

Student understanding of control of variables: Deciding whether or not a variable influences the behavior of a system

Andrew Boudreaux

Department of Physics and Astronomy, Western Washington University, Bellingham, Washington 98225

Peter S. Shaffer, Paula R. L. Heron, and Lillian C. McDermott^{a)}

Department of Physics, University of Washington, Seattle, Washington 98195

(Received 15 June 2007; accepted 13 October 2007)

The ability of adult students to reason on the basis of the control of variables was the subject of an extended investigation. This paper describes the part of the study that focused on the reasoning required to decide whether or not a given variable influences the behavior of a system. The participants were undergraduates taking introductory Physics and K–8 teachers studying physics and physical science in inservice institutes and workshops. Although most of the students recognized the need to control variables, many had significant difficulty with the underlying reasoning. The results indicate serious shortcomings in the preparation of future scientists and in the education of a scientifically literate citizenry. There are also strong implications for the professional development of teachers, many of whom are expected to teach control of variables to young students. © 2008 American Association of Physics Teachers.
[DOI: 10.1119/1.2805235]

I. INTRODUCTION

The ability to design and interpret experiments so that conclusions can be reached regarding the role of a variable in the behavior of a system is fundamental to the process of science. This paper focuses on student reasoning associated with the control of variables. The research grew out of informal observations made by the Physics Education Group at the University of Washington (UW). For many years, the group has taught special physics courses for K–12 teachers and for students who intend to major in science but lack the requisite background.^{1,2} It was observed that students in both groups struggled with many aspects of scientific reasoning, including the control of variables. The findings guided the development of *Physics by Inquiry*, a curriculum designed explicitly for these populations but appropriate for other liberal arts majors.³ A more recent and more systematic investigation included not only K–8 teachers but also students in large undergraduate physics courses.⁴ An important goal was to document student reasoning in enough detail to enable science teachers at all levels to design effective instruction.

In this paper, the term “control of variables” is used to characterize the design of an experiment in which variables are changed in certain specified ways in order to probe the effect of a particular variable on the behavior of the system. In general, a variable may be related to the result of an experiment in three ways: (1) the variable may not affect the result, (2) the variable may *influence* the result, in which case changing that variable may change the outcome, or (3) the variable may *determine* the result, in which case that variable alone can be used to predict the outcome. The way in which the variables are controlled depends on which relation is being probed. The focus in this paper is on the ability of students to design and interpret experiments to decide whether or not a variable influences the behavior of a system (cases 1 and 2).⁵

II. PRIOR RESEARCH

A primary objective of precollege science education is to enhance the ability of students to conduct scientific inquiry.

Many curricula contain activities intended to help young students develop the skills necessary to formulate hypotheses, design and perform experiments that can test their validity, and interpret the results. This approach to science instruction involves multiple aspects of scientific reasoning. Results from research indicate that many of these pose substantial difficulty for precollege students.⁶

The ability to apply the control of variables, a skill that is central to the process of science, has received much attention.⁷ Several studies have examined student ability to design controlled experiments or to evaluate experiments as either “good” (well controlled) or “bad” (not controlled or confounded).⁸ Difficulty in recognizing that several variables can contribute (additively) to an experimental result has also been documented.⁹ Some studies have shown that students have particular difficulty recognizing when the available evidence is not sufficient to decide on the effect of a variable.¹⁰

Other studies have focused on the ability of students to modify their thinking when confronted with data that conflict with their preconceptions.¹¹

Several instruments have been developed to probe student ability to reason formally.¹² Often they draw from sources such as Inhelder and Piaget’s seminal work on the development of logical thinking.¹³ Items on the control of variables are sometimes included. Standardized tests in many states contain questions involving the control of variables.¹⁴

Most prior research has focused on children. One informal study of adults that is particularly relevant to the present investigation was conducted in Israel at a special school for the preparation of primary school teachers.¹⁵ Although most teachers could design appropriate experiments, many had difficulty interpreting the results. They often could not decide how to tell whether a particular variable influenced or determined the behavior of a system.

Some of the studies we have summarized demonstrate that standard instruction often does not improve student ability to apply control of variables. In some cases it was found that instruction that explicitly addresses the underlying reasoning can have a significant effect and that targeted lessons can enable students to transfer this type of reasoning from one

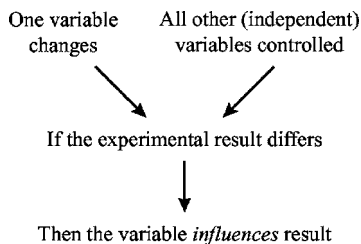


Fig. 1. Summary of the control of variables reasoning needed to test whether a particular variable influences a given result.

domain to another.¹⁶ Thus, some aspects of the control of variables seem accessible even to young students. Other aspects, however, present a greater challenge and are more difficult to teach.¹⁷

III. OVERVIEW OF THE INVESTIGATION

The present study was motivated by our interest in ascertaining the extent to which difficulties with the control of variables persist at the university level. We also wanted to probe the ability of students to reason appropriately about the control of variables in situations more complex than in most previous studies, which typically involved only two experimental trials and variables easily recognized as independent.

Various research tasks were used in this study. Each consisted of one or more questions that probed student ability to test whether or not a variable influences the behavior of a system. To answer correctly, students needed to recognize that the variable being tested should change from one trial to the next while other (independent) variables remain constant. If the outcome changes, the students should infer that the variable being tested influences the result. If the outcome is unchanged, they should infer that the variable does not influence the result, at least for the range over which the variable is changed (see Fig. 1).

A procedure different from the one we have described is needed to decide whether or not a variable that influences the behavior of a system also determines it. The variable is now kept constant while others are changed. If the behavior of the system changes, then the variable that is being tested does not, on its own, determine the outcome. If the behavior does not change, then there is evidence that the variable determines the outcome. Precollege (and university) instruction does not typically include the strategy needed to identify a variable that determines an experimental result. This aspect of the control of variables is not discussed in this paper.⁵

Data were obtained primarily through the administration of written questions at Western Washington University (WWU) and the UW. The students at WWU were enrolled in a general education physics course for nonscience majors ($N > 1000$) or in the introductory calculus-based physics course ($N \sim 100$). At UW questions were posed to students in calculus-based physics ($N \sim 200$) and to inservice K–8 teachers in a 6-week NSF Summer Institute ($N \sim 75$). Other data were collected in 1-week professional development workshops conducted by the Physics Education Group ($N \sim 150$).

The written questions were supplemented by in-depth interviews with student volunteers from a geology course for nonscience majors at WWU ($N = 5$). The science background of the interview subjects varied, but all had previously taken

high school or university science courses. The interviews helped us interpret the responses to the written questions.

IV. DEVELOPMENT OF RESEARCH TASKS

The research tasks in this study can be grouped into three categories. These evolved as the research progressed, and we learned how to elicit student reasoning more effectively. The original task assessed the ability of students to design experiments to test whether a particular variable influences an experimental result. Two subsequent tasks examined the extent to which students can interpret data that is presented to them.

Various physical contexts were used. In general, we found that student performance did not depend on the context. In the following, we present an example from each category and briefly describe the results. Discussion of student reasoning is deferred to Secs. V and VI.

Not all populations were given all three types of tasks. In most cases, students had not had special instruction on the control of variables. In some sections of the general education physics course, students received about 30 min of lecture instruction on how to test for the influence of a variable. Other classes did not have any relevant instruction. The results from classes with and without such targeted instruction were identical. Therefore, the results have been combined and are discussed together. The students in the calculus-based physics courses at UW did not have any instruction on the control of variables. The K–8 teachers in the Summer Institute at UW and in the workshops conducted by the Physics Education Group had not worked through any relevant material before taking the pretest on this topic.

A. Task 1: Ability to design experimental trials

Students were given a specific physical context and instructed to consider a specified set of variables (for example, a simple pendulum with the length of the string, the angle between the string and the vertical at the time of release, and the mass of the bob as the variables). We then asked them to design an experiment that could be used to test whether or not a particular variable (for example, the angle of release) influences the outcome (for example, the number of swings in 10 s). Students were also asked to describe how they would use the results to reach a conclusion.

We expected students to describe an experiment consisting of two or more trials in which only a single variable is changed from one trial to the next. A change in the experimental result can then be traced to the variable being tested. If the behavior is unchanged, then that variable does not play a role (at least for the particular choice of independent variables and for the range of values within which the variable was changed).¹⁸ On written versions given in the general education and calculus-based physics courses ($N \sim 500$ and $N = 25$, respectively), between 70% and 90% were able to describe appropriate experiments. This result is consistent with those from the interviews.

Although many students designed appropriate trials and mentioned the need to “control variables,” the explanations were weak. Even with substantial prompting during the interviews, students struggled to articulate the reasoning supporting the experimental design. Ideas about the causal mechanisms seemed intertwined with the ability to reason properly about the control of variables. For example, some students treated their experiments as a way to verify what

An object suspended from a spring can be made to bob up and down. An experiment is performed to investigate whether certain variables (the initial length of the spring, the distance the object is pulled at the time of release, and the mass of the object) affect the number of times the object bobs up and down in 10 seconds. The results of four trials are shown. Each trial has been repeated several times with the same results. (Other variables may play a role, but are controlled.)

Variable	Trial 1	Trial 2	Trial 3	Trial 4
Unstretched length of spring	12 in	27 in	12 in	12 in
Distance object is pulled at time of release	3 in	6 in	3 in	6 in
Mass of object	50 g	72 g	72 g	72 g

Can this experiment be used to test the statement below?

“The distance the object is pulled at the time of release influences the number of bobs in 10 s.”

If so, state which trials should be compared and explain whether the results can be used to support or refute the statement. If not, explain why not.

# of bobs that occur in 10 s	54	30	45	45
------------------------------	----	----	----	----

Fig. 2. An example of a task (task 2) in which students are asked to interpret experimental data to test whether or not a variable influences a result. In this task, the data justify a conclusion about the role of the variable being tested.

they “knew” about the role of the particular variable. Many failed to articulate the **hypothetico-deductive reasoning that underlies the use of control of variables** as a means for testing the role of a variable.¹⁹ The responses suggested that the control of variables might be a memorized algorithm for many students. In order to obtain more insight into their thinking, we designed a second task that required students to go through the reasoning needed to interpret data.

B. Task 2: Ability to interpret results when the data warrant a conclusion

In the second type of task, students were shown data from three or more trials of an experiment and asked whether a given variable influenced the behavior of the system. The data included both controlled and uncontrolled trials. To answer, students had to identify the controlled trials. In some versions of the task the trials had different outcomes, indicating that the variable does influence the behavior of the system. In other versions the outcome was unchanged, suggesting that the variable does not influence the behavior. All versions of the task yielded similar results.

Figure 2 illustrates the data from one such experiment. The number of oscillations of a spring that occur in 10 s is measured when certain variables are changed (for example, the unstretched length of the spring, the mass of the object attached to the spring, and the distance the object is pulled before it is released). Each column corresponds to a single trial with specified values for each variable. The bottom row shows the number of oscillations that occurred in 10 s.

Students were asked whether or not the given trials can be used to test the idea that the distance the object is pulled influences the number of oscillations that occur in 10 s. Trials 3 and 4, for which only the distance is changed, can be used to answer. Because the outcomes are identical, the data suggest that the distance the object is pulled does not influence the result.²⁰

This task was administered to more than 250 students in three sections of the general education physics course at WWU. Although the amount of instruction on control of variables differed in the three sections, the results did not. The question was also given to 100 K–8 teachers in the

Summer Institute and in several different 1-week workshops conducted by the Physics Education Group at UW.

As was the case for the first task, the fraction of correct responses for the students in the general education physics course was high. About 85% answered correctly, most with acceptable explanations. However, only about 45% of the teachers answered correctly. Unlike the students in the general education course, many teachers claimed that there was a need for many repetitions of the experiment. They failed to recognize that inferences could be made about the effect of the variable based on the data that were given.

Both the correct and incorrect answers provided insights into how the students and teachers were thinking. Some claimed that they could infer the influence of a given variable by using trials in which that variable did not change. Some failed to control the other variables. A single response often contained both correct and incorrect reasoning. For example, some students and teachers correctly identified the pair of trials that should be compared, but then went on to cite inappropriate comparisons as additional evidence. These latter answers led us to suspect that certain errors might be more common if there were no well-controlled trials in the given data.

C. Task 3: Ability to interpret results when the data do not warrant a conclusion

During a later phase of the investigation we designed a third type of task. In this case, the data given to the students were not sufficient to allow them to decide whether or not a given variable influences the results. The various versions of this task involved a variety of different physical situations, such as the period of a repetitive motion (for example, a pendulum), the time required for a projectile to complete its motion, or the time for a liquid to cool. A representative example is shown in Fig. 3. Students were asked if the given data could be used to test whether or not the mass of the pendulum bob influences the number of swings that occur in a given time interval. Because the data do not include two trials, in which all the variables are the same except for mass, the experiment cannot be used to isolate the effect of mass.

This task was administered at WWU to more than 950 students in 11 sections of the general education physics

A pendulum is released from rest and allowed to swing back and forth. An experiment is performed to investigate whether certain variables (length of string, mass of bob, and angle string makes with vertical at time of release) affect the number of swings in 10 seconds. (Other variables may play a role, but are controlled.) Results are summarized in the table at right. Each trial was performed several times with the same result.

Can this experiment be used to test the statement below?

“The mass of the bob influences the number of swings in 10 s.”

If so, state which trials should be compared and explain whether the results can be used to support or refute the statement. If not, explain why not.

Variable	Trial 1	Trial 2	Trial 3
Length of string	10 cm	10 cm	40 cm
Mass of bob	20 g	30 g	30 g
Angle of release	15°	30°	15°

# of swings in 10 s	16	16	8
---------------------	----	----	---

Fig. 3. An example of a task (task 3) in which students are asked to interpret experimental data to test whether or not a variable influences a result. In this task, the data do not justify a conclusion about the role of the variable being tested.

course and to 61 students in two sections of the calculus-based physics course. Only about half of the students in each population gave a correct response. Most of the rest stated that the data indicate that mass does not affect the number of swings. Performance on all such tasks was similar. Therefore, it seems that the errors the students made are primarily due to difficulties with the underlying reasoning, rather than with the specific physical context.

D. Commentary on research tasks

The performance of university students on the first two tasks is consistent with studies showing that precollege students can acquire the ability to construct controlled experiments and, in simple situations, identify unconfounded experiments. Results from the third task suggest that the first two alone do not provide a complete picture of student thinking. By presenting students with data that are insufficient to draw conclusions, the third task reveals limitations to their understanding.

V. GENERAL DIFFICULTIES WITH SCIENTIFIC REASONING

The research tasks we have described elicited a variety of modes of student thinking. Analysis revealed some general difficulties with basic scientific reasoning (see Sec. V) and other difficulties specific to the control of variables (see Sec. VI). The categories in which these difficulties are grouped are not mutually exclusive. Individual responses can often be interpreted in more than one way.

A. Failure to distinguish between expectations and evidence

Science is characterized by a rich interplay between creating explanations and testing these explanations empirically. Distinguishing between expectations and evidence is crucial. We have found that students often base their interpretation of data on their previous ideas about how a system should behave. For example, during an interview on the pendulum task, a student was asked to design an experiment to test for the effect of the angle of release on the rate of swings. This student described trials in which all the variables (mass,

string length, and angle) changed from one trial to the next. When asked how the experiment could be used to assess the effect of the angle, he explained: “The smaller the angle and the lighter the mass would lead to less swings I believe. But the shorter string would lead to more swings.”

The student was asked what could be inferred if the experiment were carried out and a different number of swings was obtained for each trial. He responded as follows: “It would confirm the hypothesis that the larger the angle and the larger the mass would lead to more swings. Although the longer the length would probably lead to less, I still think that the angle and the mass are more important than the length.” Throughout the interview the student was unable to set aside his beliefs about how the system should behave when designing an experiment and interpreting the results.

Another student, when asked if the trials shown in Fig. 3 could be used to test whether mass influences the swinging rate, initially articulated the need to hold both the length of the string and the angle of release constant. However, he went on to say that “you could probably figure out something” from the data, and inappropriately used trials 1 and 2 to make a conclusion. He explained: “The results from trials 1 and 2 don’t make sense because if mass is greater, you are going to have more swings, and here, the mass increases and the number of swings stays the same. So, I would say that the hypothesis is refuted.”

The student did not recognize that controlling the other variables is necessary in order to make an inference about mass from the data. The response suggests a prior belief about the role of mass in the behavior of the pendulum. During the interview he repeatedly commented that the data “do not make sense,” suggesting a difficulty in separating his expectations about how the pendulum should behave from his assessment of the evidence. The student seemed to regard the inconsistency between the data and his expectations as grounds for refuting the hypothesis.

B. Tendency to confuse related ideas: influence versus determine

The tendency to confuse related ideas has been documented in a variety of studies involving many different content areas.²¹ In our investigation students often were unable

to distinguish variables that can influence (affect) the outcome of an experiment from those that can determine (predict) the result.

On the third task (Fig. 3), many students inappropriately concluded that the data show that mass does not influence the number of swings. Some, like the following student, supported their answer by comparing trials 2 and 3, in which the mass is the same but the swing rate differs: “Trials 2 and 3 have the same mass but different outcomes. This proves that the mass does not influence the number of swings.”

The student seemed to recognize that the mass, which was unchanged, could not be responsible for the difference in the swing rate in trials 2 and 3. Yet he went on to conclude incorrectly that mass does not influence the rate. His statement would be correct if the word “influence” were replaced by “determines.” Thus, his error can be interpreted as confusion between influencing and determining. (Trials 2 and 3 show that mass does not determine how many swings occur in 10 s. It does not follow, however, that mass does not influence the rate.)

C. Reluctance to make inferences from data, even when warranted

On the second research task (see Fig. 2), many K–8 teachers seemed reluctant to use the experimental data they had been given. They often did not use reasoning based on control of variables.²² Instead, they tended to focus on other aspects of the experiment, as illustrated by the following: “I think more information is needed. I would test [the variable] 100 times... and find the average. The greater the trials, the greater your conclusions are probable.” Like many others, this teacher seemed to think that, in a scientific context, many trials are needed in order to draw a conclusion. Responses like this one suggest a reflexive belief that conclusions cannot be drawn from only a few trials.

Some teachers seemed to think that the data could not be used because the experiment had not been conducted “systematically.” They expressed a need for the experiment to be repeated with regular, incremental changes of the variable that was being tested. Other teachers thought that the data could not be used because of the presence of extraneous trials in which the variable being tested was not changed. The comments often suggested a naive belief that valid experiments must follow a narrowly prescribed format.

VI. SPECIFIC DIFFICULTIES DECIDING WHETHER A VARIABLE INFLUENCES A SYSTEM

Section V describes reasoning difficulties of a general nature elicited by the research tasks. The following analysis identifies difficulties specific to the application of control of variables.

A. Failure to control variables

Many student responses to the research task in Fig. 3 reflected a failure to control variables. Although most students recognized the need to compare trials in which the variable being tested was changed, some did not realize that the other variables should be held constant. Some students used incorrect reasoning to conclude that the mass of the bob does not influence the swing rate. Many of them inappropriately compared trials 1 and 2, in which both mass and angle are different while the swing rate is the same. The following ex-

ample is illustrative: “The mass of the bob does not matter, based on trials 1 and 2 having the same amount of swings in 10 seconds with trial 2 having a heavier bob.” The student apparently did not realize that the identical swing rate was consistent with two possibilities: neither mass nor angle plays a role, or both mass and angle play a role and in this case their effects cancel.

Other students used incorrect reasoning to conclude that the mass does influence the swing rate. The following example is characteristic: “Trials 1 and 3 can be used because they have different masses. The data show that the greater the mass, the higher number of swings.” This student did not seem to recognize that, because the length also changed between trials 1 and 3, length, rather than mass, might account for the different number of swings.

The tendency to draw conclusions based on confounded experiments was common for the research task shown in Fig. 3. When students were designing experiments or using data that contained well-controlled experiments (research tasks 1 and 2, respectively), this error did not arise. It seems that although some students recognized that well-designed experiments should have unconfounded variables, they tried nonetheless to draw conclusions even when more than one variable had been changed. It may be that they had learned a set of rules for the design of valid experiments but did not understand the underlying reasoning.

B. Assumption that only one variable can influence the behavior of a system

In the third task (see Fig. 3), many students correctly recognized that the results of trials 2 and 3 imply that a variable (or variables) other than mass influences the rate of swinging. However, some of these students went on to conclude that mass must not influence the behavior. One, for example, stated: “It does not seem that the mass is influential. In trials 2 and 3 the mass is 30 g yet the number of swings varies. This leads me to believe that other variables are influential.” This student’s reasoning contains the logical fallacy that if other variables influence the number of swings, then mass does not (or cannot). The statement is consistent with a failure to recognize that more than one variable can influence the result of an experiment.

C. Failure to realize that a variable must be changed to test for its influence

Some students who compared trials 2 and 3 in the third task gave responses such as: “... if mass affected the outcome, then trials 2 and 3 should have had the same outcome.” This incorrect response is consistent with a failure to recognize that in order to test whether or not a variable affects an experimental outcome, that variable must be changed from one trial to the next.

As before, the interviews provided further insight into this type of student reasoning. Two interview volunteers experienced difficulty with this task. One vacillated repeatedly between the correct response and a failure to recognize the need to change the mass when deciding whether or not it plays a role in the swinging of the pendulum: “Maybe if the mass was kept the same and you had different lengths, you could see if the mass influences the number of swings... If you kept the angle of release and the mass the same and changed the length, you wouldn’t really be able to tell if it

was the length that was influencing the number of swings, or it was the actual mass. I guess you could say that the mass does influence it.”

The student described an appropriate experiment for testing whether the length of the pendulum influences the number of swings, but gave an incorrect description of how the results should be interpreted. At the beginning of the interview, she had correctly articulated how to design an experiment to test whether or not mass influences the outcome. Yet, when asked to interpret data presented to her, she failed to apply a correct strategy. This failure is consistent with our findings from written questions administered to large numbers of students: A substantial fraction demonstrated understanding of this aspect of control of variables when designing experiments, but not when interpreting the results of experiments.

D. Rejection of experiments that include some trials that are not controlled

Some teachers had difficulty in reaching conclusions when the data contained a mixture of confounded and well-controlled trials. They often stated that conclusions could not be drawn because the experiment did not follow a prescribed form. For example, “These trials cannot be used because the variables must all be the same except [the one being tested]. In this experiment, [the controlled variable] fluctuates [in one of the three trials].”

Although the data contained trials that could be used to test whether the variable of interest influences the result, the presence of confounded data seemed to prevent this teacher from using the unconfounded trials. Teachers often indicated that the extra information invalidated the experiment and that a “more controlled experiment” must be done. Responses like these arose on all versions of the research tasks. The prevalence, however, was much lower among students in the general education and calculus-based physics courses than among K–8 teachers.

E. Complications encountered when variables do not constitute an independent set

In each research task discussed thus far, the data presented to students involved a set of variables that are independent of one another (that is, a change to one variable does not affect the others). In general, however, variables may be interrelated. For example, increasing the mass of an object while keeping the density the same increases the volume. We found in such situations that many students have serious difficulty in applying reasoning based on the control of variables.

Our initial awareness of this type of difficulty arose from informal interactions with students during instruction on sinking and floating in special content courses for preservice and inservice teachers. To investigate the nature and prevalence of these difficulties, we used a new research task. The question describes an experiment involving the sinking and floating behavior of four cylinders with the same volume but different mass. Two cylinders sink in water; the other two float. Students are asked to assess a fictitious dialogue (see Fig. 4).

Student 1 is correct, but not student 2. Mass and volume can be treated as independent variables. Because the volume has been kept the same, the behavior of the cylinders can be compared in order to infer that mass influences sinking and

Two students are discussing an experiment they have performed. Their experiment consisted of placing four cylinders of the same volume but different mass into beakers of water. Two of the cylinders sank and two floated.

Student 1: “This experiment tells us something about mass. The behavior changed when we kept the volume the same and changed the mass of the objects. So changing mass while keeping volume constant does affect sinking and floating.”

Student 2: “I disagree. To test whether mass influences the result you have to hold everything else constant. We didn’t hold density constant, which is mass over volume. So this is not a controlled experiment and we can’t say anything.”

With which student(s) do you agree? Be as specific as you can. Explain your reasoning.

Fig. 4. Task designed to probe student difficulties in reasoning with the control of variables in the context of dependent variables.

floating. Student 2 does not recognize that the change in density that results from the interdependence of mass and density is not a flaw in the experimental design. This question was administered to 230 students in the general education physics course. About 55% answered correctly, while 40% incorrectly agreed with student 2. In the following discussion, student responses are used to illustrate the most common conceptual and reasoning difficulties.

- *Tendency to treat experiments in which a dependent variable changes as being confounded.* Many students seemed to regard the cylinders experiment as being confounded because both mass and density changed. Responses like the following were typical: “I agree with student 2 because the density was different for each cylinder as well as the mass. If two different variables are changed then we can’t test for influence.” Even students who recognized the interdependence of mass and density were sometimes unable to infer the role played by mass: “I agree with student 2 because you can’t precisely say whether or not mass affects the object sinking or floating without keeping other factors constant. It is hard to tell though because density = mass/volume.” This student identifies the inherent difficulty in controlling a dependent variable. However, she fails to recognize that if mass and volume are treated as independent variables, this experiment is a valid test of whether or not mass influences sinking and floating. Such responses suggest rote knowledge of control of variables that is unconnected to the underlying hypothetico-deductive reasoning.

- *Belief that if a single variable determines the result, other variables cannot influence it.* Prior knowledge about sinking and floating seemed to prevent many students from reasoning on the basis of control of variables to make an inference about mass. Many insisted that neither mass nor volume is relevant because density is “the only variable that really matters.” One student elaborated as follows: “I agree with student 2—density definitely has an effect on whether something will sink or float. For example: a blimp filled with gas has a humungous mass, but is not very dense and will float.” This student’s answer is not based on the evidence presented to him. His reasoning does not explicitly involve the control of variables. Instead, he augments the cylinders experiment with another real world

example, the blimp, a specific case that illustrates what he already knows about the roles of density and mass in sinking and floating.

The responses of other students who drew on their prior knowledge suggest an underlying belief that because density is primary, no other variable can influence sinking and floating behavior. For example: "I agree with the second student, because density was the ultimate determining factor, mass alone cannot be contributed to the results [sic]." This student fails to recognize that the cylinders experiment provides evidence that mass influences sinking and floating. His reasoning suggests that he does not realize that even if a single variable determines the result, other variables may have an influence.²³

VII. IMPLICATIONS FOR INSTRUCTION

The process of science is often taught as a set of standardized procedures to be memorized and followed. Instruction on control of variables (if it occurs at all in a university science course) tends to follow this pattern. Students are often unable to go through the steps in reasoning necessary to deal with circumstances that differ from the particular examples used in instruction. We found that even students in the calculus-based physics course, with a relatively high level of mathematical facility, had difficulty in relating specific inferences to the experimental results that support them.

Science courses provide an opportunity to promote logical reasoning and the drawing of inferences based on evidence. The approach we have taken in helping students learn how to apply control of variables is to guide them through the required reasoning in a variety of contexts over an extended period of time.²⁴ Our research and teaching experience indicate that when instruction in reasoning is embedded in a coherent body of content, students can develop facility in a particular skill and also deepen their conceptual understanding of the associated content.²⁵ The process of going through the reasoning necessary to make inferences from experimental data helps students recognize the difference between what they know (because they have been told) and what they understand. We have used sinking and floating as the context for teaching control of variables to K–8 teachers and have found that the experience has increased their competence and confidence to do, learn, and teach science.²⁶

VIII. CONCLUSION

The research described in this paper focuses primarily on the reasoning necessary for deciding whether or not a particular variable influences the behavior of a system. Although we found that most of the students in our study could articulate the need to control variables when designing an experiment, many had difficulty with the underlying reasoning. We administered specially designed tasks to obtain detailed insights into student thinking and found that many students tried to draw conclusions about the effect of a variable even when multiple variables were changed. Conversely, some drew conclusions about whether a variable influences the behavior even when that variable had not been changed. Other students used the reasoning needed to decide that a given variable determines the behavior of a system when asked to decide whether that variable influences the behavior. Some students seemed to believe that only a single variable can influence the behavior of a system. Confusion between de-

pendent and independent variables and beliefs about the causal mechanisms relevant to a particular system also caused many difficulties.

Overall, we found that reasoning based on control of variables is challenging for students at all levels. Because the design of experiments and the interpretation of the results are central to the scientific process, it might be expected that students in calculus-based physics courses would have facility with this aspect of scientific reasoning. However, the tasks used in this study were challenging even for this population. Because most students, on their own, do not seem to reflect on the reasoning that underlies the control of variables, their difficulties need to be addressed explicitly, for example, in the laboratory component of a course.

An understanding of the control of variables is especially important for precollege teachers. Yet, even teachers who had previously taught this topic had considerable difficulty with the tasks used in this study. As described in nationally advocated standards for science education, a basic objective of instruction is for students to develop an understanding of science as a process.²⁷ They should know how they know what they know. To achieve this goal, teachers need to be able to guide students in interpreting and drawing inferences from data.

Our findings also have strong implications for the development of an educated citizenry. Data that are publicly presented in political debate are often not in a form that makes it easy to identify the effect of a particular variable. To vote responsibly, citizens need to be able to judge whether claims that are made in public discourse are consistent with the data. It is necessary not only to be able to interpret the information presented in order to draw conclusions, but also to recognize situations in which conclusions are not warranted.

ACKNOWLEDGMENTS

This paper builds on a project conducted by Lenore Hernandez for her M.S. degree at UW. Many members of the Physics Education Group, past and present, contributed to the research, curriculum development, and instruction. Deeply appreciated is the cooperation of Professor James Stewart and Professor Ajit Rupaal at WWU, who administered written questions to their students. The authors gratefully acknowledge support from the National Science Foundation under a series of grants including, most recently, DUE 0618185 and ESIE 0096511.

^{a)}Electronic mail: lcmd@phys.washington.edu

¹The courses for teachers are described in L. C. McDermott, "Preparing K–12 teachers in physics: Insights from history, experience, and research," *Am. J. Phys.* **74**, 758–762 (2006); L. C. McDermott, "A perspective on teacher preparation in physics and other sciences: The need for special science courses for teachers," *ibid.* **58**, 734–742 (1990).

²The preparatory course is discussed in L. C. McDermott, L. Pitemnick, M. Rosenquist, "Helping minority students succeed in science: I. Development of a curriculum in physics and biology," *J. Coll. Sci. Teach.* **9**, 135–140 (1980).

³L. C. McDermott, and the Physics Education Group at the University of Washington, *Physics by Inquiry* (Wiley, New York, 1996), Vols. I and II.

⁴The initial stages of this research are documented in L. Hernandez, "Preparing teachers to teach science as a process of inquiry: The development of scientific reasoning skills," M.S. Thesis, University of Washington 2001 (unpublished).

⁵An article examining student understanding of the reasoning involved in testing whether a variable determines an experimental outcome is in preparation. A. Boudreaux, P. S. Shaffer, P. R. L. Heron, and L. C. Mc-

- Dermott, "Student understanding of control of variables: Deciding whether or not a variable determines the behavior of a system."
- ⁶ An overview of some research on student ability to formulate hypotheses, design experiments, and organize and interpret experimental results, M. Njoo and T. de Jong, "Exploratory learning with a computer simulation for control theory: Learning processes and instructional support," *J. Res. Sci. Teach.* **30**, 821–844 (1993); J. Klayman and Y. W. Ha, "Confirmation, disconfirmation, and information in hypothesis testing," *Psychol. Rev.* **94**, 211–228 (1987); D. Klahr, A. L. Fay, and K. Dunbar, "Heuristics for scientific experimentation: A developmental study," *Cogn. Psychol.* **24**, 111–114 (1993); C. A. Chinn and W. F. Brewer, "The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction," *Rev. Educ. Res.* **63**, 1–49 (1993); D. Kuhn, M. Garcia-Mila, A. Zohar, C. Andersen, S. H. White, D. Klahr, and S. M. Carver, "Strategies of knowledge acquisition," *Monogr. Soc. Res. Child Dev.* **60**, 1–157 (1995).
- ⁷ T. de Jong and W. R. van Joolingen, "Scientific discovery learning with computer simulations of conceptual domains," *Rev. Educ. Res.* **68**, 179–201 (1998).
- ⁸ Z. Chen and D. Klahr, "All other things being equal: Acquisition and transfer of the control of variables strategy," *Child Dev.* **70**, 1098–1120 (1999); E. E. Toth, D. Klahr, and Z. Chen, "Bridging research and practice: A cognitively-based classroom intervention for teaching experimentation skills to elementary school children," *Cogn. Instruct.* **18**, 423–459 (2000).
- ⁹ A. Keselman, "Supporting inquiry learning by promoting normative understanding of multivariable causality," *J. Res. Sci. Teach.* **40**, 898–921 (2003).
- ¹⁰ J. Park and S. Pak, "Students' responses to experimental evidence based on perceptions of causality and availability of evidence," *J. Res. Sci. Teach.* **34**, 57–67 (1997). Also see the second article in Ref. 8.
- ¹¹ J. Park and I. Kim, "Analysis of students' responses to contradictory results obtained by simple observation or controlling variables," *Res. Sci. Educ.* **28**, 365–376 (1998). There is also a significant body of research from psychology on confirmation bias. See, for example, P. C. Wason, "On the failure to eliminate hypotheses in a conceptual task," *Q. J. Exp. Psychol.* **12**, 129–140 (1960); C. G. Lord, L. Ross, and M. R. Lepper, "Biased assimilation and attitude polarization: The effect of prior theories on subsequently considered evidence," *J. Pers. Soc. Psychol.* **37**, 2098–2109 (1979).
- ¹² A. E. Lawson, "The development and validation of a classroom test of formal reasoning," *J. Res. Sci. Teach.* **15**, 11–24 (1978).
- ¹³ B. Inhelder and J. Piaget, *The Growth of Logical Thinking From Childhood to Adolescence* (Basic Books, New York, 1958).
- ¹⁴ For example, the Washington State Assessment of Student Learning typically has questions intended to assess student ability to design controlled experiments and interpret the results.
- ¹⁵ Y. Shadmi, "Teaching 'control of variables' to primary school teachers," *Phys. Educ.* **16**, 93–98 (1981).
- ¹⁶ A. J. Ross, "Controlling variables: A meta-analysis of training studies," *Rev. Educ. Res.* **58**, 405–437 (1988). A review of relevant studies. See also Refs. 8 and 9.
- ¹⁷ Some studies have found that students tend to draw conclusions on the basis of insufficient experimental evidence. See, for example, Refs. 9 and 10.
- ¹⁸ Whether or not a variable influences the behavior of a system depends on the choice of independent variables. Consider, for example, sinking and floating in water. If the independent variables for a particular experiment are selected to be mass and volume, then changing mass while holding the volume constant affects the outcome. In contrast, if the independent variables are mass and density, then changing the mass while holding the density constant does not affect the outcome. This relatively sophisticated level of reasoning about control of variables was not expected of students in research tasks 1, 2, and 3.
- ¹⁹ A. B. Arons, *A Guide to Introductory Physics Teaching* (Wiley, New York, 1990). The importance of hypothetico-deductive reasoning is discussed.
- ²⁰ The results of the experiment do not prove that the variable (distance) does not influence the behavior of the system. It is possible that the variable was not changed sufficiently for the effect to manifest. We accepted, as correct, answers in which students stated that the variable might affect the behavior of the system it were tested over a wider range of values.
- ²¹ D. E. Trowbridge and L. C. McDermott, "Investigation of student understanding of the concept of velocity in one dimension," *Am. J. Phys.* **48**, 1020–1028 (1980); D. E. Trowbridge and L. C. McDermott, "Investigation of student understanding of the concept of acceleration in one dimension," *Am. J. Phys.* **49**, 242–253 (1981).
- ²² Teachers seemed especially hesitant to use the data in early versions of the second research task. These versions were identical to the one shown in Fig. 2 except that the teachers had not been told that many repetitions of each trial had been made. Teachers were also asked to use the data to decide whether or not a particular variable influences the outcome. The question was changed to the form described in this paper to try to focus the attention of teachers on the reasoning associated with the control of variables. In the final version, teachers were told that several repetitions of each trial had been made and the same value was obtained each time. The final version also has two parts. Teachers are first asked if the data can be used to test the idea that a given variable influences the outcome, and second, if it can be so used, to state which trials should be compared and whether the results support or refute the idea. The change to the question seemed to make a difference. On the latter version, a greater fraction of teachers applied reasoning based on the control of variables.
- ²³ We have found that situations like the one described, in which a dependent variable determines a result while an independent variable influences it, are exceedingly difficult for some students. Even extended in-class discussions with individual students on this topic are not always productive. The workshops on *Physics by Inquiry* that we conduct for high school physics teachers, graduate students in the sciences, and two- and four-year college faculty in mathematics and science reveal that these difficulties extend even to populations with substantial experience and background in math and science. Student understanding of the reasoning associated with testing whether or not a variable determines an outcome is discussed in Ref. 5.
- ²⁴ The overall approach we have taken is illustrated by the *Properties of Matter* module in Ref. 3.
- ²⁵ We have evidence that students who have studied sinking and floating in this way have a greater functional understanding of the topic than those who received more traditional top-down instruction that emphasized causal mechanisms (for example, the buoyant force). See P. R. L. Heron, M. E. Loverude, P. S. Shaffer, and L. C. McDermott, "Helping students develop an understanding of Archimedes' principle: II. Development of research-based instructional materials," *Am. J. Phys.* **71**, 1188–1195 (2003).
- ²⁶ A paper that documents the effect of an inquiry-oriented instructional approach to teaching scientific reasoning, including the control of variables strategy, is in preparation.
- ²⁷ Benchmarks, published in 1994 by AAAS Project 2061, are recommendations for what students should know and be able to do at particular grade levels.