RUNNING EFFECTIVE CLASSROOM THOUGHT EXPERIMENTS: WHAT EXPERT PROTOCOLS AND IMAGERY INDICATORS CAN TELL US[1]

A. Lynn Stephens, University of Massachusetts, USA
John J. Clement, University of Massachusetts, USA

Objectives

The purpose of this study is to explore reasoning processes that students use when running thought experiments (TEs) in science classrooms. We also focus on the use of depictive hand motions, which we regard as providing some evidence for the involvement of mental imagery in these episodes of student reasoning. We discuss our analysis of several of the episodes in terms of four forms of expert reasoning that have been associated with expert use of TEs (Clement, in preparation). We coded independently for these processes and for the occurrence of depictive hand motions. This allows us a window onto the roles imagery (and TEs) appear to be playing in student thinking during these large class discussions, where students were being asked to generate and evaluate explanatory models of phenomena.

Perspective: What are TEs and can students run them?

A number of authors have speculated that TEs make extensive use of mental imagery (Sorensen, 1992) and that both dynamic imagery and embodiment can play a crucial role (Gooding, 1992b; Nersessian, 1992; Reiner, 1998; Reiner & Burko, 2003; Reiner & Gilbert, 2000). It is generally assumed by these authors that a TE is an experiment carried out in thought and that an outcome must be predicted or inferred. However, even though a typology of TEs has been developed (Reiner & Burko, 2003) and different stages of TEs identified (Reiner, 1998), the definitions provided have not made it clear whether it is necessary that the purpose of the experiment be to test a theory; although, at times, this has been strongly implied (Gilbert & Reiner, 2000). This issue is addressed here by adopting the following taxonomy (Clement, 2002; to appear), which distinguishes between TEs in a broad sense and TEs in a narrower sense:

> *Performing an (untested) thought experiment (in the broad sense):* the act of considering an untested, concrete system* (the "experiment" or case) and attempting to predict aspects of (or underlying causes for) its behavior. Those aspects of behavior must be new and untested in the sense that the subject has not observed them before nor been informed about them.

[1] 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.

> **Concrete system* here means one involving concrete objects or experiences (and relationships between them) rather than involving abstract higher order relations only.

Performing an evaluative Gedanken experiment (in the narrow sense): the act of considering an untested, observable system designed to help evaluate a scientific concept, model, or theory—and attempting to predict aspects of its behavior.

In the latter an element of a theory is tested as it is applied to the case. Unlike earlier definitions, the broad definition above contains an *observable behavior* that can be identified in protocols. Whether TEs are considered in the broad or the narrow sense, there is some evidence that they can involve mental simulation (Clement, 1994, to appear). We will further discuss evaluative Gedanken experiments below.

Reiner & Gilbert (2000) indicate that some students can and will use TEs to find solutions to problems in cases where the problems are formulated in a way to encourage this kind of solution process. Analyses of expert protocols have yielded a set of imagery-use indicators, and evidence supports the contention that they can generate dynamic imagery in TEs as they access implicit knowledge and make it more explicit (Clement, 1994, 2003). Reiner & Gilbert (2000) believe that only a small portion of this kind of knowledge can be articulated verbally. We will provide additional evidence that addresses the question of whether TEs can involve dynamic imagery (imagery that can include strong kinematic and/or kinesthetic components). Analysis of the teaching tapes will examine whether these properties can be fostered instructionally.

Theoretical Framework

Previous research includes the structure of TEs (Brown 1986); the function of TEs in scientific thinking (Kuhn, 1977); and imagery in problem-solving in physics (Clement 1994). Giere (1988) and Darden (1991) argue that the ability to generate and evaluate mental models is a crucial aspect of science, and Nersessian (1992, 1993) and Gooding (1992b) believe that one powerful way to do this is to run a thought experiment. Research continues to indicate the importance of mental modeling in experts and students (Gentner, 2002; Nunez-Oviedo, Rea-Ramirez & Clement, to appear). There has been much research in student inquiry in science (e. g., Driver 1983) and students' use of visual and kinesthetic imagery (Hegarty, 1992; Kozhnevnikov, Hegarty & Mayer, 1999). Recent research on TEs has investigated the value of divergent, qualitative thinking methods including new knowledge obtained through TEs (Gooding, 1992a, 1992b, 1996); the nature and function of TEs in scientific discovery (Clement, in press); the importance of TEs in teaching and learning (Clement & Steinberg, 2002; Gilbert & Reiner, 2000); and the ability of students to generate and use TEs in small-group collaborative settings (Reiner, 1998; Reiner & Gilbert, 2000). However, very few of these studies investigate the role of TEs in large class discussion. TEs have more often been studied within the context of small-group sessions or individual think-aloud protocols (Clement, in press). Hammer (1995) described thought experiments in physics class discussions as one of several kinds of process skills that were exhibited by students when the teacher in his case study took care to foster an open attitude toward contributing ideas. Nunez-Oviedo (2003) has analyzed the role of TEs in the classroom from the standpoint of the apparent uses to which they are put by the teacher, but we believe there is much more that can be learned about the roles TEs play in student reasoning, and especially of the uses students can make of them in the classroom.

Methodology

We have examined a number of transcripts of classroom activity where inquiry-based methods of teaching and learning were employed. Using the definitions given above, the transcripts were coded for the presence of TEs. For selected transcripts, a detailed list of imagery indicators developed by Clement (1994) was used to code TE episodes for evidence of the presence and use of mental imagery (Stephens & Clement, 2006b). Additional transcripts were coded for the presence of depictive hand motions only. These hand motions appear to depict an image in the air, and are taken as one indication that mental imagery is being used (Clement, 1994). Here, we analyze several classroom episodes for the following factors: whether the student utterance gives evidence for the presence of a TE in the broad sense, whether it gives evidence for the use of reasoning processes associated with a Gedanken experiment; whether there is evidence for the presence of an analogy or an extreme case; and whether there are similarities between the uses that students appear to be making of the TEs and the uses experts report making of them. We are still using coding as a way to develop stable categories in this exploratory area, so in all cases coding was done jointly by the two authors and disputes were used a mechanism for refining and clarifying the coding criteria. In one of the transcripts, it was also possible to code for evidence that the students were running explanatory models. (An explanatory model is one that projects some initially hidden feature into a system that explains why the system behaves the way it does). We also note where these incidents co-occur with depictive hand motions, which will be the main form of imagery indicator considered in this paper.

Data sources

The two case studies examined here are of lengthy discussions that occurred in college preparatory physics classes that were using an innovative curriculum (Camp, Clement, et al., 1994). The classes were in a middle class suburban high school in the northeastern United States. Taught by the same teacher, the classes were from different years and the discussions were on different topics, though gravity was a factor in both. The classes were videotaped. These transcripts were selected because the discussions were animated and because a single TE had previously been identified in each. Each of the discussions lasted about 45 minutes. One of them occurred during a single class period while the other comprised the second half of one class period and the first half of the same period on the following day.

In the first transcript from which we will draw examples, the teacher wanted students to consider whether a table exerts an upward force on objects resting on its surface. A common conception prior to instruction is that inanimate objects cannot exert upward forces against gravity. The target model for the lesson was one in which objects exert normal forces that are equal and opposite to the weight of objects resting on them. The teacher began by introducing an analogy. He placed a book on his desk and called students' attention to it, then drew two figures on the chalkboard. One was a simple line drawing of a book on a table (the target), and another of a hand pressing downward on a spring (the base). He asked the students to compare the two cases. He hoped that all of the students would believe that the spring pushed up on the hand and that he could use this as an anchoring case for the lesson. It became clear that, although many of the students did believe that the spring would exert a force on the hand, a large number still did not believe that the table was exerting a force on the book. The teacher intended to introduce a number of *bridging analogies*, designed to bridge the distance between the spring/hand case and the table/book case. However, these students also preempted him, producing their own bridging cases and reasoning about them. We will call this the “Book on Table transcript;” selected quotations appear in Table 1.

The second transcript will be referred to as the “Gravity class transcript.” The class had finished a unit on density and was just beginning a unit on gravity. Common conceptions of students prior to instruction are that causes of gravity include the rotation of the Earth and/or the “downward” pressure of the atmosphere. The target model of the lesson was one in which every particle of matter pulls on every other particle. The teacher began by introducing a case designed to elicit misconceptions such as those just listed and to stimulate discussion. He drew a figure on the board (see Fig. 1) and asked the class to vote on the following: “Compared to the United States, gravity in Australia is: a little less, equal, a little bit more.”

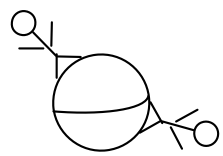

Figure 1: US/Australia Case

After the students had recorded their votes on voting sheets, the teacher opened the discussion by asking, “Just what is it that causes gravity, anyway?” It soon became apparent that, while a number of students believed that gravity was due to “everything pulling on everything else,” some of them did believe the spinning of the Earth also plays a role. What followed was a very lively discussion in which students modified the teacher-presented case in various ways to bolster their points. They also created analogous cases and new cases as they tried to convince their fellow students of their respective positions. The teacher played an almost neutral role, restating student positions, asking for clarification, and occasionally, recasting a student utterance into a slightly altered form. The teacher also presented an additional case, that of a vacuum jar, designed to stimulate discussion about the effect of air pressure on weight. He had planned to present a third case to stimulate further discussion about the effect of rotation on gravity, the case of a more rapidly rotating Earth, but his students preempted him by producing that case before he could introduce it. Fig. 2 is a flow chart of the test cases presented in the classroom discourse and is intended to be used as an advance organizer for our discussion to follow. It indicates how student cases appeared to build on each other and on the cases presented by the teacher. Jagged lines show where a case was used to argue against a statement made by another speaker. Selected quotations from throughout the discussion are in Table 2.

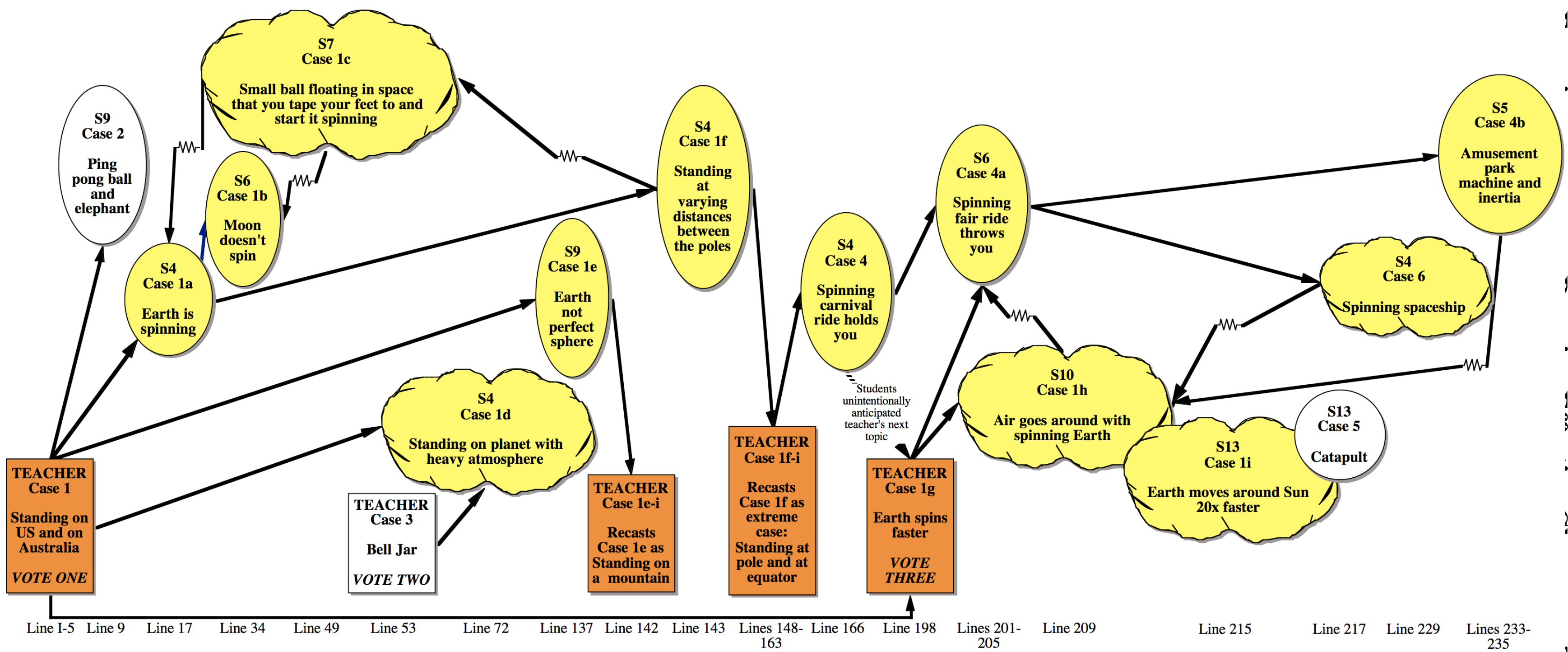

Figure 2. Flow chart of test cases in the Gravity transcript.

Below, we discuss four forms of expert reasoning (one of which is evaluative Gedanken experiments) that appear to have been occurring in these classrooms. We then present a summary of results from the coded transcripts.

Four Expert Thinking Processes

Clement (in preparation) has identified a network of thinking processes that allow experts to generate ideas divergently (brainstorming) and then to constrain these ideas in a rigorous evaluative process that is highly convergent. These processes are central to evaluating mental models and therefore are central to learning. The students in these classes appeared able to generate ideas quite freely; we are also interested in the thinking processes that the students used to evaluate their own ideas and the ideas of their classmates. The expert thinking processes that we consider are: running an explanatory model, generating analogies, generating extreme cases, and running Evaluative Gedanken experiments. These processes as defined here are not intended as exclusive categories; rather, more than one can apply to the same case in some circumstances.

In our view TE's are often run via an imagistic simulation process, and some evidence for this will be provided by the transcripts. As such, TE's can occur as part of any of the above processes or can occur alone. Gedanken Experiments are the most complex process, and can involve other processes in a variety of ways. Evidence for all of these forms of reasoning will be identified in student utterances.

<u>Running an Explanatory Model</u>

An explanatory model is a mental model of a system that projects some initially hidden feature into the system that offers an explanation for why it behaves the way it does.

In the Book on Table class, a student says the following:

> S(13): If you put something heavy on the table and it collapsed, that is because the table is not exerting enough force.

Saying that if you put something heavy on a table it will collapse, can be interpreted as running a mental model of a table. The model predicts that the table will collapse under certain conditions, those conditions being when it is subjected to a weight great enough. So far, there is nothing explanatory about this model. However, to go on to reason that the *cause* of this behavior is that the table does not exert enough force is to project the initially hidden feature of forces into the system of table + heavy object.

Running an explanatory model can involve a TE, and evidence for this occurs elsewhere in the Book on Table transcript. A student had commented that a similarity between the hand-on-spring and book-on-table situations was that the force pushing down was more than the force pushing up. Another student responded:

> S(24): At a certain point, the force pushing up on the hand and the spring, they have to be equal, because the hand can't push the spring any farther down.

This sentence alone indicates the running of an explanatory model of the hand-on-spring system, projecting forces into the hand and into the spring and predicting that those forces will be at equilibrium when the hand can no longer push the spring farther down. This does not meet the definition of a TE that we gave above because the hand-on-spring was most likely not an untested system for the student. Though she may not have thought about it in this way before, it is likely that she had some memory of how it feels to push down on a spring. However, she continued:

> S(24): And if the table under the book wasn't exerting some kind of force back on the book, then the book would go down.

The student has—we are pretty confident in this—never experienced a table that does not exert a force when under a load, but she has made a prediction for an observable effect that would be produced by such a system. She appears to be running an explanatory model of a table without the ability to exert a force, and she predicted that such a table could not hold up a book. We consider the second part of this a TE in the broad sense, though it does not match the structure of a Gedanken experiment. More will be said about that in the next section.

Analogies

For our purposes, analogies occur when a subject, in thinking about a target situation A, shifts, without being prompted, to consider a situation B (the base) which differs in some significant way from A, and intends to apply findings from B to A.

Although the teacher had introduced the cases of the hand-on-spring and book-on-table, when S(24) uttered her comments above about the system of book-on-table (the target), it is plausible that she had noted a new analogical relation between the spring (the base) and the table: both will stop things from moving. She appears to have convinced herself by running a model in the case of hand-on-spring that, in the event that the force of the spring had not matched the force of her hand, she could have continued to move her hand downward. She then used this analogical relation to generate the prediction that we have taken as evidence for a TE: in the absence of a matching force from the table, the book would go down. She appears to have run a model in a familiar system (many students have felt a "push" from a spring) and transferred her explanatory model to a system where she could not directly observe the effects (of the forces between a book and a table). If she had reached her conclusion about the forces between book and table by drawing an analogy to another system *that was also unfamiliar to her*, and, moreover, the system appeared to have been *designed by her expressly to evaluate her hypothesized explanatory model* for the interaction between book and table, this would have met the definition given above for an Evaluative Gedanken experiment. We will see examples of this below.

An interesting example of an analogy that apparently did not involve a TE, though it represented a bit of reasoning we found impressive, occurred in the Gravity class. Toward the end of the discussion about whether the daily rotation of the Earth affects gravity, a student offered this analogy:

> S13: If somebody puts me in a catapult and I go hurling two hundred feet into the air, gravity is the same, there's just another force acting.

The analogy is drawn between two *relationships*: the relationship between centripetal force and gravity on the one hand, and the force of the catapult and gravity on the other. There are important differences between the constant centripetal force of a rotating planet and the impulsive force of a catapult, and the equivalence of these forces is not asserted beyond the equivalence of their *relationship* to gravity; that is, that they are each a force acting alongside gravity but additional to it. This analogy also appears to involve the running of a model, though the model is not necessarily explanatory. It cannot be coded a TE: the student may well have played with projectiles, so there is no reason to assume that the system is untested. Also, no prediction is made for any previously unobserved behavior. However, the example is striking for its strong kinesthetic components, as the student puts himself in the place of the projectile. (For more on this strategy of refining and heightening kinesthetic imagery, see Clement, 2006; Stephens & Clement, 2006a.)

The teacher began the Book on Table class by presenting the analogous case of a hand pushing on a spring, a case deemed more likely to trigger kinesthetic memories and to be easier for the students to "run." The whole lesson was structured around a series of analogies (see the curriculum, Camp, et al., 1994; also Clement, 1993). It can also be seen from the transcripts that the teacher repeatedly mentioned to the class that he was using analogies. In at least one instance the teacher labeled a student utterance as an analogy, whereupon another student responded, "By the same analogy. . . . " and used it to generate a new prediction.

Extreme Cases

We will say that an extreme case has been run when a subject, in thinking about a target situation A, shifts, without being prompted, to consider situation E (the extreme case) where some variable from situation A has been maximized or minimized.

When reasoning about whether a table would necessarily exert a force on any object that rested on it, students began considering a "warped" table as an object that might be able to move and to act a little like a spring. A number of students seemed to feel that a warped table would exert a force. However, one student wished to reason further, and asked whether a perfectly rigid table would exert a force. When the teacher responded by referring to "those strong tables" in the back of the room, the student continued:

> S4: But if we had an ideal table that did not move at all . . . then I don't see how it could be pushing up on the book.

This extreme case could not move, by definition, and therefore could not act as a spring does. In bridging from a spring to springy objects (including the warped table) and on to rigid tables, the student had overshot the target model of a table that moves imperceptibly under a load and gone further, it appears, than the teacher had anticipated—to a table that was a Platonic ideal. This gave the student a new situation in which to run the model of (change in) movement-produces-force. Running the explanatory model in the new situation produced the prediction that the table would not be able to exert a force. Although this is, in fact, an accurate prediction, this student's extreme case had taken him beyond where the lesson was designed to go. There was nothing to prevent some students from drawing an unproductive inference, that a force was therefore not necessary to hold a book. Other students, however, correctly reasoned that, if the table were able to exert no force, the book would just fall down. As creative as this case was, we did not code it a TE. The prediction was not for an observable effect, but for the presence (or absence) of a causal factor. The observable in this exercise (absence of movement) was built into the definition rather than being a matter of prediction. This is one of a number of cases in these transcripts that would appear to argue for the value of expert reasoning processes that can use TEs, even when those processes happen to be used sans TEs.

One strategy the teacher used was to take student statements and recast them as extreme cases. In response to a question about whether gravity would change if one climbed a mountain, a student replied,

> S4: I think how far you are from the poles has more to do with it.

Although a semi-quantitative relationship is implied here, it is doubtful whether the student would have taken it further had the teacher not recast the comment,

> T: Now the other issue that you're bringing up that was kicked around some and not resolved last time was that the gravity has to do with the Earth spinning, also is another issue that was mentioned. If that's the case, let's give it a little bit of thought about what (S4) is saying. If I were to stand at the North Pole, say the pole is here and I hold on hand on the pole, how long does it take me to spin around that pole?

Once the class reached agreement that it would take one day and that the movement around the pole would be slow, the teacher continued,

> T: Let me point out, if I stand on the equator, however . . .

And a student replied,

> S: You're going real fast.

The teacher has converted the vague phrase: "how far you are from the poles" into an extreme case and the students have immediately begun to reason with it. In fact, this case continued as the topic of discussion for the next several minutes. The student who had made the original vague statement about the poles now ran the case in way that is consistent with our definition of a Gedanken experiment, and another student re-ran it with a slight refinement that produced more accurate results.

Evaluative Gedanken Experiments

Gedanken experiments, as defined above (and discussed in Clement, to appear) can be quite complex and come in many varieties. We will first consider one that is fairly simple. In the Book on Table class, a student had drawn an analogy between the book situation and a situation the class had studied earlier, that of a boat powering upstream against a current. If the current were to stop suddenly, the boat would move upstream in response. (It is not clear from the transcript whether the student was considering the thrust of the boat engine initially to be in equilibrium with the force of the current, but the problem is most simply described that way.) Another student replied,

> S4(15): But, but, by the same analogy then, if gravity disappeared, right, the force of the engine on the book, even the book would just fly off into space.

The student evaluated the explanatory model of the book resting on the table, which he must have known from experience would remain at rest. He evaluated this model by using (designing) an analogous situation that was certainly untested—the situation where gravity turns off. He then ran this untested case to generate a (theoretically) observable prediction, that, if gravity were turned off, the book would fly off into space. Interestingly, he appeared not to trust this prediction and to regard it as evidence against the table-pushing-up model. His prediction was actually scientifically sound.

A more complex example occurred in the Gravity class. Some students had suggested that the rotation of the Earth either causes gravity or contributes to it. Although several students had countered this idea, one of them by a Gedanken experiment, the proponents of the spinning model of gravity appeared not to be convinced. A student suggested the following:

> 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 going to 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.

In addition to having some powerful kinesthetic components (as anyone who has gotten queasy from being swung around can attest), this also contains an analogy to a familiar case. A ball is a familiar object to most students, and presumably, most students could realize, upon reflection, that they had never felt a gravitational pull from one. (To a physicist, this case would not be analogical, as the ball and the Earth would be seen as essentially identical, with the amount of mass the only relevant factor. However, to many students, this equivalence is not at all apparent.) In addition, this student employed the ball/Earth analogy in an extreme case. In it, the massive Earth, rotating through a complete revolution once in 24 hours, becomes shrunk to ball-size and sped up until it is spinning rapidly. This minimizes the gravitational force and maximizes the centrifugal force. After the extreme case was presented, the TE (in the broad sense) was brought into play, as the student suggested to his classmates that they imagine their feet taped to the ball. He generated a prediction from this situation (untested—unless he has previously taped himself to a ball in space!): “you’re gonna feel like you’re gonna be thrown off.” He had designed this case expressly to evaluate the spinning-as-contributing-to-gravity model. If the prediction had been for a centrifugal effect in the same direction as gravity, we could have considered this a Gedanken for which the results were nil. As it was, the prediction of a force opposite to that of gravity produced a result (in principle observable) that would tend strongly to discount spinning as an explanatory model for gravity.

Results of coding

The results of coding the two class sessions for expert reasoning processes and for co-occurrence of depictive hand motions (an indication of mental imagery) are summarized in the tables below. As an exception, we did not attempt to code the Gravity transcript for Running Explanatory Models because the discussion became so complex, with students reasoning about a number of explanatory models for different kinds of phenomena that they associated with gravity. The Book on Table transcript is coded for all of the forms of expert reasoning discussed above: Running Explanatory Models, Analogies, Extreme Cases, and Evaluative Gedanken Experiments.

In the four-part table for the Book on Table class, note the frequent pairing of evidence for running an explanatory model with coding for depictive gestures. We have coded the gestures as Shape Indicating [G-S], Movement Indicating [G-M], or Force Indicating [G-F], according to the images they appear to depict. (See the list of imagery indicators developed by Clement [1994]).

Note also the complexity of reasoning in passages where there was evidence for the use of multiple expert processes. (Although a number of these were in instances where the student reached a conclusion that did not agree with current scientific theory, these were part of the dialectic process that led students toward the target model.)

Table 1a. Book on Table

Running Model		Running Model	Running Model	Running Model	
	Analogy	Analogy			Analogy
TE	TE	TE	TE*		TE
Gestures	Gestures				
S(15): [G-M] If you think about it, when the book is on the table, the table gets warped a little bit.	S(15): [G-F] If you imagine a table built out of a balloon. . . the force pushes the book up. . .	S(24): (The forces the table and book exert on each other) have to be equal, because the hand can't push the spring any farther down. And if the table under the book wasn't exerting force on the book, the book would go down.	S(5): The table moves, you just can't measure it.	S(24): Hand is pushing down on spring, while gravity is pulling down on the book.	(jointly constructed) S(15): we build the table out of something pliable S(3): plywood S(15): cardboard, paper S(?): Bounty

* We are assuming that the student statement is not a recitation from authority.

Table 1b. Book on Table

Running Model	Running Model	Running Model	Running Model	Running Model	Running Model
				Extreme Case	
TE*	TE				TE
Gestures	Gestures	Gestures	Gestures	Gestures	Gestures
S(5): [G-M] With the bendy table, it's magnified more, so you can see it, and on the hard table, it's an insignificant amt of movement, but equal amt of force.	S(15): [G-S] Bend tiny bit, same amt of force (as bendy table)	S(5): [G-M] Atoms bend. Force is there.	S(13): The table's holding it up whether it moves or not.	S(15): [G-F] If we had an ideal table that doesn't bend at all . . . then I don't see how it could be pushing up.	S(13): [G-M] if the table didn't exert any force at all on the book, the book would just go down.

* We are assuming that the student is stating something previously unknown to him/her.

Table 1c. Book on Table

	Running Model		Running Model	Running Model	
		Analogy			
Extreme Case					Extreme Case
Gedanken			Gedanken		
TE	TE	TE	TE		
Gestures	Gestures	Gestures	Gestures	Gestures	Gestures
S(12): [G-S] [G-M] If the table is perfectly rigid, you could almost call it part of the ground. If the table isn't there, the book is going to hit the ground. But you could always dig a hole, right? It keeps falling. You're going to say the ground is moving, too?	S(14): It's got to be pushing up some way or the book would just fall down.	S(14): [G-M] It's like the river. The velocity of the engine is pushing down and the velocity of the current is pushing up. If you take the current away, then the engine; if you take the force of the table away, then the book would fall down.	S(15): [G-F] [G-M] But by the same analogy, if gravity disappeared, the book would fly off into space.	S(13): [G-F] [G-M] If you put something heavy on the table and it collapsed, that is because the table is not exerting enough force.	S(14): [G-M] Like the river going down at 3 and the boat going up at 5, then the boat is moving upstream. That's what we're saying about the elephant on the table.

Table 1d. Book on Table

Running Model	Running Model	Running Model	Running Model
			Extreme Case
	Gestures	Gestures	Gestures
S(5): But the table only has the power to just counter, it doesn't have enough to exceed and move in the other direction.	S(5): Yeah, the spring has the power to (move in other direction)	S(6): [G-M] The table has to counter the force of gravity, too, although it would never make the book jump.	S(15): [G-F] The idea of the elephant sitting on the table is consistent with the warped table theory--it punctures the table because it warps it too much.

In this table, there is evidence for student use of kinesthetic imagery and for repeated use of all of the expert reasoning processes we described earlier. Although some evidence for a TE had previously been identified in each of these transcripts, we were surprised at the amount of evidence for TEs, analogies, extreme cases, Gedankens, and explanatory models that was revealed once we began coding for it with definitions in mind. We are intrigued by the fact that these students appear able to use these expert processes.

The breakdown of the results is shown in Table 2.

Table 2a. Breakdown of Incidents: Book on Table transcript, approx. 45 minutes

Combinations of processes	# of cases identified in	# of these cases associated with depictive hand motions	# of these cases associated with TEs
Running a model alone	13	10	6
Analogy alone	3	2	3
Extreme case alone	1	1	0
Gedanken alone	0	0	0
Running a model & analogy	1	0	1
Running a model & extreme case	2	2	0
Extreme case & Gedanken	1	1	1
Running a model & Gedanken	1	1	1
TEs alone	0	0	0
Total Cases	22	17	12

Table 2b. Total Incidents: Book on Table transcript*			
	# of cases or incidents	# of these cases associated with depictive hand motions	# of these cases associated with TEs
Running a model total	**17**	**13**	**8**
Analogies total	**4**	**2**	**4**
Extreme cases total	**4**	**4**	**1**
Gedankens total	**2**	**2**	**2****
TEs total	**12**	**9**	**---**

** Many of the processes listed here were run in combination, so some cases are represented more than once in these numbers. The columns do not add. See Table 2a.*

*** Although all the Gedankens in this paper are TEs, our definitions do not require this.*

There were:

- 17 cases where the expert reasoning processes were paired with depictive gestures, including 5 cases paired with force-indicating gestures;
- 7 other cases in the transcript where depictive gestures occurred; leading to
- 24 instances total where depictive gestures occurred, many with multiple gestures, comprising
- ~53 individual depictive gestures by students in about 45 minutes of transcript.

Of the 22 cases that involved the expert reasoning processes being examined here, most involved depictive hand motions, providing evidence for the use of imagery within these processes.

In the three-part table for the Gravity class on the following pages, note the striking correlation between evidence for imagery (depictive gestures) and evidence for the running of TEs. These depictive gestures were often quite emphatic, as in Fig. 4, below. We take these kinds of gestures to be an indication that the students were using kinesthetic imagery.

A potentially useful way to think of the gesture categories is in terms of the kinds of mental imagery for which they are an indication: Shape indicating—*static imagery*; Movement indicating—*dynamic imagery*; and Force indicating—*dynamic/kinesthetic imagery*.

Fig. 4. "The Earth has more pull *on you."*

Table 3a. Gravity

Extreme Case	Extreme Case Gestures	Gedanken TE Gestures	Analogy Extreme Case Gedanken TE Gestures	TE Gestures	Analogy Extreme Case Gedanken TE Gestures
[Line 9]	[Line 12]	[Line 40]	[Line 49]	[Lines 60-64]	[Lines 72-74]
S9: If you put a ping pong ball beside an elephant, it would probably move the elephant.	S9: [G-M] You don't feel the attraction. But in order to see (movement), you'd have to have a really really tiny mass and a really really huge mass.	S5: [G-F] I just think that gravity has nothing to do with rotation, but maybe (with the rotation of the Earth) that guy (in Case 1) is trying to get thrown off the Earth. . . getting pulled at the same rate but he's also getting pushed.	S7: [G-F] I think of (rotation and gravitational force) as being two opposite forces. . . 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. . . .	S5: [G-F] [G-M] I don't see how taking air out (of the bell jar, Case 2) changes the weight of (a mass hanging from a spring scale inside the bell jar). . . . Unless the air bouys it up and that weight weighs less than air.	S4: [G-M] If you increased air pressure incredibly on all of us, like if you go to a planet with a lot more air pressure, you'd just get squished. Taking away the air pressure (in the bell jar) would make the weight lighter.

Table 3b. Gravity

TE Gestures	TE	TE	Analogy Gestures	Gedanken → TE →	refined (Gedanken) refined (TE) Gestures
[Lines 76-78]	[Line 89]	[Line 141]	[Line 166]	[Line 182]	[Lines 184-186]
S10: [G-M] Air pressure from all sides if (it was) removed would all cancel out. (The weight) would weigh the same.	S (off camera): (If there was an absolute vacuum inside the bell jar, it would not implode) if the air pressure on the outside was weaker than the force of the (xxx) of the glass.	S9: (If you climbed a mountain) the difference (in gravity) is so minute that you couldn't really measure it. I think there's a change, but. . . .	S4: [G-M] [G-F] You know that machine at the carnival that spins around and. . . turns on its side and you stay? I think that spinning will generate a lot of force and not always trying to throw you . . . in this case it's somehow helping to hold.	(from Lines 144-165) S4 (off camera): OK, say that (the rotation of the Earth) is throwing you. Then that still means that the top (North Pole) is still gonna be throwing you left and at the side (at the equator) (xxx). So your weight's gonna be different.	S4: [G-F] How could (rotation of the Earth) not have anything to do with (gravity)? If the Earth is trying to throw you off at the equator, then it will counteract the pull of the Earth on you. And at the North Pole, it wasn't trying to throw you off and the Earth has more pull on you.

Table 3c. Gravity

	Analogy		Analogy	Analogy	Analogy
					Gedanken
TE	TE	TE	TE		TE
Gestures	Gestures	Gestures	Gestures	Gestures	Gestures
[Line 189]	[Lines 201-205]	[Lines 207-211]	[Lines 213-215]	[Line 217]	[Lines 229-231]
S9: [G-M] [G-F] I'm basically taking (S4)'s position . . . when the Earth spins, it seems logical to me . . . say you're on the equator and you're going around, there's this greater force pushing you off the Earth than if you were on the pole and you're doing this little circle.	S6: [G-F] (If the Earth spun faster on its axis) I believe (your weight) would decrease. . . . At the fair they often have this . . . ride so you get thrown in a circle . . . I think that at the equator you would feel the same sort of (xxx).	S10: [G-F] [G-S] [G-M] (If the (Earth spun faster on its axis) I would say there is no change at all (in your weight). . . . The fair thing exists within the gravitational field of the Earth. . .the air goes around with the Earth, and there is no reason why the person should not be attracted only to the center of the Earth and nowhere else.	S13: [G-M] The only effect on gravity would be if all of a sudden the Earth moved around the Sun twenty times faster. . . Centrifugal force is entirely separate from gravity.	S13: [G-F] If somebody puts me in a catapult and I go hurling two nundred feet into the air, gravity is the same, there's just another force acting.	S4: [G-F] I have no idea if there's no centrifugal force in space. . . . If you have a space ship that's spinning around, someone could be standing on the outside and they're not going to get thrown off? and I find that hard to believe that the only reason we have centrifugal force is because of air or because of gravity.

It can be seen from the coding that in a number of places where students were reasoning about forces, they used hand motions that appeared to depict the forces. Fig. 4 above is an example of one such hand motion.

We are also interested to see that some students appeared to ***refine their own TEs***, as with S4 in Lines 182-186, and to ***refine the TEs of others***, as with S9 in Line 189.

Table 4a. Breakdown of incidents: Gravity transcript (42 minutes)

Combinations of processes	# of cases identified in	# of these cases associated with depictive hand motions	# of these cases associated with TEs
Analogy alone	4	4	2
Extreme case alone	2	1	0
Gedanken alone	2	2	2
Analogy & Gedanken	1	1	1
Extreme case & Analogy & Gedanken	2	2	2
TE alone	6	4	6
Total cases	17	14	13

Of the 14 cases associated with depictive hand motions, 10 were associated with force-indicating motions, an indication that they were paired with kinesthetic imagery.

Table 4b. Total incidents: Gravity transcript*

	# of cases or incidents	# of these cases associated with depictive hand motions	# of these cases associated with TEs
Analogies	**7**	**7**	**5**
Extreme cases	**4**	**2**	**2**
Gedankens	**5**	**5**	**5****
TEs	**13**	**11**	**---**

** Many of the processes listed here were run in combination, so some cases are represented more than once in these numbers. The columns do not add. See Table 4a.*

*** Although all the Gedankens in this paper are TEs, our definitions do not require this.*

There were:

- 17 cases where the expert reasoning processes were paired with depictive gestures—including 10 cases paired with force-indicating gestures;
- 10 other cases in the transcript where depictive gestures occurred; leading to
- 27 instances total where depictive gestures occurred, many with multiple gestures, comprising
- ~105 individual depictive gestures by students in 42 minutes of transcript.

All of the analogies and Gedankens and most of the TEs (broad sense) were accompanied by depictive hand motions, providing evidence that they were operating via imagery.

Findings

Developing coding criteria for so many interconnected reasoning processes is a tremendous challenge. What we have tried to do in this paper is to make an initial proposal that can be evaluated and improved and to show how such reasoning can be connected to imagery indicators. Detailed definitions were used to clarify the meaning of the term thought experiment and to identify evidence for thought experiments within the classroom episodes. We identified four different forms of expert reasoning in student episodes and have presented evidence that TEs in the broad sense were used within each of these forms of reasoning. There was also evidence for the use of TEs by themselves. Each of these forms of reasoning was frequently associated with depictive hand motions, which supports the hypothesis that mental imagery was involved, much of it apparently kinesthetic. Important similarities between expert and student uses of TEs were identified. (For instance, both experts and students have been observed modifying their TEs to enhance the use of imagery and to clarify the implications. See also Stephens & Clement, 2006a, for more on this point.)

TEs in the broad sense were more widespread in these classrooms than we expected. Our evidence supports Reiner's (1998) speculation that TEs are natural processes in science learning and her results indicating that student investigations can include TEs. In fact, far from being exotic curiosities, we conclude that TEs can play an important role in the act of sense making as students work to develop models consistent with the findings of science. Transcript excerpts quoted here provide evidence that TEs can allow students to evaluate directly the consistency of newly constructed models with what they already hold to be true about the world—by comparing their new conceptions with their own prior knowledge, though this knowledge may have been held only implicitly. The evidence for dynamic and kinesthetic imagery in the vast majority of the gestures suggests that the prior knowledge being used was perceptual-motor in character and was expressed mentally in the form of imagistic simulations. Thus we can point to evidence that imagistic simulation within Tes, on previously untested cases, was involved in students' running of explanatory models, generation of analogies, generation of extreme cases, and running of Gedanken experiments. This suggests that imagistic simulation is centrally important in the learning that took place in these classrooms.

Import

The identified similarities between student and expert uses of TEs lend support to an acceptance of their importance as a scientific reasoning process for students. TEs in science have considerable power to convince, either by exposing inconsistencies in one's conceptions or by strengthening one's conviction in an outcome, arguing for their value as a sense-making strategy (Clement, in press; Nersessian, 1993). Although we would not suggest that TEs replace real experiments in the classroom, this suggests that they can be an important complement. The widespread use and apparent effectiveness of TEs within a number of classrooms for which we have tapes suggest that it would be of value to educators to understand what goes into the running of an effective classroom TE and so we feel further research on this topic is very much needed.

References

Brown, J. R. (1986) The structure of thought experiments. *International Studies in The Philosophy of Science: The Dubrovnik Papers*, 1-15.

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. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching, 30* (10). 1241-1257.

Clement, J. (1994). Use of physical intuition and imagistic simulation in expert problem solving. In D. Tirosh (Ed.), *Implicit and explicit knowledge: An educational approach*. Ablex Publishing Corp.

Clement, J. (2002). Protocol evidence on thought experiments used by experts. In Wayne Gray and Christian Schunn, Eds., *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society, 22,* 32. Mahwah, NJ: Erlbaum.

Clement J. (2003). Imagistic simulation in scientific model construction. In R. Alterman and D. Kirsh, Editors, *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society*, 25, 258-263. Mahwah, NJ: Erlbaum.

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

Clement, J. (in press). Thought experiments and imagery in expert protocols. In L. Magnani (ed.) *Model based reasoning*. New York: Springer-Verlag.

Clement, J. (in preparation). *Creative Model construction in scientists and students: The role of analogy, imagery, and mental simulation*.

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.

Darden, Lindley (1991). *Theory change in science: Strategies from Mendelian genetics*. New York: Oxford.

Driver, R. (1983). *The pupil as a scientist?* Milton Keynes: Open University Press.

Gentner, D. (2002). Psychology of mental models. In N. J. Smelser & P. B. Bates (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 9683-9687). Amsterdam: Elsevier Science.

Giere, R. N. (1988). *Explaining science: A cognitive approach*. Chicago: Chicago University Press.

Gilbert, J. K. & Reiner, M. (2000). Thought experiments in science education: Potential and current realization. *International Journal of Science Education, 22*(3), 265-283.

Gooding, D. (1992a). The procedural turn: Or, why do thought experiments work? In Giere, R. (Ed.) *Cognitive models of science*. Minneapolis: U. of Minnesota Press.

Gooding, D. (1992b). What is *experimental* about thought experiments? In Hull, Forbes, & Okruhlick (eds.), pp. 280-290.

Gooding, D. (1996). Creative rationality: Towards an abductive model of scientific change. *Philosophica, 58* (2), 73-102.

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

Hull, D., Forbes, M., & Okruhlick, K. (eds.) *PSA 1992, 2*. East Lansing, MI: Philosophy of Science Association.

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.

Kuhn, T. S. (1977). The function of thought experiments. In T. Kuhn (ed.) *The essential tension*. Chicago: University of Chicago Press.

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. (1993). In the theoretician's laboratory: Thought experimenting as mental modeling. In Hull, Forbes, & Okruhlick (eds.), pp. 291-301.

Nersessian, N. J. (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.

Nunez-Oviedo, M. C., Rea-Ramirez, M. A., & Clement, J. (in press).

Reiner, M. (1998). Thought experiments and collaborative learning in physics. *International Journal of Science Education, 20*, 1043-1058.

Reiner, M. & Burko, L. M. (2003). On the limitations of thought experiments in physics and the consequences for physics education. *Science & Education, 12*, 365-385.

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

Sorensen, R. A. (1992). *Thought experiments*. Oxford: Oxford University Press.

Stephens, L. & Clement, J. (2006a). Using expert heuristics for the design of imagery-rich mental simulations for the science class. Proceedings of the NARST 2006 Annual Meeting, San Francisco, CA.

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