SBIR Phase I Final Report: Connecting Mathematics and Science through Data

Timothy E. Erickson, Principal Investigator
Senior Scientist, Epistemological Engineering

proposal #0060304

Technical Note. If you are reading this document on the screen as a pdf file, internal page references should be "hot"; just click on a page number (or maybe just to its left) and Acrobat will take you there.

Phase I Objectives

Here are the objectives from the Phase I proposal:

1. Identify topics in various disciplines suitable for lessons [that integrate mathematics into science classrooms through data]. This we will do in collaboration with teachers and scientists.
2. Create a series of classroom scenarios that show uses of Fathom in science education connecting to mathematics in different modes and using different sources of data.
3. Develop a list of desired enhancements to Fathom based on identified needs. Implement a selected subset of this list.
4. Develop and field test three prototype lessons for high school science using Fathom as a data analysis vehicle.
5. Learn more about the barriers that stand in the way of teachers' using technology and curriculum materials like those we will produce, and some ways of overcoming those barriers.

Summary of the Research

Let's look briefly at what we did according to the objectives in the previous section.

1. In the advisory board meeting, through extensive readings of policy and research documents as well as existing curriculum materials, and through meetings with other scientists and visits to teachers' classrooms, we identified specific topics suitable for lessons. Some of these results were as you might expect: for example, it makes sense to integrate functions into physics. But there were some surprises:
   - There are areas of learning that reside "in the cracks" between science and mathematics. We believe they are essential to basic understanding of science, but they do not appear explicitly in standards documents. We describe one in "Measurement" on page 10.
   - With our innovations in software and curriculum design, we can introduce topics that used to be too difficult or cumbersome. Some of these topics may be intrinsically interesting to students; others may help students understand scientific principles more deeply than their traditional counterparts do. See, for example, "Numerical Modeling" on page 11.
2. We have created a number of scenarios for integrating mathematics into high-school and early college science classrooms. We envision "replacement labs" that extend a typical laboratory activity to include more mathematics in the form of data analysis. We also imagine new types of data-rich problem sets and longer, multi-session units. More important than these scenarios, though, are these principles for lesson design:
   - Technology must relieve tedium for the student but must not relieve the student of thinking. In general, this means that we let the computer do the calculation, but insist that the student specify it symbolically.
   - It is data *analysis*, not data, that injects mathematics into science. Mathematics works on data: to make it understandable, and to help us interpret it. This is in contrast to the com-

mon practice of using mathematics in lectures to describe theory, and then simplifying the mathematics when you go to the lab.

- Data come from a variety of sources; some depend heavily on technology. Students need to use a range of data sources. We describe these in "Sources of Data" on page 9.
- To the extent they can, students should use mathematics as scientists do in real scientific investigations. This mathematics takes many forms, and occurs at different stages in the investigation—in more places than we had expected. We have identified many of these; see "How Scientists Use Mathematics" on page 12.
- In general, we look for opportunities for students to use *functions* where traditional curricula look at *constants*. This elevates the mathematics, and the science comes with it: students look at broader principles rather than particular cases.
- That said, many students have trouble with comparatively simple mathematics in unusual contexts; we can do a great service by highlighting these simple topics where they are essential for data analysis.

3. Our subcontractor, KCP Technologies, made four prototype enhancements to Fathom at our request: we can now get data directly from probes; we can plot more than one series at a time; we have a built-in timer; and we can write new functions as plug-ins. Those were the most urgently-needed features, and they make science development a lot easier. Working with them has clarified what else we need to do in future work in this area. See "Software Enhancements" on page 16.
4. We field-tested one large lesson ("Example Lesson: Dropping Cotton Balls" on page 4) in three high-school physics classes of varying ability. In two of the classes, we also did a short math-and-mechanics problem ("The Elevator" on page 25) that also served as proof-of-concept that with our technology, high-school physics students could solve complex mechanics problems numerically. A third field-test candidate ("Copernicus" on page 23) proved so surprisingly difficult for experienced adults (despite what seems, on the surface, to be much simpler mathematics) that it needs additional redesign before we inflict it on students.
5. We face three fundamental barriers: inadequate technology, lack of class time, and institutional inertia. So we must ensure that our materials make the technology an advantage: that students actually do and learn more science and more math, even though data analysis takes more time. You can read more about this in upcoming sections and in "Potential Commercial Applications" on page 26. And we found one unexpected "hook": you can use mathematics—good data analysis—to get good results with cruder, that is, less expensive, equipment.

Finally, when we had prototype activities and could see our emerging philosophy of what makes a good math/science/technology/data integrated lesson, we looked back at the first objective—to identify topics—and noticed a curious thing: Seen in another light, all our activities are not just about their particular topics, but really about the *nature of science itself*. This is because, in integrating data and mathematics, we generate *and test* mathematical *models*. This aligns well with some current thinkers who feel that the nature of science is what's most important for every citizen to learn. We will discuss all of these issues in greater detail.

Research Findings and Project Activities

The Phase I award is for a feasibility study. In this case, the overarching question is whether we can develop materials for high-school and college that use technology—Fathom in particular—to integrate mathematics into the science classroom more effectively through data.

The short answer is "yes." But before we continue with more detailed answers, let's summarize one of the lessons we tested with students.

Example Lesson: Dropping Cotton Balls[1]

How long does it take a cotton ball to fall? It depends on many things—most importantly, how far you drop it. And cotton balls do not drop like stones, as air resistance plays a rôle here.

At the beginning of the first session, we give the students a fake scientific paper describing a theory of cotton-ball falling. The theory (which is wrong) states that, due to air resistance, cotton balls fall a distance *s* in time *t* according to the formula $s = (1/2)kt^2$ where *k* is a constant less than *g*, the acceleration of gravity. This theory is plausible—it has cotton balls falling more slowly than rocks. And note that the theory is not about a single result, but rather about a *function*.

Before they start their experimentation, we ask students to predict what will happen, and to do so explicitly: they make their data table in Fathom and the relevant graphs, and fill them in with predicted values before they take any measurements. ("You mean, make up numbers?" they asked. You bet.) This gives them a structure for their planning, and, when printed out, gives us an assessment tool.

At last we give the students measuring sticks, stopwatches, cotton balls, and Fathom, and set them to work. Figure 1 shows some data from the field test. On their own, the data don't look very exciting, and you can't tell whether the theory is any good or not.

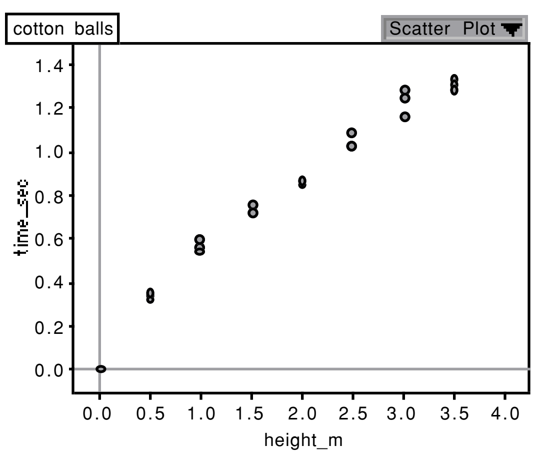

Figure 1. Sample cotton-ball data.

In Figure 2, however, we have superimposed a line and a curve. Both equations appear at the bottom of the graph; the constant **k** is a variable parameter controlled by the slider.

1. Luis Acuña, at ITCR in Costa Rica, inspired us in this activity.

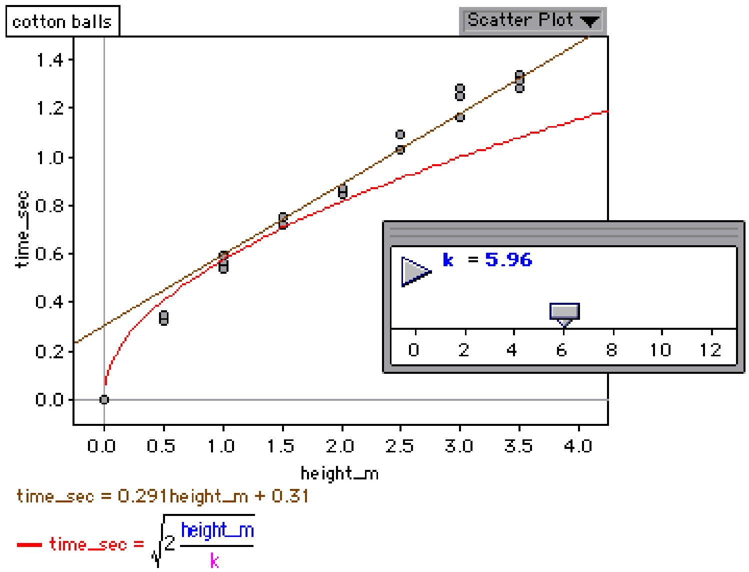

Figure 2. *The same data with a line and a curve superimposed. The line shows what may be terminal velocity; the curve—here, with a value of k=5.96—represents the theory. No value of k fits the data well.*

The students then had to make a brief report—actually a Fathom document—summarizing whether they thought the theory was correct and what evidence they could cite for their decision. The data also gave rise to a discussion about terminal velocity, what it was, and why, physically, it might occur.

The students went on (“More about Cotton Balls” on page 23) to try to improve the mathematical model for dropping cotton balls.

We enjoyed this activity because it exemplifies a lot of what we would like to see in lessons that integrate mathematics and science effectively: the curves are not “curve-fitting” exercises, but theoretical predictions being compared to data; constants that emerge (e.g., the slope of the terminal line) have real physical meanings; students use algebra for a real purpose; and they do essentially what scientists do: they look at a plausible theory, design an experiment, take careful measurements, display the data, and find out that, despite the theory’s reasonableness, they can tell it’s wrong. And the only way to tell is by doing the math.

This investigation also forces students to face many issues in that crack between mathematics and science we alluded to earlier. For example:

- Usually, in a mechanics problem, time goes on the horizontal axis. Here, it seems to go on the vertical. Why?
- To get the theoretical curve, students must not only solve the equation for *t*; they must realize that they have to.
- Measuring time is difficult. How do you account for reaction time? Students found that the shortest fall they could believably time was about half a meter.
- Some students solve for *k* and calculate a value of *k* for every drop. You can use this to invalidate the theory, but what does that value *mean* in that context?

- Is the spread between measurements at the same height due to measurement error or is it actual variation in the time of falling? Or does that matter?
- What does it mean to compare data to a function instead of to a constant?

Commentary. Lets' repeat ourselves: It is data *analysis*, not data, that injects mathematics into science. The natural venue for data analysis in science class is in a laboratory session. But traditionally, laboratory "experiments" are cookbook operations in which students collect data, usually to demonstrate some phenomenon. The "wet" work takes most of the time in the class period, so curricula usually compress data analysis into preordained calculations, written in the appropriate place in the lab notebook. The students do not decide what calculations to make; they do not compare their results to a mathematical model; they do not consider why the data and model may not match; they do not improve the model; they seldom address issues of measurement error; and they certainly don't have a chance to decide how to design the experiment in the first place.

Could students do these things? Of course they could—but we don't usually ask them to. If the lesson is not to be a cookbook, it will be somewhat open-ended. But students like to be led, and most teachers don't know how to handle open-ended lessons. Why not? One reason is the practical problem of time. Students need time to figure out how to measure something subtle or to grapple with comparing their data to mathematical models. And science teachers know that with students of varied abilities—and different levels of mathematical facility—a lesson can disintegrate when you ask students to pull that kind of mathematical load. At best, it will take a long time—an extra class session, perhaps. Do this very much, and you'll never make it to Chapter 15 by the end of April.

What can we offer for this problem? Our field tests show that, with our technology and carefully-crafted materials, high-school students can actually do real data analysis in a reasonable amount of time. What makes the difference? Fathom is fast, powerful, and easy to use. Science students become Fathom-literate almost immediately: they enter data, calculate models, and perform other analyses. Equally important are the tasks we have students do. Older curricula are written with the time limitation in mind, so real data analysis is absent or buried in an extension. But knowing that students will have the technology, and will therefore have time, we can deliberately choose tasks where data analysis substantially enhances the science.

What is the connection to mathematics education? The original, presenting problem that motivated our Phase I proposal was the prototypical science-teacher complaint that the students don't know the math—so science teachers have to teach the math "all over again." More astute teachers recognize that the students may indeed know the math, but in a different guise. For example, variables in science are seldom x and y[2], the coefficients are not likely to be small integers, and the old familiar functions are often translated to unfamiliar parts of the Cartesian plane. Technology helps the students with just this kind of difficulty. For example, Fathom plotted the data and the square-root function in Figure 2 reliably and instantly. This still requires mathematical understanding—you have to express the function symbolically and enter the data—but it lets you make lots of mistakes, and fix them, without wasting hours and graph paper. The result? Students get to use mathematics in a meaningful context.

2. I myself have seen calculus students who could easily differentiate x^2 with respect to x but were flummoxed when differentiating r^2 with respect to r.

Issues in Constructing Lessons

There are really three questions: what is the math, what is the science, and how do they interact with one another and with technology to make a better educational experience for the students?

The first two questions about math and science are interrelated. We focused our efforts during Phase I on physics and, to a lesser extent, chemistry. So the mathematics is different from if we had focused on biology or economics. Within physics, we focused on mechanics. The most obvious related mathematics has to do with functions—especially position as a function of time. However, we will see (e.g., in "Measurement" on page 10) that other areas of mathematics inevitably become important. We list a number of ways math is actually used in science ("How Scientists Use Mathematics" on page 12) as a way to help think about integrating the two disciplines.

Since we plan to use data analysis as part of this bridge, an obvious place to insinuate our ideas into the traditional high-school science curriculum is where the data are: in laboratory activities. So let's discuss the lab and how we could use more mathematics there; then we'll talk more about specific mathematics topics, and finally, we'll consider a broader science topic than physics or chemistry: science itself.

The Math in Labs. One way to make the lessons we want is to extend existing laboratory activities, making sure to do substantial data analysis that enhances understanding of the science. You could think of the cotton-balls activity above as an extension of a traditional free-fall lab for measuring the acceleration of gravity. What are those existing lessons like?

Labs in today's schools run the gamut from dismal cookbook operations to brisk activities where students efficiently assemble their equipment and take good measurements to illustrate phenomena they are studying in lecture. In a chemistry class that we observed, students were studying phase changes. They used temperature probes and CBL equipment to display how temperature changes with time on their graphing calculators as water in a test tube froze in an ice-and-salt slurry.

Such labs are successful by today's standards. But at what? Certainly at illustrating a phenomenon, which is unquestionably important.

But such lessons use little mathematics, and do very little data analysis to illuminate the science. In the phase-change lab, while students do get to see temperature as a function of time, they don't do anything with the function other than note the temperature where it's flat (zero Celsius) and observe what was in the test tube while it was there (a mixture of ice and water) before the temperature went negative. Mathematically, they're looking for a constant, not a function; one wonders how much they really needed the CBL technology: they could have looked at a thermometer once a minute and gotten the same result.

Let's consider a lab that already integrates work with functions.

We did not actually see this lab being done, but we read the students' lab reports. It was a Galileo's-ramp activity, where students time metal balls rolling down inclined planes. They sketch parabolas through the distance-time points and derive an acceleration.

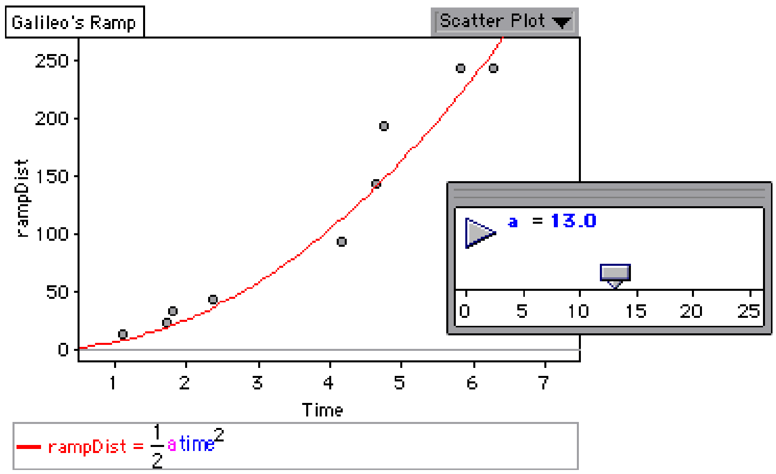

Figure 3. Galileo's ramp data taken by project staff. As in the cotton-balls investigation, the slider **a** *is a variable parameter controlling the parabola. Distance is in centimeters, time in seconds.*

This is better—in the integration-of-mathematics dimension—than the phase-change lab because of the function connection. What could we do to improve it? Here are four ideas:

- Ask students to predict the time it would take to roll a previously-unmeasured distance. Have them derive their predictions two ways: from the graph and *using algebra* with the parameters they derived. Then, of course, they try it and see what happens.
- Use simple numerical modeling (See "Numerical Modeling" on page 11.) to model incremental changes in position and velocity under constant acceleration; this is one way to see *why* the curve is a parabola.
- With technology, have students use a variable parameter for an acceleration to generate the parabola (as shown in Figure 3), so that they can see dynamically what value fits the data best.
- Once students have a parabola, have them use Fathom to create residual plots and make conjectures about what effects (e.g., measurement error, friction) are unmodeled by the function.

Returning to the phase change activity, what could we do to incorporate more mathematics? One investigation that meshes well with the chemistry curriculum is to ask how the temperature of the phase change—the freezing temperature—depends on the concentration of salt in the water. Suppose each group used a different salinity and could pool their data. Now we have a function we might be able to model—interestingly, *not* a function of time. It turns out that this very idea is already an extension to this lab as it is written. But students did not get to do it—because there was not enough time to do the data analysis.

This is typical; it is not hard to come up with ideas for extending existing physics or chemistry lab activities to integrate more mathematics through better data analysis. And there are plenty of good resources, both for physics lab *topics* (e.g., Dickens 1995) and entire lab *instructions* (the bibliography in Dickens, or, for example, Appel et al., 2000). Parallel resources exist (e.g., Bond 2000) for chemistry and other sciences.

Sources of Data. To make lessons with data analysis, we need data. It is good to think about the source of the data for the lesson, to make sure that the data will be available, to give the teacher possible alternate sources of data, and also to give students variety. Here are some that we have tried:

- Student-generated. This is the one where students take a measurement and type the data into the computer. It's essential to give students ownership of the data and the process of collecting it. If you rely only on more automatic data collection, you risk making data a mysterious black box. "Example Lesson: Dropping Cotton Balls" on page 4 is uses this kind of data.
- Static Web Data. There are lots of data on the Web, listed in tables just there for picking. Fathom is generally good at importing such data; more and more, schools have good web access.
- Dynamic External Web data. Other web sites generate pages on the fly based on forms that you fill out. A good example is the Jupiter Ephemeris Generator[3] at JPL, from which you can get positions of the Jovian satellites at any time. Erickson (2000) has several activities suitable for a science classroom that use these data. In addition, the KCP Technologies Census microdata project (Finzer 2001) has developed technology that could make retrieving such data even easier.
- Dynamic Internal Web data. By this we mean, data from a site expressly set up to be downloaded for a science lesson. This could be an internal site that collects data from all the students; then it could serve the class data to individuals for analysis.
- Special-purpose simulation. This could be a web applet or a stand-alone program that simulates some phenomenon, generating data for import into Fathom. One prototype in this project is a solar system simulator we called *Copernicus*. We will discuss that below in "Copernicus" on page 23.
- Fathom simulation. Fathom itself has extensive simulation capabilities; students can use Fathom to simulate many types of phenomena. Numerical models (see page 11) are simulations, of course. One can also use randomness, for example to simulate measurement error.
- Probeware. Probes take measurements and squirt them through an interface into a computer. This is becoming increasingly popular (we are told) in science classrooms and is (we have seen for ourselves) an exciting way to get a lot of real data to analyze, often very quickly. We should note that while it is wonderful for every student or lab group to have a probe set-up, that's not essential: if you're short on equipment, take one set of data as a whole class, then share the resulting file for analysis. "The Elevator" on page 25 is an example in this document.
- Video. If you have video of a phenomenon, you can take measurements off the screen and enter them into Fathom. This is more and more practical for curriculum developers and individual teachers. And while we would rather have students actually experience as much as possible, video has several benefits: you can stop it to measure something that happens quickly; you can use it to see something that is too inaccessible or impractical for the class to see; and you can make the thing to be studied uniform across the class (that is, no one messes up the experimental setup). "More about Cotton Balls" on page 23 has an example of what we mean here.

3. http://ringside.arc.nasa.gov/www/tools/ephem2_jup.html

Measurement. Let's return from data to the mathematics, and look at an example of math "in the cracks." When we had made some prototype activities that had the "feel" of what we had imagined, it was interesting to ask, "what mathematics do students use in these activities?" Of course, they used functions and proportion and algebra—mainstays of secondary math. But we realized they also used *measurement* in a big way.

Measurement is a major strand in most mathematics policy documents. Both the NCTM *Curriculum and Evaluation Standards for School Mathematics* (1989) and the more recent *Principles and Standards of School Mathematics* (2000) name it as one of their content standards. Of course, measurement is vital to the elementary years: students learn to use nonstandard and standard measurements, use tools correctly, understand that measurement is approximate, and use different units, both to express different ideas (e.g., square centimeters for area, centimeters for length) and different measurement systems (centimeters vs. inches). The NCTM (1989) measurement standard for grades 5–8 includes selecting appropriate units and tools, understanding derived measures and systems of measurement, and making measurements and estimates to describe and compare phenomena[4].

Read liberally, theirs is a pretty good list, especially for grade 8. But then it stops. Measurement does not appear as a standard in grades 9–12 in that (1989) document, though some concepts are folded into geometry. In the newer document (NCTM 2000), there is a measurement standard at grades 9–12, whose expectations read, in part:[5]

> In grades 9–12 all students should—
>
> - make decisions about units and scales that are appropriate for problem situations involving measurement.
> - analyze precision, accuracy, and approximate error in measurement situations;
> - understand and use formulas for the area, surface area, and volume of geometric figures, including cones, spheres, and cylinders;
> - apply informal concepts of successive approximation, upper and lower bounds, and limit in measurement situations;
> - use unit analysis to check measurement computations.

Looking at the examples (or reading between the lines) we see that NCTM is talking, understandably, about measurement in the service of mathematics. And though science education benefits from these abilities (especially unit analysis), we need more:

- Analyzing approximate error goes beyond counting significant figures or even calculating the minimum and maximum possible values based on the precision of component measurements. Where appropriate, students should analyze repeated measurements to estimate their accuracy. They need not necessarily calculate a mean and its standard error (for example), but could at least base the estimate, however informally, on data.
- Indirect and derived measurements go beyond inaccessible distances, speed, and density. How would you measure g, the acceleration due to gravity? If you're studying physics, you probably

4. NCTM 1989, http://standards.nctm.org/Previous/CurrEvStds/5-8s13.htm
5. NCTM 2000, p. 320, http://standards.nctm.org/document/chapter7/meas.htm

learned that $s = (1/2)gt^2$. It is not trivial for students to realize that they can time something falling a known distance, *do some algebra*, and figure out the acceleration. Of course, repeated measurements will help them assess the accuracy of that observation.

- The NCTM *Standards* often reduce measuring quantities other than length to knowing how to use a formula. In fact, there are many practical problems involved in making accurate measurements—for example, timing a falling object, finding the volume of something irregularly shaped, or counting the number of bacteria in a petri dish.

It seems to us that these are all legitimate areas of study: it is worth class time to give students a chance to focus on them. If we do not, there is an equity problem: if these skills are prerequisites for success later on, and they are not taught in school, only those students who get them outside of school—presumably, advantaged ones— will succeed.

Numerical Modeling. Once you have measured to get some data, you need to compare the data to something—often a mathematical model. The most obvious way to model in physical science is with explicit functions. For example, you may have a phenomenon that's linear, so you fit a line to it. The slope has some physical meaning you extract.

We also use functions to compare data to theory. But many phenomena have theories for which it's hard to figure out the function. If some differential equation describes the situation, we may not know its solution. Even if we do, we would like to avoid bringing out such difficult mathematics—especially when easier math can solve our problem: we can solve many differential equations, numerically, as difference equations to an arbitrary degree of accuracy.

This is an intriguing idea, but is it feasible that high-school students could do this? This seems an empirical question, so we field-tested three activities—a brief introduction to numerical modeling in Fathom, a complicated investigation involving air resistance ("More about Cotton Balls" on page 23), and a less-complicated but also less-structured problem ("The Elevator" on page 25)—all with considerable success given the limited time we allocated to that part of the field test[6].

The most interesting bit, however, came during a debriefing. We asked the students if they had ever solved problems that way. *No, they had not*. What was it like? *Easy. Makes sense*.

Then we pointed out that this was the end of the year when they knew physics already. Would it make any sense to learn physics using computations done this way—in little slices of time? The students replied, it would make *more* sense. They said it would have made it easier when they were first learning mechanics back in the Fall.

Aha: when you model mechanics numerically, you break time down into tiny increments, and repeatedly perform simple calculations on position, velocity, and acceleration. These yield small, step-by-step changes in those quantities which connect together to produce complex dynamical results.

6. We should note that there are lots of different ways to do this kind of numerical modeling. Since the point is to help students understand science through accessible mathematics, we chose the simplest kind that would give us answers that were good enough. One can get more accurate results (for example, by integrating with trapezoids instead of rectangles) but we think that comes at a price in understandability.

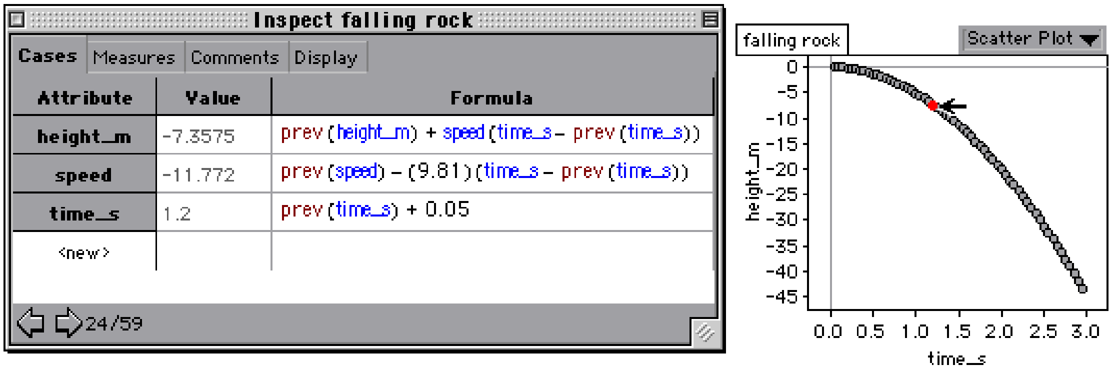

Figure 4. Numerical model of a falling rock. You can see the formulas in the Fathom "inspector" on the left. The arrow in the graph is pointing at the particular iteration that the inspector displays. Students can also see a table of the data.

But within each tiny time increment, it's easy: position is the old position plus the time step (here we have written it as **time_s – prev(time_s)**) times the speed. Speed is the old speed plus the time step times the acceleration. It's only the acceleration rule that changes: constant for gravity near the surface, proportional to position for a spring, and so forth. It only takes a tiny bit of algebra to understand a single increment.

And it's accurate. If it's not accurate enough, make the time step smaller. Here, of course, is an important connection: if you keep making it smaller, you get calculus. Wouldn't it be interesting—after all, we now have the computing power—to learn physics, numerically, when you first see algebra? How would that change how you look at the limit-taking in calculus when you finally approach that mountain?

How Scientists Use Mathematics. We have looked at measurement as an example of mathematics we may have missed, and numerical modeling as an example of math we have never tried to use with students. Let's step back a bit and ask more generally how scientists use mathematics.

The National Research Council (1995) summarizes the roles math and technology play in science:

> A variety of technologies, such as hand tools, measuring instruments, and calculators, should be an integral component of scientific investigations. The use of computers for the collection, analysis, and display of data is also a part of this standard. Mathematics plays an essential role in all aspects of an inquiry. For example, measurement is used for posing questions, formulas are used for developing explanations, and charts and graphs are used for communicating results. (NRC, p. 175)

The Phase I proposal listed a number of specific areas where we use mathematics in science. It's worth repeating them here, revised and enhanced through our work:

1. We use mathematics—or at least numbers and categories—to record data. And we use mathematical ideas, e.g., the concept of a variable, to organize our thinking and plan just what data we will collect.

2. We use mathematics (some would say computer science) to organize, clean, and transform our data, to get it ready to analyze. In one example ("The Bicycle Wheel" on page 21) we have to write a formula to tell when a spoke passes our sensor; or we might need to use relational formulas to look up data in another table; or we might have to clean up or transform our data—for

example, to look at a change rather than an absolute amount as in "Evaporation" on page 20. We often forget that this "data preparation" is an important, often time-consuming part of data analysis (Gould 2001).

3. We use mathematics to make indirect measurements. This goes beyond trigonometry for inaccessible distances. For example, knowing a relationship among variables, we first solve for the one we want to determine; then we measure the others and calculate.
4. We use mathematical functions to model phenomena. To be sure, we can calculate using formulas, e.g., to find the force of gravity on the surface of the moon. But real understanding comes when we think of the phenomenon functionally, e.g., to see how that force decreases as we get farther away. Recursively-defined functions and many types of simulations can be models as well.
5. We use mathematics to describe departures from a model. Models typically oversimplify; often we draw the most interesting conclusions from the way data depart.
6. We use mathematics to cope with variability in data. This goes beyond insisting that every measurement have a "plus-or-minus." For example, we might calculate standard errors, or repeatedly simulate an inexact measurement process. More subtly: when we fit a model to data, the criteria we use to judge the fit involve variability.
7. We use mathematics to decide if we can reject a hypothesis based on quantitative results. Often this requires only an informal comparison of data to model, but sometimes we need statistical inference in order to assess whether an observed effect could be due to chance.
8. Mathematics—logic and statistics—also helps implement our experimental designs (e.g., to control variables both informally and using multiple regression or analysis of variance).
9. We use multiple mathematical representations—graphs, of course, but also tables, charts, and formulas—to help us understand the story the data tell and to communicate with others. The more representations we have at our disposal, the more flexible we are at exploratory data analysis and communication.

These will not all be present in every activity. But they are good examples of what sort of thing we should look for as we decide what to ask students to do. Mathematics permeates the doing of experimental science. Too often, curriculum developers insulate students from that mathematics; instead, we can look for ways to put the student and the math together.

A Structure for Activities with Data

We adapted the *Sonata for Data and Brain* from Erickson (2000) for special use with science lessons. The main purpose of the Sonata form was to give teachers a structure for more open-ended problem solving. In addition, it prods students into making conjectures about what they will find when they eventually look at the data. Essentially, it consists of a single page with prompts to make a conjecture, describe the measurements, and reflect on differences between the conjecture and the results.

Our adaptation was to make the first, conjecturing phase more explicit. We ask students to invent data in Fathom before they take measurements. With these data, they set up their Fathom file, complete with its analysis, displaying what they think will happen. Then it is easy to put in the real data (by making a new column) and compare it to their predictions. Figure 5 shows some student-made graphs.

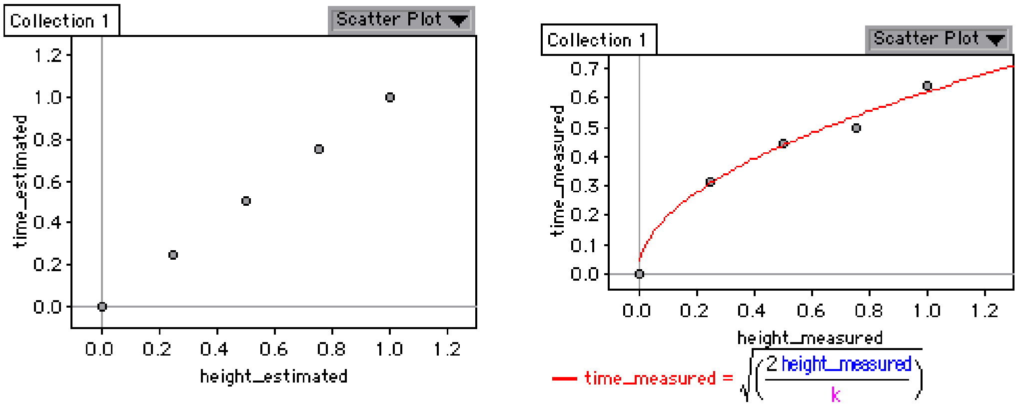

Figure 5. Graphs of one groups's predictions (left) and actual measurements (right) for the cotton balls activity ("Example Lesson: Dropping Cotton Balls" on page 4).

In practice, this not only helps students think about the situation; it also helps students plan their measurements, generally get organized, and think about data analysis. We hope it will help minimize the common situation in open-ended investigations that students take a lot of data and find out at the end that they didn't measure what they really needed to know.

We are indebted to Gerhard Salinger of NSF for a comment ("Not only should students do data analysis; they should do the analysis before they have the data.") which reinforced our intention to try this.

Broader Issues

What are some overall directions in curriculum development, and in science curriculum in particular?

At this writing, American educational reform is characterized by "tough standards" and accountability, fueled by frequent testing, particularly in mathematics and reading. This trend seems unfriendly to innovation in science education, but two factors mitigate in our favor: first, our innovations all involve technology, which people generally want more of; second, the moderating influences of more progressive education will inevitably return. One way to look at curriculum development *now* is that it is prepares us for the next swing of the pendulum.

A good example of current progressive thinking is in *Understanding by Design* (Wiggins and McTighe, 1998). They are disturbed that much current curriculum development seems to be about covering a list of topics—often, these days, an extensive list in some set of standards.

As an alternative, Wiggins and McTighe suggest what they call "Backwards Design." In their view, rather than a list of topics, we should set our sights on "Enduring Understandings"—the biggest ideas. These are more central than the items on a list of what it is important for students to know and be able to do: they should be the nuggets that have "enduring value beyond the classroom," and "reside at the heart of the discipline" (Wigging & McTighe, 1998, pp. 10–11). For example, in physics, conservation of energy is such a nugget; the formulas for kinetic and potential energy are not. Once we have identified these enduring understandings, before we start making lessons, we figure out what evidence we would need to be sure that students have accomplished what we have in mind (that is, we should attend to assessment early). Only then do we write the lessons themselves.

The key is identifying the enduring understandings. Where would we find these things? Fortunately, in science, the people at Project 2061 have been figuring them out for some time. Since *Science for All Americans* (AAAS, 1990), they have developed a compelling set of benchmarks and associated materials to help curriculum developers do their jobs coherently and productively.

Project 2061 is refreshingly brave in their choice of benchmarks. Rather than falling prey to science educators' particular tendency to put everything in, they ruthlessly keep the inessential out. Their benchmarks are models of intelligible generality rather than a list of disconnected specifics.

A Fly in the Ointment, and How to Swat It. The one area where we might differ from the benchmarks themselves in is the overall difficulty of the materials we would produce. Project 2061 explicitly intends the benchmarks to be common science-for-all goals. They state:

> *Benchmarks* specifies thresholds rather than average or advanced performance. It describes levels of understanding and ability that all students are expected to reach on the way to becoming science-literate. (AAAS 1993, page xiii)

The individual benchmarks echo this sentiment, for example, when discussing symbolic relationships:

> In Project 2061, we don't expect students to remember formulas for accelerations or parallel circuits or mass action; nor do we expect them to be able to perform algebraic manipulations or solve simultaneous equations. We do expect them to acquire an understanding of proportionality, the ability to read an algebraic formula, and to develop the ability to relate the shape of a graph to its implications for how some aspect of the world *behaves*. (AAAS, 1993, page 216)

In the face of such a statement, suggestions we will make—for example, that we could introduce numerical modeling in high school for teaching mechanics—may seem extravagant, and run foolishly counter to a document we so respect. However:

- Even though we are trying to make more quantitative science accessible to *more* students, what we produce may be "beyond the benchmark," that is, still not for *all* students.
- Our market includes college-level courses, including teacher preparation, so it is natural that some of our work would be beyond common goals for high-school students.
- In addition, since the way we use technology specifically helps students integrate their understandings of mathematics and science in new ways—ways that are impossible without the technology—our material may be more accessible than the *Benchmarks* might suggest.

What the Benchmarks Do for Us. The *Benchmarks* can help guide the project even if our alignment with them is not perfect. For one thing, their very ruthlessness can help us decide among different candidate activities or stances; for example, they remind us to concentrate on *understanding* symbolic descriptions rather than manipulating them. The 2061 *Benchmarks* are remarkable for two more things:

- They include mathematics and the social sciences within the embrace of their science benchmarks, and
- They give special attention to the process of science in general in a benchmark called the "Nature of Science."

That they even mention mathematics and how it integrates with science is unusual; many "standards" documents[7] reinforce the view that the disciplines of science and mathematics are, and perhaps should be, separate—and as incommunicado as their corresponding departments.

The social science connection is even more astounding, and suggests new directions for work outlined in Finzer (2001) and in "What About Social Science?" on page 28.

Finally, their focus on the nature of science lends legitimacy to its position as an object of study in its own right. We have long noticed that statistics and mathematical modeling "feel" like science. They revolve around questions we ask of the universe, and ask us to design experiments and plan observations to answer those questions—fitting the results into an emerging framework of understanding. The mathematics curriculum generally does not encompass much of the Nature of Science benchmark. And science curricula often start with a chapter or less on the scientific method, and leave it at that.

Making activities that real teachers will use, effective activities that have authentic nature-of-science learning objectives, would be a significant achievement. Kurth (2001) has already noted that our prototype activities are strong in that area. But how shall we make the activities so teachers will use them? Time is the problem again; we discuss ideas for solutions in "Where Does the Nature of Science Fit In?" on page 28.

Software Enhancements

The Timer. The first and simplest way KCP Tech enhanced Fathom was to develop a prototype for a software timer based on an earlier, web-based prototype from Epistemological Engineering. This is no technical miracle, but it is enormously convenient that the timer is integrated with the software, so that a keypress creates a Fathom case with an associated time code.

Students use the timer, for example, to record times in mechanics experiments. They can also use it to record times of events in other investigations. For example, students could use it to record times and categories of cars passing on a street.

This particular use suggests more a statistical study than a math/science integration. But the very existence of the timer was a proof-of-concept: could we use the "derived collections" infrastructure in Fathom to act as a portal for getting other data more directly into Fathom. Since the timer, KCP Tech has used some of the same ideas to create the IPUMS Census microdata interface (Finzer 2001) and the probeware interface, below.

Probeware Interface. Thanks to a generous contribution by Vernier Software and Technologies, we received a LabPro Interface and three probes with which to test our ideas. Using their technical documentation, KCP Tech implemented a prototype interface for two of the probes using the USB port on the Macintosh. "The Bicycle Wheel" on page 21 and "The Elevator" on page 25 are dramatic examples.

7. California's science standards, for example, mention mathematics once in the introduction and only twice in the body, in this eight-grade statement: "[Students will] apply simple mathematic [sic] relationships to determine a missing quantity in a mathematic expression, given the two remaining terms (including speed = distance/time, density = mass/volume, force = pressure * area, volume = area * height)" (CDE 2000, page 30) To be sure, they have many more statements that *imply* knowledge of mathematics (e.g., "Students know how to solve problems by using the ideal gas law in the form PV=nRT." (page 38)) but these are qualitatively different from the *Benchmarks*' deeper requirement.

Plotting Multiple Series. We also asked KCP Tech to teach Fathom how to plot more than one attribute on an axis. That is, until now, if you had cases that included, say, the time of a measurement, the voltage at one place in a circuit, and the voltage at another place in the same circuit, you had to graph those separately. Now, holding down a special key (this is a stand-in for a better interface to come) you can drag the second attribute to an axis; Fathom plots the two data series in different symbols and displays a legend at the bottom of the graph. You can see an example in Figure 6 on page 21.

This feature is particularly important in our ideas for curriculum when you consider plotting predicted values and actual measurements on the same graph.

Plug-In Architecture. We contributed part of the funds to create a "plug-in architecture" for Fathom. When fully implemented, we will be able to write small plug-in files. Users will drop these into a special folder on their disk to give Fathom new functionality. We have tested this with mathematical functions students can use to define measures and new attribute values.

Thus we will be able to sell small files that enhance users' copies of Fathom. This is good for us—it gives us another product to sell—but is also good for users, especially beginners. That is, when you start using Fathom, you will not be overwhelmed with possibilities. As you become more adept and have additional needs, you can add functionality.

For example, most math teachers will not need the probeware interface or many of the statistical distributions. A science teacher will need the probeware, but many fewer functions. On the other hand, that same science teacher may need additional functions, which she can buy.

Future Software Enhancements

We were of course unable to implement everything we would like to have done during the short Phase I period. However, we identified several important candidates for future work. We could implement many of these as plug-ins, as described above.

Time-series functions. Fathom has few functions especially designed for dealing with time series. Science activities could use them frequently. So we plan to implement functions ranging from control functions to smoothing to FFTs.

New Time-Series Oriented Derived Collections. Fathom implements samples (and other things) as derived collections. If we have time-series data where we need a new function that does not conserve the number of cases, we could use a derived collection. Some FFTs would be like this. But more important, we could implement "trigger" collections. Consider time-series data from an EKG (Vernier makes this probe, by the way). You may collect data 50 times a second, but want a case for every heartbeat. You would write a formula that determines when a heartbeat takes place; Fathom would make a new case—with whatever measures you specify—when the formula fires.

Date and Time Data Types. Fathom currently recognizes the strings "10:30" and "4 July 1776" only as strings. They have no numerical meaning at all. Yet a lot of data (especially from the Web) have time markings like those as values for the independent variable.

While there are plenty of canned routines to convert these data types, the interface question is a minefield, and needs careful design work.

Units. Measurements generally have units attached to them, and Fathom should honor that. A too-simple solution is to tag every attribute with a unit that Fathom would display on the graph axes. We should do unit algebra, so that when you multiply speed by time you actually get distance; we could also convert any "odd" units automatically.

But while we have the computational engines to do all that, we need a good design. How do users specify default units? What about overriding those defaults? Are units always present or are they a preference—or a plug-in?

Better Plot Overlays. The current "multiple series" solution (above) is a huge improvement, but it needs more work. In particular, it is impossible to plot data from more than one collection on the same graph. While this might be confusing in general, in our science materials, it would be helpful to plot data from a model on the same graph as data from an experiment, or to plot results from multiple experiments on the same graph. We imagine something like dragging one graph on top of another and having them readjust their axis bounds (and converting units) to fit together. Figure 10 on page 24 has an example of what we had in mind; the author used a complicated work-around to get the right behavior in the current version.

Plot "Adornments". Science cries out for a wide variety of elements on the graphs. A good example is error bars. KCP Technologies have long dreamed of implementing "user-constructed graphs," where users could specify screen objects that depend on calculated values. This is one solution to the error-bar problem. But in looking at our real needs based on the work in this grant, we have realized that a more modest, less general undertaking, coupled with better overlays, may be sufficient for the great majority of users' needs.

User-Created ControlText. Fathom has several places that use what we call *ControlText*—text with special numbers in it (they're blue at the moment) that do two things at once: they change when underlying values change, and you can edit them to change the underlying values. We use it, for example, as one way to set axis bounds. ControlText's dual-purpose nature is great for the user interface—it means you don't need a separate mode for changing these numbers, and yet you can see their values.

But users cannot create ControlText of their own. We'd like to make that possible. If they could, users could write reports in Fathom with statements like "The mean height of the tomato plants in the sun after **21** days was **17.3** cm. In the shade, the mean height was **11.7**," except that the numbers would actually be calculated internally, so that the text in the report would change as you added data. Here, again, the challenge is to design the interface. How does the user specify what gets inserted?

"2061" Curves. The Benchmarks (AAAS, 1993) point out that it is more important to be able to think of functions in terms of their basic shapes than to know all about their formulas. So to make activities that are still quantitative but less symbol-dependent, it would be good to include some general-shape curve-drawing tools, analogous to Fathom's existing "movable line." Students could use these to "fit" existing data or to make explicit predictions about data they will take in the future.

Networked Sharing of Data. We effectively distributed files using the computers' own capabilities during the field test, but it would be useful for Fathom itself to have networking capabilities so that different groups of students could contribute to a class data set in real time.

Field Testing at Berkeley High School

While we had hoped to try activities with students earlier in the year, we did not find a suitable classroom and teacher until May 2001. We were, however, extraordinarily fortunate to find Richard White, who teaches physics at Berkeley High School in Berkeley, California. He invited us to try activities for a total of fifteen class-hours—five hours in each of three different sections of physics.

Two of the sections were advanced placement (AP) physics; the other was "regular" physics for college-intending students. Almost all of the students were seniors. Each session was a double period totalling about 100 minutes. One of the AP classes was exceptionally capable.

That the students were, in general, so experienced, was a mixed blessing, overall positive. We do want to serve students of a wider range of ages and capabilities, and we would have liked to work with students in other disciplines. Still, this is a feasibility study. Our prototype activities are in their roughest form, so these students can give them the best chance of working.

That this all took place in May was mixed as well: we could assume that their school physics was well in hand. Interestingly, the AP students, while enormously skilled as textbook problem-solvers—they had taken the AP exam already—had done fewer labs than the "regular" class. It was therefore interesting to see where our data-oriented, more open-ended approach was new to them. Also, all of these young men and women were in their last few weeks of high school, and while we saw good attendance and solid work, some minds were clearly elsewhere.

After a brief introduction to Fathom from *Data in Depth* (Erickson 2000), we did an extended activity ("Example Lesson: Dropping Cotton Balls" on page 4) and followed it with as many extensions as we could fit in. We did the modeling extension with all classes (the first part of "More about Cotton Balls" on page 23) and a shorter problem ("The Elevator" on page 25) with the two AP sections. This last was largely to test the feasibility of approaching mechanics numerically as described in "Numerical Modeling" on page 11.

Computer Access. Having computers available is always a problem for curriculum development like ours. It's an equity issue as well; less well-off or more constrained schools either haven't enough sufficiently powerful computers or have them locked away where they can only be used easily for word processing.

Fortunately, this seems to be changing. Berkeley High—not a rich school, and with an economically diverse student population—is not alone in starting to invest in mobile computer labs. We were lucky enough to get one for the whole time we were visiting.

Ours was a cart that plugs into the wall. A printer sits atop the cart, as does an AirPort (Apple's wireless network) base station. In the cart are fifteen iBook computers with AirPort cards. We installed Fathom on the computers (25 minutes) and were ready to go. With the wireless network, students could print on the printer; and if we had plugged the ethernet cable from the hub into the RJ-45 port on the wall, the students would have had Internet access. We left it unplugged. The largest class was almost 30 students, so they worked in pairs. The smallest class was eight so they worked alone. We preferred the bustling feel of the shared computers, though there's no way to tell if it was the sharing that made their work better.

Assessment. This was a field test simply to see if the activities would be comprehensible. Nevertheless, conscious of Wiggins and McTighe's (1998) lamentations over curriculum designed without assessment in mind, we were attentive to assessment issues. There are two things that merit special mention:

- Printing is important. In "Example Lesson: Dropping Cotton Balls" on page 4, we asked students to make explicit guesses as to what the data would be like. To force them actually to do this step, and take the risk of predicting, we asked them to print their predictions and turn them in. This worked as expected in its rôle as a motivational ploy and as an advance organizer. But it gave us unexpected assessment fodder: in the middle of the activity, we had evidence on paper of the level of students' understanding and intuition about the situation.

  For example, many AP students but few "regular" students gave us predictions that eventually straightened out to show a terminal velocity. No students predicted data that showed any variation from drop to drop. And in both classes, many students produced data points that were well outside the practical range for real data collection (e.g., dropping the cotton ball 20 meters). Over the course of the year, we would expect this latter to improve.
- Screen size is important. Walking around the class, with real computer screens—as opposed to the small screens of graphing calculators—you can tell at a glance what the students are up to and who needs assistance. This helps teachers make moment-to-moment instructional decisions and evaluate individual students' progress or the overall effectiveness of a lesson.

Examples of Lessons and Problems

Here are more extended descriptions of some of the lessons and problems we have referred to elsewhere.

Evaporation

We designed this lesson before we had access to classrooms, so we did it on our own. Teachers said they thought it was completely plausible, however, for a real classroom. It's also worth describing because of its apparent simplicity—and the data analysis complications it generates.

We put 100 mL of water into each of four glass containers. The containers were of different diameters, ranging from about 10 cm (a peanut-butter jar) to about 2.5 cm (a tall graduated cylinder). Then, from time to time over the course of several weeks, we weighed each of the containers on an electronic scale that was accurate to 2 grams, and entered the weights into Fathom. We also recorded the temperature in degrees Celsius.

We made a conjecture: The wider containers would evaporate faster, possibly in proportion to the surface area; in addition, we might be able to see faster evaporation on hotter days. We imagined that humidity would play a role, but had no hygrometer.

The raw data are date, time, temperature, and weight in grams. But the weights themselves are not what we want. Instead, we want the total amount of water that has evaporated. So we need a new attribute, which we calculate as the difference between the current weight and the first weight in the series. The independent variable, time, is a problem as well. We need to combine the date and time into a single "time" variable, which we decided would be measured in decimal days.

These two calculations are simple if you're experienced, but enormously subtle if you aren't used to dealing with data. Students will need to be led though such calculations at first, and then, later, given a chance to decide for themselves what attributes need what kinds of transformations, and then to make those without hints or suggestions. This is an excellent example of the "simpler" mathematics we might realistically expect most students to be able to do.

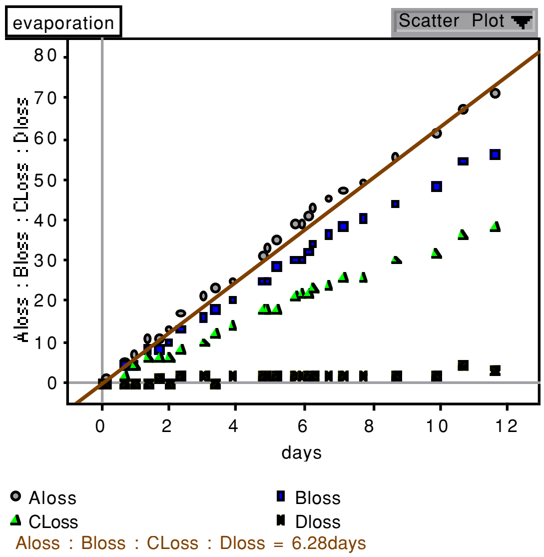

Figure 6. *Evaporation data, with a "movable line" approximating the first (largest-diameter) series. The equation at the bottom shows that it lost water at a rate of about 6 grams per day. Note that this shows the new feature of displaying more than one series on a graph.*

Figure 6 shows a graph of these transformed data; the line shows the evaporation rate for the largest-diameter container (labeled "A", so the variable denoting how much it has lost is **Aloss**). This is another piece of simple mathematics that—we know from experience—many students need practice with. The slope of the line in the figure is 6.28. What does that mean? Students may know all about slopes in math class, but to really understand that this is the evaporation rate in grams per day takes more experience—and it's useful math to be able to do.

Further analysis (not shown) suggests that the slopes of the lines are more-or-less proportional to surface area. And indeed, the temperature is a bit lower in the last few days of this period, so we think the apparent decline in slope is real.

This could have been a simple, cookbook lab: measure this, subtract that, calculate this or that slope, write it in the space. But this is a good example, we think, of an activity that could be more open-ended. With the help of the technology, middle-of-the-class high-school chemistry or biology students can grapple with some of the data-analysis issues this situation presents and, with help from the teacher and their groups, succeed.

The Bicycle Wheel

This is another investigation we did not use with students, but a good example of integrating interesting mathematics—this time, using probeware to get the data.

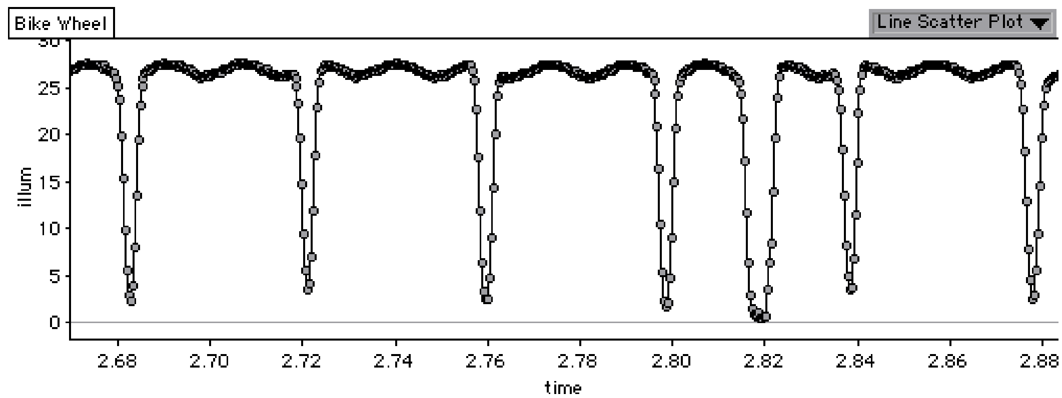

Figure 7. Light shining though bicycle spokes. Time is in seconds, illumination in lux. There is one data point every 0.0004 seconds, i.e., 2500 Hz. Notice the oscillation in the background. Also, at 2.82 seconds, the shadow of the valve.

We inverted a bicycle so that the front wheel ran free, then set up a flashlight on one side and a Vernier light probe on the other. The idea was to see whether we could detect the shadows of the spokes as they went past (we could, as in Figure 7) and ultimately, to use Fathom to predict when the wheel would stop.

At this point, we used a formula in Fathom to detect the spokes, and created a new data collection, where each case was a spoke instead of a single observation of light intensity. Then we further calculated, for each spoke, the time since the previous spoke. The left side of Figure 8 shows how that delay, **dt**, increases with time. The right shows how its reciprocal (divided by 32, the number of spokes), called **cps**, decreases with time. An important question for students is which you would rather use to predict when the wheel will stop.

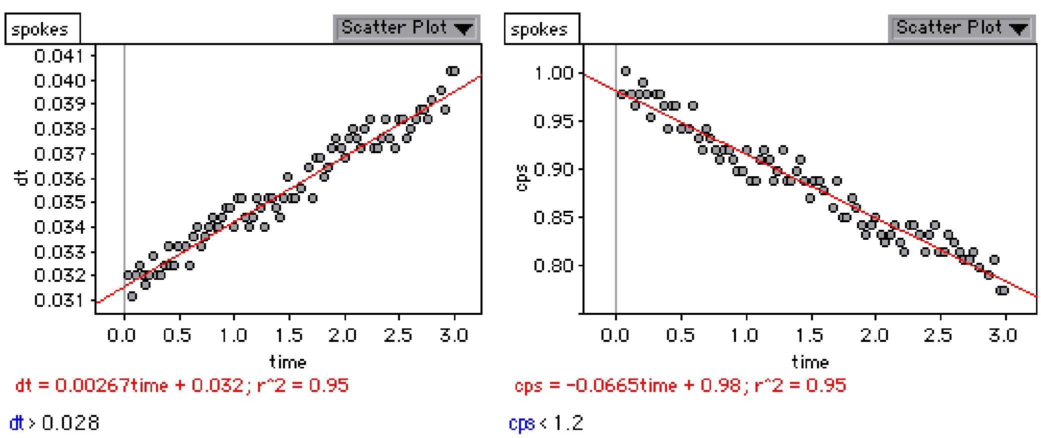

Figure 8. At left, the time between the spokes as a function of time. At right, those quantities converted to cycles per second, also as a function of time. We also show the least-squares regression line for each graph.

This situation confronts students with important mathematics. Just to get the raw data into a form they can use, they have to detect the spokes, collect these spokes into a new data set[8], convert **dt** into **cps**, and somehow (here we use a Fathom "filter") eliminate the too-small **dt** values that correspond to the

8. This will be easier with new features; see "New Time-Series Oriented Derived Collections" on page 17.

valve. But foremost, once they have made these graphs, students have to understand what the graphs mean—what the variables signify. But there are also the questions of why both graphs are so linear, how far we would have to go in the left-hand graph to detect its curvature, and finally, back in the science domain, whether the speed really decreases linearly—that is, is it right to extrapolate?—and if not, how the data will depart from the linear model.

Copernicus

We designed this lesson to use a stand-alone program as a source of data. *Copernicus* is a solar-system simulator. You are on a planet with a simulated night sky in which you see stars and other simulated planets. You control the time; as the days pass, you can distinguish planets from stars because the planets move against the stellar background. The first task is to do just that. The second is to use the motion of a superior planet (one farther from the sun, like Mars to us; and it's easy to find) to determine as much as you can about the length of its year.

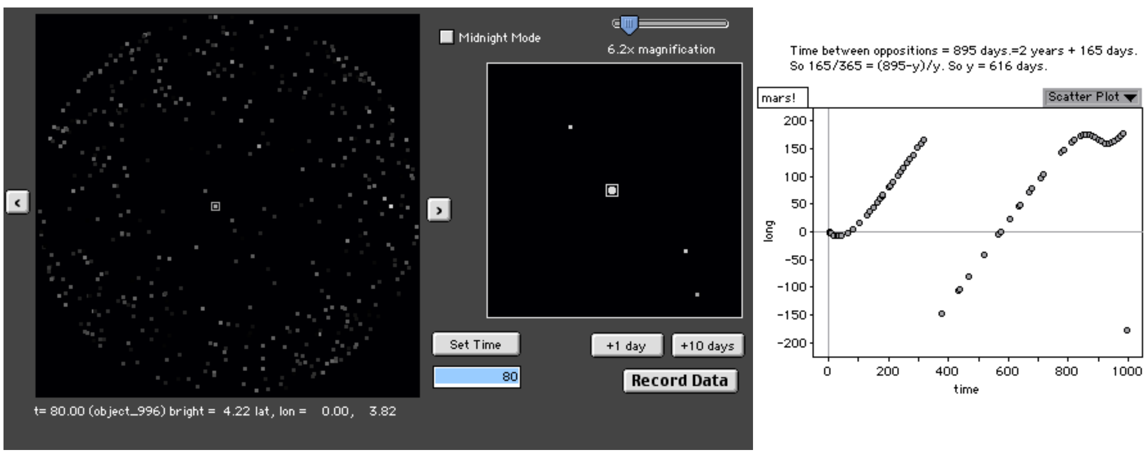

Figure 9. The Copernicus window ("Mars" is in the middle of both the sky and telescope fields) and an example of Copernicus data—in this case, longitude with respect to time—in Fathom. A Fathom text object describes the relevant calculations for determining the length of Mars's year.

When you click on an object, you measure its position in celestial latitude and longitude[9] and its brightness. You can export these data (in this prototype, by cutting and pasting) into Fathom. An example appears in Figure 9. Notice how you can see retrograde motion—where celestial longitude decreases—and that students would have to cope with longitudes that suddenly jump from plus 180 to minus 180 degrees.

The lesson succeeds as a proof of concept in its use of the technology, although the situation proved too conceptually difficult (much to our surprise) for the experienced adults who tested it.

More about Cotton Balls

We asked students what sort of model they would propose to explain the falling cotton balls; generally they suggested using the parabolic, theoretical model at first, "transitioning" to a linear model as the cotton ball reached terminal velocity. In asking students why there should be a terminal velocity, some

9. We used these rather than right ascension and declination, as the terms are more familiar.

rightly came up with the explanation that, since air resistance increases with speed, the force of air resistance would eventually equal that of gravity, resulting in constant speed.

At this point, we experimented with having the students use Euler's method to make a numerical model that incorporates air resistance *and measurement error*.

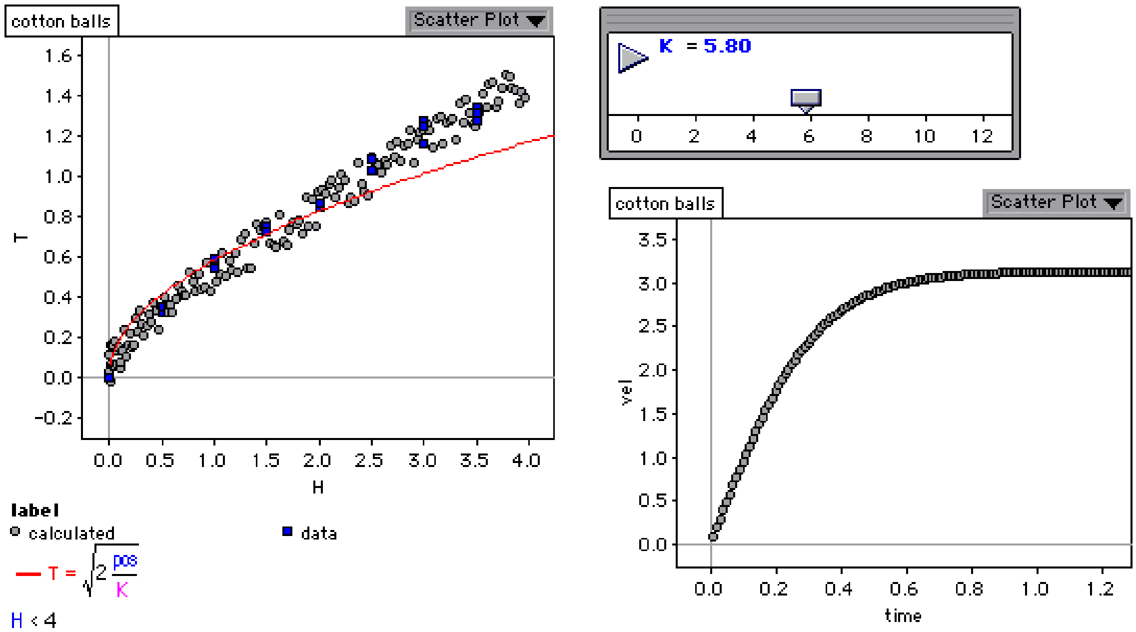

Figure 10. At left, the same cotton ball data as in Figure 1 on page 4, this time with simulated model data superimposed (with some error as well). The slider again controls the curve. On the right, you can see model values for the velocity as a function of time, showing how it approaches a terminal velocity.

This model works remarkably well, but we were not able to distinguish between models where air resistance is proportional to velocity and those where it's proportional to its square.

There is still little data at the beginning of the fall because of the problem of timing well, so we decided to experiment further after the field tests. We made a short QuickTime video of a cotton ball falling a little over a meter. Since we could look at it frame by frame, we could get consistent data every 1/30 of a second. Figure 11 shows data acquired this way.

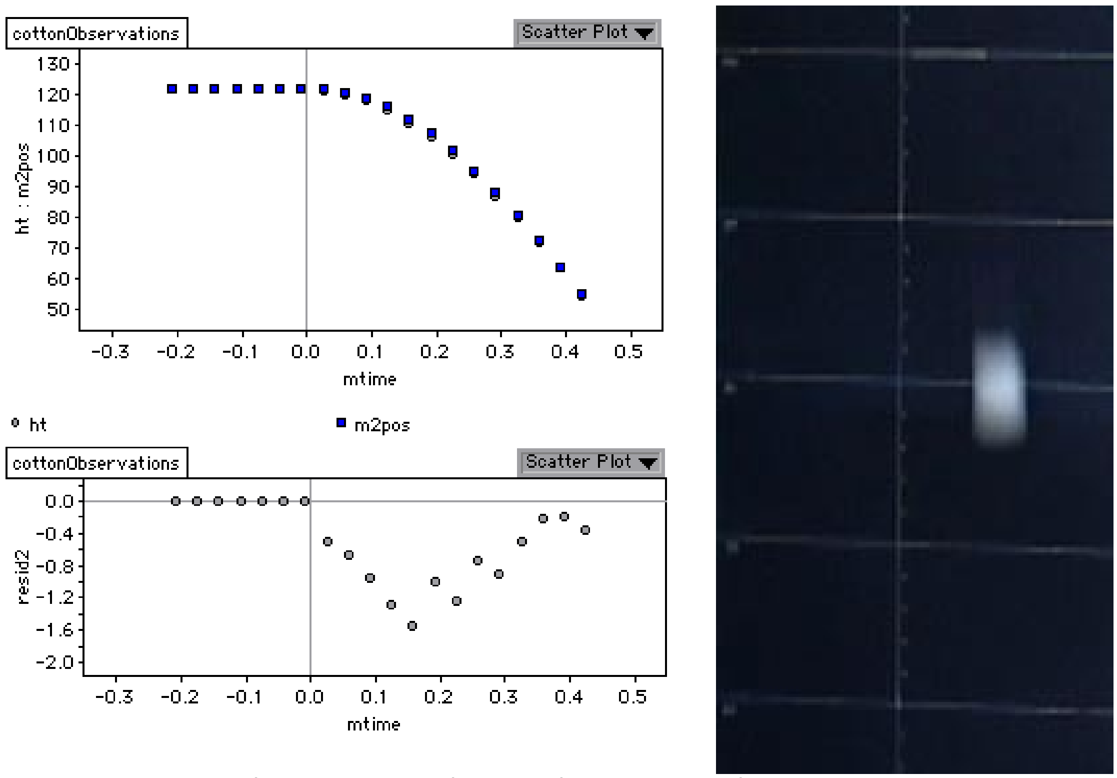

Figure 11. The height of a cotton ball as a function of time, measured from a video. The top graph shows the height as measured and according to a numerical model. Since they are nearly coincident, we also show, below, the residuals, where you can see a pattern telling us that the model, while good, still needs improvement. At right, part of a frame from the video. Horizontal lines are every 10 cm.

There are two points to showing this example here:

- Video—which teachers can create digitally or that we put on the same CD as the software—is a good medium for simple data collection. It's especially appropriate in a case like this when you need to stop the action to see what is happening.
- We can make a remarkably good model, using really rather crude equipment and measurements, for a very complex situation. Without the technology, we would never have considered posing such a problem *because the mathematics was too complex*. With Fathom, the numerical modeling is relatively easy—and it works.

The Elevator

We implemented this idea not as a whole activity but as a single problem—as you might find in a problem set at the end of an unusual chapter on mechanics. We collected the data using a Vernier force probe and then gave the computer file to the students for analysis. (This is a good example of a situation where it was not practical for students to collect the data themselves.)

We hung a force probe from a laboratory stand, and hung a coffee mug from the probe. At rest, the probe reports the weight of the mug. Then we set the whole apparatus in an elevator and, collecting 50 data points per second for 20 seconds, pressed the button and traveled two floors. Figure 12 shows the thousand points of data. The task is to figure out which direction the elevator moved—and how far.

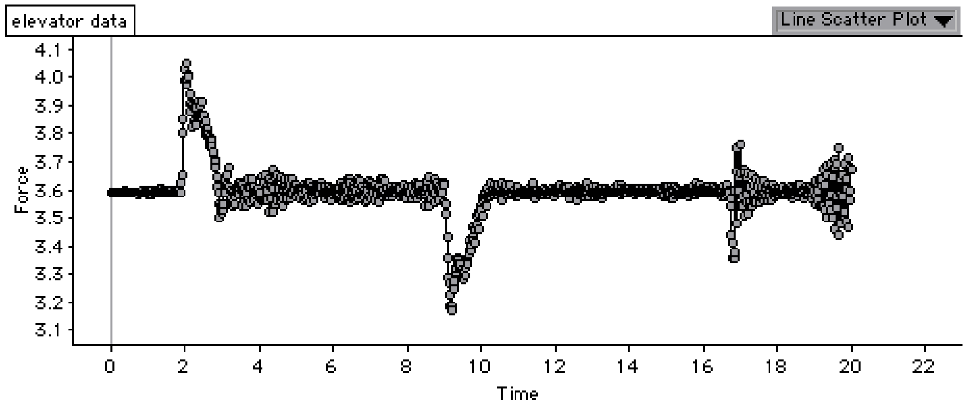

Figure 12. Force registered on a force probe from a hanging mug during a two-floor elevator ride. Time is in seconds, force in Newtons.

First you need to imagine elevator rides, and compare them to the data, to determine which way the elevator went. The cup got heavier at the beginning and lighter at the end; it went up.

To figure out how far, however, you have to set up the data for analysis, for example, by converting the force measurement to acceleration. After that, use Euler's method to integrate the acceleration, twice, to get a distance. Interestingly, it's hard to determine the mass of the cup exactly. So you make it a variable parameter, (a slider in Fathom) and adjust it. The key boundary condition to meet is that the elevator is at rest (velocity is zero) at the end. With that, we determined a height that was within 2 cm of what we got by extending a long measuring tape over the balcony between the floors.

If this is not a fluke, it speaks volumes about the power of the probe when used in conjunction with good software. We should also note that the great learning potential comes in having the students set up the numerical work rather than having an automatic integration function as a feature in the program.

Potential Commercial Applications

We expect to submit a Phase II proposal in January 2002 to continue this work. We have developed, with field-test teachers and advisors, a number of ideas about potentially viable products that would build on our results.

Replacement Labs for Physics. A book of the ten to twenty-four most commonly used physics or chemistry labs—the old chestnuts—extended and updated to use appropriate technology and to have a substantial data-analysis component. Our field-test physics teacher, Richard White, is tentatively interested in collaborating, as is his colleague, Aaron Glimme. Such a book would most likely include a CD with the software, plug-ins, blank worksheets, sample data from real students, and videos of the labs in case teachers could not use real equipment.

Such a book for physics might include traditional labs like these:

- Free-fall or Galileo's ramp: acceleration of gravity.
- Pendulum. What affects the period, and how?
- Starting and sliding friction. [Lack of] conservation of energy.

- Rotation, torque, moment of inertia, angular momentum.
- Inferring Ohm's Law.
- Parallel circuits (even though the Benchmarks don't like them).
- Optics. Distances of image and object from lens.
- Springs. Figure out $F = -kx$.
- Strength of materials: deforming a beam, breaking a bridge.
- Ripple tank interference/diffraction.
- Refraction: discover Snell's law.
- Air and water pressure, e.g., water squirting out of a column: how far does it squirt?
- Surface/volume relationships. Melting ice cubes of different sizes and shapes.

We would choose these because they are the ones teachers are doing anyway—so using these would be a minimal disturbance to the routine. It would be good, however—either in this book or in a companion volume—to suggest labs (such as the cotton balls or the bicycle wheel) that we could not do well without the technology.

Playing Experimentalist. The format of the cotton-balls lab—where the students received a fake theoretical paper that they were to support or refute with experimental evidence—is extremely attractive. One could imagine a book of these one-page "papers" with hints and suggestions for how to proceed. This would of course be more open-ended, and engage students more in the nature of science as well as their discipline-specific content.

Data Analysis Problems. We have been speaking, in this report, as if students encounter data only in their labs. But that's not necessarily so. We could publish a resource book, with CD, of problems for data analysis in physics or chemistry They would focus on skills that run from cleaning the data all the way to presenting the final report—covering many of the mathematical issues we listed above in "How Scientists Use Mathematics" on page 12. In this report, "The Bicycle Wheel" on page 21, "The Elevator" on page 25, and the video section of "More about Cotton Balls" on page 23 are good candidates.

Sciences Other Than Physics and Chemistry

Fields such as biology or earth science will have different mathematical emphases, relying less on nonlinear functions and more on proportion and statistics. There may be other areas of mathematics that are "in the cracks" between mathematics and these sciences in the same way that measurement is, that is, mathematics essential for success but not actually present in either syllabus. One candidate is what we might call "demographic thinking"—the way we can predict how a distribution evolves. Technology like Fathom is effective in coming to grips with topics like this, especially using its built-in simulation capabilities—which we have hardly touched in our applications to physics.

Still, the principles would remain the same, and the list "How Scientists Use Mathematics" on page 12 should still apply. Various advisors have suggested that a series of the "chestnut lab" books would be attractive; the "Experimentalist" and "Data Analysis Problem" books would work as well.

It would make sense to do these for Physics first, as we look for suitable like-minded collaborators with expertise in other areas to work with us on further products.

What About Social Science?

We originally intended integrate mathematics with just the natural sciences. But could we expand our ideas for into the social sciences as well? Beth Wellman, Technology Director for the California History Project, confirmed our suspicion that, while history teachers are beginning to use technology, they don't integrate it with mathematics.

Besides, what math is there in history? Quite a bit, it turns out. The *Benchmarks* (AAAS, 1993) include an entire chapter on "Human Society." Their theme of "Group Behavior" appears especially well-suited for data analysis and Fathom. It also corresponds with research in social psychology about the role of critical thinking in anti-bias education. It was clear to Allport (1954) that prejudice depends in part on faulty thinking. In particular, when people form stereotypes, they fail to realize

> an almost universal principle in respect to overlapping group differences: the differences within the same group are greater…than the differences between the averages of the two groups. (Allport 1954, page 102.)

And there's the clue to the math: differences (and similarities) between groups. This is a central issue in statistics—comparing within-group and between-group differences, and making valid inferences about those differences from data.

Can this really help in anti-bias education? It seems so: based on 1975 research, Glock et al. identified "cognitive sophistication," a disposition for critical thinking, as a key factor in avoiding prejudice. They recommended three specific kinds of instruction to combat adolescent prejudice. Two of these relate to the understanding of group differences: 1) instruction in the logic of inference, "so that youngsters can come to recognize when group differences are being falsely accounted for..." and 2) instruction to make it clear that you cannot infer individual differences from aggregate properties of groups.

This represents an intriguing possibility for the future of Fathom in schools: a chance to help teachers make mathematics more personally useful and meaningful to students, as well as a chance to open up new markets.

Where Does the Nature of Science Fit In?

In a word, everywhere. It may be that our attention to nature-of-science issues distinguishes our work from that of potential competitors. We are good at choosing the authentically scientific alternative from among myriad curricular possibilities.

That said, we envision several ways this can happen:

- Through thoughtful discussion questions and extensions in single activities, possibly flagged as focused on the nature of science;
- Through giving teachers more student-centered alternatives to traditional labs;
- Through the creation of entire units that focus on the nature of science as their main content.

This last possibility is the most exciting but the one least likely to sell in the current climate. Nevertheless, that might be changing as colleges develop quantitative literacy courses that might be fertile homes for such an idea.

References

Allport, Gordon W. 1958. *The Nature of Prejudice*. (Abridged) Garden City, NY: Doubleday Anchor. (originally published 1954)

American Association for the Advancement of Science, Project 2061 (AAAS). 1990. *Science for All Americans*. New York: Oxford University Press.

AAAS. 1993. *Benchmarks for Science Literacy*. New York: Oxford University Press.

AAAS. 2001. *Atlas of Science Literacy*. New York: Oxford University Press.

Appel, Kenneth, John E. Gastineau, Clarence Bakken, and David L. Vernier. 2000. *Physics with Computers using Logger Pro*. Beaverton, Oregon: Vernier Software & Technology.

Bond, William B. 2000. *Teacher's Guide—AP Chemistry*. New York: College Board.

California Department of Education (CDE). 2000. *Science Content Standards for California Public Schools: Kindergarten Through Grade Twelve*. Sacramento, CA: California Department of Education.

Dickens, Roderick Spaulding Jr. 1995. *Teacher's Guide to Advanced Placement (AP) Courses in Physics*. New York: College Board.

Erickson, Tim. 2000. *Data in Depth: Exploring Mathematics with Fathom*. Emeryville, CA: Key Curriculum Press.

Finzer, Bill. 2001. *Census Microdata in the Classroom SBIR Phase I Final Report*. Emeryville, CA: KCP Technologies.

Glock, Charles R., Robert Wuthnow, Jane Allyn Piliavin and Metta Spencer. *Adolescent Prejudice*. New York: Harper and Row. 1975.

Gould, Rob. 2001. *Introduction to Data Analysis*. Presentation at the Tinkerplots annual meeting, Madison, WI.

Kurth, Lori. 2001. Private communication at the annual meeting of the National Science Teachers Association, St. Louis, Missouri. Kurth is a Project 2061 staff member who graciously reviewed some prototype materials and teacher instructions in advance of the field test.

National Council of Teachers of Mathematics (NCTM). 1989. *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: NCTM.

NCTM. 2000. *Principles and Standards of School Mathematics*. Reston, VA: NCTM.

National Committee on Science Education Standards and Assessment, National Research Council (NRC). 1995. *National Science Education Standards*. Washington, CD: National Academy Press.

Wiggins, Grant.and Jay McTighe 1998. *Understanding by Design*. Alexandria, VA: Association for Supervision and Curriculum Development.