USING EXPERT HEURISTICS FOR THE DESIGN OF IMAGERY-RICH MENTAL SIMULATIONS FOR THE SCIENCE CLASS †

One of us (Clement, 2006, this volume) has derived a list of strategies from expert protocols for designing and improving central test cases and analogies using imagery; we ask whether these strategies might apply to the analogies and simulations for which we see evidence in science classrooms. The breadth of applicability of the list was investigated by examining transcripts from three classroom-activity and tutoring sessions where such cases are being described by either an educator or a student. Evidence is presented that analogies and test cases described by these teachers and students share many of the design characteristics of expert analogies and test cases. There were multiple indicators providing evidence that students were able to reason about their conceptions with the imagery from the cases; to modify and evaluate teacher cases; to modify and evaluate cases generated by other students; and to generate enhancements to the imagistic qualities of cases. We believe that this list of heuristic strategies can be honed into a list of principles teachers can use when designing analogies and simulations for their science students.

A. Lynn Stephens, University of Massachusetts
John J. Clement, University of Massachusetts, USA
Maria Cecelia Nunez Oviedo, Universidad de Concepcion, Chile

Introduction

When analyzing videos of talented teachers fostering active science learning in their students, we have been struck by the elegance of analogies and mental simulations that many invite their students to use. These appear designed to encourage the students to run their own mental models. We would like to know what goes into the design of a test case or analogy that can be effective in the classroom.

Although many researchers agree that imagistic models are a key component of science understanding for students, we are concerned that many students still appear to have difficulty learning visualizable models. Some would ascribe these difficulties to differences in spatial reasoning ability in students; however, we would like to ask whether there are principles for fostering the learning of visualizable models that are applicable to students of all abilities. One of us (Clement, 2006, this volume) has derived

---

† This material is based upon work supported by the National Science Foundation under Grant REC-0231808, John J. Clement, PI. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

a list of heuristics for the refinement of imagery from expert protocols. In the present paper, we ask whether these heuristics might apply to the analogies and test cases we see in videos of science classrooms and science tutoring sessions.

Definitions

When the list of expert strategies was derived, it was in the context of problem solving in the domains of physics and mathematics. The term *heuristics* was applied to the list, defined thusly:

Heuristic: a useful strategy that is not guaranteed to work.

This use of the term is similar to the use made of it by mathematician George Polya in his seminal work *How to Solve It* (1957). Since *heuristic* can take on somewhat different meanings in other contexts, in the present paper we often use the term *strategy* as a synonym; although in every case, we are referring to a strategy that, though it has been shown useful in some contexts, comes with no guarantees. An example is the strategy of looking for ways to break large problems into smaller ones. As every problem-solver knows, this does not always work—but it often proves useful.

We frequently use terms such as *imagistic reasoning, imagistic models, etc*. Use of the term *imagistic* refers to the possible presence of imagery.

Objectives and Theoretical Perspective

The present paper is an exploratory study

1) to investigate whether expert heuristics for the design of imagery-rich mental simulations may apply to classroom cases;

2) to discover whether there is evidence students have actually used these analogies and cases in their own reasoning.

Since this is undeveloped territory, we use in-depth case studies to map out the forms of reasoning involved.

Previous Research

Although there has been much research in students' use of visual and kinesthetic imagery and analogy, and students' ability to run mental models in science learning (Clement & Steinberg, 2002; Gentner & Gentner, 1983; Gentner & Toupin, 1986; Hegarty, 1992; Kozhnevnikov, Hegarty & Mayer, 1999; Nersessian, 1995; Nersessian, 1992; Nunez-Oviedo (2003); Reiner & Gilbert, 2000), the focus has been on how analogies and mental simulations function rather than on how they are selected or developed. Our long-term goal is to develop a theory of the role of mental simulations in science learning. Our short-term goal, addressed in this paper, is to identify strategies for designing and refining imagery-rich analogies and test cases that students can adopt for use in constructing and running their own dynamic mental models of phenomena.

Methodology

In order to investigate the breadth of applicability of the list of design strategies (see Table 1), we have examined transcripts of classroom activity and tutoring sessions where imagery was being described by either an educator or a student. We have also examined descriptions of mental simulations in the writings of well-known scientists in history. We have analyzed 19 reports or descriptions of imagery-rich mental simulations in terms of the expert strategies: 7 generated by expert scientists past and present; 7 initiated by teachers, tutors, or curriculum developers; and 5 student-initiated. Most of the simulations examined so far have been in the form of test cases or analogies created for the purpose of convincing others (as in pedagogy) or used in attempts to solve unfamiliar problems.

This paper is a companion paper to Clement (2006, this volume). The strategies derived from the expert study in that paper are shown in Table 1. These are consistent with a view described in that paper of how perceptual motor schemas can generate imagistic mental simulations.

Strategies for Choosing Cases for Imagistic Simulation

A. Find a **familiar** case close enough to something you already know to make a confident prediction.

B. Find a concrete case evoking imagery that has **kinesthetic and/or visual** components.

C. Find a case where one's predictions **agree with known constraints or accepted physical theory.*** (This may need confirmation.)

Strategies for Choosing or Improving Cases: increasing Conviction in an imagistic Simulation via Imagery Enhancement

D. To enhance the application of a schema in a simulation:

   Generate a **special case where the schema is easier to run**; e.g., by generating an **Extreme Case** that helps tap implicit knowledge in the schema by increasing the contrast in comparisons between simulations.

E. To enhance spatial reasoning or schema application:

1) Use “**markers**” (mental diagrammatic notation that makes variable differences easy to imagine)

2) Use a case that is as **simple and schematic** as possible (one that is easier to visualize or to imagine kinesthetically)

3) Imagine a case whose particular features (such as **size and orientation**) are such that the experimental variables would be **easy to see** or observe if it were a real object,

4) Imagine an object of a size and orientation such that the variables in question could be **manipulated with one’s hands**.

*The intent here is not to check against a confirmed theory answering the target question, since in that case there would be no problem and no need to use a separate simulation. Rather, the intent is to check one’s prediction for coherence with other known constraints in the target and in the area of investigation.

Table 1. Strategies for Designing and Improving Central Test Cases and Analogies Using Imagery

In the present paper, we will report the results of three exploratory case studies involving teacher-generated test cases and analogies. In each case, we inquire whether the test case or analogy is consistent with the design strategies, and map the features of the case to the strategies we think they exemplify.

CASTLE: The Air Pressure Analogy

Although the development of the CASTLE curriculum (Steinberg & Wainwright, 1993) predates the generation of Table 1, the chief curriculum designer recalls that many of the strategies listed there were employed. From transcripts of classroom tapes and tapes of tutoring sessions, we have evidence that the test cases and analogies presented in the CASTLE curriculum were used by students in the course of their problem solving and in classroom discussion. Here, we will examine a metaphor that is used globally in the curriculum: a change in air pressure used as analogous to the voltage drop across circuit components. Even though the strategies were originally developed in connection with anchoring cases that had been designed for highly specific purposes, we will map the list of strategies to the design of this global metaphor. We use data from a previous analysis of one of the tutoring sessions (Clement & Steinberg, 2002) to hypothesize that the student made implicit use of the strategies.

In the curriculum, a color code is used to represent different electric “pressure” levels (voltages). Diagrams such as those below are colored by the students to reflect their predictions of the voltage drops across circuit components. Shown are a battery lighting two bulbs for four seconds as current flows through the bulbs to charge a large capacitor. Early pressures and flows are represented in the left-hand diagram and later pressures in the right-hand diagram.

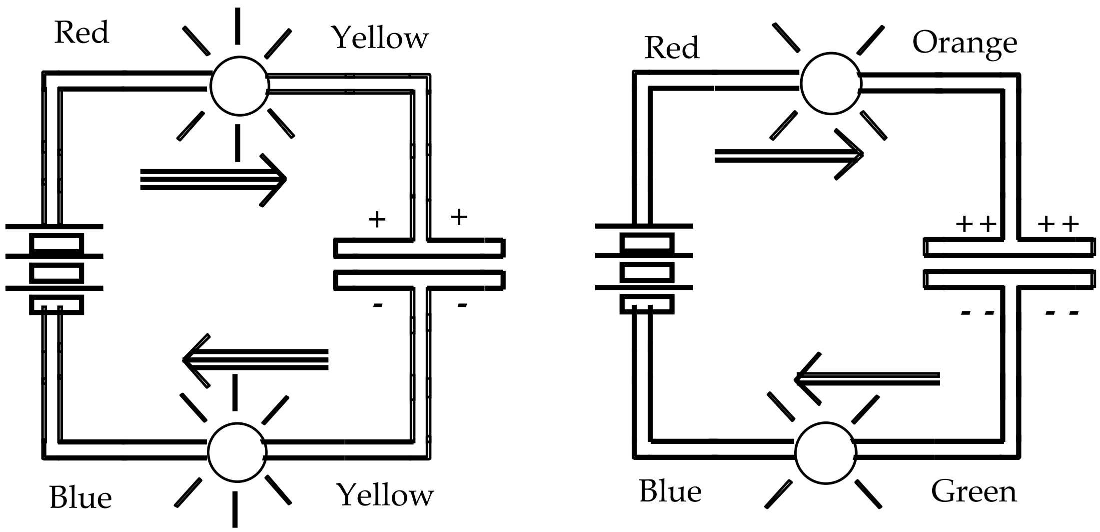


Figure 1. Colors are used in the CASTLE curriculum to represent voltage differences.

<u>The Strategies</u>

> Strategy A: Find a familiar case close enough to something you already know to make a confident prediction.

The CASTLE curriculum designers believe that air pressure is a concept more familiar or more teachable initially than voltage to most students. Evidence that this was true for the student in the tutoring session is discussed elsewhere (Clement & Steinberg, 2002). In addition to choosing that global metaphor for electric voltage or potential, particular familiar cases were chosen where students would have familiarity and thus prior knowledge to build on. One of these cases involves thinking about pressure in a car tire.

> Strategy B. Find a concrete case evoking imagery that has kinesthetic and/or visual components.

Again, the case of the tire is very concrete. Imagining changes in air pressure is believed to tap kinesthetic imagery. Evidence reported in Clement & Steinberg (2002) that the student used a kinesthetic component was her spontaneous hand motions that accompanied her verbalized ideas about pressure; she seemed to be squeezing an imaginary body of air. This happened both when she used the air pressure model directly and later in the protocol, when she imagined “electric pressure” in a circuit. Her words and hand motions gave evidence that she had transferred the imagery of air compression in a tire to charge compression in a capacitor plate. (Among evidence cited for this transfer of imagery was her phrasing, when she said “that charge, that air, whatever,” and the similar depictive gestures she used in the two cases.) To enhance the kinesthetic components of the imagery for students who have insufficient experience with air compression, the curriculum calls for students physically to manipulate air by compressing and extending air in a syringe.

> Strategy C. Find a case where one's predictions agree with known constraints or accepted physical theory.

The developers of CASTLE believe that an air pressure model will generate more predictions that agree with the behavior in actual electric circuits than will a flowing water model. Since water is not compressible, it does not match the experience of charging a capacitor where, by using a stronger battery, one can add even more charge to a plate that has already been filled with charge. The student in question initially had trouble understanding how a capacitor worked (Clement & Steinberg, 2002). However, after she understood this process, she was able to generate confident predictions for the current produced by voltage (electric pressure) differences that agreed with accepted physical theory.

> Strategy E1. To enhance spatial reasoning or schema application, use "markers" (mental diagrammatic notation that makes variable differences easy to imagine).

CASTLE uses a number of methods for enhancing spatial reasoning. Markers are used in the drawings to keep track of invisible quantities. In particular, color codes are used to indicate amount of electric pressure, and to allow students to visualize pressure changes around the circuit. (See Fig. 1.) The developers state that using markers and other visual representations together on the diagrams helps students coordinate the different variables (voltages, flow rates, and bulb brightness), enabling them to view these as descriptors of a single integrated system. In our example, we believe the student's use of markers enabled her to imagine transient processes that are much too brief to be observed in real circuits. Imagining these processes allowed her to do something that, in our experience, is unusual: she was able to predict the final states of rather complex circuits using qualitative causal reasoning (Clement & Steinberg, 2002). In comparison, in conventional classes, students are normally not required to understand how the final states come about. They are usually trained only to make a mathematical determination of those states.

> Strategy E2. To enhance spatial reasoning or schema application, use a case that is as simple and schematic as possible (one that is easier to visualize or to imagine kinesthetically).

As stated earlier, the curriculum assumes that a change in air pressure is easier for most people to visualize or to imagine kinesthetically than is a change in voltage. For instance, a kinesthetic image of a change in air pressure can be evoked by imagining squeezing tires or balloons inflated to different pressures.

> Strategy E3. To enhance spatial reasoning or schema application, imagine a case whose particular features (such as size and orientation) are such that the experimental variables would be easy to see or observe it if were a real object.

The diagrams used in CASTLE are of a size that the changes in pressure occur in areas that would be easy to see if the diagrams were physical objects. Also, the color changes would be easy to see. We hypothesize that this may foster the creation of mental imagery that also has these qualities.

> Strategy E4. To enhance spatial reasoning or schema application, imagine an object of a size and orientation such that the variables in question could be manipulated with one's hands.

The diagrams used are of a size that can accommodate students' use of markers and use of depictive hand motions over individual circuit components. For example, the student of our example is seen to move her hands over the diagram as she is visualizing the dynamics of the system.

We conclude that the development and presentation of this analogy, air pressure changes to voltage drops, is consistent with many of the design strategies from the expert list.

The Case of the Thousand-Pound Baby

This imagery-rich case was identified by one of us (Nunez-Oviedo, 2003) from the tape of a middle school classroom in a middle class suburban community, where the teacher was introducing a unit on energy. The class was videotaped and used a mixture of large and small group discussion. The innovative curriculum was Energy in the Human Body: A Middle School Life Science Curriculum (Rea-Ramirez, et al., 2004) and the unit, which they had been working on for about 6 weeks, included respiration, digestion, and circulation. The current episode occurred as the class was getting ready to talk about cellular respiration. Some students in the class held the misconception that the reason we breathe is because the air is a source of energy. In this episode, the teacher suggests to the students that if it were true they breathed in calories, by the second day after they had been born, their parents would have been holding a thousand pound baby. The evidence from the transcript and from a subsequent interview with the teacher indicates that this imagistic case was spontaneously generated by the teacher during the course of the classroom discussion.

The teacher began by asking the students about the sources of energy for our body, in response to which the students suggested food, sleep, exercise, water, and air. As a way to evaluate the idea of air as a source of energy for the body, the teacher suggested what amounted to a thought experiment involving respiration, caloric intake, and a newborn baby. In the following portion of transcript, the class was considering the suggestion that we get one hundred calories per breath.

T: You breathe 30 times per minute, and you get a hundred calories. So, in one minute, you get three thousand calories. In sixty minutes. . .

S: Eighteen zero, zero, zero!

T: Okay.

[Laughs]

The teacher then said that in twenty-four hours, one would breathe in four million, three hundred and twenty calories. The students were surprised, but the teacher was not finished. Apparently drawing on nutritional information, she equated one pound with three thousand, two hundred calories. Dividing the four million calories by this figure, she pointed out to the students that they would gain one thousand three hundred pounds per day.

T: So, if somebody says, Oh, you get energy from breathing, what evidence do you have that you don't?

S: Because you don't gain pounds.

T: Because you are not . . . over one thousand pounds, right? And, I mean, your parents, when you were born, were probably able to hold you for the first couple of hours. By the next day, you were a thousand pound baby.

[Laughs]

The teacher initially made her point with a mathematical argument, calculating a hypothetical number of calories to be gained from breathing and a consequent hypothetical gain in weight, but then she went on to introduce a thousand-pound baby. In investigating the reason for this new introduction, we note that, beyond adding humor, the new case is consistent with the use of the following strategies for imagery enhancement.

<u>The Strategies</u>

> Strategy A: Find a familiar case close enough to something you already know to make a confident prediction.

We hypothesize that this strategy was used in both parts of the illustration, before and after introduction of the baby. While calories are a fairly abstract idea, pounds are less abstract; children tend to know how many pounds they weigh.

> Strategy B: Find a concrete case evoking imagery that has kinesthetic and/or visual components.

We believe that the teacher's comment could have served as an invitation to imagine picking up a very heavy baby. This suggests that this teacher spontaneously chose to modify her imagistic case so that it had the power to evoke imagery with a strong kinesthetic component.

> Strategy C: Find a case where one's predictions agree with known constraints or accepted physical theory.

The students' prior (scientific, semantic, and/or intuitive) knowledge concerning air and energy had led them to hypothesize that air gives us energy. However, the teacher was able to lead them through a chain of reasoning into an area where their intuitive predictions were more likely to agree with accepted theory. While her mathematical argument concerning their daily weight-gain used deductive logic, it also appeared to serve the purpose of leading the students' attention from abstract quantities of energy to the more tangible quantities of weight. It is likely that the implausibility of actually picking up a thousand-pound baby tapped intuitive knowledge that was possessed by most, if not all, students in the class.

> Strategy D: To enhance the application of a schema in a simulation, generate a special case where the schema is easier to run; e.g., by generating an extreme case.

A weight gain of a thousand pounds a day constitutes an extreme example. Transferring the weight gain to a baby served to accentuate the extreme nature of the example.

> Strategy E3: To enhance spatial reasoning or schema application, imagine a case whose particular features (such as size and orientation) are such that the experimental variables would be easy to see or observe it if were a real object.

While this simulation did not appear to use spatial reasoning, it did exhibit some of the attributes of the expert methods that we have listed in this category. It would, for instance, be easy to observe the difference in a normally sized baby and a thousand pound baby.

> Strategy E4: To enhance spatial reasoning, imagine an object of a size and orientation such that the variables in question could be manipulated with one's hands.

The baby's weight would be easy to sense if it were a real object, and this weight could ordinarily be handled. These things were not true of the variables associated with the breath and internal energy delivery. In addition, a change in weight in our own bodies is normally not as easy to feel, subjectively speaking, as is the change in weight of something (or someone) we are holding.

The sections of transcript cited indicate that the case images used spontaneously by the teacher are consistent with the expert strategies in Table 1.

The US/Australia Case: Theme and Variations

In the two teaching situations discussed above, the teacher or curriculum was the primary source of key exemplars or cases. In the situation below, the teacher generates the first case, but the students take off from this to generate many more cases of their own. (These play the role of extreme cases, thought experiments, and analogies in the discussion. In a related paper, Stephens and Clement, 2006, this volume, evidence is presented supporting the idea that imagery is being used in all of the cases presented here.) We are interested here in what guides the generation of the cases--whether these teacher- and especially student-generated cases are consistent with the expert strategies in Table 1.

The high school was in a middle class suburban community, the class was college preparatory physics, and the curriculum (Camp & Clement, et al., 1994) was an innovative one. The students had just finished the subject of density and were now moving on to gravity. Common student misconceptions are that gravity is caused by the spinning of the Earth or by the downward pressure of the atmosphere. The target model was one in which each bit of matter in the universe is pulling on every other bit of matter.

The teacher began by presenting the following case to stimulate his students' thinking about gravity. He drew a simple picture on the board (Fig. 2) to indicate a round earth with one stick figure standing on the US and another on Australia, and then asked his students to vote on whether gravity in Australia is a little less, equal, or a little more than gravity in the US.

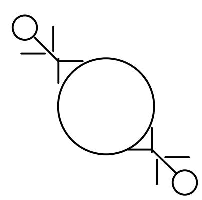


Figure 2. US/Australia Case

A number of variations of the case emerged during the discussion that followed. Rather than describe each of the variations, we diagram them to give an idea of the complexity and richness of the discussion.

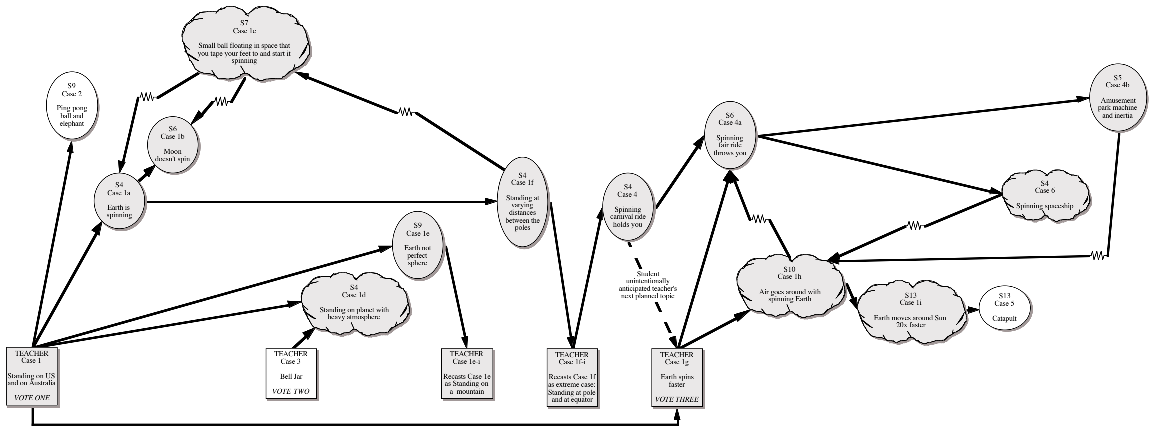
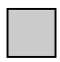
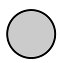
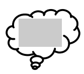
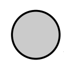
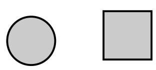


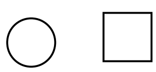
Figure 3. The Case: Theme and Variations

Key  *Teacher introduced the case or made the modification.*

 *Student introduced the case or made the modification.*

 *For present purposes, equivalent to* *. (These cases are discussed further in a related paper, Stephens and Clement, 2006, this volume).*

 *Close variations of Case 1.*

 *New cases (though they may have been inspired by Case 1).*

When mapping the list of design heuristics, we consider several of the student-generated variations.

<u>The Strategies</u>

Strategy A: Find a familiar case close enough to something you already know to make a confident prediction.

When trying to reason about the US/Australia case, the students generated variations and analogies for which they could make more confident predictions. Many cast their predictions for the original case by analogy to predictions they could more confidently make about these variations (Stephens & Clement, 2006, this volume). In particular, three students brought up fair rides and used them to reason as they generated predictions about the effect of the spinning of the Earth on the weight of a human being. (For more details about the fair ride cases, see Stephens & Clement, 2006, this volume).

> Strategy B. Find a concrete case evoking imagery that has kinesthetic and/or visual components.

When reasoning about the effect of spinning of the Earth, the students chose cases such as fair rides that make it possible to feel the effects of spinning. Evidence that they used the cases, not only logically, but to imagine *feeling* these effects is that they frequently spoke about how it would *feel* to stand on a spinning object, at times modifying their cases to produce a more intense spin. One example of this is the case of the spinning ball, labeled 1c in the diagram:

> S7: Well, in reference to rotation and gravitational force, I think of them as being two opposite forces because if you stand on . . . let's just imagine a ball floating in space you tape your feet to. And you start spinning the ball around, you're gonna feel like you're gonna be thrown off. But if it's a small ball, then the attraction between you and that little small mass is negligible so that you're just gonna feel the forces being spun around in a centrifugal force.

From the student's description, it does not appear that these kinesthetic components are a side issue, but that they play a central role in his reasoning.

> Strategy C. Find a case where one's predictions agree with known constraints or accepted physical theory.

A constraint known to these students was that gravity pulls them to the Earth. S4 generated the case of the carnival ride, saying that the outer wall would pull the rider toward the center of the ride. Although this was incorrect, it may have reflected his interpretation of this constraint.

> Strategy D. To enhance the application of a schema in a simulation, generate a special case where the schema is easier to run; e. g., by generating an extreme case.

The small ball spinning rapidly in space is also a striking example of an extreme case. This case was constructed as analogous to the turning Earth; it minimizes the effect of the pull of the mass and maximizes the effect of the spin. It appeared to allow S7 to enhance the application of the schema involved in standing on a rotating object so that it could yield a correct prediction. A number of other extreme cases were evident in the protocol, one of them spontaneously generated by the teacher. In a non-extreme variation of Case 1, a student had suggested that the amount of gravitational force a person would feel would depend on *how far* he was from the North and South Poles. The teacher turned this into an extreme case by asking the students to compare the weight of a person standing *on* the North Pole with one standing *on* the equator. This strategy appeared to work, as it triggered twenty-one lines of class dialog in response, with several students reasoning correctly about the implications of the surface velocity of a spinning Earth.

> Strategy E4. To enhance spatial reasoning, imagine an object of a size and orientation such that the variables in question could be manipulated with one's hands.

When reasoning about the Earth, these students often appeared to recreate it in the air in front of themselves as a sphere about a foot in diameter, and to animate it with their hands. One student, when pointing repeatedly to where he said a person was standing in the US and to where another was standing in Australia, did not point either to the drawing on the board or toward the general direction of Australia to the south, but to two points in the air about a foot apart.

It appears that these students, when reasoning about the US/Australia case, spontaneously used a number of the expert design strategies.

Findings

Evidence was presented that analogies and test cases described by teachers and students during activity in the classrooms studied share many of the design characteristics of expert analogies and test cases. One teacher-generated test case, a thousand-pound baby, was consistent with six of the eight design strategies. These design considerations, though used only implicitly by the teacher, are also evident in other test cases and analogous situations this teacher used. The global metaphor that was examined from the CASTLE curriculum, together with the representations given it, was found to be consistent with seven of the strategies. Student-generated variations of the US/Australia case were found to be consistent with six of the expert strategies.

There is evidence that these students could reason with cases that are consistent with the design strategies, and that this reasoning involved visual and kinesthetic components. The student in the CASTLE case study appeared to use kinesthetic imagery to reason about voltage drops in circuits. Student descriptions of kinesthetic imagery occur throughout the gravity class transcript and it appears that this kind of imagery is central to the reasoning processes of these students. (More detailed evidence indicating that students in the gravity class used dynamic imagery is presented in Stephens & Clement, 2006, this volume).

There is evidence that students can modify and evaluate teacher cases that were designed ahead of class. Much of the gravity transcript was an extended evaluation of a teacher-designed case. There were periods of time in this class when a series of modifications occurred with little input from the teacher (Fig. 3).

There is evidence that students can modify and evaluate cases spontaneously generated by other students during large class discussion. In the gravity class, the case of the fair ride underwent two modifications. There were many modifications to the case of the spinning Earth; these appear to have been generated as a way to evaluate the original student case.

Evidence was presented that some students can generate enhancements to the imagistic qualities of cases. An example is the student who enhanced the imagistic qualities of the case of a slowly spinning Earth by generating an extreme case, that of a small ball rapidly spinning in space.

Conclusion

Expert strategies for designing imagistic cases and analogies appear to have applicability to the test cases and analogies generated by curriculum developers, teachers, and students in these learning situations. In the transcripts of both the CASTLE tutoring session and the gravity class, there were multiple indicators providing evidence that students were able to reason about their conceptions with the imagery from the cases. According to reports from the teachers involved, the teacher cases analyzed here that were generated ahead of class were designed specifically either to build on or conflict with prior conceptions the teachers believed their students were likely to hold. When a teacher or curriculum designer is aware of the prior conceptions likely to be held by the student population, it may be that a honed list of expert strategies can facilitate the advance planning of analogies and test cases that have the power to stimulate imagery helpful in working with these prior conceptions in order to construct and revise student mental models.

We hypothesize that, for the teacher who presented the case of the 1000-pound baby, considerable experience in the classroom enabled her spontaneously to generate vivid imagery that students could then adopt and run via mental simulation. Likewise, it may have been his long experience that enabled the teacher in the gravity class to respond to student comments by quickly refining their imagery so that they could more easily reason with it (as when he took a student case and generated the extreme version of a person standing at the North pole and another on the equator). It may be that other teachers would appreciate guidelines in this area. This suggests pursuing the goal of honing the list of strategies into a list of principles teachers and curriculum developers can use when designing or selecting analogies and test cases to elicit and work with prior conceptions in their science classrooms.

References

Camp, C., Clement, J., Brown, D., Gonzalez, K., Kudukey, J. Minstrell, J., Schultz, K., Steinberg, M., Veneman, V., and Zietsman, A. (1994). Preconceptions in mechanics: Lessons dealing with conceptual difficulties. Dubuque, Iowa: Kendall Hunt.

Clement, J. (2006). Strategies for imagery use in expert protocols. Proceedings of the NARST 2006 Annual Meeting, San Francisco, CA.

Clement, J. and Steinberg, M. (2002). Step-wise evolution of models of electric circuits: A “learning-aloud” case study. *Journal of the Learning Sciences, 11*, 389-452.

Gentner, D. & Gentner, D. R. (1983). Flowing waters or teeming crowds: mental models of electricity. In D. Gentner & A. L. Stevens, (Eds.) *Mental Models* (pp. 99-129). Hillsdale, N. J.: Erlbaum.

Gentner, D. & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science 10*, 277-300.

Hegarty, M. (1992). Mental animation: inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 1084–1102.

Kozhevnikov, M., Hegarty, M., Mayer, R. (1999). *Students' use of imagery in solving qualitative problems in kinematics*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Nersessian, N. (1992). Constructing and instructing: the role of 'abstraction techniques' in creating and learning physics. In R. Duschl & R. Hamilton (Eds.) *Philosophy of Science, Cognitive Psychology and Educational Theory and Practice* (pp. 48-68). New York: State University of New York Press.

Nersessian, N. (1995). Should physicists preach what they practice? constructive modeling in doing and learning physics. *Science & Education, 4*(3), 203-226.

Nunez-Oviedo, M. C. (2003). *Teacher-student co-construction processes in biology: Strategies for developing mental models in large group discussions*. Unpublished Doctoral Dissertation, University of Massachusetts, Amherst, MA.

Polya, G. (1957). *How to Solve It,* 2nd ed., Princeton University Press.

Rea-Ramirez, M., Nunez-Oviedo, M., Clement, J. (2004). Energy in the Human Body: A Middle School Life Science Curriculum

Reiner, M. & Gilbert, J. (2000). Epistemological resources for thought experimentation in science learning. *International Journal of Science Education 22*(5), 489-506.

Steinberg, M.S. and Wainwright, C.L. (1993). Using models to teach electricity – the CASTLE project. *The Physics Teacher*, Vol. 31, Sept. 1993, pp. 353-357.

Stephens, L. & Clement, J. (2006). Designing classroom thought experiments: what we can learn from imagery indicators and expert protocols. Proceedings of the NARST 2006 Annual Meeting, San Francisco, CA.