes	Partial No	Contradicts Yes	Contradicts No	Avoids Yes	Avoids No
Apathetic dyad	1	3	3	1	0	4	2	2	1	3	1	3
Hurried dyad	4	3	5	2	0	7	3	4	1	6	1	6
Self-Directed dyad	3	2	5	0	0	5	2	3	1	4	0	5
Confused dyad	3	2	5	0	0	5	2	3	2	3	1	4
Methodical dyad	5	0	4	1	4	1	2	3	1	4	1	4
Solitary dyad	2	2	3	1	1	3	2	2	1	3	1	3
Totals	18	12	25	5	5	25	13	17	7	23	5	25

shape, different, would help."

Using the thinking typologies, student solutions were coded during data analysis for evidence of quality. Six codes were most common. The best solutions were coded as original (SOLorig), relevant (SOLrel), and elaborate (SOLelab). Further, high quality solutions did not use partial evidence (SOLpart), contradict the evidence (SOLcont), or avoid the problem altogether (AVOID). Table 7 catalogs the number of student solutions that met and did not meet each criterion.

One-half of the criteria were met by most solutions: originality, relevance, and non-avoidance. All dyads created solutions that were *original* and not available on any study Web page. Only the Apathetic Dyad had fewer than 50% of their solutions meet the originality criterion. The teacher was pleased with the originality of student designs, because even though many contained misconceptions and flawed reasoning, the students were reasoning for new ideas:

> Yes, a slick floor [idea] is not going to work, the building is going to slide all the way to this end and topple over, it's not going to go back the other way, but they are still thinking about what's happening to the building as the result of that wave passing through, what have other people done to prevent this, and what might I try that nobody else has tried that would help the building react with the waves, and so a lot of designs were reflective of that.

Further, most solutions were *relevant* to selected problems and did not *avoid* the underlying problems with obvious conclusions. Despite the originality of student solutions, few were *elaborated* or detailed. Such ideas would indicate students had greater depth of understanding about their problems. Most groups recorded short solution statements of 5 to 10 words.

Of the 18 original student solutions, 10 did not meet the *partiality* criterion. For example, the notion of an underground house in the Apathetic Dyad's third solution is *original.* However, it is a *partial* idea because the students apparently did not consider how earthquake waves travel through the ground, causing as much or more susceptibility to underground structures as to surface-level structures. Because only 3 of 12 *unoriginal* solutions were *partial,* the data suggest the more *original* and creative the students' solutions became, the more likely they were to fail the *partiality* criterion.

As evident in several partial and contradictory solutions, the concepts of flexibility, height, weight, and bracing were commonly misunderstood. The concept of flexibility was difficult for students to understand, as strong, inflexible structures seem more resistant to earthquakes but some flexibility is actually better. Mark and Jill, the Hurried Dyad, misinterpreted the issue by suggesting the use of nonflexible wood in a solution. Nate, of the Self-Directed Dyad, was asked if flexibility improved a column: "Yeah, it'd just pop right back into place if it was rubber." Students also exhibited naive understanding about the concepts of height and weight. Three dyads suggested the construction of short buildings as a solution, contradicting evidence for well-designed, tall buildings. Two dyads suggested the construction of wood buildings because of their light weight, contradicting evidence illustrating the use of heavy steel to support tall buildings. Bracing was another partially misconstrued topic. Ginny and Gary, the Confused Dyad, suggested placing pillows around buildings, while the Methodical Dyad suggested building an hourglass-shaped structure with external steel bracing. Pillows would likely be too weak to support a building. Steel braces are a better idea, but the Methodical Dyad's solution contradicted evidence suggesting odd- shaped buildings are prone to collapse.

Understanding during resolving. Following the trend observed during problem finding and framing, few students exhibited advanced problem understanding during problem resolving. Evidence from student design drawings and Web pages indicated their solutions were based on limited information.

In reviewing student drawings, most were not elaborated with words or close-ups to amplify various features. Only Nate and Keith (Self-Directed Dyad) and Ginny and Gary (Confused Dyad) used words to describe how their designs would help to address their problem. Other groups simply used their drawing tool to sketch a static design without visualizing earthquake forces acting on it. A few groups used labels and lines to highlight features such as "spring" or "ballast." All dyads stated that they did not change their initial ideas while drawing, except Randy, the individual who completed most of the Solitary Dyad's work: "I probably did [change my design], on some things, something more floatable or lighter to keep the house up . . . ."

Students completed the project by generating Web pages to present to their classmates. If students found and utilized other Web resources to support their new solution, they could cite and reference this evidence. Only one-half of the dyads used hyperlinks to index additional evidence from their Web pages, and only two individual students mentioned the value of hypermedia. Mark stated that Web pages were useful "to branch out for more information." Nate noted that Web pages were useful "mostly so I can make something else to connect it to, so it's a bit longer, plus if you had more than one page, it makes it look longer than it really is." Nate was unduly focused on the length of his dyad's presentation, not the value of hyperlinks to support their idea. The teacher suggested that students used their Web pages to outline, "'this is our problem, this is our solution', [not to] really explain . . . the problem."

Because of the limited problem-related information acquired by students, it is not surprising that four of six dyads chose either an original or relevant final solution from their brainstormed list, but not both. The Apathetic Dyad (Edie and Jenny) and the Confused Dyad (Ginny and Gary) chose to resolve their problems with unoriginal solution ideas, restating the work of others. For example, Ginny and Gary simply chose a steel beam as their solution. While a steel-beam solution was relevant to their problem of bracing, it was not original. Alternatively, the Hurried Dyad (Mark and Jill) and the Methodical Dyad (Ruth and Jared) selected original final solutions, but the solutions were not relevant to the dyads' selected problems. For example, Mark and Jill suggested placing springs under buildings for increased flexibility during earthquakes. While this solution was original, it did not address the dyad's selected problem by suggesting appropriate materials for buildings (i.e., brick, wood, masonry). Only Nate and Keith (Self-Directed Dyad) and Randy and Tisha (Solitary Dyad) resolved their problems with both original and relevant ideas, though as noted, Randy completed the assignment largely without assistance from his partner, Tisha. Nate and Keith created an alternative design for columns, employing layers of discs that could slide or pop-out during an earthquake instead of a single column that was shown to buckle and collapse (see Figure 6). The solution was both original and relevant.

General Content Understanding

The ordered tree instrument helped to describe a different type of understanding than the aforementioned interview and artifact data. Students were able to identify, restate, and restructure some knowledge about earthquakes and earthquake engineering, despite their convergent focus on only one subproblem of earthquake engineering (e.g., column strength, liquefaction).

Figure 6 ☐ Column Solution Presented by Nate and Keith.

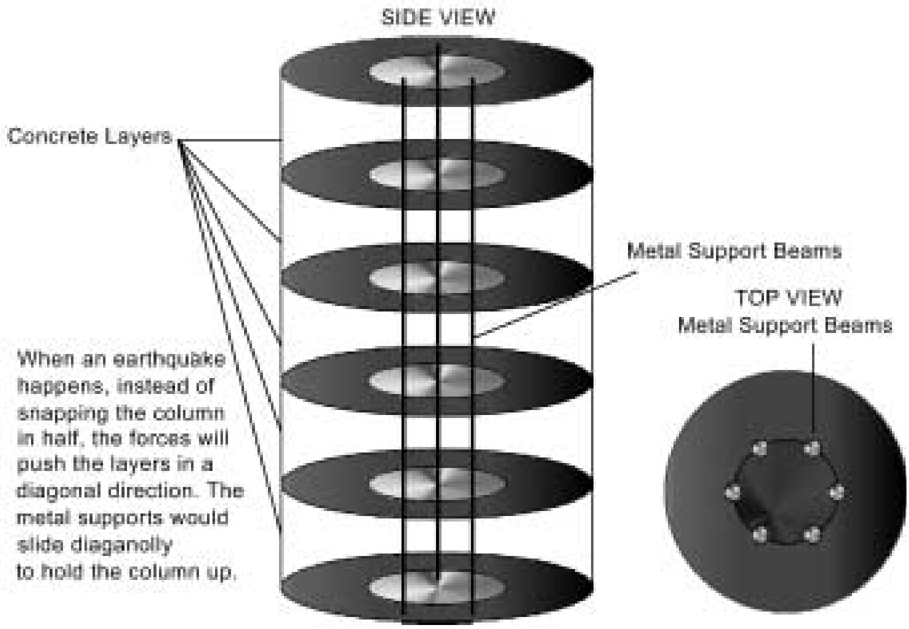


Table 8 depicts individual student performance on the ordered tree instrument. Most students chunked some terms together without identifying the nature of their hierarchical relationship. For example, Mark correctly chunked or associated primary, secondary, body, and compressional waves together on his ordered tree, but he was unable to identify hierarchical relationships between the terms. So, while most students were beginning to associate study concepts as related, few fully comprehended their complete structure. Five students went beyond simply chunking information together to arrange at least some core study concepts into hierarchical relationships. For example, Nate correctly chunked primary and secondary earthquake waves under the superordinate body-waves term. Jared correctly chunked Richter scale, magnitude, Mercalli scale, and intensity under the superordinate term, seismicity.

Students with organizing tools such as Sensemaker would ideally learn to hierarchically arrange core study concepts more effectively. Limited use of Sensemaker during the study, however, could be related to most students' inability to structure conceptual relationships after the study. Further, if Sensemaker had provided a mechanism for students to define or describe the nature of relationships (e.g., link *X* supports theory or category *Z*; link *Y* refutes theory or category *Z*), such processing might have helped students consider such patterns.

DISCUSSION

The following discussion provides rationale for why student mental models remained naive throughout problem finding, framing, and resolving, and how cooperative grouping influenced the students' work. We provide recommended modifications to the OLE to more fully realize student mental model revision, and present implications for open-ended learning in contemporary science education.

Problem Finding, Fractionating

One finding illustrated the difficulty students encountered in finding specific problems or fractionating the macro earthquake engineering problem into specific, manageable subtasks. Students with little prior knowledge may struggle to solve open-ended problems. For example, given 13 Web pages to sort using an organization tool, students with little prior knowledge of earthquake

Table 8 ☐ Conceptual ordering by the ordered tree technique

Dyads	*Individuals*	*Groups of terms chunked*	*# of hierarchical relationships*
Apathetic dyad	Edie		30
	Jenny	2	0
Hurried dyad	Mark	1	0
	Jill	0	0
Self-directed dyad	Nate	5	4
	Keith	3	1
Confused dyad	Ginny	3	0
	Gary	3	2
Methodical dyad	Ruth	5	0
	Jared	5	3
Solitary dyad	Randy	5	4
	Tisha	4	0

problems may struggle to generate categories for the evidence. They may also struggle to identify and collect relevant problem-related information from Internet or database searches.

It is important to identify what students know about an open-ended problem in order to employ appropriate tactics and strategies (Lemberger, 1995; Schauble, 1990). Individuals with schemata for a problem may develop advanced conceptual models more readily than students with limited domain knowledge because they possess a relevant conceptual facility to encode new ideas. Students with limited or no preexisting mental model for a problem may find it difficult to utilize tools strategically, for they do not know where to begin organizing or synthesizing information. The student participants likely entered this study with limited personal models for earthquake engineering and found it difficult to generate advanced models with unfamiliar concepts.

Recommendations. Students attempting to find relevant problems in complex, open-ended situations may require specific scaffolding (Hannafin & Land, 1997). Experts can create new problem representations more easily because they can recall and apply old experiences to new situations: "with an effective knowledge base the individual is capable of recognizing when new information does not fit into the existing cognitive structures" (Hoover & Feldhusen, 1994, p. 213). Students with limited or no advanced domain knowledge might benefit from simulations or analogical reasoning.

Simulations allow students to build on limited domain knowledge by suggesting pertinent variables in a scientific problem and allowing students to test and revise hypotheses. Such tools facilitate conceptual model development by helping students learn that their initial assumptions are but one interpretation "subject to falsification" (Vosniadou , 1994, p. 67). For example, a student who does not initially realize soil quality is a determining factor in earthquake safety might quickly relate loose soils with increased damage by utilizing a simulation. When simulations are used to study complex scientific phenomena (e.g., weight and density), nearly all students are able to generate rules to explain the data (e.g., the rate at which an object sinks depends more on its density than its weight) (Snir, Smith, & Grosslight, 1995). Students must be allowed to make decisions or alter variables integral to the concepts under study (White & Frederiksen, 2000).

Another mechanism to build on limited domain knowledge is encouraging students to employ everyday objects or events to reason analogically. Even students with limited background knowledge can construct crude personal theories to account for a set of events (Land, 1995). In a study by Harel and Papert (1993), students developed software with the Logo programming language to teach other students about fractions. Students employed familiar, everyday objects to represent and teach fractions to others (e.g., clocks, measuring cups, airplanes). The students gained "deeper understanding of the topics involved in their software by thinking of ways to build explanations and graphical representations for their future software users" (Harel & Papert, 1993, p. 69).

Some students in the present study employed familiar objects as part of their solution drawings, such as in the placement of trampolines under buildings and the use of hair "scrunchies" in columns to promote flexibility in earthquakes. As this conversation between the teacher and a student demonstrated, prompting to relate everyday objects to solution ideas proved beneficial:

> One of the students was saying, "well you know how the ear balances the person, it helps to keep your balance, you could use that in a sense, on top of the building, use large tanks of water, and when they're not being used to balance, they could be used as a coolant system, as a recycling coolant system or heating system." He was talking about those wave machines that go back and forth, and when the wave hit, the water would slosh way to the other side of the building, the constant weight idea, that was quite a unique and interesting idea.

This student-generated idea, we later discovered, had been implemented by engineers in Japan. Students can use mental models of familiar analog concepts to refine understanding of new, related concepts and models (Gentner et al., 1997; Glynn, Duit, & Thiele, 1995; Zietsman & Clement, 1997).

Problem Framing, Emergent Understanding

During problem framing, most students generated only partial solutions. Solutions were rarely justified on the basis of extensive evidence; students generally did not or were unable to elicit necessary data from the large database of additional Web evidence. These findings may be attributable to resource processing difficulties or the student tasks themselves.

The manner in which resources are employed may influence mental model development. Study findings suggest students collected more problem-related information when engaged in a scaffolded search involving the 13 original Web resources. Advance organizers and question prompts helped students focus on key points in this restricted resource set. In contrast, the students generally gathered less problem-related information when engaged in self-directed, framing searches through a larger hypermedia database. Students tended to browse randomly for information and use keywords ineffectively. Given that students may possess limited mental models, it may be important to initially nest student inquiry to pertinent content, particularly within disorienting hypermedia databases. Lacking an understanding of key words or concepts in a large knowledge domain and the appropriate higher-order skills to sort and organize unfamiliar information, students may have difficulty synthesizing information in large hypermedia databases (Dede, 1992). Bracketing access to information and prompting reflection are simply scaffolds for younger learners. Such tactics do not "close" an "open-ended" learning environment, as students must still engage in an inquiry process involving hypothesizing, investigation, analysis, evaluation, and revision.

Improper problem framing may be further attributed to student tasks. During problem finding and framing, students were asked to collect evidence to inform their problems and questions. This *finding* of evidence often did not lead students to *apply* evidence to their problems. Gall (1995) found the task or process for which learners engaged a hypertext database to be critical in defining conceptual understanding: Students given a browsing task performed the lowest, students given a searching task performed better, and students given a connecting task performed the best. The simple provision of hypermedia may aid knowing, but "the processes used to evolve personal knowledge are as important as the structure of the knowledge itself" (Gall, 1995, p. 130). Similarly, Dunbar (1993) found that students with goals for *finding* evidence in simulated science environments failed to generate alternate hypotheses to explain data. When student goals are to *generate* hypotheses, however, new predictions tend to be formed to account for inconsistent evidence encountered. Schauble and Glaser (1990) found that only those students, young and old, who employed hypotheses to predict and test outcomes were able to improve their models.

Recommendations. To help students find and apply information to self-selected problems, bracketing searches and emphasizing hypothesis-generation could prove beneficial. WebQuests are instructional units designed for the Internet that encourage the teacher to preselect a set of Internet resources for student research (Dodge, 2001). This limited-search approach may be more effective than nonsystematic browsing under time-restricted learning. Such strategies may also help young learners with limited preconceptions of complex problems to focus on important concepts. Students should also be asked to generate hypotheses or formulate rules for the evidence encountered (e.g., brick buildings collapse), then generate alternate hypotheses or rules when inconsistent evidence is found (e.g., only brick buildings without external bracing collapse). Such guided activities may prompt learners to actively process hypermedia to inform their predictions.

More direction or instructor scaffolding would move the learning environment described in this paper closer to "guided discovery" than "free discovery," in which students make their own choices and interact infrequently with the teacher. Free-discovery learning has been less effective than guided discovery because guidance provides students with planned exposure to the relevant features of a domain (Rogers & Aston, 1992). The problem-finding stage, during which students were prompted to answer specific questions about preselected Web evidence, was more representative of guided discovery and generated better results than students generating their own questions and free-searching through larger resource sets during problem framing.

Problem Resolving, Solution Evaluation

Student justification and defense of final solutions were limited. They rarely linked their Web pages to supporting evidence, or recorded solution advantages as notes. Students tended to select a solution early in their inquiries, and were reluctant to revise or adopt new solutions by actively seeking conflicting evidence.

The lack of feedback provided to students about their ideas may contribute to the limited understanding observed in this study. Students were occasionally provided feedback on general *processes* when those were inadequate. For example, the instructor prompted the students to read material thoroughly and to frame problems with more information when it was apparent they were behaving hurriedly. Students were not provided feedback on their *ideas*, however, until they gave their final presentations. For instance, the facilitators did not suggest strategies to overcome flawed student conceptions. If a student chose a problem that was too broad in scope, the facilitators did not focus the student on a more specific problem, since the research focus was on how students deployed the various tools to act on and refine their own interests.

Recommendations. OLEs should employ mechanisms that allow learners to externalize their thinking for peer critique, discussion, and revision. Communication may promote appropriate problem resolution by encouraging students to share their ideas with others and receive feedback related to the correctness of a solution path. Tasks in which students must verbally explain phenomena, defend arguments, and compare understanding to experts, will allow them to better understand the nature of their argumentation or their underlying theories and assumptions (Vosniadou, 1994).

In this study, students were able to clarify their understanding during class presentations and follow-up questioning. Dyads who presented their solution or hypothesis for a specific problem were often confronted by peers with different schemata for the same problem. Through peer review, the students were forced to consider flaws in their reasoning. For instance, one student asked, "What if a crack caused all the water to come in?" regarding another student's floating house design. The student designer replied, "You might want to have two stories on the house." When the teacher asked one student, "What happens when the house hits one end of the pool? Does it flip out over it?" he replied, "It would probably be anchored down to the bottom of it." Such conversations can be valuable to help students identify, then revise, inaccuracies in their designs.

To help students understand their argumentation and thinking, tasks such as presentations proved helpful, as did daily instructor scaffolding during regular classroom activity and student research. Instructors can scaffold student thinking in OLEs via such techniques as modeling, Socratic questioning, or externalizing reflection and metacognition (Sharma & Hannafin, 2001). Instructors utilizing open-ended problems should model embedded processes such as writing questions or generating hypotheses with specific examples. They should question student assumptions, suggest a closer examination of conflicting evidence, and ask students to reflect on consequences of their thoughts. Finally, they should encourage learners to check their strategies and reflect on the success of these efforts toward possible modification of tactics.

Cooperative Group Effects

Research has shown that while heterogeneous pairing of high-ability students with low-ability students generally improves the performance of the low-ability student, it may or may not detract from the performance of the high-ability student (Hooper & Hannafin, 1988; Hunt, 1992; Repman, Weller, & Lan, 1993). In this study, heterogeneously grouped dyads did not perform well when one or both members refused to participate. Both the Methodical and Self-Directed Dyads who worked cooperatively, however, performed as well or better on the ordered tree instrument than their counterparts in less cooperative dyads (i.e., the Hurried and Solitary Dyads).

Two dyads in this study were regrouped into homogeneous pairs because of a classroom argument; the consequences were negative. In their study of fifth-grade science students, Jones and Carter (1994) found low-ability students grouped homogeneously were unorganized and off-task, and competed for materials. This is consistent with the performance and behavior of the two low-ability students in the Confused Dyad, who struggled to find a relevant problem to frame. Gary frequently engaged in off-task behavior such as retyping the same word. Ginny resisted sharing the KIE computer tools with her partner. In contrast to the performance of these two low-ability students, high-ability students grouped together have been found to be efficient and organized (Jones & Carter, 1994). In this study, however, the two high-ability students in the Apathetic Dyad worked slowly and haphazardly. It seems unlikely that Edie and Jenny's performance can be attributed solely to dyad effects since each was cooperative with the other. Rather, limited interest in earthquake engineering problems may have rendered them unwilling to expend effort.

Research findings are equivocal regarding the grouping of gifted students with similar-ability and lower-ability students (Goldring, 1990; Slavin, 1991). In this case study, two gifted students, Nate and Jared, were grouped with two low-ability students, Keith and Ruth. Students in these dyads were the most cooperative among the six. They were regularly observed discussing project resources in an attempt to resolve their problems, whereas other dyads worked individually or resisted work altogether. Despite being grouped with a low-ability student, Nate's learning was apparently not slowed, as evidenced by his performance on the ordered tree which was the best in the class. Further, Nate's low-ability partner, Keith, excelled on the ordered tree with marks exceeding those of most students, including several high-ability students. Jared of the Methodical Dyad, however, might have been slowed somewhat by his grouping with the low-ability student, Ruth. He indicated that sometimes group work "slows both of you down."

Recommendations. The high-ability students in the Hurried and Solitary Dyads did not work cooperatively with or express accountability for the learning of their low-ability partners. To promote better cooperation, Putnam (1997) suggests group goals and grades should be stressed over any individual marks. Each student must be concerned with the performance of other group members, not simply self. Although group goals were stressed in this study, group grades were not. The teacher warned students that she knew who completed the work in the dyads and would grade accordingly. In addition to a lack of group grading, individual students were not held accountable for the knowledge of other group members. A "coasting" or "hitchhiker" effect in some dyads may have detracted from individual student learning and group progress, because it was possible for individual students to depend on their partners for the completion of work (Putnam, 1997, p. 68). Assigning individual students to complete interdependent tasks or to undertake complementary roles are suggested strategies for overcoming the disproportional effort exhibited by some students in coasting situations (Johnson & Johnson, 1987; Jonassen, 1996).

Finally, while most researchers suggest pairing students heterogeneously by ability, Jonassen (1996, p. 37) suggested it may be more appropriate to group students by "cognitive controls" or "relatively stable learner traits that describe how learners interact with, perceive information from, and make sense of the world." Jonassen and Grabowski (1993) suggested learners may be more global, socially oriented, and communicative, or more analytic, organized, and introverted. Pairing students heterogeneously into global-analytic groups might promote better cooperation than was exhibited by some dyads in this study.

Implications

A nagging question remains: Given the limited content understanding obtained by students in this study, as well as the emergence of only partially correct mental models, what is the justification for open-ended learning? Clearly, while a justification for OLEs must be tied to student learning, we must also recognize the value of student processes. OLEs provide students with time to practice the steps of problem solving and inquiry—processes that can be transferred to help solve future problems encountered. Depth of coverage in OLEs can help students to develop a grounded mental model that is useful to bridge to and more easily integrate concepts when and if needed. OLE goals include both high-level synthesis and evaluation of information and student self-evaluation and reflection on their own ideas and progress. OLEs allow students to actively process information by practicing strategies for finding and collecting evidence to inform a problem or support a prediction. OLEs can integrate new technologies and allow students to learn technology production skills. Finally, OLEs can include activities whereby students learn to communicate their findings to peers.

Despite students' limited understanding of a breadth of facts in this study, the processes they undertook were arguably at least as valuable as memorizing plate tectonics concepts and taking tests to assess low-level comprehension. An added benefit of open-ended learning is developing multiple, flexible skills such as those listed above. These aims are advocated by multiple government and private agencies whose recommendations for science reform since the early 20th century include teaching science to benefit everyday citizens, not to generate professional scientists with complete expertise in a discipline (Hurd, 1997).

As Hammer (1997, p. 489) pointed out, however, "the tension between inquiry and traditional content plays out mainly in decisions about the use of time: how much to devote to inquiry-oriented activities at the expense of coverage." The teacher in this study stated she was only able to participate because her school did not administer standardized science tests the year of our research. Otherwise, she would have caved to external forces, "teaching to the test" to ensure that her students had sufficient test "knowledge" to perform well. Clearly, the consequences of emphasizing learning processes over content accountability weighs heavily on teachers charged with educating children. Elected officials and government agencies are crafting legislation that increases student testing and promotes school and teacher accountability for student test scores. Such movements have caused concern for educators for more than 70 years: "[N]o system of external tests which aims primarily at examining individual scholars can result in anything but educational waste" Whitehead (1929, p. 13). Content coverage and overtesting may lead to inert knowledge that is not transferable and is quickly forgotten (Perkins & Salomon, 1988). Understanding, as Papert (1993) observed, requires time; it also requires deeper and more analytical ways of thinking.

CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

Despite focusing mostly on microlevel problems in the study, students demonstrated content understanding on some macrolevel earthquake engineering concepts. For instance, students who focused on engineering problems dealing with columns also learned about problems related to building shapes and soil strength. Unfortunately, the potential benefits of authentic constructivist problems on problem understanding remained only partially realized in the study. Most student groups were able to develop original solutions with some misconceptions and partial ideas, but student understanding of earthquake engineering problems did not improve to canonical or conventional models.

Developing advanced models of scientific phenomena is a complex process. To manage this complexity, it may help to provide students with a framework for inquiry. In this study, students were asked to generate and answer their own research questions. Study findings suggest this framework or problem-framing task was largely unproductive, as many student questions were not specific or productive, and did not lead them to continually revise and modify their conceptions of earthquake engineering problems. A more productive task might ask students to state what they know about a problem, then hypothesize or predict potential resolutions prior to accessing resources or tools. Hypotheses and predictions allow students to actively seek evidence to support or contradict a particular view, rather than passively collect information about a topic. Externalizing a student's personal mental model early in inquiry lays the groundwork for constant revision and advancement. Such approaches are in line with system affordances and teacher scaffolding that encourage students to set goals for and plan their problem solving (Laffey, Tupper, Musser, & Wedman, 1998).

Study limitations were related to limited resources, incomplete tool sets, and study methods. The hypermedia resources provided were downloaded from the Internet, somewhat constraining student searches. Students could not use keyword searches to focus on specific resources, nor access Internet resources of their own choosing, particularly human resources such as peers and problem-related experts. In addition to limited resource access, students were not provided a complete tool set. While students had access to information processing and scaffolding tools, they did not have access to the simulation or communication tools recommended previously. The ordered tree technique illustrated only how well students understand hierarchical relationships among concepts; alternative methods such as concept maps might better illustrate student understanding of network or hyper-relationships. Reitman and Rueter (1980, p. 557) noted that ordered trees "differ from multidimensional scaling in particular on two important aspects . . .: they produce a hierarchical, not spatial, representation, and thus answer questions about discrete, not graded, properties of data, such as which items function together as a chunk in memory and perception." To offset this limitation, we collected student Web-page presentations as a data source, through which students could readily express hyper-linkages to related concepts.

The present findings neither support nor refute mental model development via open-ended problems and tool-supported inquiry. Future research studies should seek to externalize student hypotheses early in the problem-solving process, and focus on how these conceptions change in light of different interventions. Do students revise their hypotheses and problem conceptions when provided with appropriate instructor scaffolding? Do students revise their hypotheses when communication tools are integrated into the learning environment for sharing ideas, reasoning with analogical ideas from student backgrounds, and regularly critiquing or commenting on one another's conceptions? Do they revise their hypotheses given opportunities to test their ideas by experimentation, even if only virtually, through simulation and modeling? Perhaps new systems will be developed, tested, and implemented as answers to these and related questions emerge. □

Kevin Oliver is an Instructional Designer and Evaluator in the Ed Tech Dept. at Virginia Tech, Blacksburg.

Michael Hannafin is a Professor and the Charles H. Wheatley, Georgia Research Alliance Eminent Scholar in Technology-Enhanced Learning, as well as Director of the Learning and Performance Support Laboratory, University of Georgia, Athens.

REFERENCES

Barman, C.R. (1996). Bridging the gap between the old and the new: Helping teachers move towards a new vision of science education. In J. Rhoton, & P. Bowers (Eds.), *Issues in science education* (pp. 155–161). Arlington, VA: National Science Teachers Association.

Baxter, J. (1995). Children's understanding of astronomy and the earth sciences. In S.M. Glynn, &

R. Duit (Eds.), *Learning science in the schools* (pp. 155–177). Mahwah, NJ: Lawrence Erlbaum Associates.

Blue Squirrel. (2000). WebWhacker. [Computer software]. Draper, UT.

Borman, K.M., LeCompte, M.D., & Goetz, J.P. (1986). Ethnographic and qualitative research design and why it doesn't work. *American Behavioral Scientist, 30*(1), 42–57.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–41.

Bruner, J.S. (1966). Some elements of discovery. In L.S. Shulman, & E.R. Keislar (Eds.), *Learning by discovery: A critical appraisal. Chicago, IL: Rand McNally & Company*.

Bybee, R.W. (1996). The contemporary reform of science education. In J. Rhoton, & P. Bowers (Eds.), *Issues in science education* (pp. 1–14). Arlington, VA: National Science Teachers Association.

Calvi, L. (1997). Navigation and disorientation: A case study. *Journal of Educational Multimedia and Hypermedia, 6*(3/4), 305–320.

Cognition and Technology Group at Vanderbilt. (1992). The Jasper experiment: An exploration of issues in learning and instructional design. *Educational Technology Research and Development, 40*(1), 65–80.

Creswell, J.W. (1998). *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage Publications.

Dass, P.M. (1999). *An STS approach to organizing a secondary science methods course: Preliminary findings.* (ERIC Document Reproduction Service No. ED443672)

Dede, C.J. (1992). The future of multimedia: Bridging to virtual worlds. *Educational Technology, 32*(5), 54–60.

Dimitroff, A. (1992). Mental models theory and search outcome in a bibliographic retrieval system. *Library and Information Science Research, 14*(2), 141–156.

Dodge, B. (2001). The WebQuest page. [On-line]. Available: http://edWeb.sdsu.edu/Webquest/Webquest.html

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science, 17,* 397–434.

Edelson, D.C., Pea, R.D., & Gomez, L. (1996). Constructivism in the collaboratory. In B.G. Wilson (Ed.), *Constructivist learning environments* (pp. 151–164). Englewood Cliffs, NJ: Educational Technology Publications.

Ennis, R.H. (1991). Goals for a critical thinking curriculum. In A.L. Costa (Ed). *Developing minds: A resource book for teaching thinking* (pp. 68–71). Washington DC: Association for Supervision and Curriculum Development.

Fisher, K.M. (1992). SemNet: A tool for personal knowledge construction. In P.A.M. Kommers, H. Jonassen, & J.T. Mayes (Eds.), *Congitive tools for learning* (pp. 63–75). Heidelberg: Springer-Verlag.

Gall, J.E. (1995). *The effect of external orienting task on learning, learner attitudes, and en-route behaviors in the use of an educational hypertext system.* Unpublished doctoral dissertation, Florida State University.

Gentner, D., Brem, S., Ferguson, R.W., Markman, A.B., Levidow, B.B., Wolff, P., & Forbus, K.D. (1997). Analogical reasoning and conceptual change: A case study of Johannes Kepler. *The Journal of the Learning Sciences, 6*(1), 3–40.

Gitomer, D.H., & Duschl, R.A. (1995). Moving toward a portfolio culture in science education. In S.M. Glynn, & R. Duit (Eds.), *Learning science in the schools* (pp. 299–326). Mahwah, NJ: Lawrence Erlbaum Associates.

Glynn, S. (1997). Drawing mental models. *The Science Teacher, 64*(1), 30–32.

Glynn, S., & Duit, R. (1995). Learning science meaningfully: Constructing conceptual models. In S.M. Glynn, & R. Duit (Eds.), *Learning science in the schools* (pp. 3–34). Mahwah, NJ: Lawrence Erlbaum Associates.

Glynn, S., Duit, R., & Thiele, R. (1995). Teaching science with analogies: A strategy for constructing knowledge. In S.M. Glynn, & R. Duit (Eds.), *Learning science in the schools* (pp. 247–273). Mahwah, NJ: Lawrence Erlbaum Associates.

Goldring, E.B. (1990). Assessing the status of information on classroom organizational frameworks for gifted students. *Journal of Educational Research, 83*(6), 313–326.

Halford, G.S. (1993). *Children's understanding: The development of mental models.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Hammer, D. (1997). Discovery learning and discovery teaching. *Cognition and Instruction, 15*(4), 485–529.

Hannafin, M.J., Hall, C., Land, S.M., & Hill, J.R. (1994). Learning in open-ended environments: Assumptions, methods, and implications. *Educational Technology, 34*(8), 48–55.

Hannafin, M.J., Hannafin, K.M., Land, S., & Oliver, K. (1997). Grounded practice and the design of constructivist learning environments. *Educational Technology Research and Development, 45*(3), 101–117.

Hannafin, M.J. & Land, S. (1997). The foundations and assumptions of technology-enhanced, student-centered learning environments. *Instructional Science, 25,* 167–202.

Hannafin, M.J., & Land, S.M. (2000). Technology and student-centered learning in higher education: Issues and practices. *Journal of Computing in Higher Education, 12*(1), 3–30.

Hannafin, M.J., Land, S., & Oliver, K.M. (1999). Open learning environments: Foundations, methods, and models. In C. Reigeluth (Ed.), *Instructional design theories and models* (pp. 115–140). Mahwah, NJ: Lawrence Erlbaum Associates.

Harel, I., & Papert, S. (1993). Software design as a learning environment. In I. Harel, & S. Papert (Eds.), *Constructionism* (pp. 41–84). Norwood, NJ: Ablex Publishing Corporation.

Heywood, J., Heywood, S., & Donovan, I. (1992). *The training of student-teachers in discovery methods of in-*

struction and learning [and] comparing guided discovery and expository methods: Teaching the water cycle in Geography. (Research in Teacher Education Monograph Series No. 1/92). Dublin, Ireland: Dublin University.

Hirtle, S.C. (1980). TIGER [Computer software]. Pittsburgh, PA: University of Pittsburgh Department of Information Science.

Hooper, S., & Hannafin, M. (1988). Cooperative CBI: The effects of heterogeneous versus homogeneous grouping on the learning of progressively complex concepts. *Journal of Educational Computing Research, 4*(4), 413–424.

Hoover, S.M., & Feldhusen, J.F. (1994). Scientific problem solving and problem finding: A theoretical model. In M. Runco (Ed.), *Problem finding, Problem solving, and creativity* (pp. 201–219). Norwood, NJ: Ablex Publishing Corporation.

Howe, A.C., & Jones, L. (1993). *Engaging children in science.* New York: Macmillan Publishing Company.

Hunt, P. (1992). *Acquisition of communication and motor skills within the context of cooperative learning groups in general education classrooms.* (ERIC Document Reproduction Service No. ED365061)

Hurd, P.D. (1997). *Inventing science education for the new millenium.* New York: Teachers College Press.

Isaacs, N. (1974). *Children's ways of knowing: Nathan Isaacs on education, psychology, and Piaget.* New York: Teachers College, Columbia University.

Jackson-Metcalf, S., Krajcik, J., & Soloway, E. (2000). Model-It: A design retrospective. In M.J. Jacobson, & R.B. Kozma (Eds.), *Innovations in science and mathematics education: Advanced designs for technologies of learning* (pp. 77–115). Mahwah, NJ: Lawrence Erlbaum Associates.

Johnson, D.W., & Johnson, R.T. (1987). *Learning together and alone: Cooperative, competitive, and individualistic learning.* Englewood Cliffs, NJ: Prentice-Hall.

Johnson-Laird, P.N. (1990). Human thinking and mental models. In K.A.M. Said, W.H. Newton-Smith, R. Viale, & K.V. Wilkes (Eds.), *Modelling the mind* (pp. 155–170). New York: Oxford University Press.

Jonassen, D.H. (1996). *Computers in the classroom: Mindtools for critical thinking.* Englewood Cliffs, NJ: Prentice-Hall.

Jonassen, D.H., & Grabowski, B.L. (1993). *Handbook of individual differences, learning, and instruction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Jones, M.G., & Carter, G.C. (1994). Verbal and nonverbal behavior of ability-grouped dyads. *Journal of Research in Science Teaching, 31*(6), 603–620.

Kelly, J. (1995). *What do they think they're doing?: Mental models of online catalog users in an academic library.* Unpublished doctoral dissertation, University of Georgia, Athens.

Kommers, P., & DeVries, S. (1992). TextVision and the visualization of knowledge: School-based evaluation of its acceptance at two levels of schooling. In P. Kommers, D. Jonassen, & J. Mayes (Eds.), *Cognitive tools for learning* (pp. 33–62). Heidelberg: Springer-Verlag.

Kyllonen, P.C., & Shute, V.J. (1989). A taxonomy of learning skills. In P.L. Ackerman & R.J. Sternberg (Eds.), *Learning and individual differences: Advances in theory and research* (pp. 117–163). New York: W.H. Freeman.

Laffey, J., Tupper, T., Musser, D., & Wedman, J. (1998). A computer-mediated support system for project-based learning. *Educational Technology Research and Development, 46*(1), 73–86.

Land, S.M. (1995). *The process of developing theories-in-action with open-ended learning environments: An exploratory study.* Unpublished doctoral dissertation, Florida State University.

Land, S., & Hannafin, M.J. (1996). A conceptual framework for the development of theories-in-action with open learning environments. *Educational Technology Research and Development,44*(3), 37–53.

Land, S., & Hannafin, M.J. (1997). Patterns of understanding with open-ended learning environments: A qualitative study. *Educational Technology Research and Development, 45*(2), 47–73.

LeCompte, M.D., & Preissle, P. (1993). *Ethnography and qualitative design in educational research* (2nd ed.). San Diego, CA: Academic Press, Inc.

Lemberger, J.S. (1995). *The relationship between a model-building problem-solving classroom and conceptual change learning.* Unpublished doctoral dissertation, The University of Wisconsin.

Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage Publications.

Linn, M.C. (1995). Designing computer learning environments for engineering and computer science: The scaffolded knowledge integration framework. *Journal of Science Education and Technology, 4*(2), 103–126.

Marzano, R.J., Brandt, R.S., Hughes, C.S., Jones, B.F., Presseisen, B.Z., Rankin, S.C., & Suhor, C. (1991). Dimensions of thinking: A framework for curriculum and instruction. In A.L. Costa (Ed). *Developing minds: A resource book for teaching thinking* (pp. 89–93). Washington DC: Association for Supervision and Curriculum Development.

McCaslin, M., & Good, T. (1992). Compliant cognition: The misalliance of management and instructional goals in current school reform. *Educational Researcher, 21*(3), 4–17.

McGregor, J.H. (1994). Information seeking and use: Students' thinking and their mental models. *Journal of Youth Services in Libraries, 8*(1), 69–76.

Merriam, S.B. (1998). *Qualitative research and case study applications in education.* San Francisco: Jossey-Bass Publishers.

Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis.* Thousand Oaks, CA: Sage Publications.

National Science Teachers Association. (1990). *Science/technology/society: A new effort for providing appropriate science for all.* (Position Statements). Arlington, VA: Author.

Naveh-Benjamine, M., & Lin, Y.G. (1991). *Assessing students' organization of concepts: A manual for measuring course-specific knowledge structures.* Ann Arbor, MI: National Center for Research to Improve Postsecondary Teaching and Learning. (ERIC Document Reproduction Service No. ED 338 124)

Naveh-Benjamine, M., McKeachie, W.J., Lin, Y.G., & Tucker, D.G. (1986). Inferring students' cognitive structures and their development using the "Ordered Tree Technique." *Journal of Educational Psychology, 78*(2), 130–140.

Oliver, K.M. (1999). *Student use of computer tools designed to scaffold scientific problem solving with hypermedia resources: A case study*. Unpublished doctoral dissertation, University of Georgia.

Oliver, K.M., & Hannafin, M.J. (2000). Student management of Web-based hypermedia resources during open-ended problem solving. *Journal of Educational Research, 94*(2), 75–92.

Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer.* New York: Basic Books, Inc.

Perkins, D.N., & Salomon, G. (1988). Teaching for transfer. *Educational Leadership, 46*(1), 22–32.

Pitts, J.M. (1995). Mental models of information: The 1993–94 AASL / Highsmith Research Award Study. *School Library Media Quarterly, 23*(3), 177–184.

Putnam, J. (1997). *Cooperative learning in diverse classrooms.* Upper Saddle River, NJ: Prentice-Hall, Inc.

Reitman, J.S., & Rueter, H.H. (1980). Organization revealed by recall orders and confirmed by pauses. *Cognitive Psychology, 12, 554–581.*

Repman, J., Weller, H.G., & Lan, W. (1993). The impact of social context on learning in hypermedia-based instruction. *Journal of Educational Multimedia and Hypermedia, 2*(3), 283–298.

Rogers, P.J., & Aston, F. (1992). Teaching method, memory and learning: An enquiry with primary school children. *Educational Studies, 18*(2), 129–149.

Rowe, A.L., & Cooke, N.J. (1995). Measuring mental models: Choosing the right tools for the job. *Human Resource Development Quarterly, 6*(3), 243–255.

Scardamalia, M., Bereiter, C., Brett, C., Burtis, P.J., Calhoun, C., & Smith, L.N. (1992). Educational applications of a networked communal database. *Interactive Learning Environments, 2*(1), 45–71.

Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology, 49*(1), 31–57.

Schauble, L. & Glaser, R. (1990). Scientific reasoning in children and adults. In D. Kuhn (Ed.), *Developmental perspectives on teaching and learning thinking skills.* New York: Karger.

Sharma, P., & Hannafin, M. (2001). *Scaffolding critical thinking in technology-mediated learning environments.* Unpublished manuscript.

Slavin, R.E. (1991). Are cooperative learning and untracking harmful to the gifted? Response to Allan. *Educational Leadership, 48*(6), 68–71.

Slotta, J.D. , & Linn, M.C. (2000). The Knowledge Integration Environment: Helping students use the Internet effectively. In M.J. Jacobson, & R.J. Kozma (Eds.), *Innovations in science and mathematics education: Advanced designs for technologies of learning (pp. 193–226). Mahwah, NJ: Erlbaum.*

Snir, J., & Smith, C. (1995). Constructing understanding in the science classroom: Integrating laboratory experiments, student and computer models, and class discussion in learning scientific concepts. In D.N. Perkins, J.L. Schwartz, M.M. West, & M.S. Wiske (Eds.), *Software goes to school: Teaching for understanding with new technologies* (pp. 233–254). New York: Oxford University Press.

Snir, J., Smith, C., & Grosslight, L. (1995). Conceptually enhanced simulations: A computer tool for science teaching. In D.N. Perkins, J.L. Schwartz, M.M. West, & M.S. Wiske (Eds.), *Software goes to school: Teaching for understanding with new technologies* (pp. 106–129). New York: Oxford University Press.

Spradley, J.P. (1979). *The ethnographic interview.* New York: Holt, Rinehart, and Winston.

Stake, R.E. (1995). *The art of case study research.* Thousand Oaks, CA: Sage Publications.

Stratford, S. Krajcik, J., & Soloway, E. (1998). Secondary students' dynamic modeling processes: Analyzing, reasoning about, synthesizing, and testing models of stream ecosystems. *Journal of Science Education and Technology, 7*(3), 215–234.

Sutton, S.A. (1994). The role of attorney mental models of law in case relevance determinations: An exploratory analysis. *Journal of the American Society for Information Science, 45*(3), 186–200.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4,* 45–69.

White, B., & Frederiksen, J. (2000). Technological tools and instructional approaches for making scientific inquiry accessible to all. In M.J. Jacobson, & R.J. Kozma (Eds.), *Innovations in science and mathematics education: Advanced designs for technologies of learning* (pp. 321–359). Mahwah, NJ: Erlbaum.

Whitehead, A.N. (1929). *The aims of education and other essays.* New York: The Macmillan Company.

Wiser, M. (1995). Use of history of science to understand and remedy students' misconceptions about heat and temperature. In D.N. Perkins, J.L. Schwartz, M.M. West, & M.S. Wiske (Eds.), *Software goes to school* (pp. 23–38). New York: Oxford University Press.

Yager, R.E. (1995). Constructivism and the learning of science. In S.M. Glynn, & R. Duit (Eds.), *Learning science in the schools* (pp. 35–58). Mahwah, NJ: Lawrence Erlbaum Associates.

Yin, R.K. (1994). *Case study research: Design and methods.* Thousand Oaks, CA: Sage Publications.

Zietsman, A., & Clement, J. (1997). The role of extreme case reasoning in instruction for conceptual change. *The Journal of the Learning Sciences, 6*(1), 61–89.