RESEARCH REPORT

Science ASSISTments: Tutoring Inquiry Skills in Middle School Students

An Interactive Qualifying Project Report

Submitted to the Faculty

Of the

Worcester Polytechnic Institute

In partial fulfillment of the requirements for the

Degree of Bachelor of Science

By:

Jacqui Richardson

May 28, 2008

Approved By:

Professor Janice Gobert

Abstract

This paper outlines the four design studies conducted in a local middle school in order to: 1) characterize how middle school students naturally approach a scientific problem, 2) develop an effective set of tutoring prompts for inquiry within a technical environment, and 3) determine the effectiveness of various prompts and scaffolding tasks at improving inquiry skills.

Table of Contents

Introduction

The Math ASSISTments Program

Computer Science Assistant Professor Neil Heffernan of Worcester Polytechnic Institute (WPI) and his colleagues at Carnegie Mellon University (CMU) are the developers of the ASSISTments program (www.assistments.org). This "intelligent tutoring" system is a web-based system for Math that provides example problems and reacts to students' inputs in real time. Math ASSISTments is presently used by 3,000 students and 40 teachers in Massachusetts with the goal of improving students' math learning. One of the measures used to assess student improvement is the Massachusetts Comprehensive Assessment System (MCAS) score for Math. There are various types of mathematical concepts covered in the ASSISTments problems to improve MCAS scores.

Within the ASSISTments system, there are three methods of guiding students through a mathematical problem: hints, buggy messages, and scaffolding. Hints and scaffolds are provided at the request of the student when they are struggling to answer a question. Each hint or scaffold gets progressively closer to the correct answer. Within the ASSISTments system, the difference between a hint and a scaffold is that hints are directed toward the current problem that the student is working on; scaffolds allow the student to see outside the context of the problem at hand by breaking down the concepts of a complicated problem to simplify the question. Buggy messages guide the student toward the right answer when they've already submitted an incorrect answer. Since the students' grades are affected when they press the hint button, they do not overuse it.

Students and teachers interact well with the program. It has become part of their weekly routine and the students have fun interfacing with the technical environment. Students are encouraged to post comments for the software developers so that the program can be tailored to their particular needs. The system also allows the teachers to track and assess students' grades and progression.

Science ASSISTments in the Making

In 2007, the ASSISTments program was extended to include science, specifically, Physical Science. Two professors at WPI, Janice Gobert and Neil Hefferenan, lead a team of graduate and undergraduate students. Professor Gobert's area of expertise is Cognitive Science and Science Education. In particular, she co-directed the Modeling Across the Curriculum project (mac.concord.org) in which she and her team developed micro-worlds for students to explore science phenomena in the areas of Biology, Physics, and Chemistry. Her expertise combined with Professor Heffernan's ASSISTments program and technical background in Computer Science form a solid team that is focused on improving inquiry skills in middle school students.

The science MCAS test currently assesses only direct knowledge. It does not measure students' inquiry performance. This project seeks to provide measures of students' inquiry and as such, this effort represents a significant advancement in science assessment. To date there have been two approaches taken to assess inquiry knowledge: short answer tests of specific inquiry skills (cf., Alonzo & Aschbacher, 2004), and hands-on paper and pencil assessment tests (cf., Gotwals & Songer). Short answer tests *can* be incorporated into large-scale standardized assessments, however, it is unclear whether these properly identify inquiry (Black 1999; Pellegrino 2001). Hands on, i.e., performance assessments, are more authentic because they require specific skills to solve real problems (Baxter and Shavelson 1994; Ruiz-Primo and Shavelson 1996). This along with the high costs associated with performing inquiry assessments, is the reason inquiry is seldom assessed in schools.

The team on the Science ASSISTments project will respond to state needs, namely, they will develop inquiry measures for science, as well as scaffolds for students' inquiry learning. Their overarching hypothesis is that teaching students' inquiry skills will support them in deep learning of science content. The Science ASSISTments team will address learning using both items from past MCAS exams as well as inquiry items that they will develop. This present paper is focused on the development of inquiry scaffolds for the science ASSISTments program.

Why Teach Inquiry?

Inquiry is important in order to perform logical scientific experiments as well as to make decisions in real-life scientific problems such as testing the affect of different variables on the outcome of an experiment. These skills are not generally developed naturally (Kuhn, 2005) and thus, many students lack these important skills. Also, these skills are not often tested in the school system due to the amount of time and effort that it takes to assess inquiry. Inquiry skills are defined as (NSES, 1996):

- *Hypothesizing*: making predictions about the outcome of an experiment
- *Conducting an experiment*: developing a method of testing the hypothesis
- *Collecting data*: correctly recording findings while performing the experiment and setting up tables of variables
- *Mathematizing*: performing calculations and writing equations to formally describe collected data

- *Interpreting data*: forming explanations that are consistent with data collected
- *Communicating*: articulating the steps of inquiry (describing results, etc.)

The long-term goal of this project is to improve inquiry skills in middle school, so that students can transfer this knowledge to other science topics, and ideally to real-world problems and improve upon content learning as well. Research suggests that these skills should be developed during middle school (Schunn, Raghavan, & Cho, 2007).The extent to which the idea of teaching inquiry skills so that they then transfer to new domains is feasible will be determined by the outcomes of tests and experiments that are to be performed in schools as part of the Science ASSISTments project.

Adapting the Math ASSISTments System for New Science ASSISTments project (Gobert, J., Beck, J., Sao Pedro, M., Heffernan, N., & Richardson, J., 2008))[1]

The Math ASSISTment system is a server-based (thin client) tutoring system implemented using Ruby on Rails, JavaScript, and XHTML/CSS stylesheets. There are several advantages to this server-based approach. A thin client implies that the software will be run through a web browser and thus not installed on their local machine. Therefore, the time and cost of installing software is virtually eliminated. We also gain greater control over content distribution since we can control what new content becomes available when for all users irrespective of their location. Updates and configuration changes are also more easily manageable. This approach also centralizes data collection efforts meaning that information is not stored on each student's machine. Finally, running in a browser abstracts from creating

[1] This section was written by Michael Sao Pedro for a larger document on the ASSISTments infrastructure

distributions for different operating systems, thus supporting Windows or Mac users. There are some disadvantages to this approach, particularly scalability and responsiveness.

The ASSISTment architecture shown in Figure 1 combats these disadvantages in multiple ways. The system is housed on a load balancing cluster with several computers and multiple application servers (called Mongrel servers) that distributes the processing workload. By adding more computers with application servers, the system can theoretically be scaled up to support more users. Additionally, smart indexing on frequently used values in the database drastically reduces the time of responding to user actions. Most importantly, the implementation pushes more of the processing onto the client.

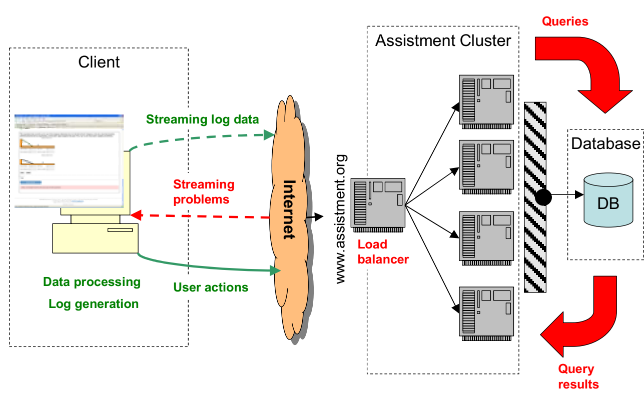


Figure 1: ASSISTments architecture

The servers are responsible for on-demand content distribution and log gathering but push data processing (such as determining the correctness of an answer) onto the client. This greatly reduces the load on the database and servers dramatically since the client does not need to query back to the server for information nearly as much. At the start of an assignment, the client requests and stores questions asynchronously while the student begins working. Network or server delays are hidden by the asynchronous communication; subsequent questions are streaming in the background. Since the content is stored on the client, the progression through an assignment is handled by the client with only log messages sent back to the server.

There are other constraints particular to Science ASSISTments we must overcome to create interactive environments that can be run client side. These environments must download quickly, be responsive, be reusable so as to create multiple kinds of questions involving them, portable, and able to log user actions, not just answers to questions. Authoring such environments to be run within a browser is difficult since browsers currently do not natively support the notion of a 'canvas' whereby one can author an animated, graphical environment with ease. Thus, a third-party environment such as Flash is required.

Our governing system design goal is to extend the ASSISTment system to handle questions involving interactive environments. This enables us to reuse the existing framework and infrastructure that manages student and teacher accounts, lets us create and assign problem sets involving these environments, send streams of problems to the client behind the scene, and asynchronously logs student actions and results. We extended the system in two ways. First, we extend the problem framework to allow Flash objects to appear inside individual problems as shown in Figure 1. Previously, only static images could appear inside problems. Second, we enhanced the tutor component to interpret the results from the Flash object and log any actions. Another benefit is that the existing rendering system with all of the correct/incorrect graphics is reused as well.

We chose to implement our interactive incline plane environment in Laszlo (www.openlaszlo.org), a new rich internet application authoring language. We chose this language for several reasons. First, it is a higher-level language can be compiled to Flash or Dynamic HTML. This affords us implementation flexibility down the line if we need it. Second, Laszlo has a number of built-in mechanisms to support drawing on a canvas and animations. Third, Laszlo is object-oriented affording us more reusability opportunities.

Computer Micro-worlds

Science ASSISTments differs from Math ASSISTments in that science problems tend to be more open-ended. In Science ASSISTments, students are given a specific goal and time to explore with one of the micro-worlds. An example of what a Science ASSISTments micro-world might look like is shown in Figure 2. This micro-world was adapted from Professor Gobert's previous work on the MAC project (mac.concod.org).

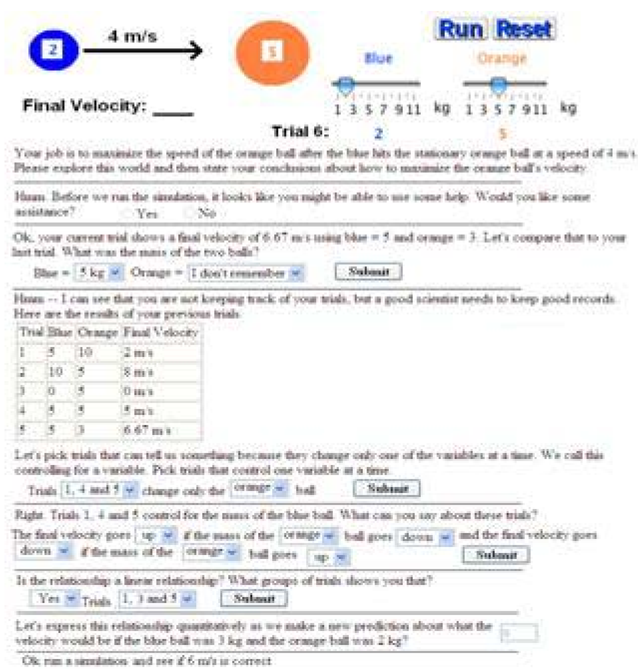


Figure 2: A mock up ASSISTment Micro-World

In this example, the student is given a specific task: "To maximize the speed of the orange ball after the blue hits the stationary orange ball at a speed of 4 m/s". The student then runs the experiment by changing the mass of the two balls to see the effect that mass has on the final velocity upon collision of the blue ball with the stationary orange ball.

Background to the Studies Conducted

David Klahr and Milena Nigam performed an inquiry test on the control for variables strategy using a wooden ramp with variable height, steepness, and run length. Klahr and Nigam' study consisted of 112 third- and fourth-grade students. The goal of the study was to compare the results of direct versus discovery learning of the control for variables strategy. Discovery learning is accomplished with no outside guidance. Direct learning in this context is the direct explanation of the control for variables strategy. Klahr's experiment addressed the common conception that "the knowledge students construct on their own is more valuable than the knowledge modeled for them" (Klahr & Nigam, 2004).

For the 'control for variables' study, Klahr and Nigam used the ramp apparatus shown in Figure 3 (a) and (b). The apparatus consisted of two wooden ramps. Each ramp had adjustable slopes on both sides, one of which was a stepped surface to impede the motion of the ball. Students were able to change the steepness of these ramps (high or low) on both sides as well as the type of ball (rubber or golf), the surface of the ramp (rough or smooth), and the run length (the ball could be released from the full length of the ramp or at the shorter length as show in Figure 3A). They were asked to determine how each of these variables affected the outcome of the experiment (how far the ball rolls).

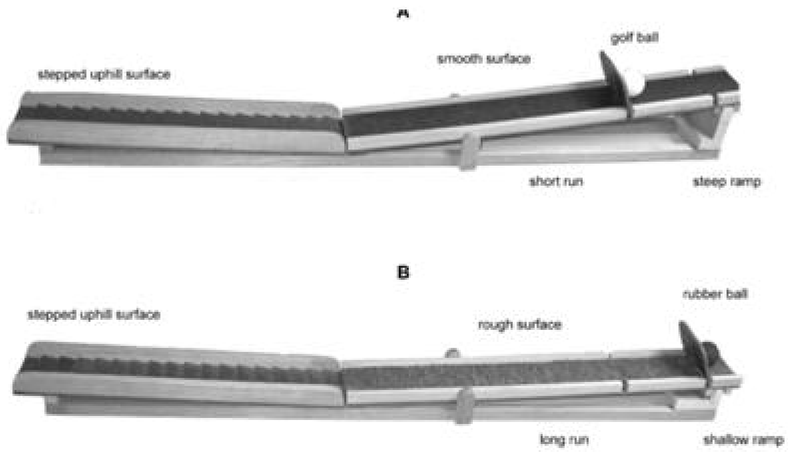


Figure 1: Klahr and Nigam's ramp apparatus

The apparatus and objective were explained to all students. They were asked to determine the effect of steepness and run length, separately, on how far the ball rolled. For the discovery learning condition, the students were not given any guidance by a teacher other than the goal of their experiment. They were told to interact with the ramp apparatus on their own and were given no feedback. For the direct learning condition teachers were present during the experiment and gave them a specific set of instructions regarding the control for variables approach. The results of Klahr's experiment revealed that students who were given direct instruction had improved significantly on their control for variable skills. Discovery learners showed little or no improvements on their ability to run a control for variables experiment. To test for transfer effects of discovery versus direct learning, Klahr and Nigam also gave the students a posttest the following week. Students were given two science fair projects created by a fellow student which they were asked to critique. One science fair poster was titled "Who has better memory: Boys or girls?" and the other was "Does the number of holes affect how far a Ping-Pong will fly?" Again, the direct learners excelled by outperforming the discovery group. More than 75 percent of direct learners mastered the control for variables strategy, whereas less than 25 percent of discovery learners achieved this mastery (Klahr & Nigam, 2004). Direct learners were able to transfer what they had learned from the previous week about the control for variables strategy in order to critique another students' work.

The ASSISTments Ramp Experiment

The Science ASSISTments team at WPI used the results of Klahr and Nigam's ramp concept to conduct their own inquiry experiments on the control for variables strategy. Four tests were conducted: two with a physical ramp apparatus and two with a simulated ramp apparatus similar to Klahr and Nigam's model. These studies were focused on discovering ASSISTments prompts to improve students' inquiry approach to the control for variables strategy.

The Physical Ramp Apparatus

The apparatus used by the WPI team is shown in Figure 4. Like Klahr and Nigam's experiment the wooden ramp apparatus had variable steepness, run length, type of ball, and surface. One major difference between the ramps was that instead of a second ramp with notches, this ramp had a wire-gridded gate at the base of the ramp that used to impede the motion of the ball. The students used each row of the grid as a unit of measurement for their experiment.

Figure 2: The ASSISTments physical ramp apparatus

The students could chose from a smooth, heavy golf ball or a rough-surfaced, light rough-surfaced ball of approximately the same size. The ramp could be set at high or low steepness by adjusting the block underneath the ramp. This feature is shown in Figure 5. It can also be seen that there are two pieces of black tape on the side of the ramp. These marks indicate the halfway and full length of the ramp where the ball could be released for variation in run length.

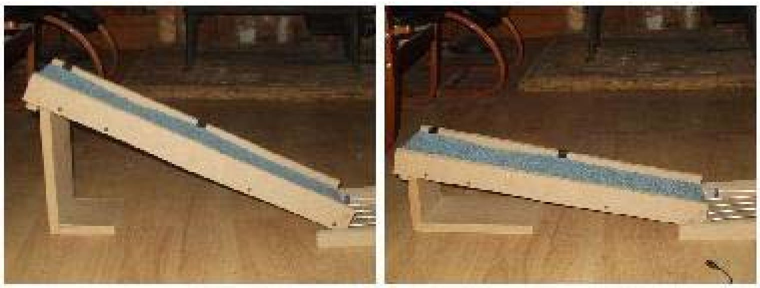


Figure 3: Variable height of the ASSISTments ramp apparatus

The students were able to choose between the smooth wooden surface of the ramp or the rough carpeted surface custom-fit to the slope. Both surfaces are shown in Figure 6.

Figure 4: Variable surface of the ASSISTments ramp apparatus

Simulated Ramp Apparatus

Finally, the ramp apparatus was simulated as an ASSISTments' computer micro-world. This simulated ramp apparatus shown in Figure 7. Like the physical model students can manipulate the steepness, run length, type of ball, and surface to see how these factors affect how far the ball rolls.

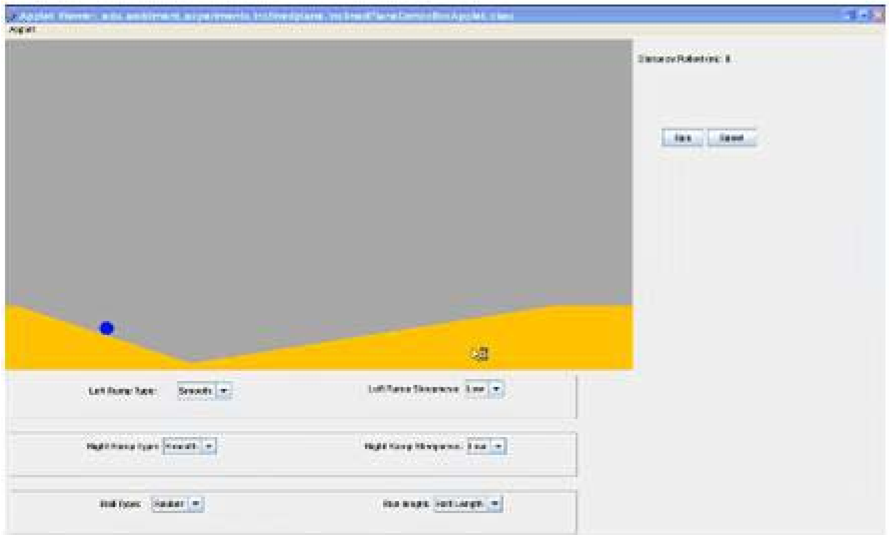


Figure 5: Simulated ramp apparatus

It can be seen from Figure 7 that there are two ramps included in the simulation. This ramp design more closely follows Klahr's model.

To run the simulation, the student adjusts the ramp settings and hits the 'Run' button. When this button is pushed the ball moves based on the laws Newtonian Dynamics and rolls down the left side of the ramp to the right ramp. The distance is then calculated in the upper right-hand corner. To start a new trial, the 'Reset' button must be hit.

Case Studies

To date, 14 middle school students from a Worcester Public School have participated in the science ASSISTments ramp experiment. Four case studies have been completed. For each study three or four participants were randomly selected, each from a different class level. These studies are explained in detail in the following sections.

The Ramp Experiment: Test Set 1

The first test set was designed to determine how students naturally approach an inquiry problem with respect to their use of the control for variables strategy. Specifically, this study was conducted to get a sense of what the students already understood in terms of inquiry and to detect common areas of weakness. Another goal of this study was to begin to test out different types of prompts. If any disorganized inquiry practices were detected, such as changing too many variables at once or failing to record data, the WPI team provided prompts with suggestions on better inquiry strategies to guide the student in the right direction. Hence this test set was interactive between the experimenters and the students.

Case Studies

Three seventh-grade students in a public middle school from Central Massachusetts constituted the sample set for the first ramp experiment. Experimenters were WPI undergraduates on the ASSISTments project team.

Method

Materials

In this study, students were presented with the physical ramp apparatus, previously described in the 'Physical Ramp Apparatus' subsection of 'Science ASSISTments Ramp Experiment', and given only blank pieces of scrap paper and a pencil to record their findings.

Procedure

Students were tested on an individual basis. They were first asked what they thought they could change on the ramp that could affect how far the ball would roll. They were then told that these factors were called variables. Next, they were asked to state a hypothesis and run an experiment to test their hypothesis on how the steepness of the ramp affects how far the ball rolls. After running the experiment, the students were asked, based on their data, to explain how steepness affects how far the ball rolls.

When the student finished testing how steepness affects how far the ball rolled, they moved on to test the effect of run length, the type of ball, and surface again on how far the ball rolled. Then they were asked to reflect back on the first experiment and say what they would have done differently if they were to run it again. This question was asked to test if they had acquired meta-knowledge about the control for variables strategy.

Prompts

The prompts in this test set were not prewritten or rehearsed. The prompts were intended to be a "gentle guide" toward improving the student's strategies. After the student was prompted, they were given more time to interact with the apparatus and revise their strategy and re-run the experiment. If the student continued to display the same type of behavior, the next prompt given was slightly more direct. For example, if the student was having organizational issues they would be prompted with a comment such as "try making a table". If they were still being unorganized in their data collection, the next prompt may have been "Here is a table. Fill in the values." This strategy of giving progressively more direct hints has shown to be successful in the Math ASSISTments project.

Data Analysis

All of the notes taken and tables made by each student during the experiment were collected. Additionally voice data was analyzed to determine on which of the 6 inquiry strands the students struggled and which prompts were helpful in improving inquiry skills, specifically the control for variables strategy. These data were collected using a voice recorder so the students could communicate their results. The inquiry in this study were the forming a hypothesis, conducting an experiment to test this hypothesis, collecting data, and analyzing the data. Control for variables strategy is a specific strategy under the larger inquiry skill of conducting experiments. In these voice recordings each of these inquiry strands was analyzed to compare any common weaknesses or strengths across students. The findings are described later in the 'Discussion of Test Set 1' section. All notes and tables for these case studies can be found in Appendix A.

Results

Participant 1

A student from a high-level class was the first subject. He was first asked how steepness affects how far the ball rolls and to run an experiment to test it. He hypothesized that if the ramp was less steep that the ball would not go as far, which is a complete hypothesis that makes mention of both the independent and the dependent variable in the experiment and also the relationship between them.

Three minutes and thirty seconds into the interview the student concluded that the golf ball goes further. When asked if that answers the question on how the steepness of the ramp affects how far the ball rolls, he replied "For the golf ball, the lower slope was better, but for the tennis ball the higher slope was better."

Approximately 5 minutes went by in which the student experimented the ramp apparatus but he was not recording his findings. He was prompted to make a table. Thirty seconds later the student said that he needed help making his table. When asked what he thought he needed to include in the table, he said that he needed the "type of ball, the incline, and the distance the ball rolled”. Seven minutes into the interview the student needed to be scaffolded again in terms of lack of productivity. The experimenter helped the student fill out the first few trials of the experiment in his table. After this interaction with the experimenter, the student started successfully recording his data on his own.

Eleven minutes and fifty-seven seconds into the interview the student was asked to answer steepness affects how far the ball rolls based on his data. He responded, “When the steepness is higher, the ball goes further.” When asked where he saw that in his table, the student pointed out that the data for the higher steepness for both the golf ball and the tennis ball were higher than the distances for the golf ball and tennis ball at the lower steepness, thus the student was clearly changing too many variables (steepness and the type of ball) at once. The student’s complete first table is shown in Table a. of Appendix A1.

Next the student was asked to run an experiment to determine how run length affects how far the ball rolls. He was asked what should be included in his table this time around. He said that he needed to include the type of ball, the run length, the distance the ball rolled, and the incline. The resulting Table is b. in Appendix A1. This table shows that the student tried the tennis ball at the high steepness at the full length and halfway point and then tried the golf ball at the low steepness at the full and halfway points. The student at this point was shown that he did not need to include the steepness of the ramp in his table. Here the student agreed that he did not need it because he already knew the effect of steepness on how far the ball rolled, indicating that his control for variables strategy had improved.

The next task was to test how surface affects how far the ball rolls. The student created Table c. in Appendix A1. This table represents a large improvement from the previous one. The student did not change the run length or the steepness of the ramp to determine the effect of surface. In fact, they did not change anything except for the surface of the ramp. It can also be seen that he took four data points for each. When asked if his experiment proved his hypothesis, he pointed out from his table that the distances for the ramp with the smooth surface were higher than the distances for the ramp will the rough surface.

Even though this table was a major improvement from the previous tables, when asked how he would run the experiment for how steepness affects how far the ball rolls if he were to rerun it, however, he went back to saying that he would have changed the run length. Since the student revisited his previous mistakes, it was evident that he did not fully understand the control for variables concept.

Participant 2

The next student was a student from a medium-to-high level class. When asked to hypothesize how steepness affects how far the ball rolls, he said "the lower it is, the less distance". He was then asked to run an experiment to test his hypothesis.

Five minutes and thirty seconds into the interview the subject was prompted to make a table. When asked what he should include in his table he said that he needed the distance, the height of the ramp, and the type of ball, and the slant of the table on which the ramp was resting. The student was then told that it could be assumed that the table was perfectly flat.

Nine minutes and forty seconds into the interview the student stated that he already understood the affect of steepness on how far the ball rolls. However he still had not recorded any data. At this point, it was clear that the student was having difficulties recording his data and he was scaffolded with an explanation of how to record the first few trials. Approximately one minute later, the student had only collected one more data point for each and was told that he should gather more data points in order to prove his hypothesis. From here the student collected five trials of the golf ball at the high steepness and then the golf ball at the low steepness. The first table is Table a. shown in Appendix A2. The table showed a great improvement in the students' concept of the control for variables strategy. Specifically, his data is consistent in displaying that at the longer run length the ball goes further.

When asked to analyze his data, the student said that "It doesn't matter how high, it matters how much speed the ball gets." His argument here was that the tennis ball at the higher steepness went further than the tennis ball at the lower steepness, but for the golf ball the data points were the same because in either case the ball went to the end of the runway. The students' answer was not scientifically accurate. However, his explanation was consistent with his data. This indicated a flaw in the design of the ramp, namely that the length of the gate to stop the ball was not long enough because in both cases, at the high and low steepness, the golf ball reached the end of the runway and the actual outcome could not be determined.

Next, the student was asked to run an experiment to determine how run length affects how far the ball rolls. When asked what variables he needed to include in his table, he stated that he needed to include the run length, the type of ball, and the distance the ball rolled. The table for run length, labeled 'b.' is displayed in Appendix A2. As can be seen from the table, data in this table were much more organized than the first attempt.

Approximately four minutes after being asked to run an experiment on run length, the student was asked to explain the results of his experiment. His response was that "if you roll it at half the distance, the ball doesn't pick up as much speed so it doesn't go as far. But if you roll it at the full distance it picks up speed and goes further." This answer was rich in terms of both content and inquiry because it was both scientifically accurate and provides a clear relationship between the independent variable (run length) and dependent variable (distance). At this time the student only had taken one data point for each and was prompted to collect more data points. The students' conclusion was the same a few minutes later after taking a few more data points, but nowhere in the explanation did the student refer to their table. Here the experimenter (a WPI undergraduate on the Science ASSISTments team) intervened by showing him exactly where his table provided evidence for the point that he was trying to make. When asked what he would change if he were to run the experiment on steepness again he responded "I would roll this ball at the low incline and then I would test it on the high steepness and collect more data points."

Shortly after the student was asked to run an experiment on how surface affects how far the ball rolls. Three minutes into the experiment he started changing the incline again. The student concluded that "the rug on the lower incline goes slower", even though he had only tried the experiment on the rough surface and not on the smooth surface. The students' table is shown in Appendix A2 and is labeled 'c'. At this point it was brought to his attention that he needed to run the experiment on the smooth surface as well to compare the results of the different surfaces. Also, it was shown by the experimenter that the incline was not necessary in their table because he already tested how the steepness affected how far the ball rolls. With this information, he created a new table shown in Appendix A2 labeled 'd.' Other than changing the type of ball, the students' data clearly shows that on the smooth wood surface the ball goes further than on the carpeted surface.

Participant 3

The third student was a student from a lower level class, but it was found later that this student was a special needs student with some learning disabilities. When asked how he thought steepness affects how far the ball rolls, he hypothesized “If the ramp is high it will gain more speed, if it’s lower it could go either way.” He also hypothesized that the golf ball would go faster “because if you throw a snowball down a hill and it gets bigger and heavier, it goes faster.”

Four minutes and twenty seconds into the interview when the student still had not recorded any data, he was prompted to make a table. When asked what he needed to include in this table, he claimed that he needed the distance the ball rolled, the height of the ramp, and the type of ball. Seven minutes and fifteen seconds into the experiment, the student had only performed trials at the high steepness. He was then prompted to try the ramp at the lower steepness to compare his results.

Approximately one minute later, he concluded that “the higher the incline the further it goes”, however, he only collected one data point for each trial. When asked if he thought he needed more data to prove his point, he stated that he already had enough. The importance of running an experiment multiple times was then explained to him and the WPI undergraduate ran a few more trials of the experiment to establish the point. The student’s first table is shown in Appendix A3 and is labeled ‘a.’

Next, the student was asked to perform an experiment on how run length affects how far the ball rolls. In his table he included the run length, incline, surface, type of ball, and the distance that the ball rolled. This is table b. in Appendix A3. It can be seen from the table that at no point does his data show the effect of run length alone. The student changed too many variables at once.

The student was then prompted to keep the surface smooth and the steepness low and only change the ramp length. With this, he went on to create a new table shown in Appendix A3 called ‘c’. The table is more organized than the previous one.

When asked to make a table on how surface affects how far the ball rolls, the student made the table labeled ‘d’ in Appendix A3. Here it can be seen that the student goes back to including the incline and type of ball. For the remainder of the time, the student argued that you need to change all of the variables at once to determine the effect on how far the ball rolls because they all matter in the experiment.

Discussion

Prior Knowledge

During the experiment all of the students mentioned in the voice recordings mentioned that they had a prior knowledge of variables from their science class. However, none of the subjects showed any previous mastery of the control for variables concept.

Forming a Hypothesis

For the most part the students did not have much difficulty stating their hypotheses. The students’ answers contained the independent variable (e.g., steepness) the dependent variable (distance rolled), and stated the affect that they thought the independent variable would have on the outcome of the experiment. The answers were also strong in terms of content. It seemed like common sense that the steeper the ramp the further the ball will roll, but as explained in the following sections the students had trouble setting up experiments to prove their hypotheses.

Conducting an Experiment and Recording Data

One of the common inquiry problems revealed in this experiment was that students struggle in setting up an experiment and recording their data. Approximately five minutes into each experiment, the student was prompted to make a table because they hadn't written anything down. When prompted to start recording their data, they still had difficulties setting up the table on their own and recording their trials. After they were scaffolded for one or two trials, however, their table-making skills were much better throughout the rest of the experiment.

All of the students did well on the first experiment on how steepness affects how far the ball rolls, but as new variables were presented to them, their tables became more complicated because they felt that they also needed to include the steepness of the ramp. However, after various prompts, the first two students started to realize why it is necessary to change only the variable which they are trying to determine the effect on how far the ball rolled.

Analyzing Data

Students struggled to analyze the data collected in their tables. Conclusions were mostly based on what they already knew about the effects of a certain variable instead of what they actually found from their experiment. Only the first student referred to his table when answering the questions. For the other two students, the information was in their table, but they insisted that the answer was obvious. For these students, the WPI student pointed out where their data proved or disproved their hypotheses.

The Ramp Experiment: Test Set 2

The second study was more tightly structured than the first. From the first data set, it was determined that students were having trouble recording their data in an organized fashion and understanding the overall control for variables concept. In the second test we sought to provide more scaffolding for these two aspects of inquiry; we also scaffolded the scientific method which involves many important inquiry skills such as being able to form a hypothesis, collecting and analyzing data, and coming up with a conclusion based on the data collected.

Participants

For the second ramp experiment three students were chosen at random from three different classes. These classes were categorized as high, medium-to-high, and medium-to-low level, respectively. A fourth student was not selected due to an early school day.

Method

Materials

For this study the students were given a laboratory notebook. The laboratory notebook, displayed in Appendix B3, included a hypothesis sheet, a results sheet, and an immediate posttest. Scrap paper was given to the students to record their data. Two prompts were written for the students to improve their inquiry strategies. The prompts can be seen in Appendix B5.

Procedure

Pretest

The pretest, given a week prior to the second test set, was the pendulum test. This test is included in Appendix B1. Students were given a brief description of what a pendulum did. Then the task was to find out what happens to the pendulum when you change the length of the string. The student was given the diagram shown in Figure 8.

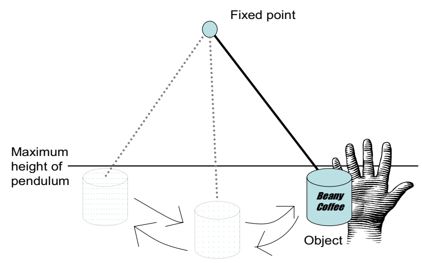


Figure 6: Picture of the pendulum experiment

The student was first asked to form a hypothesis. Next they were asked to come up with an experiment that would determine how the length of the string affects what happens to the pendulum. The student was asked to make a table as if they had actually run the experiment by making up their own data to see how the student would record their data if they actually ran the experiment. Finally, they were asked to analyze their "data" and write a conclusion based on these data.

Prompts

From the first test set it was determined that students were having trouble with recording their data and the concept of the control for variables concept. There were two prewritten prompts written to address these issues. This prompt is shown in Appendix B5. If the student was changing too many variables at once, they were prompted with the checklist displayed in B5.2. This checklist contained every variable in the ramp experiment. First the student was asked to check off all the variables that they needed in the experiment they were conducting. If they checked off too many, they were given explanations of why they did not need to include those variables.

If the student was having difficulties with organization and recording their data a multiple choice question was given. This question is shown in Appendix B5.1. In this question the student was presented with three tables. The table in answer A was incorrect because there were not enough data points taken. Also, to test how steepness affects how far the ball rolls, the type of ball does not need to be changed. This was a “curveball item” because a common mistake between the students’ in the first test set was to change the type of ball. The table in answer B was incorrect because there were too many variables being changed to see how steepness alone affects how far the ball rolls. The correct answer was the table in answer C. In this table only included the steepness of the ramp and the distance rolled, which is the minimum amount of variables needed to test the effect of steepness in a control for variables experiment. Also, in this table several trials were run, so there the data is more consistent in showing that at the higher steepness the ball rolls farther. If the student chose the wrong answer on the first try, they were given an explanation of why their answer was wrong and were asked to try again. This happened until they chose the correct answer. After they came to the correct conclusion they were given an explanation of why the answer was correct. The students were given one point for a correct answer and zero points for a wrong answer.

Hypothesis and Experiment

Students were asked to write down their hypotheses about how the steepness of the ramp, the run length, the surface, and the type of ball each affected how far the ball rolls. Next the students were asked to test their hypotheses by performing an experiment with the physical ramp apparatus. The inquiry aspect of the students’ hypotheses, in the same manner as the pendulum pretest, was graded on a scale from zero to four.

Analyzing Data

After the student conducted each experiment, they were asked to analyze the data that the students collected in their experiments and explain their results. For the first experiment they were asked to answer the question "How does steepness affect how far the ball rolls?" If there was time, they moved on to test how run length, the type of ball, and the surface each affect how far the ball rolls.

Immediate Posttest

This final section of the laboratory notebook was an immediate multiple-choice posttest. Each question was scored with one point for a correct answer and zero points for an incorrect answer. The questions for the immediate posttest can be seen in the laboratory notebook in Appendix B4.

The first two questions of the posttest focused on the ramp experiment and the control for variables strategy. These questions were designed to determine what the student had learned about variables. The first question asked for a list of the variables in the experiment to see if the student knew which factors of the ramp were actually considered variables. The second question focused more on the importance of changing one variable at a time.

The third question asked why it was important to run an experiment multiple times. This was a question was designed to test what the students had learned in the experiment from the multiple choice prompt described in the method section. This concept needed to be stressed because it was one that the student of the first test set lacked.

The last question was intended to be a transfer question to see if the student was able to approach another control for variables problem. The question was on paper airplanes. The variables that could be changed were the type of day that the airplane was flown on (sunny or windy), the type of paper used, and the shape. The student was asked to circle the best experiment that could be run to test how the kind of paper affects how far a paper airplane will soar.

Coding of Data

Voice recordings were taken for each subject and analyzed to identify where improvements in inquiry skills were yielded as well as to identify areas that need further scaffolding.

Pendulum Pretest

The students' hypotheses were graded on a scale from zero to four. One point was given if the student mentioned either the dependent or independent variable in their explanation. Two points were given if the student mentioned both in their explanation. An additional two points were given if the student provided a relationship between the dependent and independent in their hypothesis. The student was given a three for a weaker relationship and a four if their hypothesis was thorough and there was a solid relationship described between the dependent and independent variables.

The grading for Questions 2, 3, and 4 on the pendulum pretest was broken up into two separate grades. The first grade was on a scale from zero to three and was based on the strength of their answer. The second part of the grade was on a scale from zero to one. The student was given one point if their answer, whether correct or incorrect, was consistent with their previous answers. Again, these questions were graded on the students' inquiry approach not on their knowledge of the material.

Hypothesis and Experiment

Hypotheses in this section were graded on a scale from zero to two (zero being incorrect, one being partially correct, and two being fully correct).

Analyzing data

This data section was graded on three criteria: Inquiry, content, and consistency. All of these criteria were graded on a scale from zero to two. Two points were given for inquiry if in their explanation a relationship existed between the dependent variable (e.g steepness) and the dependent variable (distance) and one point was given if the student made mention of one or both of these variable but there was no connection between them. Two points were given for content for a complete scientifically correct explanation and point was given for a partially correct answer. For consistency, the student was given two points if their answer was consistent with their data and one point for partial consistency.

Results

Results of the Pretest

The results for the pendulum pretest are displayed in Table 1. Each question is graded individually and the total score out of sixteen points is calculated in the far right hand column.

Table 1: Results of the pendulum pretest

Student	Class Level	Pre1	Pre2a	Pre2b	Pre3a	Pre3b	Pre4a	Pre4b	Total Score
Student1	H	1.5	1.5	1	0	0	0	0	4
Student2	M H	2	2	1	3	1	3	1	13
Student3	M L	4	1	1	1.5	1	2	1	11.5

The question labeled 'Pre1' was the students' hypothesis. It can be seen from the table that two of the students' hypotheses were coded as weak because they did not state a relationship between the dependent and independent variables (see the grading criteria in the method section). Question Pre2, which was about setting up a scientific experiment, was also a bit of a struggle for the students, however, it can be seen from Pre2b that all students received a point for their explanation being consistent with their hypothesis. Student 2 did not move on to questions three and four of the pretest. However, it can be seen from those two questions that Student2 received full credit for collecting data and stating their conclusion.

Hypotheses

There were sixteen points allotted for the inquiry aspect of forming hypotheses and eight points allotted for content. The results of the hypotheses are shown in Table 2.

Table 2: Hypotheses results for Test Set 2

Student	Class Level	H1 I	H1 C	H2 I	H2 C	H3 I	H3 C	H4 I	H4 C	I	C
Student1	H	2	1	4	2	4	2	3	1	13	6
Student2	M H	4	2	3	1	4	2	4	1	15	6
Student3	M L	1	1	1	3	0	1	0	1	6	1

It can be seen from Table 2 the first two from the high and medium-high level classes, respectively, did well on writing their hypothesis in both the content and inquiry aspect, whereas Student 3, a medium-to-lower level student, struggled with both inquiry and content.

Analyzing Data

The results of the students' experiment conclusions are shown in Table 3. Each conclusion was graded in terms of inquiry (I), content (C), and consistency (CS). There were two possible points available for each of these three areas for each hypothesis. If the box has an 'x' in it, the student did not get that far in the experiment. The total scores for inquiry, content, and consistency are in percentages of how many points the student achieved out of the total possible points for the experiments they completed. The students were not penalized for not conducting an experiment.

Table 3: Results sheets graded for Test Set 2

Student	Class Level	C1 I	C1 C	C1 CS	C2 I	C2 C	C2 CS	C3 I	C3 C	C3 CS	C4 I	C4 C	C4 CS	I	C	CS
Student1	H	2	2	2	2	2	2	x	x	x	x	x	x	100	100	100
Student2	M H	1	2	1	x	x	x	2	1	1	x	x	x	75	50	100
Student3	M L	2	2	2	x	x	x	x	x	x	x	x	x	100	100	100

From Table 3, it can be seen that all of the students' conclusions were consistent with their data. Students 1 and 3 had conclusions that were strong in the inquiry and content aspect. Student 2, however, struggled with both the inquiry and content aspect of writing a hypothesis. An example of a strong conclusion from Student3 for steepness was "The steepness does affect the distance the ball rolls because when its steeper the ball rolls farther", whereas Student2 said "When the ramp had steepness the ball rolled further." There is a solid connection between the independent variable (steepness) and the dependent variable (distance) in the answer given by Student3. The answer given by Student2 was weaker in terms of inquiry because it was evident that the ramp had steepness, but the student did not state the effect of the high versus the low steepness on how far the ball rolled.

Immediate Posttest

The results of the immediate posttest are shown in Table 4. None of the students had time to answer the transfer question on the control for variables strategy, so the question was discarded and the total scores, calculated in the far right column, were out of three points.

Table 4: Posttest scores for test set 4

Student	Class Level	Q1	Q2	Q3	Q4	Pretest Score
Student1	H	1	1	1	x	3
Student2	M H	0	0	1	x	1
Student3	M L	0	1	1	x	2

Table 4 shows that the student from the high-level class scored the highest on the posttest. Student 2 did not display much knowledge of variables after doing the experiment, which is determined by questions Q1 and Q2. All of the students answered correctly on Q3, which was the importance of running an experiment multiple times.

Discussion

Pendulum Pretest

Interesting pretest data was yielded by Student 2. The student's answers are included in section B2 of Appendix B. The students' hypothesis was that the pendulum "would keep going for a longer time and would go further than it did before". Although this hypothesis is not well constructed and the idea is not scientifically accurate, the rest of the student's pretest was rich in terms of inquiry. For example, the student was able to effectively come up with an experiment to test her hypothesis, come up with data for her experiment, and come to a conclusion based on her data. Their answers were consistent throughout the pretest.

Hypothesis

The hypotheses for this experiment where much easier to analyze because they were written instead of recorded. Students already had prior knowledge on writing hypotheses and of all the inquiry strands this seems to be of the least concern.

Experiment and Prompts

As expected from the first study, the students were not recording their data as they collected it. The prompt to address this issue was the multiple choice question shown in Appendix B5.1. In this question the students were presented three tables and asked which best represented how steepness affects how far the ball rolled. This question worked out very well. All three subjects were given this prompt. Two of the students chose ‘A’ as their first answer. This answer was included in the question because it had two common mistakes from the first experiment: there were not enough data points taken and it included the type of ball which was not necessary to test the effect of steepness. One student’s first choice was B, which changed too many variables at once. The correct answer was C. Once the students were given this prompt, they were able to construct their own tables based on the table in answer C.

The checklist prompt for the control for variables strategy was also effective in teaching the control for variables strategy. All of the students checked off at least one extra variable and when given the explanation of why they did not need to include those variables, they went on to simplify their tables. It is not clear, however, whether they were just following instructions or if they entirely grasped the control for variables concept because when asked to conduct another experiment, they seemed to make the same mistakes. In general, the prompt was not well written and needs to be refined to provide a better explanation of how to go about a control for variables problem.

Analyzing data

The conclusions for the experiments were good well done Students 1 and 3. However, Student 2 struggled with writing the conclusions for the two experiments they ran. This was surprising because this student's was awarded fifteen out of sixteen possible inquiry points for the strength of their hypotheses. One of the conclusions for Student 2, for the steepness of the ramp, was "When the ramp had steepness the ball went faster and further." This was a weak conclusion because it was obvious that the ramp had steepness, but the student did not specify whether a high or low steepness caused this outcome. The other conclusion for surface was "It affects how far the ball rolls because with the rug it slows down, with the soft surface makes it go faster." This conclusion was weak because the question was how steepness affect how *far* the ball rolls, not the affect on the ball's speed. In either case, the student does not refer back to her data, indicating a difficulty in analyzing her own data.

Immediate Posttest

Only one student, the honors student, mastered the concept of control for variables and the importance of running an experiment multiple times. This student completed all three questions correctly. Student 2, only the other hand, did not seem to gain any knowledge of variables in the experiment. Student 3 only struggled to list the variables in the experiment

It was unfortunate that the students did not have time to answer the transfer question of the posttest. The question seemed very wordy and was rewritten for the fourth test set. This is described in detail in that section.

The Ramp Experiment: Test Set 3

The third test set was the students' first interaction with the simulated ramp model shown described in the subsection 'Simulated Ramp Apparatus' in 'The ASSISTments Ramp Experiment'. Since this was a new environment, Test Set 3 was very similar to Test Set 1 in that the goal of the test set was to see how the students interacted with the simulated ramp in order to see what they did naturally in terms of inquiry. However, the tasks in this test set were slightly more structured like the second ramp experiment because from the second study it was determined that a structured approach was more effective in teaching the control for variables strategy. Voice data and videos of students' interactions with the simulated ramp apparatus were collected and analyzed for each case study.

Case Studies

Four students were subject to Test Set 3 from each class level.

Method

Materials

Materials for this experiment included the simulated ramp apparatus previously described and shown in the 'Simulated Ramp Apparatus' subsection of 'The ASSISTments Ramp Experiment' section. Since the first and the second studies revealed that students were having trouble recording their data, blank tables were provided that included only blank rows and columns, none of which were labeled.

Procedure

Students were tested individually. First the students were asked "How does the Left ramp steepness affect how far the ball rolls?" (Recall, the simulated ramp apparatus had a left and a right ramp steepness that could be altered). Students were asked to state their hypotheses out loud and then to test their hypotheses by collecting some data and recording it in the tables that were provided. If the students had difficulties filling out the tables, the column labels were provided by the experimenter. Then if they were still struggling, they were scaffolded for the first few trials of the experiment.

Direct Instruction

After the students completed their first table, the data was analyzed on the spot by the WPI team. Since Klahr and Nigam's study proved direct instruction to be an effective method of teaching the control for variables strategy, if the student did not conduct the first experiment correctly meaning they were changing too many variables at once, they were given a direct explanation of how to go about solving the problem. In this prompt, they were told in what way their table was incorrect. They were then given a table that kept all variables the same except for the steepness of the left ramp and were told to collect three data points for the ramp at the high steepness and three data points for the ramp at the low steepness. After doing this the students were asked to compare these results with the results of their first table.

Progress Worksheets

After completing the direct instruction worksheet the students were to go on to test how run length, surface, and type of ball affect how far the ball rolls (in no particular order) to see what they learned about the control for variables experiment. After each experiment, the team analyzed the students' tables and detected and corrected any further flaws in their strategies.

Results

Participant 1

The first student was randomly selected from a high-level class. When asked how steepness affects how far the ball rolls, the student replied “If the right ramp steepness is low it’s easier to go through and if the right ramp steepness is high the ball will not have enough momentum to get up the hill”. He went on to try and test this hypothesis.

When asked what he had found from his trials, the student conclusion was the exactly the same as his hypothesis. As can be seen from the table labeled ‘a’ in Appendix C section C1, the students’ data was not consistent with his conclusion. The student changed the right ramp steepness as well as all other variables. Also, the student ran an experiment on the steepness using the right ramp not the left ramp, as asked. In either case there are no two trials seen in his table that validate this conclusion.

Next the student was given the methodology behind the control for variables strategy and was asked to collect three data points at the high left ramp steepness and three at the low steepness. When asked why he thought it was important to collect multiple data points, he responded “Because it might not always come out the same.” Upon completion of this exercise the student was asked to give his conclusion on how steepness affects how far the ball rolls. When asked how his original table differed from the one given, the student replied, “That table (pointing to his own table) was random. This table (pointing to the example table) kept all the same variables except left ramp steepness.” In this explanation the student showed the ability to compare data and detect two different control for variable approaches.

For the next task the student was asked to run an experiment to test how run length affects how far the ball rolls. In his next table ‘b’ in Appendix C1, the student did not show much improvement. The student changed every variable in the experiment.

For a second time the student was given an explanation of the control for variables strategy. They were again given exactly how to run the experiment and were told to perform three trials at the full run length and three trials at the half run length. When asked if he knew why the numbers were coming out more consistent when he kept everything the same except for the run length, the student said “I changed too many variables.” However, in the student’s third table called ‘c’ in Appendix C1, again the student showed no signs of improvement or mastery of the control for variables strategy.

Participant 2

The second student was chosen from a medium-to-high level class. When asked to run an experiment on how steepness affects how far the ball rolls the student started recording his data in a straight line. This can be seen in the first trial of Table ‘a’ in Appendix C2. The student was then told that making columns may be a more organized way of recording data.

After the first experiment was performed, the student concluded that, “Left ramp steepness is more important than right ramp steepness.” The students’ explanation was inconsistent with his data because in his table he had changed all of the variables and there were no distinct patterns that developed.

The student was then given the explanation for control for variables and was told to take three data points at the high steepness and three data points at the low steepness. The student was not effectively able to analyze the new data because he was having trouble forming an explanation for what this new table showed him and was given further explanation of exactly where the table proved that at the higher steepness the ball rolled further.

Next the student was asked to perform an experiment on how run length affects how far the ball rolls. He hypothesized that “If you do it halfway, it’s going to go up more. This table is called ‘b’ in Appendix C2. In this table the student included only the run length and the distance. Three trials were performed at the half length and three trials were performed at the full length. Every other variable was kept the same throughout the experiment. When asked if his data was consistent with their hypothesis, he stated “I thought half way it would go further, but at the long run length it went further. The numbers were higher.” This statement implied that the he was now effectively analyzing his data and forming a conclusion that was consistent with his hypothesis.

When asked to move on to test how the ball type affects how far the ball rolls the student created a similar table. He only included the ball type and distance and kept every other variable consistent. This table is labeled ‘c’ in the Appendix C2.

Participant 3

Student 3 was randomly selected from a class categorized as medium-to-low. Unfortunately, voice and ramp data were lost for this student. However, the students’ written data was collected and analyzed. Her tables are included in Appendix C3. In this section these tables are analyzed and described in the best detail possible.

For the experiment on left ramp surface, the student created the table labeled ‘a’ in Appendix C3. It can be seen that the table has no particular order to it. The student changed every variable of the ramp apparatus.

Next the student was given a table left ramp steepness and the distance rolled and kept everything else the same. She was asked to perform three trials with the rough surface and three trials with the smooth surface. In her new table, she was much more consistent. With the rough surface the ball traveled 51, 49, and 52 meters. With the smooth surface the ball traveled 95, 102, and 99 meters. Her data now showed a clear relationship between the left ramp steepness and the distance.

For the experiment on how left ramp steepness affects how far the ball rolls, she only performed one trial. From the written notes and materials it seems that the student may have had more difficulties performing the experiment on left ramp steepness because another table was drawn up for her explaining how she would run an experiment to test the left ramp steepness.

Next the student created a table to show how the type of ball affects how far the ball rolls. This is Table 'b' in the Appendix C3. It can be seen from the table that the student kept the steepness and the surfaces the same throughout the experiment but changed the type of ball as well as the run length. The student ran two trials with the rubber ball at the half length, two trials with the golf ball at the full length, and one trial with the rubber ball at the full length. Since the student was changing run length as well, it was difficult to see the effect of the type of ball alone.

Participant 4

Student 4 was chosen from a lower-level class. After running the first experiment on how steepness affects how far the ball rolls, the student concluded, "It depends on how you put it. If you put it full length it will go more far, and if you do the half length, it won't go that far." The student's conclusion for the effect of steepness only made mention of the effect of run length on how far the ball rolled.

After listening to the control for variables explanation and running the experiment three times at the low steepness and three times at the high steepness, the student was asked how steepness affects how far the ball rolls. She pointed out that for the first trial (at the high steepness), the number was higher and for the second trial (at the low steepness), the number was lower. She said “The higher steepness the faster it goes.” In comparison to her original table, she said “I had more than one type and I was scattered.” The student’s first table is labeled ‘a’ in Appendix C4. It can be seen that the student changed every variable in the experiment, indicating that she had little or no prior knowledge of the control for variables strategy.

When asked to run an experiment on how the type of ball affects how far the ball rolls, the student hypothesized that “the rubber ball is made out of rubber and is bouncy so the ball will go higher.” The student’s table for this experiment is shown in ‘Student4: Table 2’. In this table the student changed the left ramp surface, the left ramp steepness, and the type of ball but caught her mistakes when explaining her results. She asked for the chance to try again. The new table is shown in ‘Student4: Table 3’ in the Appendix. In this table the student kept every variable the same, except for the type of ball. She performed two trials with the rubber ball and two trials with the golf ball and concluded that, “The golf ball went a lot further”.

When asked to run an experiment to test how run length affects how far the ball rolls, the student generated the data shown in the table labeled ‘Student4: Table4’. In this table it can be seen that the student kept all variables the same and performed three trials for the run length at the full length and three trials for the run length at the half length. Her conclusion was that, “When it is full, it will go farther and go more length and when it was half it will go smaller.” This showed that she was able to effectively analyze the data in her table.

Discussion

The results of Study 3 did not reveal any additional findings beyond those of the first two studies. Students, in general, were not having too much trouble stating logical hypotheses, but they did struggle in setting up the experiment, in testing their hypotheses and, in recording their data. All students needed to have the columns of the table set up for them so that they could correctly record their data. As expected and discovered from the first two studies, none of the students showed a prior mastery of the control for variables strategy. Again the conclusions for the first trial were not based on data collected.

The interesting aspect of Test Set 3 was the direct instruction and the students' responses when asked to compare their original tables to the ones set up for them. When asked to reflect back on their original table, the students realized that were changing too many variables at once making it harder to see the affect on the variable they were trying to test. The new tables, which were set up for them, were much clearer and the outcome was more salient to the students. Two of the students, Students 2 and Student 4, seemed to master the control for variables concept.

Ramp Experiment: Test Set 4

For the fourth test set the simulated ramp apparatus was used along with a laboratory notebook similar to the one used in Study 2. Like the one from the second test set, it was more structured and organized. The notebook also included a direct explanation of the control for variables strategy similar to the scaffolding from Study 3, since it was determined from Klahr and Nigam's experiment and presumed from ASSISTments ramp experiment Study 3 that the direct approach is more effective in teaching the control for variables concept. However, in this experiment the experimenters did not provide any guidance to the students other than asking some think-aloud questions from the laboratory notebook. The laboratory notebook was written so that the student did not need any assistance from a live tutor. Also each section of the laboratory notebook was strictly timed and recorded. The think-aloud responses were graded.

Participants

Four students from a Worcester Public Middle School participated in the fourth study. Each of the students was chosen at random from different classes categorized by their levels. Students from high, medium-high, medium-low, and low level classes were chosen.

Method

Materials

The students in this study were given the simulated ramp apparatus and the laboratory notebook shown in Appendix D2. This laboratory notebook consisted of an immediate pretest, a worksheet designed to see how they approached the problem naturally, a direct explanation of the control for variables strategy, three worksheets designed to see if the explanation was effective in teaching the control for variables concept, and an immediate posttest.

Procedure

Immediate Pretest

The students were given 10 minutes to complete the immediate multiple-choice pretest. This pretest is included in Appendix D1. Questions were graded either as being correct or incorrect. For a correct answer, the student was given one point. If the students answer was incorrect the student was not awarded any points.

The first two questions on the pretest addressed the importance of running an experiment multiple times. The first question was about what the validity of data and being able to replicate an experiment. The second question was about repeating an experiment multiple times as well as analyzing data from a table.

The third question was about the order of operations for conducting an experiment. In this question, there was an unorganized list of steps in an experiment. The student was asked to put these steps in a logical order. This assessed the students' skills to conduct a logical experiment.

The last question on the pretest was a modified version of the last question on the posttest for the second test set. Recall that this question was never assessed because the question was lengthy and students did not have time to complete it. In this test set, the question was written in a briefer way. This question was included in the pretest to determine whether students had prior knowledge of the control for variables strategy.

<u>Lab Notebook</u>

WORKSHEET 1

The students were given 5 minutes to complete the first worksheet in the lab notebook. This worksheet was designed to see how students naturally approached a control for variables problem. First the students were asked in a think aloud question to state their hypothesis on how they thought steepness would affect how far the ball rolls.

Following the hypothesis, the students were given a brief explanation on how to use the simulated ramp apparatus. They were given a blank table with labeled columns and asked to run an experiment to see how steepness affects how far the ball rolls and record their results in the table. The experiment was followed by a multiple choice question on how steepness affects how far the ball rolls and a think aloud question to say why they choose that answer.

DIRECT INSTRUCTION OF CONTROL FOR VARIABLES

Next each student was given 20 minutes to read a direct explanation of the control for variables strategy and move on to three more worksheets similar to the first. The definition of a variable was given followed by a brief explanation of the control for variables strategy.

An example table, produced by the WPI team, was given to show how steepness affects how far the ball rolls. This was same experiment that the student had just performed. Following this table was a think aloud question. The student was asked to think aloud while comparing the table that they created in the previous worksheet to the one shown. This question was included to see if the students could determine what they did incorrectly in the first worksheet. The question was not graded, but the students' responses are included in the 'Results'and 'Discussion' sections. Then they were asked to answer how steepness affects how far the ball rolls based on the new table and explain their answer in a Think Aloud (Questions 9 and 10 in the laboratory notebook).

PROGRESS WORKSHEETS

The next three worksheets were almost identical to the first, except the student was asked to determine how run length, the type of ball, and the surface affect how far the ball rolls. This section of the lab notebook was designed to see how well they understood the control for variables explanation. Again, students were asked to perform a specific task, record their data in the table, and answer a question based on the results of their experiment.

Immediate Posttest

The last section of the workbook was an immediate posttest in which students were given 10 minutes to complete. All questions were multiple-choice and graded in the same manner as the pretest. The students were given one point only for the correct answer to each question.

The first question on the posttest, Question 20, addressed the importance of variables in a scientific experiment. This question was aimed to test what students had learned about variables during the experiment.

Question 21 was supposed to mirror the first two questions of the pretest to determine what the student had learned about the importance of running an experiment multiple times. This is one of the concepts covered in the direct explanation of the control for variables strategy in the laboratory notebook.

The third question was intended to be a transfer task to solve a control for variables problem. The students were asked how they would determine if a vaccine for birds was effective. The correct answer was to vaccinate half of the sample space, do not vaccinate the other half, and expose all of the birds to the disease.

The last problem was another data analysis problem similar where the student had to examine experimental results recorded in a table to determine how the length of a string will affect the motion of a tire in a tire swing application.

Coding of Data

<u>Lab Notebook</u>

WORKSHEET 1

The hypotheses were graded the same way the hypotheses in Test Set 2 were graded: on a scale from zero to four. Students were given one point if they made mention of the independent or dependent variable in the experiment, two points if they mentioned both the independent and dependent variables, three points if they included a vague or incomplete relationship between the two, and four points for a complete and thorough hypothesis.

The multiple choice question Q6a.was a right or wrong answer and the student was given one point only if their answer was correct. The think-aloud response following this question Q6b (labeled 'Question 7 in Appendix D2.1) was graded on the consistency of the student's explanation with the table they created. The student was given one point if they were consistent and no points if they were inconsistent.

DIRECT INSTRUCTION FOR CONTROL FOR VARIABLES

Think Aloud Questions (9 and 10 in the laboratory notebook) were coded in the same way as questions 6 and 7, except in this case consistency is defined as whether or not the student referred to the example table in their response.

PROGRESS WORKSHEETS

Grading for the questions in these sections were exactly the same as in Worksheet 1. Questions 11, 14, and 17 were think-aloud questions asking the student to state their hypotheses and were graded based on the criteria for Question 5, Questions 12, 15, and 18 were multiple choice questions based on the results of the experiment and graded based on the criteria for Question 6, and Questions 13, 16, and 19 were Think-Aloud questions about their results and are graded based on the criteria for Question 7 from Worksheet 1.

Results

Immediate Pretest

The results of the immediate pretest are shown in Table 5.

Table 5: Pretest results for Test Set 4

Student	Class Level	Q1	Q2	Q3	Q4	Pretest Score
Student1	H	1	0	0	1	2
Student2	M H	0	0	1	0	1
Student3	M L	1	0	0	1	2
Student4	L	0	0	1	1	2

None of the students answered the Question 2 correctly. This was a question focused on the importance of running an experiment multiple times and analyzing data from an experiment. Three of the students answered the control for variables Question 4 correctly. There were no distinct patterns for the other two questions. Two of the students answered Question 1 correctly on the importance of replicating an experiment and two of the students answered Question 3, the order of operations, correctly.

Lab Notebook

WORKSHEET 1

All of the students' hypotheses on how they think steepness will affect how far the ball rolls were correct. The results are shown in Table 6. Each student received full credit in the inquiry and content aspects of the first hypothesis.

Table 6: Hypotheses for how steepness affects how far the ball rolls

Student	Class Level	5I	5C
Student 1	H	x	x
Student 2	M H	4	2
Student 3	M L	4	2
Student 4	L	4	2

Although the students answers were correct, none of the students when asked to perform an experiment on how steepness affects the distance the ball rolls, performed the experiment correctly on the first try. Of the six variables that the students were given, each student changed at least four.

Scoring for Questions 6 and 7 are shown in Table 7. Note that Question 6 is renamed as 6a and Question 7 is renamed as Question 6b.

Table 7: Data analysis for Worksheet 1

Student	Level	6a	6b
Student1	H	1	x
Student2	M H	1	0
Student3	M L	0	1
Student4	L	1	1

Three of the four of the students answered the multiple choice question on how steepness affects how far the ball rolls correctly. Only two of the students, those who referred to their table when coming up with their conclusion, received a point for consistency.

DIRECT INSTRUCTION OF CONTROL FOR VARIABLES

After the first worksheet students read the direct instruction worksheet for the control for variables strategy. The students were then asked to compare the experiment that the WPI team ran to the experiment that they just conducted. The results of these explanations were not scored and are described in further detail in the Discussion of Test Set 4 section.

The students' data analysis for this table, however, was scored and is shown in Table 8.

Table 8: Data Analysis of the WPI team experiment

Student	Class Level	9a	9b
Student1	H	1	1
Student2	M H	1	0
Student3	M L	1	1
Student4	L	1	1

All of the students answered questions 9a correctly, which corresponds to the multiple-choice Question 9 of the laboratory notebook. Three of the four students' answers to Question 10, labeled as 9b in Table 8, were consistent with the table given.

PROGRESS WORKSHEETS

Worksheets 2, 3, and 4 were similar to Worksheet 1 and were scored accordingly. Questions 11, 14, and 17 in the laboratory notebook are the students' hypotheses for the last three experiments. Scoring for these questions for inquiry and content are shown in Table 9. It can be seen from Table 9 Student 1 did not state his hypotheses in the Think-Aloud questions, so there is no data to analyze for that student.

Table 9: Hypotheses for the progress worksheets

Student	Class Level	11I	11C	14I	14C	17I	17C	I	C
Student1	H	x	x	x	x	x	x	x	x
Student2	M H	4	0	x	x	x	x		
Student3	M L	3	2	2	1	x	x		
Student4	L	3	2	x	x	x	x		

For the most part, the hypothesis remained fairly strong. For Question 14, however, Student3 failed to state the relationship between the dependent and independent variables in the experiment.

Questions 12a and 12b (Questions 12 and 13 in the laboratory notebook), 15a and 15b (Questions 15 and 16), and 18a and 18b (Questions 18 and 19) were similar to questions 6a and 6b on Worksheet 1 and scored in the same manner. The results are shown in Table 10.

Table 10: Data Analysis Results for the Progress Worksheets

Student	Class Level	12a	12b	15a	15b	18a	18b
Student1	H	1	1	1	1	1	1
Student2	M H	1	0	x	x	x	x
Student3	M L	1	1	1	1	x	x
Student4	L	1	1	x	x	x	x

It can be seen from the Table every student answered all of the multiple choice questions that correctly on how run length, the type of ball, and surface affect how far the ball rolls. Most of the explanations were also consistent with the data they collected. For example, for Question 12b Student3 said that both the short and the long run lengths had the same effect on how far the ball rolls. Even though her data was not correct the values in her table were close together.

Immediate Posttest

The results of the immediate posttest are shown in Table 11. The scores of the posttest are not much of an improvement from the results of the pretest, however, the scores of two students that mastered the control for variables strategy increased by one point and the scores of the two students that struggled with the control for variables concept decreased by one point each.

Table 11: Posttest results for Test Set 4

Student	Class Level	Q20	Q21	Q22	Q23	Posttest Score
Student1	H	0	1	0	1	2
Student2	M H	0	0	0	0	0
Student3	M L	1	0	0	0	1
Student4	L	1	1	0	1	3

None of the students answered Question 22 correctly on the posttest, which was intended to be the control for variables transfer task. There were no distinct patterns for the rest of the questions in the posttest.

Discussion of Test Set 4

Immediate Pretest

None of the students answered the second question of the pretest correctly. This question was confusing because the question was not about providing a conclusion for an experiment, but about the fact that the table data was invalid because there was not enough data collected. Altogether, the question was not well written and will be either discarded or rewritten in the future.

Three of the four students answered the control for variables question correctly, which indicated some prior knowledge of the control for variables strategy. This was not very surprising because in the second test set all of the students said that they had already been introduced to variables in their science class.

An interesting question came from one of the students who approximately two minutes into the pretest asked, “Is there a right or wrong answer?” So for at least one student, there was some level of test anxiety.

Lab Notebook

WORKSHEET 1

For the first worksheet, every student changed at least four of the six variables given to them. Even though they did not have the data to support their conclusion, the students insisted that they had discovered the answer and that their data proved it. For some it seemed too obvious that if the ramp is steeper that the ball will roll further. One student commented that “A” makes more sense. The higher the ball, the faster it goes.” When asked if they saw those results in their table, all students answered “yes”, however, they did not give a deeper into the explanation of how they had proven that the higher the steepness is the further the ball will go.

DIRECT INSTRUCTION OF CONTROL FOR VARIABLES

After the explanation of variables and the control for variables strategy, students were asked to compare their table to the one given. At this point two out of four of the students realized what they had been doing wrong. One of the students said, “They kept all these things the same.” When asked, “Do you think you changed too many things in your experiment?”, the student said “Yes.” Another student answered, “The values are different.” When the experimenter asked them “What about keeping the variables the same?”, the student replied, “I didn’t do that.”

PROGRESS WORKSHEETS

Two of the students of the four mastered the control for variables strategy because in the following worksheets the students kept every variable the same except for the variable in which they were trying to determine the effect of. One of the improvements is shown in Tales a. and b. in Appendix D4.

One of the students may have misinterpreted the explanation because in the following worksheets, they changed every variable except for the variable that they were trying to determine the effect of. The other student followed the directions too closely because from there on they only changed the type of ball and the steepness of the ramp even when determining the effect of run length on how far the ball rolls.

When asked "How does run length affect how far the ball rolls?' every student answered correctly. Two of the students said that they saw from their data that the longer the run length the further the ball rolls and actually had the data to back it up. Another student was stuck on the common sense aspect of the problem that it is just obvious that if you release the ball from the full length that it will go further.

A common problem in the last three worksheets was that the right approach was taken, but there was not enough data collected. The student would perform only two trials and feel that they were ready to move on and answer the questions. The last three worksheets of one of the students are shown in Tables a., b., and c., in Appendix D5. It can be seen that this student mastered the control for variables concept, but only two trials were performed for each experiment.

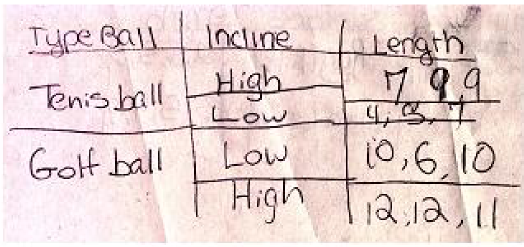
Immediate Posttest

The major disappointment of the posttest was that the students did not answer Question 22, the transfer task for the control of variables concept, correctly. Upon reflection, the question was a difficult and may be yield more useful data in the future once the students have completely mastered the control for variables concept. It also may have been too difficult for seventh graders in general. It may have been a better assessment to use a simpler question that mirrored Question 4 of the pretest more closely.

Also, the pretest and the posttest were more general inquiry tests and were not specific enough to the material covered in the laboratory notebook. Hence, it was harder to see if the laboratory notebooks improved students' inquiry skills. Fortunately, the tables in the progress worksheet produced valuable information that was a good measure of this.

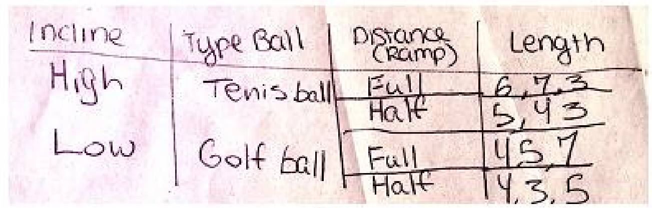
Appendix A: Test Set 1

A.1 Student 1 Data



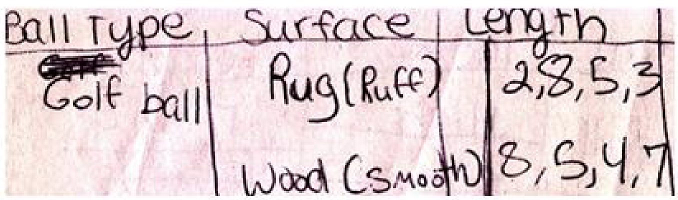
Type Ball	Incline	Length
Tenis ball	High	7,9,9
	Low	4,5,7
Golf ball	Low	10,6,10
	High	12,12,11

a.



Incline	Type Ball	Distance (Ramp)	Length
High	Tenis ball	Full	6,7,3
		Half	5,4,3
Low	Golf ball	Full	4,5,7
		Half	4,3,5

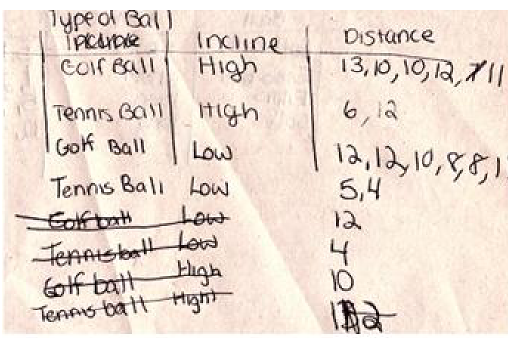
b.



Ball Type	Surface	Length
~~Golf~~ Golf ball	Rug (Ruff)	2,8,5,3
	Wood (Smooth)	8,5,4,7

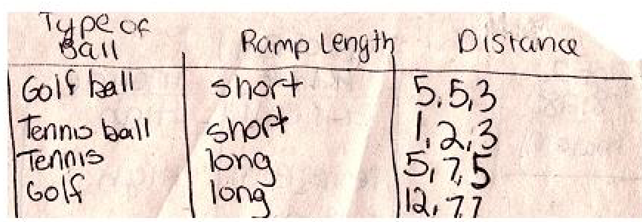
c.

A.2 Student 2 Data



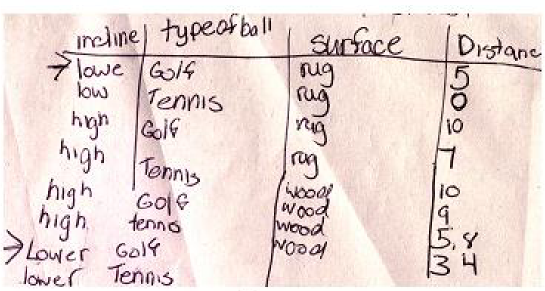
Type of Ball	Incline	Distance
Golf Ball	High	13,10,10,12,~~7~~11
Tennis Ball	High	6,12
Golf Ball	Low	12,12,10,8,8,11
Tennis Ball	Low	5,4
~~Golf ball~~	~~Low~~	12
~~Tennis ball~~	~~Low~~	4
~~Golf ball~~	~~High~~	10
~~Tennis ball~~	~~High~~	12

a.



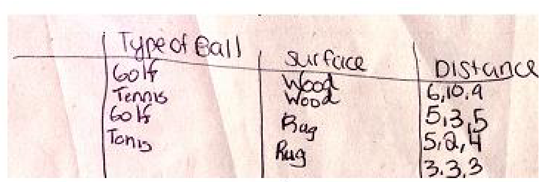
Type of Ball	Ramp Length	Distance
Golf ball	short	5,5,3
Tennis ball	short	1,2,3
Tennis	long	5,7,5
Golf	long	12,7,7

b



incline	type of ball	surface	Distance
lowe	Golf	rug	5
low	Tennis	rug	0
high	Golf	rug	10
high	Tennis	rug	7
high	Golf	wood	10
high	tennis	wood	9
Lower	Golf	wood	5,4
lower	Tennis	wood	3 4

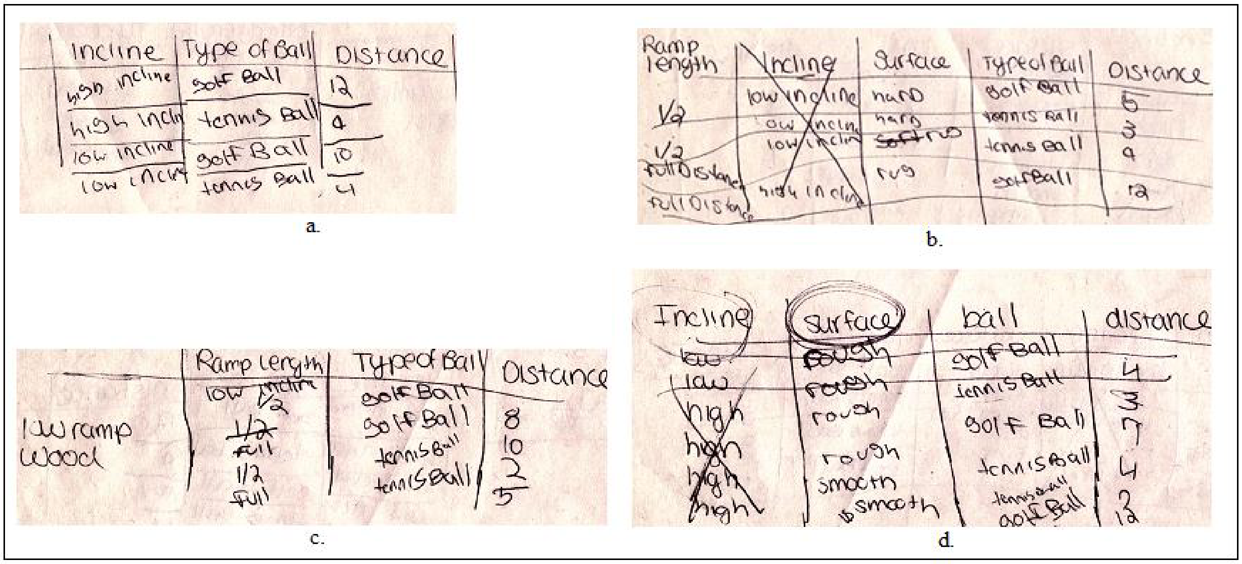
c.



Type of Ball	Surface	Distance
Golf	Wood	6,10,9
Tennis	Wood	5,3,5
Golf	Rug	5,2,4
Tonis	Rug	3,3,3

d.

A.3 Student 3 Data



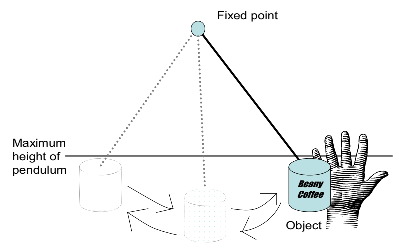
Appendix B: Test Set 2

B1. Pendulum Pretest

Student #___________________

A **pendulum** is an object that hangs by a string attached to a fixed point. A pendulum can swing freely back and forth as it hangs from the string. The figure below shows an example of a coffee can that is made into a pendulum by attaching it to a fixed point using a string.

When the coffee can is released from its starting place, it will swing down in the direction shown by the solid arrows. When it reaches the bottom it will swing back up and reach a maximum height. Then it will stop for a small moment and begin to swing back towards its starting place. The coffee can will keep swinging back and forth until someone stops it.



Your question is to find out what happens to the pendulum when you change the **length of the string**.

Note: You are not performing the experiment itself. Instead, you are designing an experiment and using your own understanding of the results to come to a conclusion.

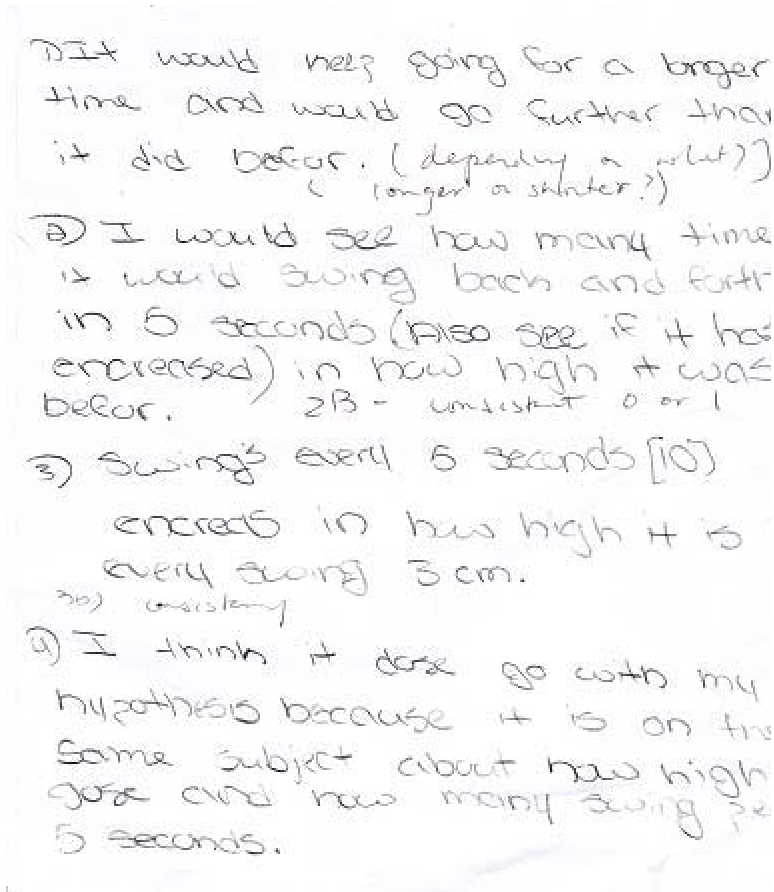
1. Write a **hypothesis** that states what you think might happen when you change the length of the string. **Don't worry if you think your hypothesis is incorrect!**

2. Describe how you would **conduct an experiment** to test your hypothesis.

3. Pretend you can run this experiment many times. **Make up data** showing the results of the pendulum swinging many times, each time with different lengths of string.

4. Look at your data. Do they prove or disprove your hypothesis? Write down your **conclusions** from looking at your data.

B2. Sample Pretest



1) It would keep going for a longer time and would go further than it did befor. (depending on what?] (longer or shorter?)

2) I would see how many time it would swing back and forth in 5 seconds (Also see if it has encreased) in how high it was befor. 2B - unconsistent 0 or 1

3) Swing's every 5 seconds [10]

encreas in how high it is every swing 3 cm.

3a) consistency

4) I think it dose go with my hypothesis because it is on the same subject about how high gose and how many swing pe 5 seconds.

B.3 Lab Notebook

B.3.1 Hypotheses Worksheet

Hypotheses

How does steepness affect how far the ball rolls?

____
____
____

How does run length affect how far the ball rolls?

____
____
____

How does surface affect how far the ball rolls?

____
____
____

How does the type of ball affect how far the ball rolls?

____
____
____

B.3.2 Results Worksheet

Results

How does steepness affect how far the ball rolls?

How does run length affect how far the ball rolls?

How does surface affect how far the ball rolls?

How does the type of ball affect how far the ball rolls?

B4 Posttest

Multiple-Choice Posttest

QUESTION 1: What were the variables in our experiment?

a. Steepness
b. Steepness, run length, the type of ball, and surface
c. Steepness and the distance the ball rolled
d. The ramp

QUESTION 2: What is the most important thing about variables when doing a science experiment?

a. Variables can change
b. You can change variables one at a time to find out how that one variable affects the outcome of your experiment
c. You can change variables all at the same time to see how they affect the outcome of your experiment
d. Variables do not change

QUESTION 2: What is the most important thing about variables when doing a science experiment?

a. Variables can change
b. You can change variables one at a time to find out how that one variable affects the outcome of your experiment
c. You can change variables all at the same time to see how they affect the outcome of your experiment
d. Variables do not change

QUESTION 3: Why is it important to run the same experiment multiple times?

a. So we can collect enough data to support our conclusion.
b. So we can follow the scientific method.
c. So we can see how many variables you can change.
d. It is not important. One trial can tell you how steepness affects how far a ball rolls.

QUESTION 4: Sam wants to run an experiment on how far a paper airplane can fly. There are a few variables that could affect how far his airplane soars: how windy the day is, what kind of paper he uses, and what shape he makes it. What is the best experiment that Sam can use to test how the kind of paper affects how far his airplane flies?

a. Sam should make two paper airplanes of the same size and shape and try flying them on a windy day to see how far they go
b. Sam should make one paper airplane and try flying it on a windy day. Then fly the same airplane on a sunny day. He should record his results in a table.
c. Sam should make two paper airplanes that are the same shape but use a different type of paper. Then he should go out on a sunny day and measure how far they go.
d. Sam should make one paper airplane and fly it on a windy day. Then he should wait for a sunny day and make another paper airplane that has a different shape and type of paper than the first plane.

B5. Prompts

B5.1 Prompt 1

Johnny created three tables of data. Which one best describes the effect of steepness on how far the ball rolls?

A)

Height of Ramp	**Type of Ball**	**Distance Rolled**
High	Golf Ball	12
Low	Golf Ball	9
High	Tennis Ball	8
Low	Tennis Ball	3

B)

Height of Ramp	**Run Length**	**Type of Ball**	**Distance Rolled**
High	Half	Golf Ball	9,7
Low	Full	Golf Ball	7,7
High	Half	Tennis Ball	3,4
Low	Full	Tennis Ball	5,4

C)

Height of Ramp	**Distance Rolled**
High	12,10,11,7,9
Low	10,9,11,8,4

B5.2: Prompt 2

□ Steepness

□ Run Length

□ Type of Ball

□ Surface

□ Distance

Appendix C: Test Set 3

C1. Student 1 Data

Table Title: ① How does steepness affect how far the ball rolls? Student

left ramp steepness	right ramp steepness	Left Ramp Type	Right Ramp type	Ball Type	Run Length	Distance rolled		
High	High	Rough	Smooth	Golf	Full Length	285 meters		
High	Low	Smooth	Smooth	Rubber	Half Length	202 meters		
Low	High	Rough	Rough	Golf	Full Length	59 meters		
Low	High	Rough	Rough	Rubber	Half Length	41 meters		

a.

Table Title: ② How does run length affect how far the ball rolls?

Left Ramp Type	Right Ramp Type	Left Ramp Steepness	Right Ramp Steepness	Run Length	Ball Type	Distance		
Smooth	Smooth	High	Low	Full	Golf	482		
Smooth	Smooth	High	Low	Half	Golf	237		
Smooth	Smooth	Low	Low	Full	Rubber	151		
Smooth	Smooth	Low	Low	Half	Rubber	71		
Rough	Rough	High	Low	Full	Golf	349		
Rough	Rough	High	Low	Half	Golf	157		

b

Table Title: Surface of Right Ramp

left ramp Type	Run length	Right Ramp type	Right Ramp steepness	Left Ramp steepness	Ball type			
Smooth	Full	Rough	Low	High	Golf			
	Half	Smooth		Low	Rubber			
	Full	Rough		High	Golf			
	Half	Smooth		Low	Rubber			
	Full	Rough		High	Golf			
	Half	Smooth		Low				

c.

C2. Student 2 Data

Table Title: ① How does left ramp steepness affect how far the ball rolls?

[illegible]	Distance	340m	[illegible]	Right ramp [illegible]	[illegible]			
left ramp type	right ramp type	ball type	left ramp height	Right ramp height	Run length	Distance		
rough	Smooth	golf	Low	Low	Full	106m		
Smooth	Smooth	golf	High	Low	Full	513m		
rough	Smooth	rubber	High	Low	half	163m		
Smooth	Smooth	rubber	low	low	half	26m		
Smooth	Rough	golf	High	High	half	127m		

a

run length	Distance
half	96m
Full	179m
half	90m
full	173m
half	87m
full	164m

b.

L.Ramp	smooth
R.Ramp	Smooth
Ball	Golf
L.Ramp	Low
R.Ramp	High

c.

C3. Student 3 Data

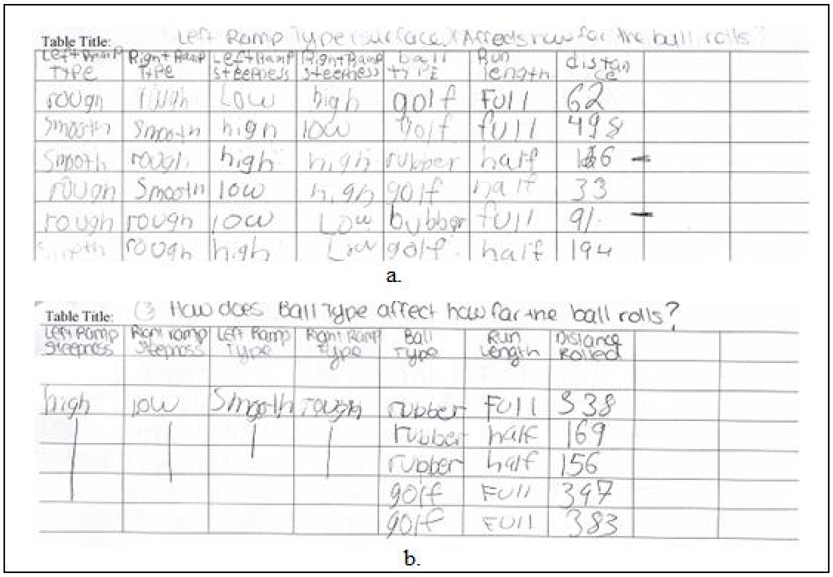


Table Title: Left Ramp Type (surface) Affects how far the ball rolls?

Left Ramp Type	Right Ramp Type	Left Ramp Steepness	Right Ramp Steepness	Ball Type	Run length	distance		
rough	High	Low	high	golf	Full	62		
Smooth	Smooth	high	low	golf	full	498		
Smooth	rough	high	high	rubber	half	166 −		
rough	Smooth	low	high	golf	half	33		
rough	rough	low	Low	bubber	full	91 −		
Smooth	rough	high	Low	golf	half	194		

a.

Table Title: (3) How does Ball Type affect how far the ball rolls?

Left Ramp Steepness	Right Ramp Steepness	Left Ramp Type	Right Ramp Type	Ball Type	Run Length	Distance Rolled		
high	low	Smooth	rough	rubber	Full	538		
|	|	|	|	rubber	half	169		
|	|			rubber	half	156		
|				golf	Full	397		
				golf	Full	383		

b.

C4. Student 4 Data

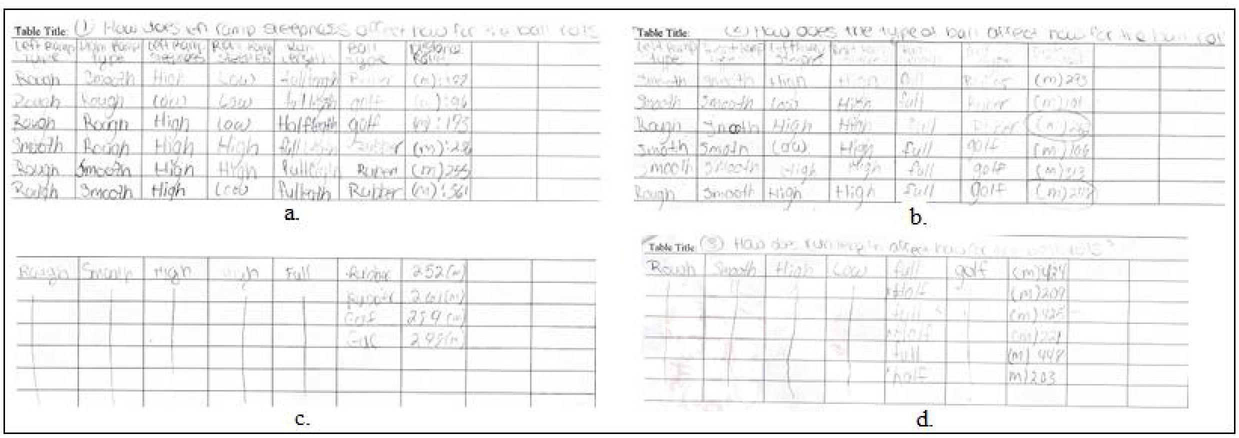


Table Title: (1) How does left ramp steepness affect how far the ball rolls

Left Ramp Type	Right Ramp Type	Left Ramp Steepness	Right Ramp Steepness	Run Length	Ball Type	Distance Rolled		
Rough	Smooth	High	Low	Full length	Rubber	(m): 192		
Rough	Rough	Low	Low	Full length	golf	(m): 96		
Rough	Rough	High	Low	Half length	golf	(m): 193		
Smooth	Rough	High	High	Full length	Rubber	(m): 28[illegible]		
Rough	Smooth	High	High	Full length	Ruber	(m) 255		
Rough	Smooth	High	Low	Full length	Rubber	(m): 361		

a.

Table Title: (2) How does the type of ball affect how far the ball rol[ls]

Left Ramp Type	Right Ramp Type	Left Ramp Steepness	Right Ramp Steepness	Run Length	Ball Type	Distance Rolled		
Smooth	Smooth	High	High	full	Rubber	(m) 293		
Smooth	Smooth	Low	High	full	Ruber	(m) 101		
Rough	Smooth	High	High	full	Ruber	(m) 2[illegible]		
Smooth	Smooth	Low	High	full	golf	(m) 166		
Smooth	Smooth	High	High	full	golf	(m) 212		
Rough	Smooth	High	High	full	golf	(m) 2[illegible]		

b.

Rough	Smooth	High	High	Full	Rubber	252 (m)		
|	|	|	|	|	Rubber	261 (m)		
|	|	|	|	|	Golf	259 (m)		
|	|	|	|	|	Golf	299 (m)		

c.

Table Title: (3) How does run length affect how far the ball rolls?

Rough	Smooth	High	Low	full	golf	(m) 424		
|	|	|	|	half	|	(m) 209		
|	|	|	|	full	|	(m) 425		
|	|	|	|	half	|	(m) 221		
|	|	|	|	full	|	(m) 448		
|	|	|	|	half	|	(m) 203		

d.

Appendix D: Test Set 4

D1. Pretest

Pretest

(10 minutes)

Question 1 (Multiple Choice)

A biologist in a laboratory reports a new discovery based on experimental results. If the experimental results are valid, biologist in other laboratories should be able to

a. Repeat the same experiment with a different variable and obtain the same results
b. Repeat the same experiment and obtain the same results
c. Perform the same experiment and obtain different results
d. Perform the same experiment under different experimental conditions and obtain the same result

Question 2 (Multiple Choice)

A student hypothesized that lettuce seeds would not germinate (begin to grow) unless they were covered with soil. The student planted 10 lettuce seeds under a layer of soil and scattered 10 lettuce seeds on top of the soil. The data collected are shown in the table below.

Data Table

Seed Treatment	Number of Seeds Germinated
Planted under soil	9
Scattered on top of soil	8

To improve the reliability of these results, the student should

a. Conclude that darkness is necessary for lettuce seed germination
b. Conclude that light is necessary for lettuce seed germination
c. Revise the hypothesis
d. Repeat the experiment using a larger sample size

Question 3 (Multiple Choice)

Jan will conduct an activity to measure how long it takes a marble to fall 20 centimeters when placed in different liquids. Four steps of the activity are listed below.

A. Set up the equipment to do the activity
B. Graph the data

C. Figure out how to do the activity so as to measure the times fairly
D. Drop the marble into different liquids and record the times

In what order should Jan conduct these steps?

a. A D C B
b. C A D B
c. C D A B
d. D C B A

Question 4 (Multiple Choice)

There are a few variables that could affect how far a paper airplane will soar:

- The weather
- The type of paper used
- The shape

If Sam wants to know how the type of paper used affects how far a paper airplane soars, he should:

a. Make two paper airplanes of the same size and shape and try flying them on a windy day to see how far they go.
b. Make one paper airplane and fly it on a windy day. Then fly the same airplane on a sunny day. He should record his results in a table.
c. Make two paper airplanes that are the same shape but use a different type of paper. Then he should go out on a sunny day and measure how far they go.
d. Make one paper airplane and fly it on a windy day. Then he should wait for a sunny day and make another paper airplane with a different shape and type of paper.

D2. Lab Notebook

D2.1 Worksheet 1

Worksheet 1

How does steepness affect how far the ball rolls?

(5 minutes)

Question 5 (Think Aloud)

Hypothesis: How do you think steepness will affect how far the ball rolls?

Data: Adjust the left and right ramp steepness, the left and right ramp surface, the type of ball, and the run length from the drop down menus. Then hit run and record you data. To perform your next trial, hit the reset and start over.

Trial	Right Ramp Steepness	Right Ramp Surface	Type of Ball	Left Ramp Steepness	Run Length	Left Ramp Surface	Distance

Question 6 (Multiple Choice)

From you data, how does steepness affect how far the ball rolls?

a. When steepness increases, the ball rolls further.
b. When steepness decreases, the ball rolls further.
c. Both the high and the low steepness have the same effect on how far the ball rolls.
d. None of the above. The run length, type of ball, and surface need to be considered.

Question 7 (Multiple Choice)

Briefly explain the results of your experiment.

D2.2 Direct Instruction Worksheet

Variables

(20 minutes)

The factors we change to determine how far the ball rolls such as steepness, run length, the type of ball, and surface are called **variables**. To find out how one of these variables affects the outcome of our experiment (for example, how far the ball rolls), it is important that we keep all the variables the same and only change the one we want to know about. That way we can find how only that one variable affects the outcome of our experiment. This process is called "*control of variables*".

Trial	Right Ramp Steepness (same)	Right Ramp Surface (same)	Type of Ball (same)	Left Ramp Steepness (changes)	Run Length (same)	Left Ramp Surface (same)	Distance (m)
1	High	Smooth	Rubber	High	Half	Smooth	223
2	High	Smooth	Rubber	Low	Half	Smooth	125
3	High	Smooth	Rubber	High	Half	Smooth	226
4	High	Smooth	Rubber	Low	Half	Smooth	130
5	High	Smooth	Rubber	High	Half	Smooth	225
6	High	Smooth	Rubber	Low	Half	Smooth	128

Question 8 (Think Aloud)

How did your table differ from this one?

Control of variables means we only change one variable during our repeated experiments to test our hypothesis. Look at the table. We only changed steepness for each trial and kept all other variables the same. Note: It wouldn't have mattered if we used the lighter ball, at the full distance, with a bumpy surface as long as *they stay the same during the experiment.* Also note that it is important to try the experiment multiple times. *The more data you collect, the more accurate you can assume you results to be.*

Question 9 (Multiple Choice)

From you data, how does steepness affect how far the ball rolls?

a. When steepness increases, the ball rolls further.
b. When steepness decreases, the ball rolls further.
c. Both the high and the low steepness have the same effect on how far the ball rolls.
d. None of the above. The run length, type of ball, and surface need to be considered.

Question 10 (Multiple Choice)

Briefly explain the results of your experiment.

D2.3 Progress Worksheet 2

<u>Worksheet 2</u>

How does run length affect how far the ball rolls?

Question 11 (Think Aloud)

Hypothesis: How do you think run length will affect how far the ball rolls?

Data: Adjust the left and right ramp steepness, the left and right ramp surface, the type of ball, and the run length from the drop down menus. Then hit run and record you data. To perform your next trial, hit the reset and start over.

Trial	**Right Ramp Steepness**	**Right Ramp Surface**	**Type of Ball**	**Left Ramp Steepness**	**Run Length**	**Left Ramp Surface**	**Distance**

Question 12 (Multiple Choice)

From you data, how does run length affect how far the ball rolls?

a. When run length increases, the ball rolls further.
b. When run length decreases, the ball rolls further.
c. Both the high and the low run length have the same effect on how far the ball rolls.
d. None of the above. The steepness, type of ball, and surface need to be considered.

Question 13 (Multiple Choice)

Briefly explain the results of your experiment.

D2.4 Progress Worksheet 3

Worksheet 3

How does the type of ball affect how far the ball rolls?

Question 14 (Think Aloud)

Hypothesis: How do you think the type of ball will affect how far the ball rolls?

Data: Adjust the left and right ramp steepness, the left and right ramp surface, the type of ball, and the run length from the drop down menus. Then hit run and record you data. To perform your next trial, hit the reset and start over.

Trial	Right Ramp Steepness	Right Ramp Surface	Type of Ball	Left Ramp Steepness	Run Length	Left Ramp Surface	Distance

Question 15 (Multiple Choice)

From you data, how does the type of ball affect how far the ball rolls?

a. The heavier ball rolls further.
b. The lighter ball rolls further.
c. Both the heavy and the light ball go the same distance.
d. None of the above. The run length, steepness, and surface need to be considered.

Question 16 (Multiple Choice)

Briefly explain the results of your experiment.

D2.5 Progress Worksheet 4

Worksheet 4

How does surface affect how far the ball rolls?

Question 17 (Think Aloud)

Hypothesis: How do you think surface will affect how far the ball rolls?

Data: Adjust the left and right ramp steepness, the left and right ramp surface, the type of ball, and the run length from the drop down menus. Then hit run and record you data. To perform your next trial, hit the reset and start over.

Trial	Right Ramp Steepness	Right Ramp Surface	Type of Ball	Left Ramp Steepness	Run Length	Left Ramp Surface	Distance

Question 18 (Multiple Choice)

From you data, how does steepness affect how far the ball rolls?

a. On a bumpy surface the ball rolls further.
b. On a smooth surface the ball rolls further.
c. Both the bumpy and the smooth surfaces have the same effect on how far the ball rolls.
d. None of the above. The run length, type of ball, and steepness need to be considered.

Question 19 (Multiple Choice)

Briefly explain the results of your experiment.

D3. Posttest

<u>Posttest</u>

(10 minutes)

Question 20 (Multiple Choice)

What is the most important thing to remember about variables when doing a science experiment?

a. You can keep them all the same and record the results of the experiment in a table.
b. You can change them one at a time to find out how that one variable affects the outcome of your experiment.
c. You can change them all at the same time to see how they affect the outcome of your experiment.
d. They don't change.

Question 21 (Multiple Choice)

Why is it important to run the same experiment multiple times?

a. To collect enough data to support your conclusion.
b. For organizational purposes.
c. To see how many variables you can change.
d. It is not important. One trial can tell you how steepness affects how far a ball rolls.

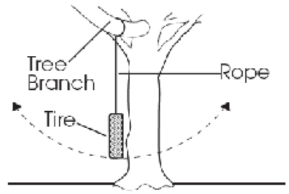
Question 22 (Multiple Choice)

Which statement describes the best procedure to determine if a vaccine for a disease in a certain bird species is effective?

a. Vaccinate 100 birds and expose all 100 to the disease.
b. Vaccinate 50 birds, do not vaccinate 50 other birds, and expose all 100 to the disease.
c. Vaccinate 100 birds and expose only 50 of them to the disease.
d. Vaccinate 50 birds, do not vaccine 50 other birds, and expose only the vaccinated birds to the disease.

Question 23 (Multiple Choice)

A class investigating the motion of a tire swing collected the data in the table below. The students were able to draw conclusions about the factors that affect the motion of a swing. Two students from the class decide to use the class data to build a different-size tire swing in their backyard. They build the tire swing shown in the diagram.



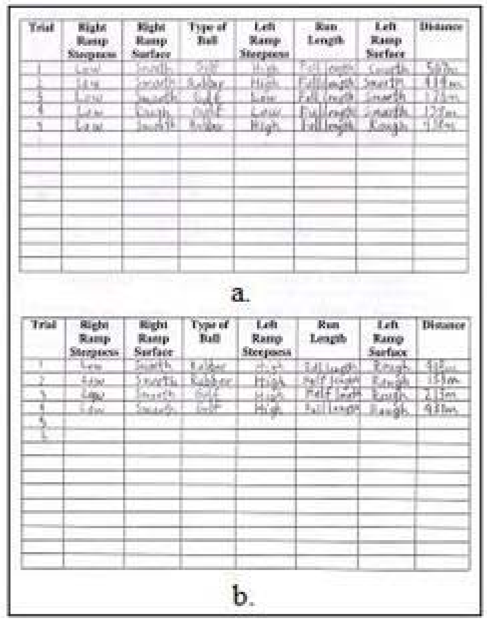
Tire Swing

Tire Swing Investigation Data			
Swing	Length of Rope (meters)	Mass of Tire (kilograms)	Time it Takes for the Tire Swing to Move Back and Forth Once (seconds)
1	2	10	2.8
2	2	20	2.8
3	4	10	4.0
4	4	20	4.0

After testing the swing, they decide that they want to make it swing faster. Based on the data from the class investigation, what could the students do to make their tire swing move back and forth faster?

a. Use a shorter rope
b. User a longer rope
c. Use a less massive tire
d. Use a more massive tire

D4 Sample Tables 1



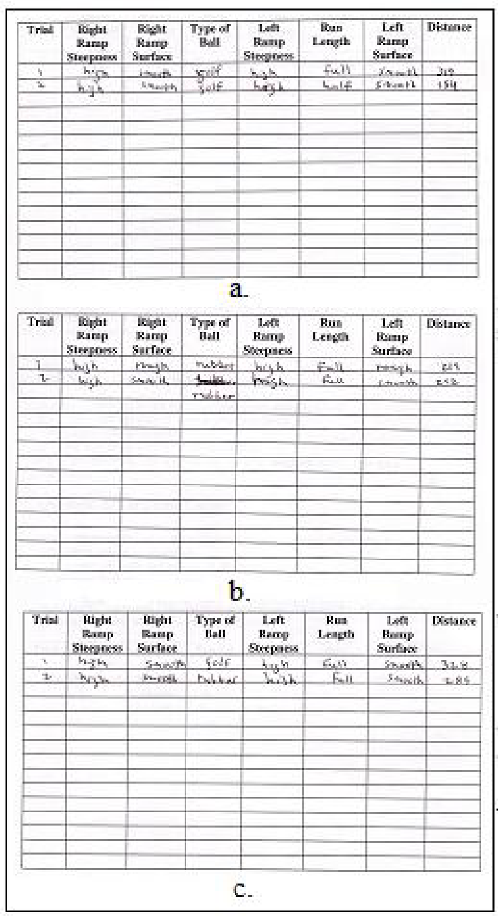
Trial	Right Ramp Steepness	Right Ramp Surface	Type of Ball	Left Ramp Steepness	Run Length	Left Ramp Surface	Distance
1	Low	Smooth	Golf	High	Full length	Smooth	[illegible]
2	Low	Smooth	Rubber	High	Full length	Smooth	[illegible]
3	Low	Smooth	Golf	Low	Full length	Smooth	[illegible]
4	Low	Rough	Golf	Low	Full length	Smooth	[illegible]
5	Low	Smooth	Rubber	High	Full length	Rough	[illegible]

a.

Trial	Right Ramp Steepness	Right Ramp Surface	Type of Ball	Left Ramp Steepness	Run Length	Left Ramp Surface	Distance
1	Low	Smooth	Rubber	[illegible]	Full length	Rough	[illegible]
2	Low	Smooth	Rubber	High	Half length	Rough	[illegible]
3	Low	Smooth	Golf	High	Half length	Rough	[illegible]
4	Low	Smooth	Golf	High	Full length	Rough	[illegible]
5							
6							

b.

D5 Sample Tables 2



Trial	Right Ramp Steepness	Right Ramp Surface	Type of Ball	Left Ramp Steepness	Run Length	Left Ramp Surface	Distance
1	high	smooth	golf	high	full	smooth	[illegible]
2	high	smooth	golf	high	half	smooth	154

a.

Trial	Right Ramp Steepness	Right Ramp Surface	Type of Ball	Left Ramp Steepness	Run Length	Left Ramp Surface	Distance
1	high	rough	rubber	high	full	rough	[illegible]
2	high	smooth	~~golf~~ rubber	high	full	smooth	[illegible]

b.

Trial	Right Ramp Steepness	Right Ramp Surface	Type of Ball	Left Ramp Steepness	Run Length	Left Ramp Surface	Distance
1	high	smooth	golf	high	full	smooth	328
2	high	rough	rubber	high	full	smooth	[illegible]

c.

References

Alonzo, A. & Aschbacher, P.R. Value Added? *Long assessment of students' scientific inquiry skills*. Presented at the Annual Meeting of the American Educational Research Association. San Diego, CA, April.

Baxter, G. P., & Shavelson, R. J. (1994). Science performance assessments: Benchmarks and surrogates. *International Journal of Educational Research, 21*, 279-298.

Black, P. (1999). *Testing: Friend or Foe? Theory and Practice of Assessment and Testing*. New York, NY: Falmer Press.

Gobert, J., Beck, J., Sao Pedro, M., Heffernan, N., & Richardson, J. (2008). Scaffolding and Logging Students' Scientific Inquiry within Science ASSISTments program. To be presented at the International Conference of the Learning Sciences. June 23-28th, Utrecht, The Netherlands.

Gotwals, Amelia W. and Songer, N.B. (2005). The Symbiosis of Cognition, Observation, and Interpretation in an Assessment System for BioKIDS. In S. Barab and K. Hay (Eds.) *The Proceedings of the International Conference of the Learning Sciences.*

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science*, *15*(10), 661-667.

Kuhn, D. (2005). *Education for thinking*. Harvard University Press, Cambridge, MA.

National Committee on Science Education Standards and Assessment (1996). National Education Standards, Washington, DC: National Academy Press.

Pellegrino, J., Chudowsky, N., & Glaser, R. (2001). Knowing What Students Know: The Science and Design of Educational Assessment. Washington, DC: National Academy Press.

Ruiz-Primo, M.A. & Furtak, E. (2004). *Informal assessment of students' understanding of scientific inquiry.* A paper presented at the American Education Research Association annual meeting, San Diego, CA.

Schunn, C., Raghavan, K., & Cho, NB. (2007). *Domain-general Learning Accelerators in Middle-School Science.* Paper to be presented at the annual meeting of the American Educational Research Association, Chicago, IL. April 9-13.