Bringing atoms into first-year physics

Ruth W. Chabay and Bruce A. Sherwood
Center for Innovation in Learning and Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

(Received 17 May 1999; accepted 15 June 1999)

We argue that thermal physics should not be treated as a separate topic in introductory physics. The first-year calculus-based college physics should offer a modern, unified view of physics representative of the contemporary scientific enterprise. It should focus on the consequences of the central fact that matter is composed of atoms, and on the process of modeling physical systems. Such a focus is more interesting and relevant to students than a repetition of a purely classical treatment. We give an example of a course that emphasizes physical modeling of phenomena in terms of the atomic nature of matter. Thermal physics is woven into the entire course and is fully integrated with classical and semiclassical mechanics.

I. INTRODUCTION

We argue that at the introductory level, an important opportunity is lost if thermal physics is taught as a separate, self-contained topic. Our thesis is that it is important to introduce students to a modern, unified view of physics in the introductory college-level course, and to avoid the compartmentalization of topics and techniques that usually occurs. This course is the only opportunity to convey this unified picture to the many students who will take no further physics courses, and this course is an optimum time to convey a unified view to the small number of potential physics majors who will spend the rest of their college careers studying specialized topics in significant depth.

An appropriate goal for the introductory college sequence is that it reflect the contemporary physics enterprise authentically, while still extending students' knowledge of classical principles. The pursuit of this goal requires the design of a course quite different from what has come to be the standard introductory course. In particular, thermal physics should not be treated as a separate topic, segregated from other portions of the course by special vocabulary, different symbol conventions, and a focus on processes such as quasistatic expansions which are unlike anything students encounter in the rest of the course. We need to move beyond the 19th century and allow thermal physics to make contact with contemporary physics, especially focusing on the central fact that matter is composed of atoms.

A focus on the atomic nature of matter offers the opportunity to integrate thermal physics with mechanics, and to stress the connection between microscopic and macroscopic descriptions and analyses. Discussions of physical principles can involve the properties of real matter, instead of focusing solely on idealized macroscopic, material-independent situations. Rather than emphasizing algebraic manipulation in sanitized exercises, the course can offer opportunities for students to explain and predict complex phenomena, reasoning from an atomic model of matter and a small number of fundamental principles. In this process students need to make approximations, to idealize complex systems, and to think through the consequences of a particular model—activities that are central to physics but are absent from the traditional curriculum.

By addressing these issues we may also address other serious problems with the introductory university physics sequence. First, many college freshmen already have a reasonable foundation in classical mechanics from a high school physics course, so their first college physics course is essentially a repeat of their high school course. This repetition engenders boredom and often fails to improve students' understanding of physics. Second, many students (and many instructors) have little interest in 17th-century mechanics and are disappointed that the typical college introductory physics course makes little contact with contemporary scientific issues and models. Third, although physicists prize the elegant unity of physics, this unity is obscured for students in the introductory sequence by the presentation of mechanics and thermodynamics as two separate and largely unconnected sciences.

II. PHYSICAL MODELING BASED ON THE ATOMIC NATURE OF MATTER

Many possible introductory curricula could achieve the goals outlined above, and we urge colleagues to consider creating new courses along these lines. As a concrete example, we offer a brief description of an alternative introductory calculus-based physics sequence which we have developed at Carnegie Mellon University. The course, entitled ''Matter & Interactions,'' emphasizes atomic-level models of real matter and its interactions. We refer to the first semester of the sequence as ''modern mechanics:'' classical and semiclassical mechanics applied to atomic-level interactions of real matter. Because this approach is unfamiliar, students are less inclined to dismiss it as a repetition of familiar material. The topics and methods of analysis are significantly more authentic in representing the nature of contemporary science, and permit students to explore some aspects of intriguing modern concepts. Macro-micro connections are continually made throughout the curriculum, to stress that mechanics and thermal physics are one science rather than two. The second semester course, covering electricity and magnetism, and waves and physical optics (based on the classical analysis of the interaction of light and matter), introduces the more abstract concept of fields while retaining a strong emphasis on the atomic nature of matter and the interaction of fields with matter. The entire sequence emphasizes physical modeling and the explanation and prediction of complex phenomena. By modeling, we mean the process of idealizing a complex system, making explicit approximations and simplifying assumptions, and applying fundamental principles to predict or explain the behavior of the system.

A course emphasizing a modern view of matter addresses the needs of the many students for whom the first-year physics course will be the last physics course that they take. Atoms, vibrational and rotational energy, and discrete energy levels are the fundamental concepts that students of chemistry, biology, materials science, and even electrical engineering will need in their further studies. Real materials, including their deformation and their electromagnetic interactions, are important to engineering students. Physics students are particularly intrigued by modern concepts. Potential physics majors frequently report that their interest in physics was stimulated by reading about quantum mechanics in high school. We want to encourage this interest, instead of requiring students to defer it until they have run the gauntlet of 17th, 18th, and 19th century science.

In this paper we report on our experiences in teaching this course, emphasizing in particular the role that ideas from thermal physics play in the course. At this writing we have taught the first-semester modern mechanics course twice. The second semester E&M course is significantly more mature, though we are in the process of revising it to exploit students' deeper experience with concepts introduced in the first semester. Prepublication drafts of textbooks for both semesters exist (the second is a significantly revised version of the published text *Electric & Magnetic Interactions*[1]).

III. MODERN MECHANICS

The first semester course emphasizes physical modeling, atomic-level description and analysis of the interactions of real matter, and qualitative reasoning to support quantitative calculations. This modern mechanics course also engages students in significant computer modeling of physical systems. The one-semester course and the accompanying textbook, *Matter & Interactions*,[2] contain four major sections:

(1) Newton's laws
(2) Energy
(3) Multiparticle systems
(4) Kinetic theory and statistical mechanics.

Although this sequence of topics sounds traditional, the emphasis and details of implementation differ significantly from those of traditional courses. We will describe those aspects of the curriculum that are traditionally considered thermal physics, though in our curriculum the boundary between thermal physics and mechanics is deliberately blurred as much as possible. Other aspects of the curriculum are mentioned where necessary to provide adequate context.

A. Newton's laws

Issues of thermal physics are not addressed directly in this 3–4 week introductory segment of the course, but the study of model systems lays the foundation for subsequent discussions. The most important of these is the harmonic oscillator, which figures prominently in models of solids and gases. Students explore the behavior of masses on springs experimentally, numerically (by numerical integration of Newton's second law), and analytically. They then make the key transition to a semiclassical model of a solid as a lattice of balls and springs, first by measuring Young's modulus for a real material (aluminum), then translating this result into a value for the effective strength of the interatomic ''spring'' representing the force between aluminum atoms. Further exploration of the simple ball-and-spring model leads to a numerical prediction for the speed of sound in aluminum, which agrees well with measurements.

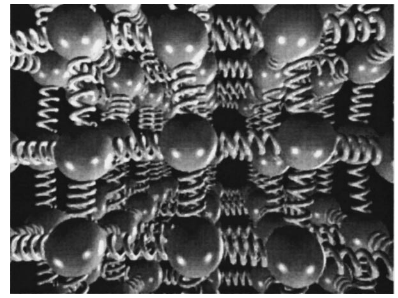

Fig. 1. A frame from a computer-generated movie of a ball-and-spring model of a solid.

Students were asked to reflect on the important aspects of this work. The following reflections are representative of what students wrote:

Student 1: ''In my opinion, the central idea in this chapter was to learn that atoms bonded to each other can be thought of as two balls connected to one another with a spring. Once we understood this concept, we could apply the models of springs from the macroscopic world to the atomic level, which gave us a general idea of how things work at the atomic level. Understanding that gave us the ability to predict vibrational frequencies of diatomic molecules and sound propagation in a solid. It is absolutely amazing how we can use very simple concepts and ideas such as momentum and spring motion to derive all kinds of stuff from it. I truly like that about this course.''

Student 2: ''I believe that the most important concept in this chapter was how the micro world can be used to analyze many aspects of the macro world. For example, the method of sound traveling through solids being comparable to vibrations traveling through a spring. I never actually considered it in such a fashion before. However, considering it this way, we were able to predict a spring constant on the molecular level as well as predict the speed of sound in that material.''

B. Energy

The second section of the course deals with energy. Energy is a central concept in the course, especially because it is the bridge connecting mechanics and thermal physics. Often in traditional courses the concept of energy introduced to help solve mechanics problems appears to be a different concept from the energy concepts discussed (using different symbol conventions) in thermodynamics or chemistry. If we consider energy at the atomic level, these dichotomies can be resolved.

The ball-and-spring model of a solid continues to figure prominently, providing a framework for reasoning about the thermal energy of a solid, thermal energy transfer, and friction. A high point of the course has been a class mostly devoted to talking about the thermal energy of a solid. A computer-generated QuickTime movie of a $5\times5\times5$ array of balls connected by springs is shown to the class and left looping on the computer projection screen (Fig. 1).

Even relatively sophisticated students are intrigued by the movie, and surprised by the realization that atoms in a solid

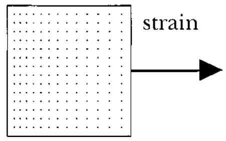

Fig. 2. A block pulled by a wire has a gradient of strain.

are in constant motion. They begin to ask questions about how the movie was created, and quickly recognize that they themselves would be capable of writing a program to model this system–both the physics and the implementation of the model are within their reach. Stimulated by the movie, students become deeply engaged in discussing various aspects of energy and of energy transfer. They already have had significant experience with the harmonic oscillator, and they can identify both spring energy and kinetic energy in the model of a solid, and see vividly that in our model, thermal energy is just ordinary energy but at the atomic level. One thread of the discussion is to think through what would happen if a solid with higher average energy were placed in contact with the solid portrayed in the movie. The discussion touches on the probabilistic aspects of energy transfer, because it is evident from the complex motions in the movie that an atom in the adjoining solid might sometimes gain energy rather than lose it. Students identify the average energy as likely to be proportional to temperature. We tell them that it is the average kinetic energy that turns out to be proportional to the temperature in many circumstances, but we do not attempt to be too formal at this point in the course.

We draw the students' attention to the fact that when the movie loops back to the first frame, there is a noticeable glitch. We ask them what would have to be done to make the movie loop smoothly. They recognize that all the balls in the model solid would have to return to their original positions and momenta, and that even in a small system of 125 balls this return has an inconceivably small probability. Looking ahead, this discussion sets a context in which questions about randomness and reversibility occur naturally.

The discrete nature of energy in atomic systems is the most accessible of the quantum aspects of matter, and discrete electronic energy levels are familiar to students from their previous chemistry courses. We build on this background and connect it to physics by discussing the quantization of vibrational energies in the harmonic oscillator, which is a new idea to the students. A careful treatment of the quantized harmonic oscillator, including absorption and emission of photons, lays a foundation for later study of the specific heat of solids and diatomic gases as a function of temperature.

C. Multiparticle systems

The section of the course devoted to multiparticle systems opens with a discussion of the complex sequence of events at the atomic level when a block is pulled by means of an attached wire. The initial tug sends a disturbance at the speed of sound down the wire and into the block. Waves propagate throughout the system. Even the final steady state is complicated, involving a gradient of strain throughout the wire and block, which corresponds in the microscopic ball-and-spring model to a gradient of the stretch of the interatomic ''springs'' (Fig. 2).

Given this complexity, why is it that we can nevertheless say something rather simple about the motion of the block? This question motivates a discussion of the role of the multiparticle form of Newton's second law involving the total momentum and the role of the center of mass. Significant emphasis is given to the separation of kinetic energy into translational kinetic energy plus kinetic energy relative to the center of mass (vibrational and rotational). Among other advantages, this treatment facilitates later discussions of diatomic gases.

A significant pedagogical innovation is the way in which students learn how to deal with ''pseudowork.''[3,4] Previous attempts by the authors to teach this topic were not very successful, apparently because students had difficulty making the crucial but subtle distinction between two different equations that look seductively similar (the pseudowork-energy equation, which is really a momentum equation, and the true energy equation). In the new scheme, we introduce the notion of the ''point-particle system,'' a system with the total mass of the real system, with all of this mass concentrated at the location of the center of mass of the real system, and subjected at that point to all of the forces that are exerted at various locations on the real system. There are no longer two energy-like equations. Rather there are two different systems, the real system and the point-particle system, both subject to the usual true energy equation. For the point-particle system the true energy equation is very simple: the net force does work along the displacement of the point particle, resulting in a change of the translational kinetic energy; there are no other energy forms in the point-particle system. In contrast, the true energy equation for the real system involves translational, vibrational, and rotational kinetic energy, other forms of energy including potential energy, and additional forms of energy transfer including heat and radiation. The work done on the real system is calculated in terms of the actual work done by each individual force, instead of the pseudowork done by the net force.

The former effortful distinction between two similar-looking mathematical equations has been replaced by an easy visual distinction between two different and different-looking physical systems. On homework and exams we saw significantly improved student performance in applying these concepts to multiparticle systems. Homework problems include humans jumping and skating. A nagging paradox related to sliding friction is resolved by multiparticle analysis and a mesoscopic model of the surfaces in contact.[5]

D. Kinetic theory and statistical mechanics

The climax to the course is the section on statistical physics, which brings together atomic models of matter with energy concepts to derive, in a very concrete and understandable fashion, the concept of entropy.

First there is a fairly traditional treatment of kinetic theory of gases. Because students are quite familiar with momentum, multiparticle systems, and atomic-level description and analysis, this topic is easily accessible to them. Kinetic theory is used to introduce notions of statistical averages and distributions. In keeping with the microscopic orientation of the course, we view density as a primary quantity, and pressure as a macroscopic consequence.

The second portion of this section is on statistical mechanics. This treatment is based on the recommendations of Moore and Schroeder.[6] Students use a computer to calculate explicitly the number of ways to arrange energy in two Ein-

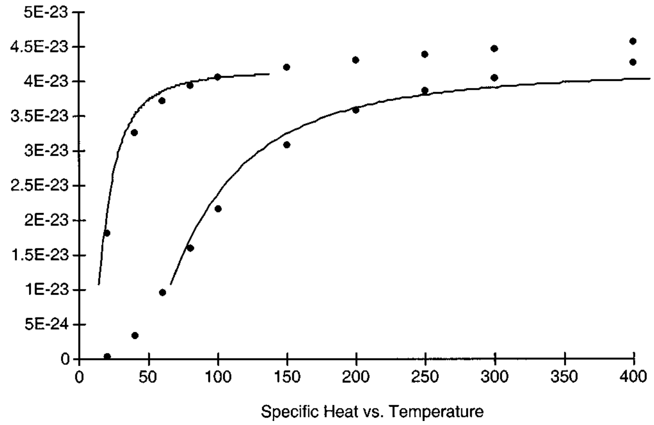

Fig. 3. Display from a student's program showing a fit to experimental specific heat data (per atom) for aluminum and lead.

stein solids in thermal contact. An Einstein solid is a ball-and-spring model of a solid (and therefore very familiar to our students) with the approximation that each atom is an independent three-dimensional oscillator, with discrete harmonic oscillator energy levels (familiar from earlier work on the quantized harmonic oscillator). Students quickly see in their own computed graphs that Ω (the number of ways of arranging the energy between the two solids) is strongly peaked. We then define entropy S as $k \ln \Omega$, and the second law of thermodynamics emerges from these considerations.

It is intuitively plausible to students that when two blocks come to equilibrium, the quantity that is the same about them is their temperature. Students can observe from their own computations that the slopes of the entropy versus energy plots for two objects are the same when the most probable situation is reached, so this consideration motivates a definition of temperature in terms of entropy. Using the Einstein solid model, students compute the temperature for one object as a function of energy, and then the specific heat as a function of temperature (Fig. 3).

Simple measurements of the specific heat of water in a microwave oven[7] provide concrete experience with the concept of specific heat, and add meaning to the final computation of the specific heat of a metal as a function of temperature. Students fit their theoretical curves to literature values of the specific heat of aluminum and lead as a function of temperature by adjusting a single parameter—the effective interatomic spring constant. Excellent results are obtained using values based on those obtained earlier in the course from a microscopic interpretation of Young's modulus. Mechanics and thermal physics merge here into a seamless whole: a student measurement of Young's modulus leads to a prediction of the speed of sound in a solid and a prediction of the specific heat of a metal as a function of temperature. The Boltzmann factor is obtained by considering a single oscillator in thermal contact with a large Einstein solid. The Boltzmann factor is used to analyze various aspects of diatomic gases.

As observed by Moore and Schroeder with their own students, this way of introducing entropy and the second law of thermodynamics has enormous pedagogical advantages. Like them, we observe that our students find these usually difficult concepts surprisingly accessible because the whole treatment is very concrete and explicit. On quizzes, exams, and final exams, we have been pleased to see students use these ideas successfully. A student who encountered the concept of entropy in a subsequent course in another discipline commented that he was very glad he had studied entropy first in our physics course, because he was able to make sense of the terse, macroscopic treatment in a detailed way that would not otherwise have been accessible to him.

A final topic treats heat engines as an application of the entropy principle rather than as an introduction to it. It is unnecessary to wander in the thicket of ideal gas calculations, because the second law of thermodynamics applies to any system, not just ideal gases (a fact that may be obscured in the usual treatment of heat engines). This chapter closes with a treatment of the efficiency of finite-time heat engines.[8]

The following four problems from a final exam given in this first-year course reflect the integration of thermal physics with other concepts:

Problem 1. A disk with a string wrapped around it is pulled with a constant force. At a particular instant the center of the disk has moved a distance d, and an amount of string L has unwound off the disk.

(a) At this instant, what is the speed v of the center of the disk? Explain carefully what principles you use and why.

(b) At this instant, what is the disk's angular speed ω? Explain carefully what principles you use and why.

Problem 2. A hot bar of iron glows a dull red. Using our simple model of a solid, answer the following questions, explaining in detail the processes involved. (Note that our simple model of a solid does predict a line spectrum for this system.)

(a) What is the energy of the lowest-energy spectral emission line?
(b) What is the approximate energy of the highest-energy spectral emission line?
(c) Predict the energies of two other lines in the emission spectrum of the glowing iron bar.

Problem 3. A system consisting of four iron atoms (object 1) initially has 1 quantum of energy. It is brought into contact with a system consisting of two iron atoms (object 2) which initially has 2 quanta of energy.

(a) Using the Einstein model of a solid, calculate and plot $\ln \Omega_1$ versus q_1 (number of quanta in object 1), $\ln \Omega_2$ versus q_1, and $\ln \Omega_{\text{total}}$ versus q_1 (put all three plots on the same graph). Show your work and explain briefly.

(b) Calculate the (approximate) temperature of the objects at equilibrium. State what assumptions or approximations you made.

Problem 4. You are on a spacecraft measuring $8\,\text{m} \times 3\,\text{m} \times 3\,\text{m}$ when it is struck by a piece of space junk, leaving a circular hole of radius 4 mm, unfortunately in a place that can be reached only by making a time-consuming spacewalk. About how long do you have to patch the leak? Explain what approximations you make in assessing the seriousness of the situation.

IV. RUNNABLE MENTAL MODELS

In addition to its contemporary flavor, an advantage of an atomic model of matter is that it provides students with more options in reasoning about physical phenomena; they are not restricted to macroscopic or continuum reasoning in all cases. Because an atomic level model is often mentally ''runnable,'' mechanistic causal reasoning is possible by mental simulation. For example, a constraint-based continuum argument says that because $PV = nRT$, if the temperature of a gas is kept constant but the volume of its container increases (isothermal expansion), the pressure must decrease. A mechanistic model of the same process considers collisions of gas molecules with the receding piston. Because the average speed of the molecules decreases in such a collision, the average kinetic energy decreases, and the temperature starts to fall. This decrease in turn leads to an inflow of energy from the surrounding bath to raise the temperature back to the bath temperature. In the continuum argument, the result is constrained by the relationship of two or more global variables, and involves reasoning only about initial and final states. In contrast, reasoning from a mechanistic mental model involves ''running'' the model in one's head, and observing the consequences. An auxiliary benefit is a concrete sense of the process by which the system moves from one state to another.

Psychologists have found[9] that reasoning from runnable mental models is often more natural than is constraint-based reasoning. Eylon and Ganiel have shown that the lack of microscopic mental models makes it very difficult for students to reason about certain kinds of problems in electricity and magnetism.[10] Thacker, Ganiel, and Boys[11] have found that students who have a sense of mechanism at the microscopic level perform better in reasoning about RC circuits than do students who have studied them only on a macroscopic level. It is reasonable to expect that working with a model of matter at the atomic level would allow students to gain a sense of mechanism which can help make sense of thermal phenomena. Further, discussions uniting macro and micro viewpoints lead naturally to a consideration of some of the approximations and idealizations that must be made in order to reason about real systems.

V. COMPUTER MODELING

Computer modeling plays an important role throughout the course, and students write programs of various kinds, especially numerical integrations of Newton's laws. No prior programming experience is required. In the first week we spend two 50-minute periods teaching the students a subset of the cT programming language[12] adequate for modeling physical systems. By the end of the second period students begin to animate a spacecraft coasting to the Moon, using a simple numerical integration method.

The impact of an atomic model of matter on the understanding of thermal properties of matter is significantly enhanced by computer modeling, which offers the possibility of concrete visualization of abstract quantities. For example, students add energy calculations and graphs to their previous numerical integration of the spring-mass system and explore visually the relationships among kinetic, potential, and total energy. The computational treatment of energy, entropy, and heat capacity of small Einstein solids gives students a very concrete sense of the microscopic meaning of entropy.

VI. REFLECTIONS ON THE MODERN MECHANICS COURSE

We hope that the details of our modern mechanics curriculum have shown the extent to which this course engages students in physical modeling of phenomena, and the extent to which mechanics and thermal physics are thoroughly intertwined throughout the course. We want to engage students in an activity that is central to contemporary science: the attempt to explain a broad range of phenomena using a small number of powerful fundamental principles. In the modern mechanics course these fundamental principles are the momentum principle, the energy principle, and the angular momentum principle. To these principles are added some basic ideas about the atomic nature of matter, including discrete energy levels. This course is to a significant degree about models of real, atomic matter such as aluminum or lead solids, and helium or nitrogen gas. This treatment is in contrast to traditional mechanics courses where the only property of the object that matters is typically its mass, shape, or temperature.

We have given arguments for the advantages of combining the study of mechanics with the study of thermal physics. To these arguments we would add another: The study of mechanics as a separate subject is incomplete and inconsistent. To take a very simple example, if you push a block across a table at constant speed, the traditional mechanics analysis would imply that no work is done on the block, and that there is no change in the energy of the block. But the block gets hot! Within the framework of traditional mechanics there is no resolution of the seeming paradox. Only if one takes seriously the actual nature of real matter can one understand what is going on.

A freshman math major who had just completed the first semester of the traditional introductory physics course made a very perceptive comment about that course. He remarked that Newtonian mechanics was presented essentially as an axiomatic system, closed within itself. In our E&M course, with its emphasis on the atomic nature of matter and on the boundaries of classical theory with relativity and with quantum mechanics, it became clear to him that Newtonian mechanics is in no way a closed system relative to the real world. He commented that he valued axiomatic treatments in math courses, but that such a treatment was inappropriate for a physics course. He's right.

ACKNOWLEDGMENTS

We thank Tom Foster for useful discussions. This work was supported in part by the National Science Foundation, Grant No. DUE 9554843.

[1] R. Chabay and B. Sherwood, *Electric & Magnetic Interactions* (Wiley, New York, 1995), preliminary second ed. (Wiley, New York, 1999). For additional information see http:// cil.andrew.cmu.edu/emi.html

[2] R. Chabay and B. Sherwood, *Matter & Interactions*, preliminary ed. (Wiley, New York, 1999). For additional information see http:// cil.andrew.cmu.edu/mi.html

[3] B. A. Sherwood, ''Pseudowork and real work,'' Am. J. Phys. **51**, 597–602 (1983).

[4] A. Arons in this issue.

[5] B. A. Sherwood and W. H. Bernard, ''Work and heat transfer in the presence of sliding friction,'' Am. J. Phys. **52**, 1001–1007 (1984).

[6] T. Moore and D. Schroeder, ''A different approach to introducing statistical mechanics,'' Am. J. Phys. **65**, 26–36 (1997). A specific realization of these ideas is found in the statistical mechanics section of the textbook by T. Moore, *Six Ideas that Shaped Physics/Unit T: Some Processes are Irreversible* (WCB/McGraw-Hill, Boston, 1998).

[7] A take-home experiment developed by Vidhya Ramachandran at Carnegie Mellon University.

[8] F. Curzon and B. Ahlborn, ''Efficiency of a Carnot engine at maximum power output,'' Am. J. Phys. **43**, 22–24 (1975).

[9] P. N. Johnson-Laird, *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness* (Harvard U. P., Massachusetts, 1986).

[10]B. Eylon and U. Ganiel, ''Macro-micro relationships: The missing link between electrostatics and electrodynamics in students' reasoning,'' Int. J. Sci. Ed. **12**, 79–94 (1990).
[11]B. A. Thacker, U. Ganiel, and D. Boys, ''Macroscopic phenomena and microscopic processes: Student understanding of transients in direct current electric circuits,'' Physics Education Research—A supplement to the Am. J. Phys. **67**, 525–531 (1999).
[12]See http://cil.andrew.cmu.edu/ct.html

SCIENCE VS. MATH

Bennett wouldn't have been happy with no mention of Mary and Henry. That kind of problem was for the pure mathematician: Consider object A, which is six units greater than object B, whereas object B has a measure two thirds that of A. No, Bennett was a scientist, and he wanted to begin in the dirt and debris of the real physical world. But he took pleasure in sifting through that world, distilling it, cleansing and purifying it until he was left with a single mathematical equation of inescapable solution.

Alan Lightman, *Good Benito* (Pantheon Books, New York, 1994), p. 63.