A Simple Exercise about Awareness and Analysis of Error

Neil Andrew Wells
Department of Geology
Kent State University
Kent, Ohio 44242
e-mail: NWells@kentvm.kent.edu

ABSTRACT

Variability on the order of 1-2 mm between standard 30-cm (12-inch) rulers permits a simple but effective class exercise on the analysis of error, including the distinction between accuracy and precision.

Keywords: Education – undergraduate; miscellaneous and mathematical geology.

Despite its importance, error analysis is not a subject that fills students with enthusiasm. They naturally prefer to focus on certainty rather than on what might be wrong with their work, and they can be demoralized rather than challenged by the message that errors, uncertainties, and complexities lurk everywhere. Exercises about evaluating and managing sources and magnitudes of error are therefore best kept simple. Success is favored when a) the data and the means of collecting it are straightforward and easily understood, b) the data are collected by the students themselves, c) the students are working well within the bounds of their competence and with the expectation of little or no error, and d) the ramifications of the error are obvious.

In the process of brow-beating my geomorphology students into taking pains to measure features on maps as precisely and accurately as possible, I came to realize that standard plastic and non-engraved metal rulers are surprisingly variable, with differences between rulers of 1 or 2 mm over 30 cm distances (Table 1). This variability permits a short and simple in-class exercise on error that nonetheless leaves a strong impression on the students.

In one of my "map math" labs, I ask the students to measure the distance between the farthest corners of two X's typed about 25 cm apart (using the corners avoids difficulties in pinpointing thick centers). I ask them to try to estimate the measurement to the nearest 0.1 mm, although I warn that their measurements are probably only precise to the nearest 1/3 or 1/4 mm. I then ask them to turn the paper top-to-bottom and to measure again, only this time using a magnifying glass to read the ruler, which, I tell them, could result in identical values but could also quite legitimately result in slightly different values (turning the page and using the magnifying glass helps them think that the second measurement may be legitimately "different"). I ask, and write on the board, how many people got the same value both times, how many got differences of 0.1 mm, 0.2 mm, and so on. Most people report differences of 0.2 mm or less.

We calculate the class's average error (usually around 0.1 mm) and discuss reproducibility, and thus what level of precision they can hope to achieve (0.2 mm usually seems attainable). We also briefly discuss whether they think that their second measurements were more accurate, and why the pairs of measurements might differ – besides straightforward mismeasurement and errors in estimation, the thickness of the ruler's tick marks allows some ambiguity in placing and reading the ruler, and both the switch from naked-eye viewing to magnification and changes in eye position may allow some parallax.

Then I ask the students to tell me their second measurements, which I tabulate on the board ("How many people got 24.91 cm? ... 24.92?...," and so on). Measurements typically span about 1.5 mm, or about 0.6 percent of the entire measurement, which is considerably less accurate than the students were led to expect from the relatively high precision. Student surprise leads to a broader discussion of possible sources of inaccuracy. For example, some students may have under-reported their earlier variations because admitting to a variation seems like admitting to a mistake. Old rulers that are chipped or faded and poor-quality rulers with thick or blurred markings can be difficult to read. Rulers whose readings start exactly at their left edge can be in error if the edge is miscut or worn away. Thick and raised-edge rulers can also be difficult to read because of problems of parallax. As far as I can tell, there is no noticeable variation between photocopies produced in a single run on our copier. Overall, however, it seems that the most significant variations are created during manufacture of the rulers.

To learn more about variabilities from manufacturing, I wrote to a major American manufacturer of high-quality rulers. The information below comes from a helpful and detailed reply that I received from a company representative who requested personal and corporate anonymity in this article. It appears that the variability results mostly from variations in ambient temperatures during manufacture. The company representative told me that no plastic ruler can guarantee absolute accuracy and that even etched stainless-steel rulers (as opposed to engraved ones) can be inaccurate to about 1 mm in 12 inches. The representative said that molded plastic rulers may shrink variably during molding because mold temperatures are controlled only within plus or minus 10^{o}F and because the ambient temperature of the production room may additionally vary. Printed and laminated rulers are made by laminating clear plastic over a printed sheet by using a press that heats the sheets of plastic to an almost molten state, and these will also vary with temperature changes. For "scholastic-grade" rulers, the company in question tests all new batches of raw material, inspects a portion of the

Ruler ID	Brand and Style	Type	Country	Difference from standard in mm	Remarks
1	1A	Molded plastic	China	-0.1	same shipment as standard
2	1B	Molded plastic	China	-0.5	#'s 2 & 3 = same shipment
3	1B	Molded plastic	China	-0.5	#'s 2 & 3 = same shipment
4	1C	Molded plastic	UK	0.0	compared over 12 inches
5	1D	Molded plastic	UK	0.0	
6	2E	Molded plastic	China	-0.6	diff. shipment from #'s 7 & 8
7	2E	Molded plastic	China	-0.4	#'s 7 & 8 = same shipment
8	2E	Molded plastic	China	-0.4	#'s 7 & 8 = same shipment
9	3F	Molded plastic	Hong Kong	0.0	
10	4G	Molded plastic	USA	-1.3	
11	5H	Molded plastic	?	-0.2	
12	6J	Molded plastic	USA	+0.7	
13	6K	Laminated plastic	USA	-0.3	years younger than #14
14	6K	Laminated plastic	USA	-1.1	years older than #13
15	6L	Laminated plastic	USA	-0.4	
16	6M	Etched metal	China	0.0	
17	7N	Etched metal	Japan	-0.1	
18	8P	Etched metal	USA	-1.0	
19	9Q	Etched metal	USA	-0.2	
20	10R	Engraved metal	USA	0.0	compared over 12 inches
21	11S	Engraved metal	USA	-0.2	compared over 12 inches
22	12T	Wood	China	-0.3	#'s 20 & 21 = same shipment
23	12T	Wood	China	-0.5	#'s 20 & 21 = same shipment

Notes:

My standard ruler was equivalent to 29.997 cm on 15-cm and 1-ft Mitutoyo dial calipers, although differences in remeasuring 10 rulers with calipers ranged up to 0.4 mm, due to parallax problems and misplacement.

The mould for 2E is almost identical to the one for 1A but is more worn.

Rulers 13 and 14 were measured over 29 cm only.

Rulers 17, 20 and 21 had some 1/64" markings.

* To test reproducibility, I remeasured 12 rulers the next day and 14 another day later: relative to the first set, 11 remeasurements matched exactly, 10 differed by 0.1 mm, and 5 differed by 0.2 mm, so average reproducibility = 0.08 mm and was never worse than 0.2 mm. Differences apparently relate to problems of parallax and estimation.

Table 1. Differences between 30-cm lengths on several rulers using one ruler of Brand 1, type A, as a convenient but arbitrary standard. Readings are estimates, but are reproducible to 0.2 mm or less. (*See note directly above.)

output against steel-engraved standard rulers, and keeps the manufacturing process within fairly tight tolerances, thereby achieving an accuracy of plus or minus 1/32 inch (.79 mm) over 12 inches, equivalent to plus or minus 0.26 percent. (Of course, when comparing two rulers, one adds the absolute values of the uncertainties on both rulers, which returns us to possible differences of up to 1.5 mm or so over 30 cm.) The company's quality control provides a textbook example of determination and control of sources of variation.

Lastly, I ask the students to calculate what a specific uncertainty amounts to in terms of meters on the ground, for example, if they were measuring a 20 cm distance on a 1:125,000 map with a 0.5 percent uncertainty.

I have run this exercise four times. Initially, I asked students to exchange rulers or match them up in order to see the variability directly, but this can focus their attention on specific similarities rather than maximal differences. Small groups of less than a dozen students may demonstrate unusual distributions of variations, so it might be helpful to have on hand results from a larger class or previous classes, although using other students' results is less convincing. We sometimes talk about strategies for avoiding error, such as always using the same ruler when comparing measurements, starting 1 cm along the ruler if its end is worn, and so forth. We have also talked about why the thicknesses of the markings on the ruler and the lines or points on the map and the student's ability to estimate tenths of millimeters create proportionally more error (percent error, or less accuracy) when measuring over short distances rather than longer ones. I do not recommend particular brands or models because variation between production runs seems to swamp variations between models, and I would have to analyze patterns of variations in many hundreds of rulers to address relative quality fairly. I also side-step the question of "what is the exact correct answer?": caliper measurements reported under Table 1 suggest that my arbitrary "standard ruler" is insignificantly incorrect, but these measurements entail other uncertainties and inaccuracies. I prefer instead to focus on assessing acheivable levels of precision and accuracy, what can be done when accuracy is incompletely known, and how to set up work to minimize error.

Debriefing 32 students from the last exercise, by asking them to restate what they learned and comment on the exercise, produced 11 no comments, 17 students who were enthusiastic and impressed at the variation (4 of whom, however, didn't really get the point), 3 negative comments (knew it, didn't get it, or bored by it), and one who thought it interesting but without practical significance.

In this light, a suggestion by a reviewer could lead students to think more about how error works by asking them to construct or work out a worst-case scenario, where errors compound as quickly and calamitously as possible. Imagine, for example, that two students are planning a railroad route over flat terrain, using rectified aerial photos and with partial coverage by a photocopied map, and they need to calculate the length for a bridge over a river that is not covered by the map. They find two houses on the photo and the map. They measure the distance between the houses on the map as 12.4 mm, on the photo as 5.6 mm, the width of the river as 1.0 mm, and the length of 1 km on the map's scale bar as 40.8 mm, although the respective measurements should have been 12.6 mm, 5.4 mm, 1.2 mm, and 40.6 mm. To keep things simple, all the errors are 0.2 mm and are attributed entirely to errors in reading the ruler at that level of precision. Also, two errors are overestimates and two are underestimates. If we could expect to measure a 20 cm interval on the aerial photos with an error of 0.2 mm, we might anticipate an error of 1 percent, equivalent to only 1 m per km. In the students' work, calculations for the length of the bridge should be set up as follows:

$$1.2\ mm\ \textit{river width on photo}\quad \frac{12.6\ mm\ \textit{between houses on map}}{5.4\ mm\ \textit{between houses on photo}}\quad \frac{1000\ m\ \textit{on ground}}{40.6\ mm\ \textit{on map}}\quad 69.0\ m\ \textit{river width}\,.$$

However the students calculated:

$$1.2\ mm\ \textit{river width on photo}\quad \frac{12.4\ mm\ \textit{between houses on map}}{5.6\ mm\ \textit{between houses on photo}}\quad \frac{1000\ m\ \textit{on ground}}{40.8\ mm\ \textit{on map}}\quad 54.3\ m\ \textit{river width}\,.$$

Clearly, their planned bridge is about one-fifth too short. Said less fairly, although no error amounted to more than 0.2 mm, their total error ballooned to the equivalent of 213 m per km. They evidently needed larger maps and aerial photos and longer reference distances, in order to reduce their percentage error.

Conclusion

Overall, the proposed exercise in error analysis offers an opportunity for instructing the students that some error is inevitable, even in seemingly simple measurements, and that error is best handled by expecting it, quantifying it, and determining its significance, its sources, and its patterns or behavior. Then they can decide whether the probable error is tolerable or avoidable, or whether they can or should improve the precision and accuracy of their measurements.

Acknowledgments

I am very grateful for information on the manufacture of rulers provided by a company representative who wishes to be anonymous. I also thank my colleagues Rachael Craig, Don Palmer, and Dan Holm for helpful comments on the manuscript, Tim Miller and Ute Dymon for loans of metal rulers, Rod Feldmann for dial calipers, and Physics technical staff Dean Burkhart and Josh Brechtelsbauer for use of engraved metal rulers and another pair of calipers.

About the Author

Neil Wells is an associate professor of geology at Kent State University, with degrees from Albion College, Ohio State University, and the University of Michigan. Lately he has been working on landscapes in Madagascar and small alluvial fans in NE Utah.

Food for Thought

Throughout most of the twentieth century, public school educators have pressed – usually successfully – for the inclusion of ever more non-academic materials in the curriculum, while the counter-pressure for more academic rigor, "back to basics," and the like, has come primarily from laymen. As laymen have urged more emphasis on teaching mathematics, science, languages, and other traditional academic subjects, educators have promoted such personal concerns as nutrition, hygiene, and "life adjustment" in an earlier period, or sex education and death education more recently, along with such social crusades as environmentalism and the anti-nuclear movement, or such exotic topics as the occult. While the particular subjects that are fashionable change over time, what has been enduring is the non-academic thrust of the professional educators.

Thomas Sowell, 1993, Inside American Education – The decline, the deception, the dogmas: Toronto, The Free Press, 368 p. (from p. 32).

Annual statement of ownership, management and circulation (required by 39 U.S.C. 3685)

1) Title of publication: Journal of Geoscience Education
2) Date of filing: 1 October 1996
3) Frequency of issue: January, March, May, September, November
4) Location of office of publication: Robert Christman, c/o Western Washington University, 516 High Street – Geology Department, Bellingham, Washington 98225-9080
5) Location of business office of publisher: P.O. Box 5443, Bellingham, Washington 98227-5433
6) Publisher: National Association of Geoscience Teachers, Inc.
7) Owner: National Association of Geoscience Teachers, Inc., Editor: James H. Shea, University of Wisconsin-Parkside, Box No. 2000, Kenosha, WI 53141
8) Known bondholders, etc.: none
9) Extent and nature of circulation:

		A	B
a.	Total number copies:	2800	2800
b.	Paid circulation		
	1. Through dealers, etc.:	0	0
	2. Mail subscriptions:	2498	2665
c.	Total paid circulation:	2498	2665
d.	Free distribution by mail:	4	4
e.	Free distributon outside of mail:	0	0
f.	Total free distribution:	4	4
g.	Total distribution:	2502	2669
h.	Copies not distributed		
	1. Office use, left over, etc.:	298	131
	2. Return from news agents:	0	0
i.	Total:	2800	2800
	Percent paid and/or requested circulation (9c x 100)/ 9g	99.8%	99.8%

Column A: Average no. of copies each issue during preceding 12 months. Column B: Actual no. of copies of single issue published nearest to12 monthsI certify that the statements made by me above are correct and complete.

(Signed) Robert A. Christman, Executive Director

TABLE 1. Differences between 30-cm lengths on several rulers using one ruler of Brand 1, type A, as a convenient but arbitrary standard Readings are estimates, but are reproducible to 0.2 mm or less*.

Notes:
* I remeasured 12 rulers the next day, & 14 another day later:-
relative to the first set, 11 remeasurements matched exactly,
10 differed by 0.1 mm, and 5 differed by 0.2 mm,
so average reproducibility = 0.08 mm & was never worse than 0.2 mm.
Differences apparently relate to problems of parallax and estimation.
My standard ruler was equivalent to 29.997 cm on 15-cm & 1-ft Mitutoyo
dial calipers, although differences in remeasuring 10 rulers with calipers
ranged up to 0.4 mm, due to parallax problems & misplacement.
The mould for 2E is almost identical to the one for 1A, but is more worn.
Rulers 13 & 14 were measured over 29 cm only.
Rulers 17, 20 & 21 had some 1/64" markings.

ID Brand Type Country Difference Remarks
& Style in mm
1 1AMoulded plasticChina-0.1same shipment as standard
2 1BMoulded plasticChina -0.5 #'s 2 & 3 = same shipment

Annual statement of ownership, management and circulation (required by 39 U.S.C. 3685)

Title of publication: Journal of Geoscience Education

Date of filing: 1 October 1996

Frequency of issue: January, March, May, September, November

Location of office of publication: Robert Christman, c/o Western Washington University, 516 High Street – Geology Department, Bellingham, Washington 98225-9080

Location of business office of publisher: P.O. Box 5443, Bellingham, Washington 98227-5433

Publisher: National Association of Geoscience Teachers, Inc.

Owner: National Association of Geoscience Teachers, Inc., Editor: James H. Shea, University of Wisconsin-Parkside, Box No. 2000, Kenosha, WI 53141

Known bondholders, etc.: none

Extent and nature of circulation:

Average no. of copies each issue during preceding 12 months

a. Total number copies: 2800 2800

3 1BMoulded plastic China-0.5 #'s 2 & 3 = same shipment

4 1CMoulded plasticUK 0.0compared over 12 inches

5 1D Moulded plastic UK 0.0

6 2E Moulded plastic China-0.6 diff. shipment from #'s 7 & 8

7 2E Moulded plastic China -0.4 #'s 7 & 8 = same shipment

About the Author

Neil Wells is an associate professor of geology at Kent State University, with degrees from Albion College, Ohio State University, and the University of Michigan. Lately he has been working on landscapes in Madagascar and small alluvial fans in NE Utah.

8 2E Moulded plastic China -0.4#'s 7 & 8 = same shipment

9 3F Moulded plastic HongKong 0.0

10 4GMoulded plasticUSA -1.3

11 5H Moulded plastic ? -0.2

12 6J Moulded plastic USA +0.7

13 6K Laminated plasticUSA -0.3 years younger than #14

14 6K Laminated plasticUSA -1.1 years older than #13

15 6LLaminated plasticUSA -0.4

16 6MEtched metal China 0.0

17 7NEtched metal Japan -0.1

18 8PEtched metal USA -1.0

19 9QEtched metal USA -0.2

2010REngraved metal USA 0.0compared over 12 inches

2111SEngraved metal USA-0.2compared over 12 inches

22 12TWood China-0.3#'s 20 & 21 = same shipment

23 12TWood China-0.5#'s 20 & 21 = same shipment