Toward a Set of Lessons on Accuracy and Uncertainty in Measurement for Future High School Teachers†

James J. Madden (LSU Math)

with the assistance of

David Kirshner (LSU Math-Ed), Ed Lamy (Math Teacher, Catholic High)
Julia Hiatt (LSU Math-Ed undergrad), Kevin Zito (LSU Math & Math-Ed grad)

Introduction

These notes were prepared in summer 2002 in order to report on work done during the 2001/2002 academic year at LSU in a project entitled "Mathematics for Future Secondary Teachers". The project involved four teams—I shall call them lesson-study teams—at four institutions in Louisiana. The teams set out to produce and test experimental curriculum modules—or lesson packages—intended for use in college-level mathematics classes for pre-service secondary teachers. The project, which is still in progress, has a research component. It seeks to answer the question, "What can a group of people working diligently under the conditions created by the project and pursuing this goal actually produce?" This report gives a preliminary answer.

Each lesson-study team included a mathematician, a math educator, a high school teacher and one or more pre-service secondary math teachers. The members of the LSU team appear as authors of this report. Our team met several times in summer and fall of 2001, and several more times in spring and summer 2002. We piloted lessons informally on a few different occasions at different levels: undergraduate, graduate and high school.

Our lesson package is intended to elaborate the theme of measurement error and uncertainty. We feel that this topic is appropriate because of its intrinsic importance in high school science, because of the difficulties that many high school and college students have in dealing with measurement uncertainties and because of the many rich and interesting connections that the topic has in mathematics and in statistics. The theme of measurement uncertainty does not, however, figure prominently in existing high school mathematics curricula. We did not consider this to be a serious disadvantage, since we wanted to plan a lesson package of small size and limited scope in any case.

Our lesson-study team returned again and again to the basic theme of how errors in the factors of a multiplication affect the accuracy of the product. As the mathematician on the team, I sometimes felt impatient to get to some more esoteric topics. However, 20 years experience teaching mathematics shows, if anything, that it is foolish for me to insist upon an agenda for learning that is determined solely by my interests. In writing this report, I have had plenty of opportunity to follow mathematical trails where I please.

† This work was supported by NSF-0087892 and LaCEPT CRP (1999–2000)-1703, Amendment Number 3

We felt that work of the kind that we were doing would require a new format for reporting. We view our work as an attempt to marry serious mathematics with serious mathematics pedagogy in a way that would be interesting to a mathematician and to a mathematics educator, and would be useful to a high school teacher, to a teacher-educator and to a student preparing to become a math teacher. There are very few examples of work that meets all these criteria.[1] The reader may judge the extent to which we have succeeded. We would be delighted to hear readers' reactions and to make use of readers' recommendations in future revisions of this work.

We have divided this report into three sections. The first section, "Mathematical Discussion", is itself a mixture of points of view. We hope it will contain parts that might be used as text for lessons at various levels, and also contains purely mathematical discussions. The second section, "Pedagogical Discussion", considers "what the standards say", as well as some comments on where the high school curriculum is headed. (See also Appendix I.) We give a very brief report on related math-ed research and we include a theoretical discussion of a fictional lesson. The last section, "Lesson reports and student work", is self explanatory. We have tried a number of things in classrooms, and have acquired a sense of how students at different levels tend react to tasks of various kinds that involve the theme of how measurement uncertainty affects calculations.

What are the prospects? We set out to develop something on the scale of a lesson or short sequence of lessons. For a number of reasons, I now believe that the proper scale for a portable curriculum module is a unit—a sequence of lessons and activities that might take several weeks. The structure of the standards-based NSF curricula for K-12 bears this out. Many of these are based on coherent, "unit-sized" chunks. A unit should involve one or more significant themes, and would consist of many interconnected activities. A good organizational principal could be to start with a significant problem, and develop the resources needed to solve it.

We could broaden our focus sufficiently to support a unit-sized project by integrating more themes from statistics. This would fit very nicely with the historical origins of statistical ideas—the central limit theorem was first devised in connection with the problem of reducing measurement error by combining observations. Computing with uncertainty seems like a very reasonable theme to include in such a unit along with this theme. We are beginning to plan such a unit.

[1] But there are *some*. The work of Dick Stanley (Berkeley) and Al Cuoco (Educational Development Center) come particularly to mind.

I. Mathematical Discussion

A. Accuracy and uncertainty in measurement.

Measurements are approximate. The numerical data we record when we make physical measurements of quantities such as length, position, mass or temperature are almost never perfectly certain. Inaccuracies arise in the measurement process from the limited precision of the instruments that we use and as a consequence interferences from sources both known and unknown. No matter how exacting we may be when we set about determining a quantity that lies on a continuous scale, the numerical figure that we ultimately arrive at is, at best, an approximation. With better equipment and more refined procedures we may narrow the uncertainty, reducing it to hundredths, thousandths or millionths of a unit, but we can never eliminate it. Except when we have *declared* an object to be a standard, and therefore to possess a certain measure exactly, we can never know a physical dimension with perfect precision.

Measurement reports should include uncertainty estimates. When we report a measurement, the figure that finally stands as our best guess need not and should not stand alone. It does not represent all we know. When we complete a measurement we always acquire substantially more knowledge than just the "best guess", since we have been engaged in the process of obtaining that guess. If we have employed a procedure that we have used many times before, then we are likely to have a sense of how much accuracy we can count on, how close we probably are to the "true" value, or at least to the value that might be obtained if other, more time-consuming and expensive procedures were employed. In some cases, we may be able to estimate our uncertainty by repeating a measurement over and over and observing the pattern of variation that occurs. It is not unusual to find that the results fall into a simple, bell-shaped distribution, a shape that results from the net effect of numerous random influences. There are well-established conventions for reporting such distributions.

It may be just as important to know the uncertainty as it is to know the measurement itself. If the old scale in my bathroom says I've gained 6 pounds, I may brush it off, knowing that the readings I get tend to change by several pounds depending on where I place my feet. On the other hand, if the scale at my gym says I've gained then I worry, since I know that the gym scale is a reliable instrument. When I report my weight to my trainer, he wants to know if I've gotten the figure from my bathroom scale or from gym scale. If the figure from my bathroom scale shows a large gain, he may just ask me to reweigh myself. If the gym scale shows a gain, then I can expect consequences. Far more serious examples arise regularly in engineering. The strengths of materials and the stresses and strains to which a structure may be subject are all estimates and are all subject to a degree of uncertainty. The engineer's task is to create a design that will resist failure despite the uncertainties. To be able to guarantee reliability, the engineer needs to know more than the estimated values of the strengths and strains—she also needs to know how confident she can be that they are in a given range. Similar examples from industry, medicine and finance abound. The point is, a measurement may be quite useless unless it is accompanied by an estimate of its accuracy.

When measurements are used as the starting point for a calculation, the final results that come out of the calculation have an accuracy that is limited by the accuracy of the input data. If we wish to know how far to trust the output, we must keep track of how the uncertainties in the input affect the calculation. Linear functions afford great simplicity. An error in the input to such a function produces a proportionate error in the output. The proportion is always the same. Different inputs with equal errors produce outputs with equal errors. For any other kind of function this fails. The proportion between input error and output error varies depending on the input, and it may be quite difficult to track.

Calculations performed by machines present even more problems, since they may be subject to uncertainties from sources other than the inputs. Calculators and computers typically do not represent numbers with perfect accuracy. A pocket calculator, for example, may display 12 or 13 digits, but no more. Larger computers use what is called "floating point arithmetic". Each number must be stored in a part of the computer's memory of some limited size—typically 64 bits. This means that the computer can use only 2^{64} or roughly 18.4 billion billion distinct numbers. After each step of a calculation, the computer must round off the result to one of the numbers that it can store. Although this allows enough accuracy for a lot of purposes, there are many kinds of computation in which the errors that accumulate due to rounding quickly overwhelm the computation and render its results meaningless. Finding robust algorithms—algorithms, that is, that are resistant to small uncertainties—is a major engineering challenge.

Let us summarize the main points. Measurements are almost always approximations. Anyone who uses measurements is likely to find necessary to analyze precision, accuracy, and approximate error, and particularly must be able to account for the way in which uncertainties in data may influence the certainty of conclusions drawn from those data. When data are recorded and processed electronically, the influence of errors becomes increasingly subtle and complex.

B. Naive and sophisticated ways of talking about uncertainty.

In developing lessons on uncertainty, the students and teachers that we worked with often felt that the idea of an approximate measurement was somewhat slippery. For example, consider the sentence:

$$\text{To within a tenth of a mile, my house is 2.3 miles from the store.} \tag{1}$$

Is the symbol "2.3" to be interpreted as an exact number? Or is it a reference to something that is vague and inexact—not a number but a fuzzy something? These questions are not intended to raise philosophical issues but to suggest a certain kind of puzzlement that we actually observed when we posed problems involving approximate data.

If I write "22.3 inches" and "33.5 inches" on the board and identify these as approximations—accurate to within a twentieth of an inch—of the length and width of a table and then ask students how well 747.05 (the product of 22.3 and 33.5) approximates the area, they may wonder if it is actually permissible to multiply the figures 22.3 and 33.5 given that they are only approximations. At some point in the discussion, a student is likely to ask whether 22.3 means *exactly* 22.3, or possibly something that is merely close. And if these figures are symbols for approximations rather than true numbers, then are there special rules for multiplying them? These are merely indications of the kinds of questions that I actually dealt with when discussing measurement with college students. There are many other ways in which students struggle to represent the meaning of “approximation”. For many people (of all ages) the idea is not pinned down precisely. It floats.

There is something interesting going on here. College students did not have problems accepting the idea that measurements are approximate, and the idea that errors in the input to a calculation might produce errors in the output seemed very natural to them. But if asked to determine the accuracy of the answer to a multiplication problem knowing the accuracy of the factors, many were stymied. What accounts for this?

Students deal effectively with approximations in routine, daily contexts. *Add about a cup of flour. Write a paper about 20 pages long. Take about an inch off the top. The room is 12 feet, 3 and 1/4 inches wide, as best as I can measure.* Statements like these are exchanged every day, and they are effective in achieving their desired ends in the contexts where they are used. No one stops to analyze them or pick their meaning apart. If asked to explain what an approximation is, college students provide a hazy answer. Here is a statement that was offered by one of the undergraduates in our lesson-study group:

> *An approximate number is an imprecise value that is practical enough to work with yet not accurate theoretically—a substitute value for a number, that is acceptable to work with.* (Written 8/9/01.)

In mathematics, we deal with questions that demand more care and more precision than everyday tasks normally require and than everyday language normally delivers. To meet these demands, mathematics creates an artifice, a specialized language with new terms, finer distinctions and rules that are more rigid than those in natural language. Understanding this and agreeing to use the artifice is part of what it is to learn and to do math.

In mathematics, statements about approximations have a definite, specialized logical structure. The mathematical language agrees with the vernacular up to a point, but often questions that can be understood in the vernacular cannot be answered without making the specialized terms and structures explicit. It may be difficult for students to see that everyday comfort with a term like “approximate” does not equip them with the special conceptual tools needed in mathematical contexts. Speaking about uncertainty in a mathematical way involves certain specific conventions. Actually, there are different sets of conventions that are used in different mathematical contexts when more or less refined treatments of uncertainty are needed, and we will examine these below.

C. The "basic representation".

The conventions suggested in this section set up a very simple mathematical model for measurement situations. When one is sufficiently comfortable with the model, it can be very useful. The main point is to create a terminology that distinguishes among the specific quantities that are relevant in describing and thinking about measurement errors, recognizing their roles and and assigning them variable names. We will introduce a scheme for creating algebraic representations of features of the measurement situation, and we will use the following terms: **true value**, **measured value**, **error** and **accuracy**. The next paragraph says how we plan to use these words.

A quantity that is to be measured is assumed to have a specific real number value, called the **true value**. An attempt to measure it produces another real number called the **measured value**. If the true value is t and the measured value is m, then the **error** is $m - t$ and the **absolute error** is $|m - t|$. We say that a measurement m of a quantity with true value t has **accuracy** δ if the absolute error is less than δ.[‡]

If m is a measurement of a quantity with true value t, then the following statements all have exactly the same meaning:

i) the measurement has accuracy δ,
ii) $|m - t| < \delta$,
iii) $-\delta < m - t < \delta$,
iv) $t - \delta < m < t + \delta$,
v) $m - \delta < t < m + \delta$,
vi) $m \in (t - \delta, t + \delta)$,
vii) $t \in (m - \delta, m + \delta)$.

Using this convention, we can rephrase sentence (1) as follows:

$$\text{The distance from my house to the store lies in the interval } \left(\tfrac{22}{10}, \tfrac{24}{10}\right). \tag{2}$$

The "basic representation" is actually several representations, as the list above shows. Some of the alternatives are easier or more natural for students to latch on to. For example, in attempting to answer a question similar to, "How accurately can I determine the area of a table if its dimensions are 22.3 inches and 33.5 to within a twentieth of an inch?" one college student examined the products 22.25×33.45, 22.35×33.55 and many of the products in between (such as $22.26 \times 33.46, \ldots, 22.27 \times 33.47, \ldots, 22.34 \times 33.54$). On being questioned, the student revealed that she had translated the accuracy statement into a statement about membership in certain intervals and had decided to look at a sampling of numbers from the intervals involved in order to develop an understanding of the situation. This student was using the "basic representation".

A variant of the "basic representation" uses intervals that might not be symmetric. What I know about a number x might be encapsulated by the statement that x lies in some interval I, where I might be open, closed, half-open, half-infinite, *etc.* For example, the statement $x \in [2/3, 3/4)$ might be all I know about x.

[‡] We are treating δ as the radius of an inteval of *certainty.* More sophisticated error models allow for intervals in which the level of confidence is not 100%. See §I.E.

Representing approximate data by means of intervals is the basis for the disciplines known as "interval arithmetic" and "interval analysis". These are systematic attempts to develop substantial parts of elementary mathematics using intervals in place of numbers for the purposes of analyzing computations with approximate data and/or providing robust algorithms. See Alefeld and Herzberger 1983 or Neumaier 1990; both books contain extensive bibliographies.

In examining *i)*–*vii)*, mathematicians will be reminded of the notation that occurs in the δ-ϵ definition of continuity. If f is a function and t is a number, then the statement

$$f \text{ is continuous at } t$$

means

> given any $\epsilon > 0$ (no matter how small), we can ensure that $|f(m) - f(t)|$ is less that ϵ by requiring of m nothing more than just that it be a sufficiently accurate measurement of t.

D. Significant digits.

The conventions for significant digits are built on the fact that when we read the digits of a measurement from left to right, there is point at which the digits cease to be meaningful. For example, if I write my height as 175.26 centimeters, then the 6 in the hundredths place is meaningless, for even if I stand absolutely still, my height bobs up and down by many hundredths of a centimeter due to my breathing, the beating of my heart and the reflex motions by which I maintain balance. The 2 in the tenths place, in contrast, is uncertain but *not* meaningless. It is true that if I measure my height at different times of the day, then I see some variation. Yet, although the 2 is somewhat uncertain, 175.2 is always closer to my height than 174.8 or 175.6. So, the 2 carries some information.

In a measurement, one can usually identify a last meaningful digit. (Generally, the last meaningful digit is also the first digit about which there is some uncertainty; but it cannot be completely uncertain, for then it would be meaningless.) Based on this idea, we can make the following definition, which pretty well describes how the term "significant digit" is used in a wide variety of contexts:

Definition. *The* **significant digits** *in a measurement start with the first non-zero digit you encounter as you read the digits from left to right and they continue to the last meaningful digit.*

For example, the significant digits in my height (expressed in centimeters) are the 1, the 7, the 5 and the 2. My height in centimeters has four significant digits.[2]

[2] If I were to express my height in kilometers, it would be .001752. The zeroes between the decimal point and the 1 do not count as significant; there are still 4 significant digits. When recording a measurement, we write all the significant digits, and no more—unless, of course, it is necessary to write some zeros as place-holders. If I express my height in microns (millionths of a meter), I write: 1,752,000. Here, the last significant digit is in the thousands place. (Note that zeros are significant when they occur after the first non-zero digit, but not after the last meaningful digit.)

Some textbooks go further and suggest additional rules. A workbook for introductory college chemistry that I consulted contains the following passage, describing how the results of measurements should be reported.

> *The conventions governing uncertainty in any numerical measurement are as follows:*
>
> 1. *The last place in the number is where the uncertainty lies.*
> 2. *The uncertainty is* ± 0.5 *of these units.*
>
> *Therefore, in the number 30.6, the uncertainty is* ± 0.05 *(since the last place is the tenths place).*

This book tells us that if a measurement is reported to be 2.5 centimeters, then we can assume that the true value is between 2.45 centimeters and 2.55 centimeters. Similarly, if a measurement is reported to be 1.00 units, then the true value of the measured quantity is between 0.995 and 1.005.[3] This rule makes an explicit connection between the significant figures convention the first convention for expressing uncertainty, that I called above the "basic representation." Note that significant figures conventions lead to the possibility that someone might use a symbol such as "3.14159" to refer *not* to a number, but to what the standard representation would treat as an interval—in this case, $(\frac{3141585}{1000000}, \frac{3141595}{1000000})$. This would seem to set the stage invite confusions such as those mentioned at the beginning of §B—*i.e.*, the thought that 2.3 is not a number, but a fuzzy, number-like entity.

In some situations, part 2 of this rule is reasonable. For instance, if we are working with numbers that we have rounded off to the nearest hundredth (say), then the uncertainty is exactly what the convention says it should be. Another instance: a trained technician using a gauge with decimal gradations can generally "eyeball" one digit beyond the finest gradations. (When measuring with a ruler that has a mark every centimeter, the technician can usually estimate to the nearest millimeter.) So here again the convention that the error is less than half a unit in the last significant place is reasonable. However, in some cases significant figures cannot be used in conformity with this rule. For example, if we observe the rules for significant figures, we must express the measurement that appears in sentence (1) as 2.3, since the "3" is significant. But the actual distance is *not* known to be between 2.25 and 2.35. Here is another example:

Example 1. Using a ruler marked in inches, I have measured the length of a piece of wire to be 37.13 inches. Although this is my best guess as to the real length of the wire, I am confident of this figure only to the nearest hundredth of an inch. In other words, I am certain only that the true length is between 37.125 inches and 37.135 inches. Now, suppose I want to express the length in centimeters. The conversion from inches to

[3] Note that the two zeroes are important in this expression. If we wrote a measurement as 1.0, we would mean we were certain only that the true value was between .95 and 1.05.

centimeters is accomplished by multiplying by 2.5400, because 1 inch = $\frac{254}{10000}$ meters.[4] In terms of centimeters, what can I claim about the wire? If I go with the best estimate, 37.13 inches, and convert to centimeters, I get 94.3102 cm. Now, are all the digits of this figure meaningful? Clearly not. I have already acknowledged that I may have erred by as much as .005 inches in my measurement. Expressed in centimeters, my error may be as much as .01270. The following table shows exact conversions of my best guess, as well as the lower and upper limits of certainty:

$$\begin{aligned} 37.125 \text{ inches} &= 94.2975 \text{ centimeters} \\ 37.130 \text{ inches} &= 94.3102 \text{ centimeters} \\ 37.135 \text{ inches} &= 94.3229 \text{ centimeters} \end{aligned}$$

Thus, if I am sure that the wire is between 37.125 inches and 37.135 inches then I am also sure that it is between 94.2975 millimeters and 94.3229 millimeters. The significant figures are 94.31. *However*, the true value is *not* guaranteed to be in the interval $(94.305, 94.315)$.

Problem. Suppose we know that quantity Q is between 2.142 and 2.152. The width of the interval in which Q lies is .01, and we can write $Q = 2.147 \pm .005$. Can I express what I know about Q using the significant figures convention? Why or why not?

Many introductory chemistry books also provide rules for determining the accuracy in the output of a numerical calculation when approximate data is used as input. The rule for multiplication, for example, is typically stated as follows:

> *When measurements are multiplied, the answer can contain no more significant figures than the least accurate measurement.*

It is well to note that this rule does not tell you how many significant digits you can trust in the answer, but only sets an upper limit. The rule that gives the exact number of digits that can be trusted is very clumsy to state.

Example 2. Suppose we want to approximate π^2 to the nearest 100,000th. Note that $\pi = 3.14159265358979 32\ldots$, thus π to the nearest 100,000th is 3.14159. Let us square this approximation: $(3.14159)^2 = 9.869587728$. We cannot trust more than 6 digits, so the digits beyond the 100,000th place (the second "8") ought to be discarded. Rounding off to the nearest 100,000th, we obtain 9.86959. However, the true value of π^2, written out to the 9th decimal place is 9.869604401. Rounded to the nearest 100,000th, this is 9.86960.

The significant digits conventions are prominent in many chemistry courses. If examined scrupulously from a purely mathematical point of view, they appear clumsy at best. Significant digits conventions are useful as rough rules of thumb. Attempts, such as those quoted above, to make them into a rigid system of rules are misguided, and probably arise from attempts to create items for tests and quizzes.

---

[4] The conversion factor is exact. In 1959, the national standards laboratories of the English-speaking nations agreed to adopt

$$1 \text{ inch} = \frac{254}{10000} \text{ meters}$$

as the definition of the inch. (By international agreement in 1983, a meter is the length traveled by light in vacuum during 1/299,792,458 of a second.)

E. NIST guidelines for evaluating and expressing uncertainty.

The National Institute of Standards and Technology (NIST) is a sub-body of the Technology Administration of the United States Department of Commerce. NIST maintains standards for measurement and provides calibration services for research and technological enterprises in the United States. NIST Technical Note 1297, entitled *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, (Taylor and Kuyatt, 1994) is intended to provide a uniform set of directions for expressing the uncertainty of NIST measurements. It was based on the *Guide to the Expression of Uncertainty in Measurement* prepared by the International Organization for Standardization (Geneva, Switzerland, 1983).

Underlying the NIST guidelines is the notion that when a systematic attempt at measuring a given quantity is made, much of the knowledge that results can be summarized by assigning levels of confidence to various intervals. In other words, a person who has made a measurement will be prepared to affirm with some level of confidence that the measurand lies within some particular interval. Actually, the person may have various levels of confidence associated with several different intervals. For example, a laboratory might be prepared to assert with a confidence of 50% that a quantity lies between 1.234 and 1.235, and to assert with greater and greater confidence that the quantity lies in larger and larger intervals—*e.g.*, with 90% confidence that it lies between 1.23 and 1.24, and with 99.999% confidence it lies between 1.21 and 1.26.

In reporting a measurement with uncertainty, NIST recommends reporting the best estimate of the measurand as well as the so-called **combined standard uncertainty**. This is the best estimate of the standard deviation of the distribution of measurements. If the distribution of measurements is approximately normal, then such information translates into statements of confidence for various intervals by well-known properties of the normal distribution. If the distribution of measurements departs from normality, then this is significant information that should accompany the measurement report.

The combined standard uncertainty results from assembling the all the relevant uncertainty estimates. NIST classifies evaluations of measurement uncertainties as **type A** or **type B** according to how they are obtained. Type A uncertainty estimates are derived from series of independent observations. Such observations generally vary. Appropriate statistical methods applied to a large number of observations provide an estimate of the standard deviation of the distribution of all the results that would be expected if the observations were continued. Type A uncertainty is that which is evaluated by such statistical methods. Type B uncertainty estimates are those that are based on anything other than the statistical analysis of a series of observations. They are derived from many sources, including general knowledge of the materials and instruments used, manufacturers' specifications, the accuracies assigned to reference data, *etc.*

The information that is supplied with a measurement report should be adequate to guide its use. The NIST recommendations create a uniform set of expectations for reporting uncertainty and in this way facilitate communication. The standards do not limit the information that metrologists might share. Depending on the context and intended use, there may be reasons to report many other things.

More and more sophisticated representations of the state of knowledge one may have about a measurable quantity are possible. Kacker (2000) makes a connection between the NIST guidelines and Bayesian statistical analysis. A lengthy discussion of the Bayesian understanding of what it means to measure a parameter may be found in Sivia (1996).

II. Pedagogical Discussion

A. What should a unit on measurement and uncertainty accomplish?

The goal of our lesson-study team is to produce a treatment of the topic of measurement that will be useful in the undergraduate preparation of high school mathematics teachers. This leads immediately to several questions: *What should students and teachers know about measurement, error and uncertainty? When and how should they learn? What skills should they develop, and how should they develop them?*

Standards for high school mathematics and for the preparation of high school mathematics teachers do not put a great deal of emphasis on measurement. It is a topic that tends to be more prominent in the lower grades. In the NCTM Standards, emphasis on measurement fades in the upper grades. The NSF-sponsored high school math curricula that are based on the NCTM standards do not stress the topic of measurement, nor does the CBMS recommendations for the preparation of mathematics teachers. A review of what these document say can be found in Appendix 1. It is important to consider this in fitting the ideas presented here into a curriculum for future high school teachers.

Nonetheless, understanding accuracy and uncertainty becomes increasingly important in science courses as the grades advance, and these topics begin to present significant challenges to students. It is well-documented that undergraduates often have a poor conceptual understanding of measurement uncertainty and that they experience significant difficulties in finding proper interpretations of uncertainty statements and in evaluating the confidence of conclusions drawn from approximate data. For a useful set of references on this topic, see Deardorff (2002) and the references cited in §II.B, below.

The basic themes of accuracy and precision in measurement and calculation take center stage in the method of "significant figures", a prominent topic in high school chemistry. Most introductory chemistry books include a laborious treatment of the rules for determining what figures are significant, and chemistry students put significant efforts into memorizing and practicing them. As we have pointed out, the rules that many books present are not mathematically rigorous and not intended to be so. They are rules of thumb that have exceptions, and users need experience and good judgment to be able to use them in meaningful ways. In an apparent attempt to reduce the rules to a set of regimented procedures, some books unfortunately wind up presenting a system that is subtly flawed and incorrect from a mathematical point of view.

Historically, the basic tools for modern statistics arose in the development of mathematical methods for the combination of observations to obtain precision. The most important contributors were Legendre (1752–1833), Laplace (1749–1827) and Gauss (1777–1855); see Stigler 1986. A treatment of measurement error and uncertainty for future high school teachers that failed to take advantage of this great historical connection or to explore the rich mathematics that lies in this direction would be incomplete. At the present time, we are just beginning to think about how to include some of this material, and we expect to do much work on this.

When one attempts to analyze way that input errors effect the output of a calculation, one very quickly becomes involved in problems that have interesting and often deep connections with basic algebra, calculus and statistics. For example, one proves the continuity of a function by showing that output of any desired accuracy can be obtained from approximate input, provided that the error in the input is sufficiently well-controlled. If the effect of error on the output of a differentiable function of several variables needs to be estimated, partial derivatives are the tool of choice. Partial derivatives appear in formulae for the propagation of errors that have important applications in metrology and statistics; see Taylor and Kuyatt, 1994.

B. Brief review of math-ed research on round-off error.

The mathematics education literature in the past two decades has placed considerable emphasis placed on students' estimation abilities and strategies (Hanson & Hogan, 2000). Generally speaking, the research that has been done on estimation is classified into three categories: numerosity estimation, measurement estimation, and computational estimation (Sowder, 1992), with round-off being a strategy for computation. Generally speaking, students at the elementary and secondary level find computational estimation "surprisingly difficult" and "feel uncomfortable with the estimation process" (Hanson & Hogan, 2000, p. 484). Similar difficulties have been noted at the postsecondary level, with respect to students' abilities to estimate computational products of decimal or whole numbers (LeFevre, Greenham, & Waheed, 1993; Levine, 1982).

In the lower grades, students are taught to round off numbers to some specified degree of accuracy (*e.g.*, the nearest ten, unit, tenth, or hundredth), and they are also taught to use rounding as a strategy for computational estimation. Rounding itself hasn't attracted much interest of researchers. Indeed, the focus of almost all of this previous research on computational estimation has been on students' strategies (*e.g.*, reformulation, translation, and compensation) for estimating answers to computational problems. Within this set of concerns, the problem that we are addressing in this report—namely, how students represent measurement uncertainty and use their representations to keep track of the effects of uncertainties on computations—has not previously been taken up by math education researchers.

C. Additional remarks on pedagogy.

In this section, we describe how students have reacted to certain problems that require analyzing how uncertainties in data influence the certainty of the results of calculations based on those data. Our focus will be on the conceptual opportunities that are afforded by certain problems based on this theme, and the extent to which our lesson study team was able to create situations that led students to explore these opportunities.

Inaccuracies in the data that serves as input for a calculation may affect the output in a number of distinct ways. When output is a linear function of input, output accuracy determined solely by the accuracy of the input. Many very common functions, however, are nonlinear. For such functions, output error depends upon *both* the value *and* the accuracy of the input. The following problem gives a very simple instance of this:

Problem 1. *Suppose a rectangular piece of sheet metal has been measured, yielding estimates $\overline{\ell}$ for the length and $\overline{w}$ for the width. The measurer knows with certainty that $\overline{\ell}$ and $\overline{w}$ are strictly within a quarter of an inch of the true length and width. How accurately can he estimate the true area of the piece of metal?*

I will be making remarks on the variety of ways that problems of this sort can be posed and the many different approaches that people with different mathematical backgrounds and experiences use to deal with the variants. Some of what I will say is based on observing students in a junior-level mathematics class. Discussions that occurred among the members of my lesson study-team provided another source of information. I do not think that these represent a large enough or diverse enough source of information to begin to draw conclusions about the range of responses that might come from the many potential audiences for a lesson or unit on this topic. The class I taught and the team I was on had their idiosyncrasies, as any small groups must. I did not record the team's discussions but have relied on my memories of some key moments, and in portraying them I have taken some literary liberties. At this stage, therefore, creating some imaginative hypotheses may be all we can do.

Let me describe two approaches to Problem 1. Observe that both approaches use the "basic representation" in one form or another.

Approach A. *If $\overline{\ell}$ and $\overline{w}$ are overestimates by full quarter inch, then the true area is $(\overline{\ell}-\frac{1}{4})(\overline{w}-\frac{1}{4})$. If both are underestimates by a full quarter inch, then the true area is $(\overline{\ell}+\frac{1}{4})(\overline{w}+\frac{1}{4})$. Since the measurer knows with certainty that both errors are less than a a quarter inch, he knows with certainty that:*

$$(\overline{\ell}-\frac{1}{4})(\overline{w}-\frac{1}{4}) < \textit{true area} < (\overline{\ell}+\frac{1}{4})(\overline{w}+\frac{1}{4}). \qquad (A.1)$$

Approach A has some features that were common in solutions proposed from undergraduates. In problems like this, they had a tendency to look at what happens when the errors took extreme values. Since the area function $A = \ell w$ $(\ell, w \in (0,\infty))$ is increasing in both variables, this is a reasonable strategy to employ here. It is well to note, however, that there are many calculations in which the maximum output error does not come from the maximum input error. The mathematician in me would expect the discussion included with this approach to pay some attention to this point. The student work discussed in §I.C meets this expectation to some extent, since it acknowledges that checking the extremes is alone no assurance.

Approach B. *Let ℓ and w stand for the true length and true width respectively. Also, let $E_\ell := \overline{\ell} - \ell$ and $E_w := \overline{w} - w$ represent the errors. Then,*

$$\textit{area estimate} = \overline{\ell}\,\overline{w} = (\ell + E_\ell)(w + E_w) = \ell w + \ell E_w + E_\ell w + E_\ell E_w, \qquad (B.1)$$

and

$$\textit{true area} = \ell\, w = (\overline{\ell} - E_\ell)(\overline{w} - E_w) = \overline{\ell}\,\overline{w} - \overline{\ell} E_w - E_\ell \overline{w} + E_\ell E_w. \qquad (B.2)$$

Let $E_A := \overline{\ell}\,\overline{w} - \ell w$ represent the error of the estimated area. Using the assumptions on the size of E_ℓ and E_w and $(B.1)$, we get

$$|E_A| = |\ell E_w + E_\ell w + E_\ell E_w| < \frac{\ell + w}{4} + \frac{1}{16}. \qquad (B.3)$$

Similarly, from $(B.2)$ we get:

$$|E_A| = |\overline{\ell} E_w + E_\ell \overline{w} - E_\ell E_w| < \frac{\overline{\ell} + \overline{w}}{4} + \frac{1}{16}. \qquad (B.4)$$

Approach B as more typical of the way I (the mathematician in our lesson-study group) thought about problems like this. I didn't hesitate to give names—here ℓ w, E_ℓ, E_w and E_A—to quantities that were implicit in the problem.

My willingness to create names may be a significant difference between the habits of mind that I've developed over the years, and the kind of thinking that the undergraduates are more comfortable with. I have observed this difference in numerous contexts. When faced with a routine math problem, I easily discern the distinct quantities that might be isolated and attended to, and I have a good intuitive ability to choose ones that will prove useful. Where I confidently and cheerfully introduce symbols to name variables, less experienced students hesitate. They do not treat algebraic symbolism as a helpful medium for expressing ideas. The most unfortunate go to it as a last resort and find it burdensome.[5]

There is one important point at which Approach A is arguably more informative than Approach B—at least when the two approaches are developed no further than they have been here. $(B.4)$ allows us to deduce that $\overline{\ell}\,\overline{w} - (\frac{\overline{\ell}+\overline{w}}{4} + \frac{1}{16}) < \ell\, w$. $(A.1)$ implies the stronger statement that $\overline{\ell}\,\overline{w} - \frac{\overline{\ell}+\overline{w}}{4} + \frac{1}{16} < \ell\, w$. We trim down interval containing the true area by $\frac{1}{8}$ square inch. Of course, the stronger conclusion is implicit in $(B.2)$. We've lost it by trying to limit the size of the error rather that trying to say what we know about the area in some other way. On the other hand, Approach B produces insights we don't obtain

[5] The comparisons and contrasts between expert and novice behavior outlined by Dreyfus and Dreyfus (1986) seem to capture the differences well.

from A. For example, if we subtract $(B.2)$ from $(B.1)$, we get the following interesting expression of the error in the area:

$$E_A = \frac{\ell + \overline{\ell}}{2} E_w + \frac{w + \overline{w}}{2} E_\ell.$$

III. Lesson reports and student work

A. The decimal problem.

In Summer 2001, our lesson-study team was very much taken by the following problem, which led to a very illuminating discussion.

Problem 1. *If two numbers are known with sufficient accuracy so that 5 digits after the decimal point are known with certainty, but the digits that follow (in reading left to right) are uncertain, then with how much accuracy can the product be determined?*

In many ways, our team's experiences with this problem influenced everything we did subsequently. Perhaps we were hoping somehow to reproduce the learning experience that we ourselves enjoyed quite by accident when we stumbled on this riddle.

In Fall 2001 in a junior-level college number theory class attended by many intending secondary-school math teachers, I attempted a lesson that consisted simply of posing the above problem and then directing students to work together on it in small groups. Most groups failed to understand the problem, and used their time either to puzzle about what it meant, or to solve a problem of their own invention that they believed to be the intended question. The stumbling block lay in not understanding precisely what the uncertainty statement meant. Upon recognizing the students' problems, I discussed and illustrated the ideas of accuracy and uncertainty, and presented the following problem and its solution:

Problem 2. *I know two numbers N_1 and N_2 out to two decimal places:*

$$N_1 \cong 6{,}791.22 \qquad N_2 \cong 2{,}021.01.$$

The decimal digits of N_1 and N_2 beyond those shown are unknown to me. If I attempt to approximate the product N_1N_2 by the product of the decimal expressions, then what digits in the product

$$(6{,}791.22)(2{,}021.01) = 13{,}725{,}123.5322$$

can I trust—that is, which digits in the expression 13,725,123.5322 *must necessarily coincide with the digits in the decimal expression for N_1N_2?*

Note that in this problem, there are two ways of interpreting the uncertainty, one corresponding to rounding, the other corresponding to truncation. If, for example, 2.3 is the result of rounding x to the nearest 10th, then $x \in [2.25, 2.35)$. On the other hand, if 2.3 is the result of truncating a decimal expansion of x, then $x \in [2.3, 2.4]$. I pointed this out and we discussed both of the possible meanings. I demonstrated on the board how to find the interval to which N_1N_2 must necessarily belong under either interpretation of what it means to know the numbers to two decimal places. In doing so, I was intentionally modeling a use of the "basic representation".

After this, I presented a longer and more carefully worded version of Problem 1. (This version is quoted in Example 2.b, below.) As homework, I asked students to write up their understandings of the problems we had considered, and to provide a solution to the first (or its more elaborate rephrasing). Here are some examples of student writing. The first example probably shows the clearest insight, yet it also shows that the student started out with an incorrect opinion.

Example 1.

1. *I know two numbers to 5 places after the decimal point. I multiply them. How many decimal places in the product can I be sure of?*
   Answer: 5 decimal places; after that we cannot be sure of [them] because the errors from the 2 original numbers can influence it.

2. *I am sure of the first two decimal digits:*

$$N_1 \cong 6{,}791.22????$$

$$N_2 \cong 2{,}021.01????$$

so

$$6{,}791.22 \leq N_1 \leq 6{,}791.23$$

$$2{,}021.01 \leq N_2 \leq 2{,}021.02$$

and

$$\underline{13{,}725}{,}123.5322 \leq N_1N_2 \leq \underline{13{,}725}{,}211.6546.$$

Homework—Conclusion

In case of the second question, we have noticed that we can be sure of none of the decimal places of N_1N_2. In fact, only the 5 most significant digits are what we can be sure of. So the answer I gave above to the first question is wrong. The number of correct decimal places of the product depends not only on the number of decimal places we know, but on the size of the integer part of the two original numbers as well. For example, if

$$x = 0.00001\ldots \quad \text{and} \quad y = 0.00001\ldots,$$

then

$$0.0000000001 \leq xy \leq 0.0000000004\ldots,$$

and we know the first 9 decimal places of xy for sure. But if

$$x = 10{,}000.00001\ldots \quad \text{and} \quad y = 10{,}000.00001\ldots,$$

then

$$100{,}000{,}000.2000000001 \leq xy \leq 100{,}000{,}000.4000000004\ldots,$$

and we cannot be sure of any value after the decimal point of xy. The size of the integer parts of the two original numbers can influence the number of correct decimal places of the product in such a way that the larger the two numbers are, the lesser the number of decimal places of the product that we can be sure of. So the correct answer to the first question may be something like: "It depends on the exact values of the two numbers given."

Example 1 is interesting because the student reveals that he rejected his initial answer to the first problem after considering the second. The student applied the representational resources he had acquired in thinking about the second problem to the first.

The next example is amusingly enthusiastic. Yet, unlike the first student, who was able to use a form of the "basic representation" very effectively, the second seems to be unable to bring representational resources effectively to bear. This student's work comes in two parts, one written immediately after the class, and a second piece written during the following week. It's quite interesting that he is able to recapitulate the use of the "basic representation" with the numerical example (Problem 2), but he cannot transfer it to the more general context, as the first student did.

Example 2, part a.

Statement: *I know two numbers to 5 decimal places. I multiply them. How many decimal places in the product can I be sure of?*

Summary:

There initially appeared to be some ambiguity as to the actual number of decimal places of the two numbers to be multiplied. Some students felt that the numbers were only as long as five decimal places, while other maintained that they were only sure of 5 decimal places, although there could be more. The latter case proved to be what was asked for.

There were many techniques used to draw some interesting conclusions. For instance, many of us chose to rely on the concept of significant digits that we were introduced to in high school chemistry. Others of us chose to employ the technique of counting the decimal places of the original multiplied pair of numbers, and using that total to justify the total decimal places of the product.[6] *For example, if two numbers of two decimal places were multiplied with each other, then the product would have four decimal places.*

The best method of illustrating the proper amount of certain decimal places involves using the product of some arbitrary approximations. The approximate method involves the following:

[6] The student is referring to the rule used in the school multiplication algorithm: when multiplying two numbers, take the the sum of the number of digits after the decimal point in each; place the decimal point in the product so that the number of digits to the right of it is that sum.

Let: the numbers

$$N_1 \cong 6{,}791.22$$

$$N_2 \cong 2{,}021.01$$

be the numbers to be multiplied. Since we know the above numbers to two decimal places, we know the absolute least values that the approximate values can be. The upper limit of the above approximations cannot exceed the sum of the last certain decimal place plus 1, because all of the decimal places from the third decimal place to ∞ would be less than the one that we add to the second decimal place. We therefore get the following inequalities:

$$6791.22 \leq N_1 \leq 6791.23,$$

$$2021.01 \leq N_2 \leq 2021.02.$$

Looking at the smallest possible product and the largest possible product (obtained from using the smallest possible approximations and the largest possible approximations, respectively), we obtain:

$$13{,}725{,}123.532 \leq N_1 N_2 \leq 13{,}725{,}211.65.$$

From the above results, I deduce that the first five digits are the only digits that I am certain of, because they are the only digits that match.

Example 2, part b.

Problem: Two quantities have been measured and I have been given two decimal expansions, one for each of the measurements, and I have been assured that in each expression the integer part and the first five digits after the decimal point are correct. I have been warned, however, that all digits after these are uncertain. Then, with what accuracy can I determine the product of the two quantities? In other words, if I multiply the two expressions together, which digits in the product can I trust, and which digits will I be uncertain of due to the uncertainty in the expressions I was given?

Summary of Findings: In class last Thursday, we noted that this latter proposition is significantly longer than the initial one. The introduction of more details was aimed at clearing up some of the ambiguity of the first proposition. I am not, however, convinced of the above propositions's clarity, because the more details added brought more questions and confusion. (So much so, that we spent the rest of our in-class time discussing the myriad possibilities of the questions.) Alas, I am still uncertain, as I expect many of my classmates to explore this question in several of its alternate directions.

One source of uncertainty that we noted relates to the methods used in calculating the given numbers. Were they rounded or truncated? It makes a difference! (Just consult with the banks who may round or truncate depending on which methods work in their favor.)

Note: I assume that this is why scientists keep explicit notes on their experiments and observations because someone may want to reproduce their findings in the future. Instruments used, calibration standards, laboratory temperature, air pressure, humidity, etc.

play an important role in measurements of size and quantity. If a company wants me to do my absolute best, then it should provide me with the best tools possible, including knowledge. I digress.

Another related source of uncertainty is in considering whether or not my calculator rounds or truncates. I could be compounding may inaccuracies unknowingly.

Both rounding and truncating are inaccurate to some extent, because truncating involves "throwing away" some quantity of measure, while rounding may also "throw away" quantities at times, and "add" to calculations at other times. It seems, though, that the cumulative effect of truncation results in the underestimation of measure, because quantities are always being discarded. Furthermore, while it is possible that the same can be said of rounding, the opposite case is true also—rounding may have the cumulative effect of adding or overestimating measurements. In still a third possible case, rounding up and down (alternately) may produce more realistic results (or moderate results), because of a continuous cancelling effect from both overestimating and underestimating. Though, none of these cases are assured.

I conclude that certain conventions or standards must be agreed upon and adhered to in order for our question not to exceed its still cloudy meaning, as well as about 5 more pages of more concise statements.

The next example was produced by a student who had had experience teaching high school. Again, it comes in two parts, one written immediately after the class, and a second part written during the following week. In the first part, the student has not found a way of representing the problem that he felt comfortable with. In the second part, he has recalled some version of the rules for significant digits, and used this as a means for solving the problem and providing some additional insight as well.

Example 3, part a.

This class had a theme of demonstrating how a mathematics problem could be interpreted in different ways, offering different solutions. Dr. Madden posed the question, "I know two different numbers to five decimal places beyond the decimal point. How many decimal places in the product can I be sure of?"

After contemplating the solution, students were asked to give an answer. Most of the solutions for one reason or another indicated that the product could be certain to ten decimal places. This solution took into account both the number of decimal places obtained when multiplying two five-decimal place numbers by hand, or when trying to calculate the area of a square with sides 0.00001 *units by* 0.00001 *units.*

Dr. Madden posed a similar question using two approximated numbers, each having two decimal places of certainty. When the range of values of the two numbers is considered, the number of decimal places known for certain in the product of these two numbers is none. This shows that the interpretation of a mathematical statement can often affect the outcome of the solution.

Example 3, part b.

Problem:

Two quantities have been measured and I have been given two decimal expansions, one for each of the measurements, and I have been assured that in each expression the integer part and the first five digits after the decimal point are correct. I have been warned, however, that all digits after these are uncertain. Then, with what accuracy can I determine the product of the two quantities? In other words, if I multiply the two expressions together, which digits in the product can I trust, and which digits will I be uncertain of due to the uncertainty in the expressions I was given?

Claim:

For the problem as stated, one can only be certain of the first 5 digits beginning from the left of the product. Otherwise, one would have to know how many places are in the integer part of the original measurements in order to be certain of more that 5 digits in the product.

Example 1: If one truncated the measurements A and B in order to obtain a product:

Determine the level of certainty of the product of 6521.01012 . . . *and* 324.02235 . . .*:*

$$6521.01012 \leq A \leq 6521.01013,$$
$$324.02235 \leq B \leq 324.02236.$$

$$\textit{Lower product} = 2112953.02346$$
$$\textit{Upper product} = 2112953.092$$

Note that there are 7 digits of which one can be certain in the product, equal to the number of significant digits in measurement B minus one (the last digit is assumed to be uncertain).

Example 2: If one rounds the measurements A and B in order to obtain a product:

Determine the level of certainty of the product of 6521.01012 . . . *and* 324.02235 . . .*:*

$$6521.010115 \leq A \leq 6521.010125,$$
$$324.022345 \leq B \leq 324.022355.$$

$$\textit{Lower product} = 2112952.98923 \sim 2112953.0$$
$$\textit{Upper product} = 2112953.05768 \sim 2112953.1$$

Note that again, there are 7 digits of which one can be certain in the product, equal to the number of significant digits in measurement B minus one due to its uncertainty.

This scheme of determining the number of digits of which one can be certain will give similar results no matter how many integer digits there are in each measurement. Since the number of integer digits can vary, depending on the measurements that are made, this number is often more important in determining the number of digits of certainty in a product than the number of decimal digits. The number of digits of certainty can be determined by counting the number of certain digits (the last digit in each measurement is assumed to be uncertain) in each measurement to be multiplied, then using the minimum of those numbers of digits, count the number of digits that is certain in the product.

This method will obtain the same results no matter how many integer digits there are, or whether rounding or truncating is chosen as the method to estimate the measurements.

These three examples include enough to "stake out" the conceptual territory that students covered. The remaining student papers bear out the following points:

1. Students had difficulty in absorbing the point of the original problem and in appreciating how uncertainty figured in it.
2. Students were comfortable with the idea of representing uncertainty by membership in an interval, but their ability to use this representation varied.
3. Many students were aware of the significant digits conventions.

B. A lesson involving perimeter and area.

Our lesson-study group designed a lesson entitled "Round-off Error, Perimeter and Area". The intent of the lesson was to induce students to think about the contrasts between the ways that input errors affect linear and non-linear computations. We tried the lesson in a graduate mathematics education course at LSU in June 2002. The seven students in the course were young adults pursuing careers in mathematics or mathematics education. Most students had at least a bachelors degree in mathematics with several also having advanced degrees.

The lesson began with a brief general discussion of measurement and estimation. The major points established were: (1) that measurements are always approximations, and (2) either implicitly or explicitly, a given measurement generally is accompanied by some estimate on the error bound. After the introduction, we divided the students into groups of two or three and presented, in sequence, the four problems below. Students were instructed to do all computations mentally.

Problem. *You are doing some remodeling of your home. You are in a rush to get to the building supply store to purchase building materials—crown molding and carpet—for the small utility room. You have no time to find a measuring tape, so going toe-to-toe, you pace off the dimension of this square room getting a result of $5'$. Assuming your measurement is correct to the nearest foot.*

1. *Estimate the minimum length of crown molding you should buy to ensure you'll have enough to go around the entire room.*
2. *Estimate the minimum number of square feet of carpet you'll need to buy to ensure that you have enough to cover the entire floor.*

Later that day you decide you're going to fix up the master bedroom as well. Again, being in a rush, you decide to just pace off the dimension of this square room and you get a result of $30'$. Assuming, again, that your measurement is correct to the nearest foot,

3. *Estimate the minimum length of crown molding you should buy to ensure you'll have enough to go around the bedroom.*
4. *Estimate the minimum number of square feet of carpet you'll need to buy to ensure that you have enough to cover the floor of the bedroom.*

We expected that if students rushed through these problems, they might underestimate the amount of carpet needed in 4. We hoped that this would occasion some thoughts or discussion through which students might develop a better conceptual understanding. However, the students did not make this error, and the discussion that was generated was not memorable. After this, we asked the students to consider a variant of the decimal problem that was discussed above. They took this home and returned the next day to share some thoughts about it. As a mathematical observer of the lesson, I saw some modest conceptual content in the what was said, but overall, I was not convinced that this lesson was appropriate for this audience.

Our lesson-study group had worked on designing this lesson with a lot of enthusiasm, and by this, the team members seemed to have learned a lot. I think the team hoped to share some of what they had experienced. In this situation, it didn't work out that way. Perhaps the most useful lesson to draw from this is the difficulty of designing a new lesson, and the necessity of testing. What seems good in the design room may not work in the classroom as expected.

C. Perimeter and area in a high school lesson.

We present here a lesson designed and piloted by Ed Lamy, the high school teacher on our lesson-study team. The lesson plan below is a revision of the lesson plan Ed used in a 9th-grade class on April 29, 2002. It has been modified very slightly to incorporate the experience gained from the pilot. The main change was to include the suggestion that the discussion part of the lesson be extended and that additional discussion should follow the assignment. (The lesson has not been piloted in the form give here.) A description of the pilot lesson follows the lesson plan.

The Cost of Inaccuracy
A 9th-Grade Math Lesson

by Ed Lamy

Summary. This lesson introduces students to the idea that measurements are approximate, and has them engage in an activity that shows how imprecisions in data may affect calculations. The activity draws attention to the contrasting ways in which errors in measuring the edges of various squares affect the calculation of perimeter and the calculation of area. Students receive a homework assignment in which they are asked to explain this contrast.

Prerequisites. Students should be able to compute area and perimeter of rectangles, and should be able to give a clear explanation of the formulae or area and perimeter. Students should be accustomed to expressing their mathematical ideas in writing and in diagrams.

Materials needed. For the demonstration in **Part A.** 3), you will need several rulers of different quality, *e.g.*, a cheap foot-ruler marked in inches, a yardstick, a carpenter's tape measure.

Scheduling. A 50-minute class is barely enough time for this lesson. A better strategy might be to use two 50-minute sessions on consecutive days. The first period could be used to get students to the point where they could work meaningfully on the homework assignment. Their writing can then form a starting point for an extended discussions in the second session.

Part A. Discussion. (*10 minutes)*

1) *The teacher may open the lesson by mentioning the following points:*
   - Math allows us to study natural phenomena that occur around us and enables us to make predictions of future occurrences. In this setting, math requires that we use numbers derived from measurements, such as weight, length and time.
2) *This sets the measurement as the main theme. The teacher now sharpens the focus by pointing out that:*
   - All measurements are approximate.

   To make this point more vivid, the class may consider the example of a person's height.

- How tall are you? Using the most accurate ruler available, how tall are you? Is this exact? Can your exact height be determined?

3) *The teacher might enlist a student volunteer to help with the demonstration and measuring his/her height a using successively more precise rulers. In summary:*
   - All measurements are ultimately approximations. The better the equipment, the better the approximation.

   The class should not have much trouble feeling comfortable with this idea. Therefore, this part of the lesson need not take too long. Students will be ready to proceed to part **B** *when they recognize that measurement uncertainty will be a center of attention.*

Part B. Activity. (*20 minutes)*

1) *The teacher explains that the following activity is intended to demonstrate how errors in measurement may introduce errors into calculations. Point out that the situation is over-simplified in order to allow students to focus on the main point.*
   - Consider a home remodeling in which two rooms need to be outfitted with baseboard and carpet. There are no doors in the rooms and both rooms are square. You have the ability to measure the rooms to the nearest foot, but no better. You find the smaller room to be 5 feet by 5 feet and the larger room to be 30 feet by 30 feet. (These are your measurements, not necessarily the true size.)
2) *Ask the students to try to use mental skills to answer the following questions:*
   - How much baseboard should you buy for the smaller room. You want to purchase enough to be sure you can do the job in spite of whatever error your measurement might have, yet no more than is absolutely necessary in order to be sure o having enough. *(22 feet)*
   - How much *carpet* would you need for the smaller room? Same constraints as for baseboard. *(30.25 feet)*
   - How much baseboard would you need for the *larger* room? (Same constraints.)
   - How much carpet would you need for the larger room? (Same constraints.)
3) *Two feet of "extra" baseboard is needed for both the small room and the large room. The sequence of questions might lead some students to believe that the amount of "extra" carpet needed for the large room would be the same as for the small room. Determine if any students have made this assumption. Allow students to think through the questions again, using any resources (paper and pencil, calculators) that they wish.*

Part C. Discussion. (*10 minutes or more)*

Have students discuss the following question:

- The amount of "extra" baseboard that we needed for both rooms was the same, but the amount of "extra" carpet that we needed increased when we went from the small room to the large room. Why?

Part D. Assignment.

- Homework: Write a short essay explaining why the size of the room matters when computing the extra materials needed for area, but not for perimeter. Use a sketch in your explanation.

Ed reported on how the lesson went when he tried it on April 29. (Unfortunately, none of the other members of the lesson-study team were able to be with him.)

> *I spent the first ten to twelve minutes establishing why we study math—to explain and predict the everyday phenomena that take place naturally around us. Mathematics in these settings requires us to use numbers that are measurements of some kind, hence there is imprecision. I asked a volunteer how tall he was. He said five feet eight inches. I drew on the board an expanded representation of a ruler with increments marked at 7, 8, and 9. I showed them that his $5'8''$ might actually be $5'7\frac{3}{4}''$ if the measurement instrument were more accurate. I continued to play with this accuracy issue, helping to lead them to the conclusion that his "exact" height is deceptively hard to pin down, and is dependent on the increments used to measure. I summed with questions like: What dimension is it reasonable to measure the distance from my house to school? How about the distance from Earth to the moon? Etc.*
>
> *I asked them to open their minds to an overly simplified example of the problem inherent in measurements, i.e., there is always imprecision. Then I proceeded to lay out our perimeter-area lesson. The students had cleared their desktops and offered answers to the requested estimates of baseboard length and carpet area using mental arithmetic. I documented their guesses. Many did follow our predicted pattern: 2 extra feet of baseboard needed for both rooms, so* 5.25 *square feet of extra carpet needed for both rooms. One student, confident in his mental multiplication skills, offered a nearly correct estimate for the carpet in the* 30×30 *room. When I asked how he arrived at his answer, he said he did* 30.5×30.5 *in his head. I seized this opportunity to go for the aha! moment. How can this be?! I summarized the pattern:* 2 *extra feet of baseboard needed for the small room,* 5.25 *square feet extra carpet needed for the small room;* 2 *extra feet of baseboard needed for the big room,* 30.25 *extra square feet of carpet needed for the big room. Can this be right?*
>
> *With five minutes left in the period, I had them take out a piece of paper and I asked them to explain why the amount is the same for perimeter, but not for area. I asked them to use a sketch in their explanation. I really wish they could have had more time for the written work, but I think you'll find some of the responses illuminating.*

The students' papers were variable. At least half seemed to indicate a good grasp of the meaning of the question that was posed and an understanding that the contrast between addition and multiplication was an important factor. A couple of papers contained comments that seemed like the beginnings of complete explanations. For example, one student wrote:

Example 1. *Why is there a difference between area but not perimeter? Because with area you multiply* $L \cdot W$ *to get the square footage. To get perimeter you just add the sides together. When getting perimeter all you need is the size of the side & add* .25 [sic]. *To get the area you must multiply the sides* $+.5$ *which* $=$*'s something different for each side.*

It's not possible to determine what this student meant without questioning him, so I won't try to guess. A paper like this, however, could be a very useful tool for a teacher, since what the student has written can—with some small modifications—be made into the beginning of a very good account.

Example 1, as it might be revised by teacher. *Why is there a difference between area but not perimeter? Because with area you multiply* $L \cdot W$ *to get the square footage, while to get perimeter you just add the side lengths together. When finding the largest possible perimeter that could occur if you have a measurement* m*, all you need is the measurement plus the maximum error, which is* $m + .5$*. The actual perimeter must be less than* $4(m + .5) = 4m + 2$*. To get the largest possible area, you must square the measurement plus the maximum error:* $(m + .5)^2$*. The amount by which this exceeds the square of the measurement* m^2 *will be different for different measurements.*

If the student would agree that this is what he meant, then the teacher could ask for a clarification of the last assertion. This leads to a consideration of the equation

$$(m + .5)^2 = m^2 + m + .25.$$

The point to observe is the way the student's own words lead into to some good mathematical talk at a level of precision just a little higher than what the student himself has produced. When a teacher finds a performance that approximates a desired model, it represents a great pedagogical opportunity. It may take some significant mathematical intelligence, however, to be able to spot such opportunities, and see the path to follow in reshaping the student's performance. Here are some other student responses that seem to provide similar opportunities:

Example 2. *If a half of a foot is added to ... [each side], a total of 2 feet will be added [to the perimeter], but if half of a foot is added to the lengths in the area, since the number will be squared, the answer will vary.*

Example 3. *Because if you have a 5-ft. room with nearest foot you have* $\approx$ *.5 ft. to work with. ... You add all sides so if you work with .5 you add* $.5 \times 4$ *to the ... perimeter [of the smaller room]. With area you multiply 2 sides to get it. ...*

Some other papers show students who may possibly be "barking up the wrong tree".

Example 4. *The perimeter is a smaller number and not as big so room for less error. The area is a larger number and contains more space so there is more room for error when calculating and rounding.*

Example 5. *The calculation of area has more error because the per[imeter] measured before has error. If there is error in the first measure (per.), then there will be bigger error in area.*

Example 6. *Difference in area and not perimeter because the area is filled in but the perimeter is at the outside of the area.*

References

- Götz Alefeld and Jürgen Herzberger. (1983). *Introduction to Interval Computations.* Academic Press.
- H. L. Dreyfus & S. E. Dreyfus,(1986). *Mind over machine: the power of human intuition and expertise in the era of the computer.* New York: Free Press.
- Deardorff, Duane L. (2002). Measurement Uncertainty Instructional Resources. `http://www.physics.unc.edu/~deardorf/uncertainty/`
- Hanson, S. A., & Hogan, T. P. (2000). Computational estimation skills of college students. *Journal for Research in Mathematics Education*, 31(4), 483-499.
- Hilton, P., & Pedersen, J. (1986). Approximation as an arithmetic process. In H. L. Schoen, & M. J. Zweng, (Eds.), *Estimation and mental computation. The 1986 Yearbook of the NCTM* (p. 73). Reston, VA: NCTM.
- Kacker, Raghu. (2000) *An Interpretation of the Guide to the Expression of Uncertainty in Measurement.* NIST Special Publication 500-244. National Institute of Standards and Technology. U.S. Government Printing Office. Washington D.C.
- LeFevre, J. A., Greenham, S. L., & Waheed, N. (1993). The development of procedural and conceptual knowledge in computation estimation. *Cognition and Instruction*, 11, 95-132.
- Levine, D. R. (1982). Strategy use and estimation ability of college students. *Journal for Research in Mathematics Education*, 13, 350-359.
- Neumaier, Arnold. (1990). *Interval methods for systems of equations.* Cambridge University Press.
- Sowder, J. (1992). Estimation and number sense. In D. A. Grouws, (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 371-389). New York, NY: Macmillan.
- Sivia, D. S. (1996). *Data Analysis: A Bayesian Tutorial.* Clarendon Press, Oxford.
- Stigler, Stephen M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900.* Harvard University Press.
- Taylor, Barry N. & Kuyatt, Chris E. (1994). *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results.* NIST Technical Note 1297. National Institute of Standards and Technology. U.S. Government Printing Office Washington: 1994.

James J. Madden
Department of Mathematics
Louisiana State University
Baton Rouge, LA 70808-6710
(225) 578-1580
`madden@math.lsu.edu`

APPENDIX I

A. The NCTM Measurement Standard

Measurement is one of the five content standards presented in the NCTM's *Principles and Standards for School Mathematics 2000*. As with all the standards, there is a general statement that appears in Chapter 3 and there are further, grade-specific details supplied in subsequent chapters. The general measurement standard reads as follows:

> *Instructional programs from prekindergarten through grade 12 should enable all students to:*
> - *understand measurable attributes of objects and the units, systems, and processes of measurement;*
> - *apply appropriate techniques, tools, and formulas to determine measurements.*

The discussion of the measurement standard in Chapter 3 includes the following statement:

> *Understanding that all measurements are approximations is a difficult but important concept for students. They should work with this notion in grades 3–5 through activities in which they measure certain objects, compare their measurements with those of the rest of the class, and note that many of the values do not agree. Class discussions of their observations can elicit the ideas of precision and accuracy. Middle-grades students should continue to develop an understanding of measurements as approximations. In high school, students should come to recognize the need to report an appropriate number of significant digits when computing with measurements.*

At the high school level, discussed in Chapter 7, the measurement standard becomes:

> *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.*

The Chapter 7 discussion of the measurement standard includes the following passage:

> *High school students should be able to make reasonable estimates and sensible judgments about the precision and accuracy of the values they report. Teachers can help students understand that measurements of continuous quantities are always approximations. For example, suppose a situation calls for determining the mass of a bar of gold bullion in the shape of a rectangular prism whose length, width, and height are measured as 27.9 centimeters, 10.2 centimeters,*

and 6.4 centimeters, respectively. Knowing that the density is 19,300 kilograms per cubic meter, students might compute the mass as follows:

$$\begin{aligned} \textit{Mass} &= \textit{(density)} \cdot \textit{(volume)} \\ &= \left(19300\,\frac{\mathrm{kg}}{\mathrm{m}^3}\right) \cdot \left(\frac{1}{10^6}\,\frac{\mathrm{m}^3}{\mathrm{cm}^3}\right) \cdot (1821.312\,\mathrm{cm}^3) \\ &= 35.1513216\,\mathrm{kg} \end{aligned}$$

The students need to understand that reporting the mass with this degree of precision would be misleading because it would suggest a degree of accuracy far greater than the actual accuracy of the measurement. Since the lengths of the edges are reported to the nearest tenth of a centimeter, the measurements are precise only to 0.05 centimeter. That is, the edges could actually have measures in the intervals 27.9 ± 0.05, 10.2 ± 0.05, *and* 6.4 ± 0.05. *If students calculate the possible maximum and minimum mass, given these dimensions, they will see that at most one decimal place in accuracy is justified.*

B. The measurement standard in the NSF high school curricula

Robinson and Robinson[7] provides in-depth curriculum summaries for each five of the NSF standards-based secondary curricula:

1. Contemporary Mathematics in Context: A Unified Approach (CPMP)
2. Interactive Mathematics Program (IMP)
3. MATH Connections: A Secondary Mathematics Core Curriculum
4. Integrated Mathematics: A Modeling Approach Using Technology (SIMMS)
5. Mathematics: Modeling Our World (ARISE)

No reference to the measurement standard is made in any of the summaries, and the terms “certainty”, “uncertainty” and “significant figure” are never used. “Confidence intervals” are mentioned only in connection with statistical sampling. Portions of the summaries that are relevant to the topic of measurement uncertainty are reproduces below:

Interactive Mathematics Program (IMP), Probability and statistics, year 1:

Students will be able to predict and compute a mean for a set of data. They will be able to construct frequency bar graphs, including those that group data into intervals along the horizontal axis. They will investigate measurement variation and normal distribution as they pursue their objective to design and conduct experiments that consider the effect of a given variable on an outcome. They will explore the dispersion of data by finding the range of variation for ordinary and rare events in a normal distribution, by comparing sets of data that have the same mean but a different spread, and by examining several methods to measure data spread. They will be able to calculate the standard deviation for a set of data, and compare and arrange data sets in terms of their spread.

[7] E. Robinson and M. Robinson. (1999). Guide to Standards-Based Instructional Materials in Secondary Mathematics, First Draft Edition. `http://www.ithaca.edu/compass`

Interactive Mathematics Program (IMP), Probability and statistics, year 4:

> [Students] *will be able to construct confidence intervals and determine margin of error. They will also understand the terms confidence level and confidence limits. Moreover, they will understand the popular use of the term margin of error. In addition, given some information about confidence intervals (e.g. confidence level and margin of error) they will be able to determine sample size.*

Mathematics: Modeling our World (ARISE), Mathematical reasoning/modeling/logic, course 2:

> [Students] *will discuss the imprecision caused by hand measurement and electronic drawing utilities and understand that even small errors can become significant. They will discuss the value of deductive proof (such as it eliminates imprecision caused by physical measurement) and the "method of generalization" (basically, solving a general problem by utilizing variables rather than a specific problem involving actual numerical measurements.)*

(The paragraph just quoted suggests significant connections to the themes we explored in our lesson-study group.)

C. The CBMS recommendations for mathematics teacher education.

The CBMS recently published a thoughtful set of recommendations on the mathematical education of teachers.[8] The measurement standard is mentioned in this document, but the weight of the discussion concerning it is in the elementary grades and attention fades in the higher grades. In Chapter 7, "The Preparation of Elementary Teachers", there are several places where the approximate nature of measurements is mentioned. In Chapter 8, "The Preparation of Middle Grades Teachers", the following statement is all that is said about the approximate nature of measurements:

> [Teachers of middle school] *must also understand that measurements of continuous quantities are approximate, and that the need for greater or less accuracy influences the choice of the instrument selected for measuring.*

In chapter 9, "The Preparation of Secondary Teachers", the word "measurement" occurs only twice, and never in connection with error or uncertainty.

---

[8] Conference Board of the Mathematical Sciences. (2001). *The Mathematical Education of Teachers.* American Mathematical Society, Mathematical Association of America.

APPENDIX II

Measuring error as a percent ... of what?

Introduction. Error bounds in measurements are often given as percentages. For example, a seller might guarantee a length of rope to be "100 feet plus-or-minus 5%". Off hand, this would seem to guarantee:

i) the length of the rope is between 95 feet and 105 feet.

In many cases, this may be exactly what's meant, but there are other possibilities. The 5% might apply to the rope's true length, and the intended guarantee might be:

ii) 100 feet lies between 95% and 105% of the rope's length.

In some situations it might not be clear which interpretation is intended. For example, if a thermometer is guaranteed to give readings that are accurate to within 5%, then does this mean that the temperature of the environment is between 95% and 105% *of the thermometer reading* or that the reading is between 95% and 105% *of the temperature of the environment*?

The following problems are intended to lead you through a consideration of these issues.

1) I am testing thermometers to determine if they are 5% accurate by examining the reading they give when I hold them in boiling water. I know the water has a temperature of exactly 100° Celsius. Should a thermometer that reads 105.25° be rejected? How does the answer depend on the interpretation of the meaning of "accurate to ±5%"?

2) Are there interpretations of "accurate to ±5%" that are other than those mentioned in the discussion? Let t stand for the true temperature and m stand for the measured temperature. Some possible interpretations include:

a) $(.95)t < m < (1.05)t$,
b) $(.95)m < t < (1.05)t$,
c) $t - (.05)m < m < t + (.05)m$,
d) $m - (.05)t < t < m + (.05)t$,
e) $|m - t| < (.05)m$,
f) $|m - t| < (.05)t$.

3) Statements of accuracy can be viewed as guarantees about the value of a quantity t. Let us replace the 5% appearing in *e*) and *f*) with an arbitrary parameter ϵ. There are two kinds of guarantees about the value of t: 1) t is such that $|m - t| < \epsilon m$, 2) t is such that $|m - t| < \epsilon t$. In each case, the guarantee is has two parameters, m and ϵ that need to be supplied in order for a statement about t alone to result. Another way to talk about this situation is to define two sets:

- $Y(m, \epsilon) := \{ t : |m - t| < \epsilon m \}$
- $Z(m, \epsilon) := \{ t : |m - t| < \epsilon t \}$.

The problem I want to pose to you is open-ended. I ask you to explore the relationships between these sets.

Discussion. Suppose I say that t (which I assume to be positive) is known with an accuracy of $\pm\epsilon$, where $0 < \epsilon < 1$. This leaves open the question of whether the error is bounded by ϵt or by $\epsilon \times$ (the estimate). If m denotes the estimate, then the former condition is

$$|m - t| < \epsilon t,$$

while the latter is

$$|m - t| < \epsilon m.$$

To examine the relationship between these two conditions, note that if $0 < m, t$, then

$$\begin{aligned} |m - t| < \epsilon t \quad &\iff \quad -\epsilon t < m - t < \epsilon t \\ &\iff \quad (1-\epsilon)t < m < (1+\epsilon)t \\ &\iff \quad \frac{m}{(1+\epsilon)} < t < \frac{m}{(1-\epsilon)} \\ &\iff \quad (1-\epsilon)\frac{m}{(1-\epsilon^2)} < t < (1+\epsilon)\frac{m}{(1-\epsilon^2)} \\ &\iff \quad |t - M| < \epsilon M, \quad \text{with} \quad M = \frac{m}{(1-\epsilon^2)}. \end{aligned}$$

In terms of the sets that we introduced in **3)**, we have:

$$Z(m, \epsilon) = Y\left(\frac{m}{(1-\epsilon^2)}, \epsilon\right).$$

To tie this back to the problem we started with, observe that

$$|t - m| < \frac{1}{20}t \quad \iff \quad |t - M| < \frac{1}{20}M,$$

where m and M are related by:

$$400m = 399M.$$

Concluding remark. If we know that an estimate m is in error by no more that a given percentage of the true (but unknown) value t, then we can fashion new estimate M, of which we can assert that it is in error by no more than that *same* percentage of this known (but not necessarily true) value M. The new estimate can be computed from a knowledge of the old estimate and the given percent. This seems paradoxical, since knowing that an error is bounded by a percentage of a number that is not known exactly seems much less informative than knowing that an error is bounded by a percentage of a known estimate.