Importance of Data in Teacher Teaching: Two Case Studies

Somnath Sarkar and Richard Frazier

University of Central Missouri, Missouri, USA

sarkar@ucmo.edu, frazier@ucmo.edu

We describe two case studies where science teachers are engaged in treating data gathered from scientific investigations and making inferences. The nature of the treatment of data raises some concerns. We analyze the teacher's treatment of anomalous data and make some recommendations to the faculty who are engaged in science teacher preparation.

Keywords: Anomalous data, Data analysis, Pattern recognition, Scientific investigation, Inquiry

Introduction

The National Science Education Standards emphasizes that "inquiry into authentic questions generated from student experience is the central strategy for teaching science" (NRC, 1996). Investigating such questions involves gathering both quantitative and qualitative data and interpreting them for making reasonable inferences which may lead to additional experiments. The purpose of this paper is to examine two cases where teachers and their students perform experiments, generate and interpret data. Our analysis of these two cases identifies some critical issues that lead us to make certain recommendations on what must be addressed when science teachers are prepared to acquire, analyze and interpret data specifically anomalous data (Chin & Brewer, 1993; Lin, 2007; Nott & Smith, 1995).

We have been involved in a 3 year teacher professional development program to promote scientific inquiry in the middle school science curriculum. These two examples stems from the project. The first case comes from the work presented by a teacher during one of our follow-up meetings and involves classroom data. The second case represents work by a group of four teachers during our summer professional development training.

Case I

Mr. Smith was teaching his middle school students about the conservation of mass. He instructed them to place an approximate amount of baking soda in a balloon and to pour an approximate amount of vinegar in a flask. The students determined the mass of the balloon and the flask containing the vinegar on a triple beam balance. Students placed the balloon on the top of the flask tightly and then the baking soda was dispersed into the flask. As the vinegar and baking soda reacted, the balloon inflated. Students weighed the entire system again as soon as the reaction appeared to be over (end of fizzing) and the balloon stopped inflating any further. Following table shows representative student data form one class.

Initial mass of the system in grams before reaction	Final Mass of the system in grams after reaction	Initial mass of the system in grams before reaction	Final Mass of the system in grams after reaction
224.9	222.7	216.3	214.2
209.0	206.4	350.4	346.7
231.7	229.0	243.7	240.7

Table 1: Student data

The students wrote, "The average difference is about 2 grams. We think that the difference is caused by small holes in the balloon. Even though we can't see them, they seem to let gases through over time. If we had a better way to capture the carbon dioxide, it might be the same mass before and after. Would a different kind of balloon like Mylar will work better?"

This experiment was conducted in four sections of sixth grade and similar data was obtained. Since the data did not verify the conservation of mass, Mr. Smith and his students explained the loss of mass in terms of leakage of gas from the balloon and experimental errors. No further exploration of the loss of mass was undertaken but Mr. Smith provided his classes with a different reaction and a different container to illustrate the conservation of mass. In this next experiment, the students added iron powder to a syringe, pushed the plunger all the way down and then withdrew oxygen gas from a balloon that had been filled from an oxygen gas cylinder. The syringe was then capped, weighed, set aside for thirty minutes for iron powder to react with oxygen, and the students weighed the system. The following table shows representative student data form one class.

Initial mass of the system in grams before reaction	Final Mass of the system in grams after reaction	Initial mass of the system in grams before reaction	Final Mass of the system in grams after reaction
11.1	11.2	11.0	11.6
13.1	13.0	11.3	11.2
10.6	10.9	10.0	11.4
10.8	10.1	10.8	11.1
11.0	10.9	10.8	10.8
10.8	11.0	Average11.1	11.2

Table 2: Representative classroom data from the conservation of mass experiment

This experiment was conducted in four sections of sixth grade and similar data was obtained by students. In all cases, the average starting mass and the average ending mass was calculated. In two of these sections, the starting average and the ending average was the same, in one class it was 0.1 g higher in the ending mass, and in the fourth class it was 0.3 g higher. The classes concluded that the average starting mass and the average ending mass were close enough together to support the law of conservation of mass. The variation that occurred was explained by mistakes in measurement and procedure.

Our response to the data: The reaction of baking soda and vinegar

We decided to examine the data from the first experiment in more detail since explanations for this unexpected outcome were proposed but they were never tested. The student data pointed to a loss of mass in all experiments. The loss of mass was systematic but the variation in mass loss was random because the procedure followed by the students did not include any specific amounts of starting materials. Therefore, from the experimental data, the loss of mass could not be correlated with any other variables. The student data that did not illustrate the conservation of mass led us to examine the procedure in order to find the origin of the systematic loss.

In our investigation, five samples of baking soda were weighed in paper cups in 2 g increments up to 10 g, and each one was transferred into a 9-inch balloon. For each sample of baking soda, 150 ml of white vinegar (5%) was poured into a 250-ml Erlenmeyer flask, and the balloon was fixed to the mouth of the flask without mixing the reactants. The initial mass of the system was measured on a triple beam balance. The baking soda from the balloon was dispensed into the vinegar allowing the reaction to take place. While the entire system remained on the balance, the balloon inflated and the scale progressively shifted suggesting that the system was getting lighter. After the reaction subsided the final mass was obtained. The circumferences of the inflated balloons were determined by using a cloth tape measure stretched consistently around each balloon assuming that the shape is spherical. All volumes were calculated using mathematical formulae for a sphere. In order to test the teacher's explanation that the loss of mass was due to leaks, each system was set aside and weighed again after 45 minutes. No change in the mass of the system was observed during this time. We concluded that that leakage did not contribute to the observed loss of apparent mass. The following table shows our experimental data.

Mass of baking soda in *grams*	Mass of system in *grams* before reaction	Mass of system in *grams* after reaction	Apparent loss of mass in *grams*	Balloon circumference after inflation in cm	Calculated volumes in cm^3
2	275.90	275.25	0.65	28.0	371
4	272.07	271.00	1.07	37.4	884
6	248.41	246.49	1.92	42.2	1270
8	277.40	275.00	2.4	46.1	1656
10	263.31	260.27	3.04	49.8	2087

Table 3: Data from the conservation of mass experiment

This experimental design helped us correlate the loss of mass with the amount of baking soda used and the mass of loss with the volume of the balloons. The correlation shows that this is a systematic error and the apparent loss of mass is because of buoyant effect.

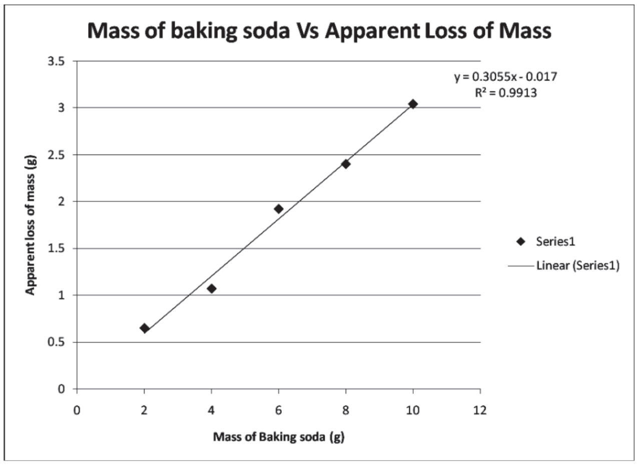

Figure 1: Graph of mass of baking soda & apparent mass loss

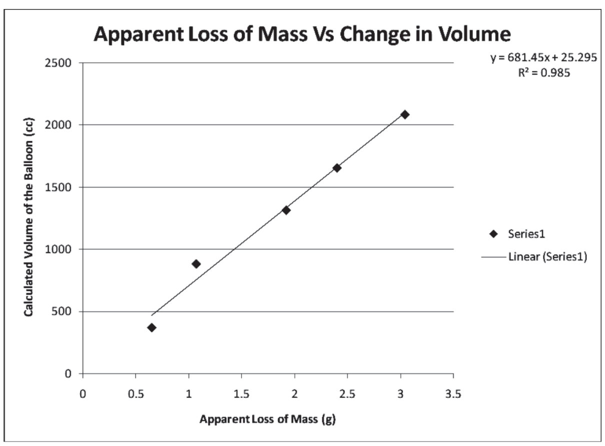

Figure 2: Graph of apparent mass loss with the change in volume

Second experiment involving the reaction of iron and oxygen in a syringe

In the second experiment, two classes observed no loss of mass on the average, and two classes observed the beginning and the final average mass to be close enough to explain the data in terms of experimental error. Even though it was realized by one of the classes that there might not be a sufficient number of data points to make a reasonable conclusion from the average, the data from all the classes were not aggregated. We decided to aggregate the data by carrying out a frequency distribution from all the experiments. Our histogram from the frequency distribution of the mass difference between "after the reaction" and "before the reaction" is shown below.

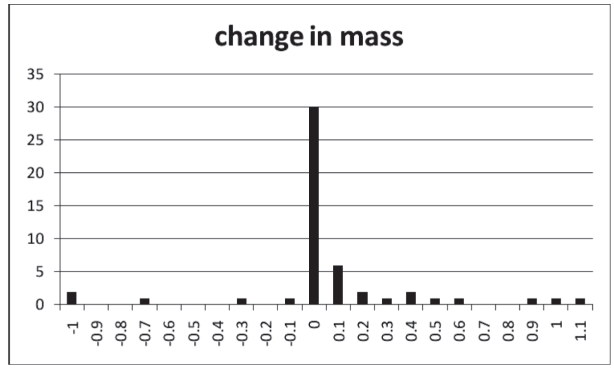

Figure 3: Histogram showing the difference in mass for each experiment

The histogram clearly demonstrates that 30 students out of 51 students did not observe any change in mass. A total of 37 students were within 0.1-0 range which is within the measurement limit of a triple beam balance. Also the histogram shows the outliers. Average of these mass differences show a value of 0.05 g which is within the limit of the experimental error. We did not perform any other statistics such as a t-test on this data because it is beyond the level of middle school science but such a graphical representation clearly delineates the validity of the law of conservation of mass.

In neither experiment the data from all four classes were examined in adequate detail to discern the overall distribution. Only averages before and after the reactions were used to conclude that leaks had occurred in the first and that mass had been conserved in the second experiment. More important than the average loss of mass in the first experiment is the fact that all results indicated a loss of mass. This suggests a systematic effect, especially when the results challenge a well-established law. Leakage of the balloon could produce an observed change in mass and can be a systematic error but the assumed origin of the systematic error was not tested. In accepted scientific practice surprising results call for a thorough analysis of variations and check for systematic and random errors. Mr. Smith explained to us that he assumed that leakage was the problem in the first experiment. He considered switching to a Mylar balloon assuming that no leakage will take place because it is made of a different material. He did not follow through this experiment which would have similar apparent mass loss due to buoyant effect. Instead he used a syringe as a device to house a reaction between oxygen gas and iron filings. By changing the size and the material of the reaction vessel, he inadvertently rendered buoyancy a negligible part of the mass measurement. Mr. Smith saw the use of the balloon as a device to safely collect the gas. When the data did not conform to his expectation, he inferred that the leakage of the balloon was the source of the problem. He did not test whether the balloon leaked before drawing the conclusion. The leakage within the time frame of the experiment could have been easily ruled out by weighing the system at the end of the experiment. We wonder what would have happened if Mr. Smith had noticed that leakage was not the problem. Would that observation lead him to question his operational definition of mass as an absolute value that can be determined on a balance?

CASE II

During our professional development workshop on conducting field-based inquiry, teachers were assigned the task of developing an investigation where temperature was one of the variables. This activity followed an initial survey of a study area where a number of environmental parameters had been measured. Teachers were directed to formulate their own investigative procedures. One group was interested in demonstrating that temperature varies with altitude. They visited the study site, took measurements, and prepared a report for their peers in the workshop. The direct quotes from the teacher's report are presented below in italicized format. Our comment follows.

"Our goal for this study was to see if we could determine a variation in temperature of soil with differences in elevation similar to one sees climbing to altitudes on mountain ranges".

The statement is a reasonable starting point for scientific investigation. The phrasing of the statement, however, suggests an illustration rather than an inquiry. We wonder if the teachers thought of their actions as a search for a "teaching activity" where a relationship between elevation and temperature could be clearly demonstrated rather than an investigation of that relationship as it applied to the particular time and place.

"Materials needed: A GPS tracking device to assure we stayed on a straight line approach for our samples, and to identify altitude, a Compost thermometer to assess soil temperatures at depths of 1" (surface) and 6" (sub-surface), an air temperature indicator to compare with air temperatures. Our procedure was to select eight sites, beginning at the 850 feet elevation of the former hotel and work downward to approximately 800 feet, comparing the temperature differences of the soil in sunlit and shaded areas, with that of the surrounding air".

The procedure described here can be accomplished. It uses a tracking device to measure altitude and a thermometer to measure soil temperature at certain intervals. But the details of the procedure are not linked well to the question. For example, the reasons for choosing 1inch and 6 inch soil depth for temperature data collection are not explained. The question does not mention any relationship to air temperature but air temperature is being collected.

"It looked like our experiment was going to be a success, at least as far as gathering data".

This statement was made after completing the measurements at the first site. We think the teachers are referring to the fact that the equipment they used was adequate to obtain data. Framing the sentence in terms of success supports the possibility that the teachers' view of the activity is directed toward illustrating a point rather than investigating the relationship between the elevation and the soil temperature at the study site.

"As it was, the data we gathered held no true relationship to each other, or to actual conditions, and when we tried to graph the results, instead of getting what we expected: a nice graph curve showing the temperature increasing as the altitude decreased, we got such a wide variety of figures that graphing proved to be of little or no value. A textbook case of "scientists"proving a point by not being able to prove their point".

The following table shows the data collected during the study.

Altitude in feet	**Sunlight readings soil**		**Shade readings soil**		**Air temperature**
	1" depth	**6" depth**	**1" depth**	**6" depth**	**93.2⁰F**
850.5	99⁰F	92⁰F	No shaded reading		
844	86⁰F	78⁰F	79⁰F	76⁰F	94.3⁰F
841.9	82⁰F	75⁰F	76⁰F	74⁰F	92.1⁰F
839	82⁰F	76⁰F	75⁰F	74⁰F	101⁰F
838	87⁰F	78⁰F	76⁰F	73⁰F	96.3⁰F
805	79⁰F	76⁰F	78⁰F	78⁰F	94.3⁰F
778	84⁰F	77⁰F	78⁰F	76⁰F	100.3⁰F
769	91⁰F	77⁰F	82⁰F	76⁰F	100⁰F

Table 4: Teacher collected data at the study site

The questions arises what does the teacher mean by "*the data we gathered held no true relationship to each other, or to actual conditions*." Our interpretation of "true relationship" is that the temperatures did not relate to elevations in the manner they wanted to illustrate. We think "or to actual conditions" means that since the data did not provide the expected relationship, the particular site and time were not suitable for proving the intended point. When the teachers write, *"that graphing proved to be of little or no value,"* we take their meaning to be that the graph they obtained could not be used to demonstrate the expected relationship between elevation and temperature. The graph matched the actual conditions and therefore "a needed re-examination of the original expectation" was not the view shared by these teachers. The comment that the graph has "no value" is problematic. As long as the measurements were made carefully and the graph constructed correctly, the graph shows what it shows and, therefore, has scientific value. This mismatch between how the teachers conceived of graphs and why graphs are constructed relates to the important questions about the purposes of inquiry and investigation in school science. Graphing here is being considered as an illustration and not a tool. The graph shows a scatter plot and indicates that the variation of temperature with altitude is random within this small range.

The procedure yields data that suggests some patterns among the temperature of the soil at shaded and sunny sites, and air but not with respect to changes in altitude. The teachers, however, do not make a data-driven conclusion that within the range of elevation and prevailing weather conditions that a relationship between temperature and altitude cannot be determined even though such a conclusion addresses their initial goal. The fact that a testable question and reasonable data are present but from the data the question is not answered points to the importance of teacher "beliefs" about experiments and their purpose in school science.

Recommendation to the Teachers Involved in Science Teacher Training

The problem involves a level of details we use in our language in science teaching. If we include all the clarifier, the discourse becomes cumbersome and confusing. If we rely on shorthand where various students have different levels of understanding, the shorthand can mislead. Should we not then train science teachers to gain a mind set to question whether they have considered all related variables and their magnitude effect on the experiment? We will never be able to take them through all possible aspects or scenarios, but we must provide the opportunity where teachers are trained in situations where the data may surprise them and ask them to take necessary action. In science teacher preparation and development, teachers should be involved in discussion on how to understand the nature of the anomaly. (Language reason, magnitude effect, generalized laws and theories, variables and details, systematic and random error, precision of measurements), and how to respond to these circumstances. Following are some key suggestions for science teacher educators.

- Present cases to teachers that involve anomalous data from authentic inquiry and discuss the process of analysis including pattern recognition, types of possible errors, magnitude effect, variables, and strategies to take action.
- Practice trial and adjust instead of trial and error. Teach error analysis
- Teach appropriate use of demonstration and verification labs and how those relate to opportunities for genuine inquiry
- Emphasize that data should drive experiment. Teach various tools of data analysis
- Train to pay attention to student results and ideas and act upon them
- Design experiments that include refining procedure and perform them to solve a problem
- Promote clear communication and good writing in science

References

Chin, C.A., & Brewer W.F. (1993). Factors that influence respond to anomalous data. Proceedings of the *Fifteenth Annual Conference of the Cognitive Science Society,* 318-323. Hillsdale, NJ: Erlbaum.

Lin, Jer-Yann. (2007). Responses to anomalous data obtained from repeatable experiments in the laboratory. *Journal of Research in Science Teaching*, *44*(3), 506-528.

National Research Council. (1996). *National science education standards*. Washington DC: National Academy Press.

Nott, M., & Smith, R. (1995). "Talking your way out of it", "rigging" and "conjuring": What science teachers do when practicals go wrong. *International Journal of Science Education, 17*(3), 399-410.