Comparing two activities' effectiveness improving reasoning with multiple-variable information

Rebecca Rosenblatt, and James Perrone

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

We report on studies showing large differences in student ability to use, and reason with, certain data. Many students incorrectly assume there must be dependence between the axes of any graph irrespective of whether the data suggests a relation. Additionally, students have difficulty with variables and their role in the experiment. A majority of the errors are consistent with failure to control variables or illogical reasoning from the data. We developed and tested two different one-hour group activities to improve student understanding. One activity was laboratory-based and focused on control of variables and experimentation. The other was recitation-based and focused on logical reasoning and data manipulation. Both activities improved student ability to draw conclusions and answer questions about the data in graph or pictorial form. The relative effectiveness of the activities suggests that both logical reasoning and control of variables are important for students working with this data.

I. 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 interpreting graphs have been explored in several ways [5, 6]. Some research has been done combining these problem areas to observe student reasoning with control of variables and graphed data [7, 8]. In these studies, we focused specifically on observations of student reasoning with graphed and pictured data and on conclusions about the effect one variable has on another variable.

In this paper, we will review our previously published findings [8] about student reasoning with single-trend data and graphs (see Figure 1), expand and generalize these findings to all single-trend data, discuss new findings pertaining to student reasoning with multitrend graphs and with experimental trial data sets, outline the instructional techniques used in two different activities designed to improve students' skills in this area, and discuss the effectiveness of these activities and their implications. (This research was done post instruction, between the 9th and 12th weeks of the semester, with 250 traditionally taught algebra-based mechanics students.)

II. STUDENTS' REASONING DIFFICULTIES

Here the main issues are summarized; the evidence supporting these claims is presented in the following subsections. **A.** Students struggle with questions about a variable which was unchanged in the graph and/or pictorial data and with data that is not controlled, i.e. too many variables are changed at once. **B.** On multiple trend graphs, the data is controlled but students make reasoning mistakes with the data because students do not properly control their observations or account for changes in multiple variables. **C.** With more elaborate pictorial experimental trial data sets, student mistakes resemble the mistakes made with simpler graphs and pictures.

Students make persistent and consistent logical mistakes concluding "since B changes and A does not, then A must not affect/change B" or "since A changed and B changed, A

FIG. 1. Pictures on left show single-trend data sets. Graphs show the "same" data. (Students did not see graphs and pictures together.)

must affect B." They also make conflicting claims that some data implies A affects B while other data on the same graph implies that A does not affect B. These findings suggest students' errors are rooted in their understanding of control of variables and ability to reason in a correct logical fashion.

A. What is known from single-trend graphs and data sets

Fig. 1 shows a few single-trend pictures and graphs. Previous research published in PERC 2014 showed that students tend to have a hard time interpreting the effect of mass on pendulum period with the horizontal and vertical graphs shown in the top two graphs of Fig. 1 [8]. Around 75% of students were correct on the horizontal style of graph, which shows that mass does not affect the period, and about 30% were correct for the vertical style of graph, which shows that length affects period but does not show how mass affects period.

Our new more extensive research indicates that what is re-

ally confounding the students is not the horizontal or vertical nature of the graph but rather the unchanged value of the x, y, or z variable. When a variable is left unchanged, necessary for controlled experimentation on another variable, around 60% of students will incorrectly claim that this unchanged variable does not affect other variables using the logic, “If x doesn’t change and y does, then x doesn’t affect y.” To clarify what is meant with an example, when students are asked about the length variable’s effect on the period in the top horizontal style graph shown in Fig. 1, only 30% to 40% of students will respond correctly. This error rate is comparable to that observed for the unchanged mass variable in the vertical style of graph shown beneath it. Our research group has now observed this response pattern for unchanged variables in many different graph styles, with many different variables both concrete and abstract, and over multiple semesters.

In addition to having difficulties interpreting graphs, students use the same type of incorrect reasoning for unchanged variables when given simple pictures of the data like those shown next to the graphs in Fig. 1. For example, about 60% of students will claim that this data shows that length does not affect a pendulum’s period when shown picture A, and that mass does not affect the period when shown picture B. This consistency of responses indicates that the difficulties students are having are not graph-related but rather stem from incorrect reasoning with the data [8].

The last key difficulty students have with interpreting single-trend data is students’ failure to account for controlled vs. uncontrolled data. This is exemplified by students’ responses to questions about data like that shown in the bottom picture and graph in Fig. 1. Many students fail to consider the effect of unasked-about variables. The result is that only 30% to 50% of students correctly respond to questions about uncontrolled data. The patterns of similar responses and student reasoning explaining those responses for uncontrolled data and unchanged variables suggests that students’ difficulties are not just bad logical reasoning but are also connected to a failure to understand, and/or use, control of variables.

B. What is known from multitrend graphs

The next topic investigated was student interpretation of multitrend graphs. Because multitrend graphs are inherently controlled, but the way students consider the data does not have to be, such graphs are a good test for the conclusion that bad logic and/or control of variables is causing these misinterpretations of the data. If students’ difficulties did not persist into this new domain, the conclusion would likely be wrong.

An eighteen question quiz with abstract multitrend graphs questions was created and its validity established by several methods including think aloud interviews. One of the graphs is shown in Figure 2. Students were asked about the relationships among the variables and what certain selected data comparisons told the students about those relationships.

Findings from this new quiz corroborate previous findings

FIG. 2. An example of a multitrend graph with a linearly increasing relationship between pods and widgets. (An example prompt is - Circle or Indicate: None, One, ... All of the answers as you think most appropriate. Which data comparison(s) would suggest that the # of pods changes, or affects, the # of widgets? a) A and B; b) A and D; c) A and C; d) B and C; e) B and D; f) C and D.)

of student misinterpretations of unchanged variables and uncontrolled data. Around a third of students can correctly indicate the data comparisons which allow them to reach conclusions for the relationships among the three variables (pods, bagets, and widgets). Students tend to use comparisons that either do not account for the importance of fixing the third variable - e.g. using data comparison between A and D to support the claim that pods affect widgets - or that fix one of the variables they want to compare and then change the wrong variable - e.g. using data comparisons between A and C or B and D to make claims about pods effect on widgets.

Also, students made additional logical errors interpreting the data that could not be observed with single-trend graphs but are consistent with errors in that domain. Students claim that certain data comparisons indicate that two variables affect each other while other comparisons indicate the same two variables do not affect each other. The data comparisons chosen in these relationships are consistent with the unchanged variable reasoning issues seen in single-trend data (e.g. students respond: comparisons of A and B & C and D indicate pods affect or change widgets, comparisons of C and B suggest pods do not affect widgets.) Overall, 56% of questions were answered in this illogical fashion for multitrend graphs. 42 of 50 students did this at least once for the different questions asked. Students often do not seem to be aware that it is contradictory to claim that some data, in these simple data sets, suggest a does-not-affect relationship and other data suggest a does-affect relationship.

C. What is known from pictured experimental trials

A final test to these conclusions was made. It was thought that having questions that more closely resembled real experimental labs would give the students cues to use control of variables and this would improve students’ scores and reasoning. Two quizzes were created using the topics of static friction and light bulbs. (Students learned about friction but not light bulbs in the course, thus, comparing the scores assessed

TABLE I. Correct response comparisons for single-trend.

Experiment Type	Variable X	Graphs/Pics	Friction	Bulbs
Controlled	Changed	85%	66%	52%
Controlled	Unchanged	35%	42%	48%
Uncontrolled	Changed	40%	55%	32%
Uncontrolled	Unchanged	–NA–	44%	35%

the importance of the physics topic vs. abstract reasoning.)

As before, students were given a cover page explaining the variables that can be changed in any situation, how they could change (e.g. floor material can vary between carpet, metal, and wood), and the measured variable (the force readout). Students were asked to determine if the data on each page suggested that a variable affects, does not affect, another variable or if it has an undetermined affect. (An example prompt is - The data suggests that: a) the mass does affect, i.e. change the force to just get it moving; b) the mass does not affect, i.e. change the force to just get it moving; c) neither a) nor b).)

Table I presents this data and Figure 3 shows two examples of friction trial sets. There are several things of note in this data. 1. Students responses to questions about the graphical data were very similar to responses for these experimental trial sets. The only deviation of note is for situations involving a controlled data set and a changed variable. Students do worse on these more complex experimental trials. 2. Students do better on questions about friction than about light bulbs. 3. Consistent with previous results, students perform worse on uncontrolled experiments and unchanged variable questions than for controlled and changed ones.

III. THE TWO INSTRUCTIONAL MODIFICATIONS

From these studies, it can be seen that student difficulties interpreting unchanged variables and uncontrolled data are persistent, consistent, and prevalent. All evidence shows they are caused by logical reasoning issues and/or control of variables difficulties. The next step was to see to what degree these difficulties could be remediated with an instructional activity. By creating two activities, the importance of logical reasoning vs. control of variables could be investigated.

Students were randomly assigned to do either a recitation-like tutorial on logical reasoning or a hands-on lab on experimental design and control of variables during their regularly assigned lab period. Both activities started and ended with individually done pre and post tests. The pretest was a set of single-trend graphs and pictorial questions, which took 10–15 minutes to complete. Then the students completed one of the two activities, tutorial or lab, in small groups. The groups worked at their own pace taking 50–70 minutes. Then they did the posttest which took 30–40 minutes. The posttest consisted of the pretest questions, additional graphs and pictorial questions, and the friction and bulbs quizzes described above.

FIG. 3. Two examples of experimental trial sets for static friction. Students were asked to make conclusions about surface material, mass, and bottom area's effect on the force readout. The left trial sets could be a controlled-changed question if asked about mass and controlled-unchanged if asked about floor material or surface area. (Students were not given the labels controlled or uncontrolled.)

A. The tutorial on logical reasoning

This activity started with the students reading a short one page explanation, with included examples, of what it means for variable A to affect (or not affect) variable B and when it is acceptable to invert a statement of “A affects B” to “B also affects A”. Then students were asked to create an example and state a general rule.

Next, students were given a data table including periods of a simple pendulum and several variables: length of string, mass, material of the bob, and size of the spherical bob. From this data, students were asked to draw conclusions about each variable and its relationship to the period. Students were asked to explain their reasoning, point out evidence supporting their choice, and to note any contradictory evidence.

The students then answered a set of “student dialog” questions that directly addressed students’ conceptual and reasoning difficulties. Students were asked whether they agreed or disagreed with each comment and write to explanations of why (e.g. Sue: I think the material type does not affect the period. If you look at brass you get lots of different periods for the brass so that must mean that material type does not affect the period. Chris: I think your conclusion may be right but your reasoning is wrong. If you look at brass you get lots of different periods because the length changes so that tells you length affects period but not that material does not.)

Students were asked to graph some of the data from the pendulum table. After this, they were asked to create a rule for what data looks when graphed if it shows that a variable affect another variable and to create a rule for what data looks like when graphed if a variable does not affect another.

Finally, students were asked to examine a multitrend graph, to transform the data from the graph into a table, and explain the relationships between the variables providing evidence for their conclusions.

B. Experimental control of variables reasoning

This activity started with each student designing an experiment to determine what effect mass and surface area have on static friction. Then students were brought together for a pendulum demonstration. Students were asked to come up with possible variables that would affect the period. The students identified mass, length, angle of release, and shape of the object as possible variables. Students were presented with two comparison pendulums and asked what conclusions, if any, they could make from those two pendulums. After class discussion, the periods of the two pendulums were demonstrated and conclusions discussed. Several different demonstrations were done in this manner with the students offering suggestions before each demonstration on how to change the experiment to account for specific variables. The demonstrations covered multiple options including: a variable has no effect, a variable does have an effect, uncontrolled experimental set ups (e.g. varying mass and length), and unchanged variables (e.g. varying length but asking about mass).

Next, students combined their friction experimental designs, made changes if needed, and carried out their experiments with force probes and/or track tilt. Finally, they graphed their data and presented their findings to the class.

C. Results of the instructional modification

The overall finding is that both modifications improved the skills tested and did so in approximately equal amounts for each instructional activity. This data is presented in Table II.

For single-trend graphs and pictorial data, the greatest improvement was in students' ability to correctly identify uncontrolled experiments and to understand that these experiments will not convey information about the relationships among the variables (an 11% gain). Students improved the least on controlled experiment and changed variable questions (a 6% gain), but this was still statistically significant ($F = 4.1$, $p < 0.001$, effect size of 0.29). The two instructional activities were not significantly different from each other in overall improvements.

On the friction and light bulbs quizzes, students performed significantly better than they did without the instructional activities. ($F = 5.1$, $p < 0.001$, effect size of 1.1 for friction and $F = 7.4$, $p < 0.001$, effect size of 1.6 for bulbs.) The two instructional activities are not significantly different from each other for either the friction or light bulbs quizzes.

(There was a difference between the means of 0.7% on the friction quiz and 2.4% on light bulbs quiz with the experiment instruction being slightly larger on both.) What is especially notable is that students are scoring the same on the light bulbs quiz and the friction quiz despite the fact that they are more familiar with friction and despite direct use of friction in the lab-based modification. This suggests that students were abstracting the skills the activities addressed and not simply learning the physics of the specific scenarios.

In addition, the fact that students are only 60% correct immediately after the hour long intervention indicates the misunderstandings students have on this topic are resistant to change and may not be easily reformed.

TABLE II. Average correct response percentages after modified instruction. Parenthesis show gains over pretest/previous results.

Type	Variable X	Graphs/Pics	Friction	Bulbs
Controlled	Changed	91% (6%)	89% (23%)	91% (39%)
Controlled	Unchanged	50% (8%)	60% (18%)	61% (13%)
Uncontrolled	Changed	64% (11%)	55% (0%)	53% (21%)
Uncontrolled	Unchanged	–NA–	65% (19%)	67% (32%)

IV. CONCLUSION

Here several different experiments are described which show that many students in this algebra-based class have difficulty working with data to determine the relationships among different variables. Specifically, students struggle with questions which address a variable that was not changed (controlled in order to look at other variables), with the importance of consciously constraining their observations of the data, and with the implications of data that is not controlled. A majority of the errors are consistent with failure to control variables and illogical reasoning. Also presented are two different one-hour group activities to improve student skills - one focused on control of variables and the other focused on logical reasoning and data manipulation. Since both activities improve student understanding and skills on these types of unchanged variables and uncontrolled experiment questions, this is additional evidence that both logical reasoning and control of variables play important roles in these questions of graphical and pictorial data interpretation.

[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 Ed. 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).
[7] A. M. Bogdan and A. F. Heckler, *PERC Proceedings* (2013).
[8] R. J. Rosenblatt, *PERC Proceedings* (2014).