

Comparing Student Ability to Reason with Multiple Variables for Graphed and Non-Graphed Information.

Rebecca Rosenblatt

Department of Physics, Illinois State University, Normal, IL 61790

Abstract. I present results from a two-year study classifying student difficulties reasoning with graphed data. I show large differences in student ability to use certain graphed data. Namely, students struggle more with graphs in which the variables have no relation or an unknown relationship than with typical graphs where the data shows a relationship. I expand on this finding and present evidence for three deep issues with graphical/data-driven reasoning. First, many students incorrectly assume there must be dependence between the axes of any graph whether or not the data suggests a relation and whether or not it was a controlled experiment. Second, students have issues using a legend to infer information about a third variable. Third, by comparing pictorial reasoning responses to graphical ones, it is evident that students have deeper logical reasoning issues such as, “if x doesn’t change and y does, than x doesn’t affect y.”

Keywords: Student difficulties, graphs, control of variables, reasoning, physics education research
PACS: 01.40.Fk

INTRODUCTION

Student difficulties reasoning with multiple variables and reasoning with control-of-variables are both well-documented topics in physics and the sciences [1, 2, 3, 4]. In addition, student issues with graphs have been explored in several ways [5, 6]. However, little research has been done combining these problem areas to observe student reasoning with control-of-variables and data. And, there is even less when looking at graphed data. Here I will present several findings in this area from an ongoing project exploring a small part of this section of student knowledge and reasoning. These findings indicate that students are good at interpreting the main trends on a graph only if the question asked is: a) solely about the two axis, b) the graph comes from a controlled experiment, and c) the graph has information that shows a relationship. These three things tend to be true of commonly-seen graphs but are not generically true for science graphs. Student difficulties with graphs that do not fit these three categories show a disturbing deficiency in their reasoning sophistication and the science education being provided.

MAIN RESEARCH QUESTION

This article examines student ability to correctly answer questions regarding a set of graphed data and a small experimental setup describing and constraining the data collection. The items required students to determine whether the graphed information was consistent with a variable either a) affecting, b) not affecting, or c) having an unknown affect on another variable. Figure 1 illustrates two common situations where students use this reasoning. It also shows four of the

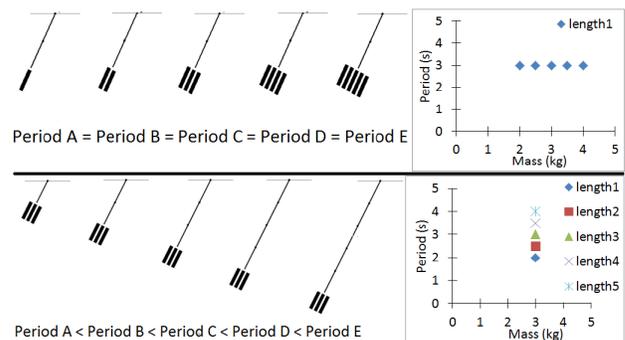


FIGURE 1. Pictures on left shows a control-of-variables example. Graphs on the right show the “same” data. The top situation indicates mass does not affect the period but does not say how length affects the period. The bottom indicates length affects the period but does not say how mass affects the period. This reasoning proved difficult for students in both picture and graph form. (Students did not see graphs and pictures shown together.)

items used in Experiment 3, two graphs and two pictures.

Presented here are three different experiments (5 different conditions) which used randomized groups to study student reasoning processes and the significance of certain features of the graphs. This was done by comparing the effects of changes to the graphs, the tasks, and the item wordings on student ability to correctly answer the questions.

The data presented was collected over three semesters during the last weeks of the algebra-based mechanics course at Illinois State University. For a majority of these students this is a required course as part of their biology or geology degree. Experiment 1 and 2 data was collected via paper quizzes given at the beginning of lab and took the students 20-30 minutes. Experiment 3 data was collected in separate

sessions that students completed on computers in the education research space during proctored times and took 25-40 minutes. For each experiment, students were given a participation grade and appeared to take the questions seriously.

EXPERIMENT 1: EXPLORING

The Experiment Design

A 32 item quiz was given to 54 students. The quiz was designed to make both within - and between - student comparisons, which allowed for a wider, more exploratory first study. Between students, all 32 items were either 2 or 5 data points. Within student, graphed axes varied among eight physical quantities – four sets of axes – and four data trends – increasing, decreasing, flat, or vertical. Lastly, students were asked to analyze both the general relationship of the data and to analyze the expected relationships in the values of extrapolated data on either side of the plotted data. The quiz also was given in reverse order to control for effects of question order and/or learning as students worked.

Constraining the Experiment

Unlike the picture in Figure 1, where the third variable is very clearly exhibited on the graph, the graphed data was constrained by a brief text above each question. It read: “For each experimental trial the student only varied one condition. The student has plotted the measured values for you. The student did not always plot the data in the best way and the student did not always collect a lot of data. However, it is your job to determine what functional relationship(s) between the variables are supported by the data.” It is actually very important for there to be some description of how the data was collected in order for certain reasoning to occur with a set of data. Without this clause all relations between the data are suspect because a controlled experiment was not necessarily done.

Findings

The analysis of this data showed several interesting things. First, while students did not significantly improve their performance as they worked through the quiz, suggesting that students were not learning from the quiz questions, there was a significant benefit to student ability to interpret the graphed data from a greater number of data points being provided on the graph. Second, there was no main effect of the axes labels, i.e. quantities being graphed. Third, students were better at analyzing general data relationships than comparing unplotted data values. Lastly, while students were quite good at monotonically increasing and decreasing data trends, they struggled with the less standard, and trickier, trends represented by the horizontal and vertical data.

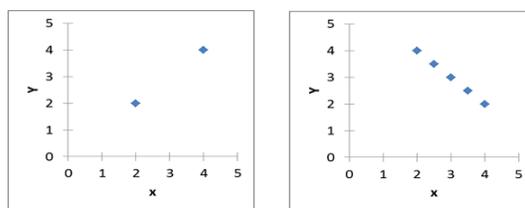


FIGURE 2. Example of 2 data-point and 5 data-point graphs.

Students do Better with More Data Points

Students were significantly better at answering questions about graphs with five data points vs. two data points like the graphs shown in Figure 2. The overall difference in correctness is small at 9%, but this is an effect size of 0.62 ($F = 4.13, p < 0.047$). However, there was no significant effect of question order nor an interaction between order and data point number. (A two-way ANOVA was done to analyze this data.) Nor was there any effect of learning seen through comparing the first few and last few items on the quiz. These different findings show that students are able to benefit from the between-data comparisons when analyzing a graph, but they did not improve in their ability to do this on a graph-to-graph basis as they went through the questions.

Identifying Relations is Easier than Comparing Values

There were two main tasks the students were asked to do to analyze the graphs. They were asked a multiple choice question about either the general trend or the relative values of unplotted data to either side of the graphed data. Each multiple choice option applied to a graph they were shown equally often. (See Table 1 for item wording examples.)

Considering the data presented in Table 2, students are better in general at determining the relationship than comparing the values for unplotted data. From analyzing correlations between items and written student reasoning, there are two main causes for this. One is simply that students were making random mistakes with the greater than less than arguments as they went through the value questions. The other is

TABLE 1. Examples of the wording for relation and value items.

Relation: Which of the following best describes the relationship between the variables X and Y?

- a) Proportional (i.e. when x increases, y increases)
- b) Inverse (i.e. when x increases, y decreases)
- c) Independent of each other (i.e. x does not affect y)
- d) There is not enough information to determine a relationship

Value: Which of the following best describes the expected relationship between the Y values for $X = 1$ and $X = 5$?

- a) $Y(x = 1) > Y(x = 5)$
- b) $Y(x = 1) < Y(x = 5)$
- c) $Y(x = 1) = Y(x = 5)$
- d) A, B, and C are all possible

TABLE 2. Correct response percentages sparsed by data trend and value vs. relation

Data Trend	Relation item	Value item
Up (Proportional)	95%	79%
Down (Inverse)	93%	75%
Horizontal (No)	79%	60%
Vertical (Unknown)	31%	63%

that the a, b, c, d answer pattern is different for the relation and value questions and the respective graphs for upward and downward trend data. This was done intentionally to distinguish students who were looking at the trend from those considering the values. These students did not make random errors but rather put a) for all monotonically increasing data trends and b) for all decreasing trends. When analyzing the data as a whole, there is no main effect of value vs. relation abilities, but there are interesting interactions with the different graph trends, particularly for the horizontal and vertical, “less standard,” data trends.

Students Struggle with Less Standard Data Trends

As can be seen from Table 2, students struggled the most with horizontal and vertical data trends, which most students do not have as much experience dealing with and thus are considered “less standard”. To be clear, a horizontal data trend is created by a horizontal set of data points, which for every “x-horizontal variable” the “y-vertical variable” is the same. This trend indicates the horizontal axis variable does not affect the vertical axis variable. (This of course assumes that all other variables are fixed.) Likewise a vertical trend is where the horizontal variable was not changed but the vertical variable did change because some other causal variable changed. This data indicates that the student cannot determine a relationship for how the horizontal variable affects the vertical variable (see Figure 3 for examples).

Students were not bad at determining the relation for horizontal data, 79% correct, but this picture gets a little less clear when realizing that only 60% could describe the relationship for unplotted data values. However, what is perhaps even more perplexing is that many students, 63%, could correctly say that for unknown values of X the relationships for the Y values were not known but could not match this with not knowing the relationship for the variables. Instead, they usually indicated that they were independent.

Following these results, Experiments 2 and 3, were conducted to determine why students were having issues dealing with vertical variable graphed data. Experiment 2 hypothesized that students did not realize there was a third variable that was changing in an otherwise controlled experiment, and thus they were not able to correctly reason because they were missing this important piece of information. This hypothesis proved to be incomplete and Experiment 3 revealed that many errors made were actually reasoning-related logical issues.

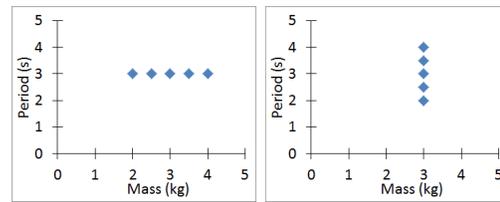


FIGURE 3. Examples of less standard graphs showing no - left graph - and an unknown - right - relationship. (Of course this is true only if the data was collected via a controlled experiment.)

EXPERIMENT 2: CHANGING THIRD VARIABLE SALIENCE

For this experiment the quiz instructions were modified to make it clear that the experiment included three variables. For example, for period-mass graphs students were told: “A student is experimenting with a simple pendulum to see what the relationship is between the mass on the bob, the length of the string and the period of each oscillation. For each experimental trial the student varied either mass or length and measured the change in the period.”

Of the forty-three students tested, twenty-one students were given graphs which used a legend to show the length changing like those graphs shown in Figure 1. Twenty-two students had to infer the length was changing when the mass stayed constant but period changed like those in Figure 3.

Surprisingly, using the legend to increase the salience for the length of the string made no difference in students’ ability to correctly interpret these horizontal and vertical graphs. Students averaged 83% correct for horizontal graphs and 27% vertical. (This compares to 79% and 31% from Experiment 1.) Subsequent interviews showed that many students did not have enough experience with a legend to use it without instructions. (A few students were even confounding the legend icon in the horizontal case as “outlier data”.)

Student interviews also showed evidence of mistaken student logic. For instance, about 10% of students *incorrectly* explained that the horizontal data indicated that period depends on mass by saying, “For every mass, I know what the period is going to be.” Similarly, students assumed that the vertical data demonstrated that there was no relationship between the x and y axes because, “If x doesn’t change and y does, than x doesn’t affect y.” These comments suggested that there were logical reasoning issues causing the students difficulties. Consequently, another experiment was created to allow for graphical issues to be distinguished from reasoning issues.

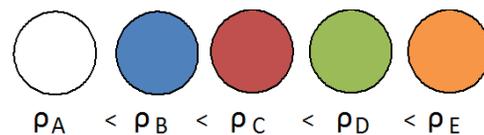


FIGURE 4. Example of a shape, color, density picture item.

EXPERIMENT 3: REASONING WITH PICTURES AND GRAPHS

This experiment required students to consider both pictures and graphs that showed multiple variable data and asked students to explain the relationships between that data. This allowed for reasoning to be separated from the graph.

Students were asked Graphs Questions for variables A, B, and C; period, mass, and length; and X, Y, Z as well as Picture Questions about density, shapes, and colors; period, mass, and length; properties A, B, and C. 36 students were randomly assigned to one of two conditions, half did the three picture categories first and the three graph categories second, and the other half did the three graphs categories first. (Examples of picture questions are shown in Figures 1 and 4.) Before each section, students were given a brief explanation of the activity, the graph legend, and were given an opportunity to ask clarifying questions.

In addition to the picture vs. graph comparison, this quiz – which has increasing, decreasing, vertical, and horizontal items – had two major changes from Experiment 2. It required students to determine not only if the horizontal variable affects the vertical variable but also vice versa, and if the legend affects the vertical variable. This put the three variables on an even level. It also asked questions where two variables would change simultaneously making it impossible to tell the effect of any one variable on another. These three changes showed several interesting results.

Students did significantly better when comparing axis variables, 45% correct over all, than when comparing legend to vertical axis 32%. There was no significant difference in items asking how the horizontal variable affects the vertical vs. how the vertical affects the horizontal. Also, overall graphs and pictures were not different than each other 42% and 40% correct respectively (see Table 3). This data strongly suggests that many student issues with the horizontal and vertical graph are reasoning - not graphical - issues.

In addition, when this is factored by whether the situation was controlled or uncontrolled, i.e. if the legend variable changed, two interesting facts are revealed. First, students did 15% better examining graphs than pictures when the experiment is controlled, but 6% better examining pictures than graphs when uncontrolled. This suggests that the pictures are better than the legend at helping the students see and reason with three variables, but the graphs are a better at helping the students see trends. It also shows that although 80% to 90% of students correctly identified increasing and decreasing trends in Experiment 1 this does not demonstrate graphical understanding of the relationship between the variables. Many students reported upward trend data as causally linked no matter what was happening to the third variable. This allowed them to be correct only on controlled items. Whether in picture or graph form, students were attending mostly to whether the variable in question changed and were not accounting for the importance of a controlled finding.

TABLE 3. Correct response percentages with standard error

Item Type	All	Controlled	Not Controlled
Graphs	42.3 ± 4.7%	65.1 ± 4.5%	28.0 ± 2.8%
Pictures	39.1 ± 3.8%	50.1 ± 5.1%	34.4 ± 3.4%

CONCLUSION

The three experiments reported here show several main facts. First, students do significantly better at recognizing data trends with 5 data points than 2. Second, for these experiments there were no strong effects of the variables being graphed. Third, students were better at analyzing general data relationships than comparing values for unplotted data. Forth, while students were quite good at the standard increasing and decreasing data trends on graphs this is only true for controlled set ups that change just the two main axes' variables. Fifth, because of students' equally poor graphical and pictorial reasoning with horizontal and vertical data, it seems that students are not good at the more sophisticated reasoning required to consider data where one variable does not change but the second does change. Their responses are consistent with reasoning issues seen in logic curriculum, and seem to be of the same ilk.

There are several studies that could stem from these findings, including: investigating if calculus-level students have similar reasoning issues; allowing physical manipulation of objects and testing whether the logic errors remain; and considering different types of graphs that use functional forms instead of, or in addition to, the data points. Also, this has several instructional implications; foremost, a need for the inclusion of data, graph, and reasoning experiences other than causal dependence. Additional instructional implications await more information about the persistence of these issues with instruction and across different situations.

ACKNOWLEDGMENTS

I would like to thank Dr. Epa Rosa and the undergraduate students who worked with me on this project: Christine Ewald, Chris Rosenzweig, and Jamie Svetich.

REFERENCES

1. A. E. Lawson, *Journal of Research in Teaching Science* **15** (1), 11–24 (1978).
2. L. C. McDermott and the Physics Ed. Group at the Univ. of Washington, *Physics by Inquiry* Wiley, New York, (1996).
3. A. J. Ross, *Review of Educational Research* **58**, 405–437 (1988).
4. A. Bodreaux, P. S. Shaffer, P. R. L. Heron, and L. C. McDermott, *Am. J. Phys.* **76**, 163–170 (2007).
5. R. J. Beichner, *Am. J. Phys.* **62**, 750–762 (1994).
6. D. E. Meltzer, *Am. J. Phys.* **73** (5), 463–478 (2005).