Inquiry Space: Using Graphs as a Tool to Understand Experiments

Introduction

In our experience, high school students often see graphs as the result of an assignment, a final product to be constructed correctly at the end of a lab and handed in. Students learn to plot points in order to display the results of a scientific experiment or they attempt to associate equations with the proper curves. Scientific inquiry requires students to develop both process skills (Roth & Roychoudhury, 1993) concerning graphing, and critical thinking skills based on the use of graphical analysis (Ben-Zvi, 2004) as a tool to understand scientific outcomes. Rarely are students asked to use graphs to inform decisions during scientific inquiry investigations, to use graphs as a data analysis tool, or to problem-solve scientific issues. In this paper, we address what happens when students encounter graphs while using an integrated data analysis environment to conduct data analysis during physics experiments. We investigate two research questions:

(1) Whether and how did elements within the software environment appear to facilitate student reasoning about patterns in data? about how graph features related to their experimental results?

(2) Whether and how did students make use of the different kinds of graphs and other data representations in the software?

This paper describes work from the *InquirySpace* (IS) project. IS, a materials development/research project, offers students authentic science inquiry experiences complete with data collection and analysis using both hands-on and virtual experiments. The project has created software, curriculum, and educational strategies to extend the ways in which students can engage in inquiry-based experimentation in high school physics classrooms. The set of browser-based activities supports students in reasoning about multiple experimental runs and emergent patterns in the data by the coordinated use of several features embedded in the software environment. The IS curriculum requires students to complete multiple experimental runs that generate time-series data and then to define variables and design graphs to represent their data. In this process, students are confronted with expected and unexpected variability both within and across runs.

During the classroom observations reported here, we saw instances of students using graphs to think about variations in data even when they did not appear to have a clear understanding of individual features of the graphs.

Also, we noted a number of instances of increased engagement and interesting student reasoning in response to anomalous patterns in graphs. From a large data set of screencasts of student activity, we selected for analysis episodes in which students were responding to representations of their data on the screen so that we could inquire into the nature of those responses. The intention was to provide a window into how students were reasoning about and using the various data representations incorporated into IS.

Theoretical Approach

Students have trouble understanding graphs. A wealth of research in the 1980s and 1990s demonstrated that children and adults, even undergraduate physics students, have trouble interpreting and constructing graphs and connecting them with physical concepts and real-world phenomena (e.g., Beichner 1994; Clement, Mokros, & Schultz, 1986; McDermott et al. 1987; review by Shah & Hoeffner 2002). Shah, Mayer, & Hegarty (1999) found that undergraduates often failed to glean the author's intended message when asked to describe what they saw in graphs from middle school textbooks on U.S. history and often failed to describe trends. Difficulties students have with interpreting kinematics graphs representing acceleration or velocity have been particularly well documented (Beichner, 1994; Clement 1989; Thornton & Sokoloff, 1990). More recent research investigates possible cognitive factors associated with difficulties in interpreting kinematics graphs (Kozhevnikov et al., 2002) and the extent to which teaching assistants are familiar with these difficulties (Maries & Singh 2013). Friel, Curcio, & Bright (2001) discuss issues with student use of statistical graphs and write, "Very little is known about the relationship between the development of graph comprehension and the practice of creating graphs within the context of statistical investigations."

Students bring resources to their classroom experience, including pre-instructional theories and views about graphs. Shah & Hoeffner (2002) reported, in their pilot study of middle school students' conceptions of graphs, that most recognized the communicative function of graphs. In addition, students bring their intuitive responses to data patterns (Chinn & Brewer 1993) and their pre-instructional theories, which include accurate aspects (Masnick, Klahr, & Morris, 2007). On the other hand, Masnick et al. found that *graphs were less commonly viewed as a tool for thinking about data*. This is consistent with our own experiences; we have observed many students responding to graphs as a problem to be solved, the endpoint of their explorations, not as a tool that could be useful in their inquiries. However, we will argue that the resources students bring can be leveraged to help them change their view of graphs. For instance, with the right activities, students will begin to view graphs as a tool to detect unusual features in their data, as suggested by Spence & Lewandowsky (1990).

Students can reason about variation in data even in the absence of domain knowledge. Masnick et al. (2007) presented evidence from both theory-driven and data-driven contexts suggesting that, when learning to separate signal from noise in experimental data, children and adults can use a "bootstrapping pattern" in which their mostly accurate theories guide their investigations and the resulting data then influence their theories. These investigators concluded that even in the absence of domain knowledge on which to base theoretical explanations, students from third grade to college were aware that there is variation in data and that this was something for them to consider in their reasoning. Rubin, Hammerman, & Konold, (2006) found that a statistical visualization tool appeared to enable middle and high school teachers who did not know statistics to reason about statistical issues. According to Ben-Zvi (2006) and Paparistodemou & Meletiou-Mavrotheris (2008), when using a statistical visualization tool, students from third grade up were observed using a process of "informal inference" (Rubin et al. 2006) to reason about signal and noise and other types of variability in data, and this occurred without their having been taught the mathematics of statistical inference.

Visualization tools can enhance learning in science. Studies by Finkelstein et al. (2005) and Klahr, Triona, & Williams (2007) indicated that in some situations, physics and physical science students who had used simulations had pre-post gains as great or greater than students who had conducted real experiments. Lee, Linn, Varma, & Liu (2010) showed in their cohort comparison study that visualization-rich inquiry can be more effective than typical instruction when learning complex science topics.

Use of sensors can help with graphing misconceptions and statistical issues. Thornton & Sokoloff (1990) showed that use of motion sensors by undergraduate physics students were associated with dramatic improvements on velocity and acceleration test questions. Mokros & Tinker (1987) demonstrated in a longitudinal study that graphing misconceptions could be removed with the use of sensors. Brasell (1987) showed that for high school physics students, a single class period with a motion sensor resulted in improved comprehension of distance and velocity graphs as compared with a pencil and paper graph construction control group. However, a delay in displaying the data inhibited nearly all of the learning, which indicates the importance of simultaneous or near-simultaneous display of initial graphical information from sensors. Such simultaneous display is a feature of IS.

However, the use of multiple external representations can be problematic for students. Ainsworth, Bibby, and Wood (2002) showed that students receiving mathematical and visual forms of feedback on their mathematical calculations initially did not perform as well as those receiving a single form of feedback. They believed that students were not learning to translate between the different representations. Cognitive load theory (Sweller, 1994; Sweller & Chandler, 1994) contends that instructional designs that require learners to simultaneously assimilate multiple elements of information can impose a heavy extraneous cognitive load. However, Ainsworth (2006) suggests that computers can support translation between representations. These results raise a question of how students would respond to the multiple data representations in IS.

In summary, there is a variety of literature that demonstrates difficulties students have with understanding patterns in data and in interpreting graphical data representations, although there is evidence that a coordinated use of visualization tools and sensors can help. On the other hand, the need to integrate multiple representations of information can impose an additional cognitive load. This leads us to ask whether students were able to use the IS representations to reason about their data, and if so, what about those representations appeared to facilitate their reasoning.

Research Design and Methods

The InquirySpace Environment

The IS project is a materials development/research project designed to help students achieve the skills necessary to do independent inquiry. IS is unique in that it integrates the following into a single, browser-based environment: 1) *sensor data collectors* (software developed by the project to collect data from probeware), 2) *computational models* (created with the use of NetLogo[1] and the Next-Generation Molecular Workbench[2]), and 3) *a powerful visual data exploration tool for analyzing data* (Common Online Data Analysis Platform[3]). Therefore, students not only can explore data from models, but can compare these data with real-time data from their own hands-own

[1] NetLogo is a multi-agent programmable modeling environment. http://ccl.northwestern.edu/netlogo/
[2] The Next-Generation Molecular Workbench is powerful, award-winning software that provides visual, interactive computational experiments for teaching and learning science. http://mw.concord.org/nextgen/#interactives
[3] Common Online Data Analysis Platform (CODAP), is a browser-based data exploration and visualization tool. http://concord.org/projects/codap

experiments, captured by sensors and fed into the modeling environment. Students are guided to use this browser-based environment by *a curriculum created by the IS team* that comprises a scaffolded sequence of increasingly open-ended computer-based activities using games, sensors, and computer models to develop student inquiry skills. The final activity in the IS sequence requires that students work in groups to explore a real or virtual system. Students are asked to design an entire experiment, from framing the question to reporting their findings, using multiple experimental runs and describing the patterns they discover that relate outcomes to input parameters. To simplify integration into traditional course instruction, the activities are designed as suitable alternatives to standard treatments of common physical science content such as motion, oscillation, and friction.

Several kinds of data representations are used in an integrated fashion in IS activities. Data display software uses data generated by real or virtual student experiments to produce “run-time” graphs (Figure 1b). These are generated in real time, helping students map their experiment to a graphical representation of data obtained from the experiment. Students make multiple runs for different values of one or more parameters that influence the outcome. Each run-time graph is of an individual experimental run and can be generated by a computer model-based experiment or by a real experiment with the use of sensors. Students choose whether to export some or all data from a given run to a data table (Figure 1c), a second type of representation. Data from the table can then be represented in a “time-series” graph (Figure 1d). Time-series graphs appear similar to the run-time graphs except that they are not produced in real time, they represent only the student-selected data from each run, and they can show multiple runs on a single graph. One or more outcome measures (e.g., distance traveled, terminal velocity) are computed from each run.

These outcome measures can be used to generate the third kind of graph, a “parameter space” graph (hereafter referred to as a “P-space” graph; see Figure 1e). Each point in a P-space graph represents one outcome measure derived from a single run plotted against the value of a parameter that was used during the run (e.g., terminal velocity of a run vs. size of parachute that was used). A P-space graph contains the relationships that are often sought in experiments: how something a student controls (the parameter) influences the outcome of the experiment. These relationships are often expressed as equations and the goal of instruction is often to have students find—or more commonly, verify—the equations. Science teachers may think that an equation *is* the goal and the explanation. The IS activities, however, largely avoid the equations and focus on a different kind of understanding.

IS has, as its core, a data analysis tool (Common Online Data Analysis Platform or CODAP). Success with the IS activities relies heavily on students' ability to analyze their data by reading and interpreting graphs and data tables accurately. In this environment, a graph is a representation of student-generated data and students have to make sense of it. For the purposes of this paper, when we are talking about students *interpreting graphs*, we mean *data analysis*.

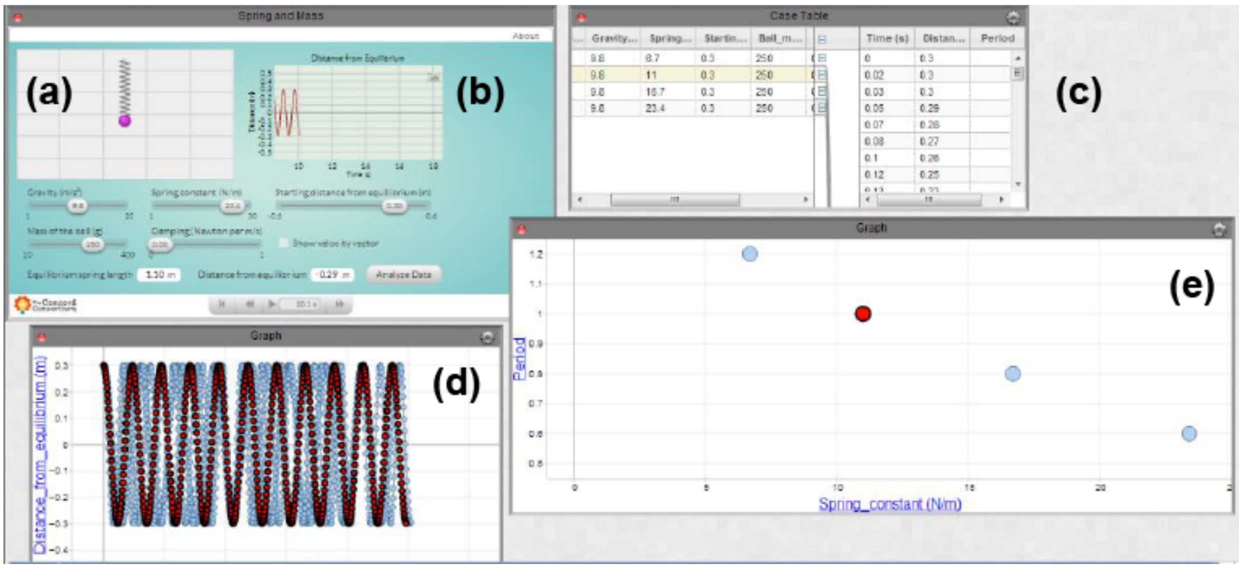


Fig. 1 The integrated IS environment. (a) Spring-Mass Model, (b) run-time graph, (c) data table, (d) time-series graph, and (e) P-space graph with one point selected (red). The window integrates an IS computational model and data display window (a and b) and data representations from CODAP, a data analysis platform (c-e). It can also integrate contributions from sensors (probeware) and include instructional videos created by the IS team.

Student Interaction with the IS Environment: An Example

Figure 1 shows the IS environment for a simulated spring-mass system. Students can conduct multiple experimental runs by varying gravity, the spring constant, the starting position and mass of the ball, and/or damping. When students finish a run, they can select part or all of their data by selecting a portion of the run-time graph (Figure 1b) and clicking the Analyze Data button. This exports their selection to the data table (Figure 1c) and adds it to the time-series graph (Figure 1d). In this instance, students have conducted four runs, varying the spring constant parameter for each run. In the data table, students have added a column for the period of oscillation of the system and inserted values for the period for each run, which they have estimated from examining the time-series graph. Students can drag variable labels from the table to the axes in the graphing area and the software will create a plot between those two variables. In Figure 1, students have created two graphs: the time-series graph (distance vs. time) and a P-space graph (period vs. spring constant, Figure 1e). The students have clicked one point in the P-space graph, the red (dark) point. In response, the corresponding row in the data table, the second row, has become highlighted and the corresponding run in the time-series graph has turned dark red.

IS displays the same integrated set of windows for the hands-on Spring-Mass experiments with sensors except that there is no computational model (Figure 1a), and the run-time graph (Figure 1b) shows data collected from the sensors.

Activity Sequence, Participants, and Method of Analysis

The Lesson Sequence

Participating classes completed a sequence of nine activities: a pre-test, the Ramp Game, two hands-on Spring-Mass activities, a Spring-Mass activity with a computer model, a Parachute activity with a computer model, a hands-on Parachute Activity, an independent project, and a post-test.

The Ramp Game, developed in response to the need for basic graph literacy, is a one-period computer activity designed to help students explore the tools embedded in the IS environment. It guides students through manipulating the model of a car traveling down a ramp and creating time-series graphs and P-space graphs. The game is designed so that interpreting the P-space graph helps students earn a higher score. (See Appendix A.)

In the Spring-Mass and Parachute activities (approximately 45 minutes per activity; see Figures 1 and 7), students complete computer model experiments and hands-on experiments where data are collected by sensors. Students are guided through the use of sensors, the graphing and data analysis tools, and the experimental process. The goals of the Spring-Mass activities are to explore the effect of mass, gravity, spring constant, and starting position on period. In the Parachute activities, students explore the effects of mass and parachute size on the terminal velocity of a parachute. Each of these five activities scaffolds students to explore the parameters, conduct one or more experiments, analyze data, explain their results, and summarize their conclusions. (For more on these activities, see Appendices B and C.)

Finally, students use the skills they have learned to undertake an independent project of their own design (approximately three class periods) by either extending the earlier experiments or using available classroom resources to invent new experiments (Appendix D).

Participants

The classroom studies described here were conducted in two suburban schools located in the northeastern part of the United States. Students in the study were approximately 87% white, 5% Asian, 4% Black and Hispanic. Approximately 10% were on free and reduced lunch. Schools were chosen for their proximity to the project development site and for the teachers' willingness to replace three to four weeks of school instructional time with the material. Teacher A taught four physics classes (mixed 11th and 12th graders) that followed block scheduling, where each class met every other day for 80 minutes. Teacher B had five classes (three freshman physics classes and two senior honors physics classes) following a rotating schedule with most classes lasting 51 minutes and a long block once per week of 80 minutes. 157 students worked in 60 small groups. In all classrooms, students completed the pre- and post-tests individually and worked in groups of two to three on the game and activities. All student work in the IS environment was collected electronically. Students also created brief (2 to 5 minutes) *student screencasts* at the end of each activity in lieu of traditional lab reports. In addition, we asked the teachers to choose small groups of varying ability from each class for more intensive observation. Teacher A chose two small groups per class while Teacher B chose three small groups per class. These groups had a different kind of screencast software running in the background when they were using their computers; their voices and all their onscreen actions were automatically recorded throughout the entire class period for every lesson in the sequence. We call these full-period recordings *research screencasts*.

At least one IS staff member was present in every class, observing the student groups and taking observation notes. During the implementation, teachers generally focused students on what needed to be covered and then interacted with student groups as needed. There were few full class discussions or demonstrations. IS staff helped troubleshoot software issues and sometimes intervened to instruct students on how to use the software, to help with data analysis, or to discuss physics content.

Data Sources and Method of Analysis

For the purposes of this qualitative study, the period-long research screencasts and observer notes proved most informative. We collected 219 research screencasts from 21 small groups of either two or three students. From this large corpus of data, we used purposeful sampling (Glaser & Strauss, 1967, 1970) to identify lessons and specific episodes for analysis. Our purpose was not to test existing hypotheses, nor to establish frequencies of phenomena, but to generate rich descriptions of student activity (Clement, 2000). We wanted to identify episodes that could provide a window into how students were reasoning about and using the various data representations in IS. For this kind of analysis, it is appropriate to select for analysis those portions of data that contain the types of activity to be described (Miles & Huberman, 1994). Before conducting our observations, we selected five lessons to focus on for this study: the three spring-mass activities, the parachute simulation, and the Independent Experiment. After the lesson sequence was conducted, we used classroom observation notes to select 3 of the 21 small groups for qualitative analysis. Criteria for choosing the groups were: they must have had complete sets of research screencasts, represent varying ability in physics according to their teachers as evidenced by their prior work in these physics classes and supported by our observations, and they must have articulated their thinking sufficiently to allow for analysis. The three groups chosen represented both teachers, were from three different classes, ranged from freshmen to seniors, and had an even number of males and females. These 3 groups produced 25 screencasts during the 5 lessons of interest. To address research question 1, we were interested in screencasts where students discussed their data patterns. To address research question 2, we wanted screencasts in which at some point students had multiple representations of their data on screen (numerical data tables, different kinds of graphs), were focusing on them and giving some indication of their thinking. 15 of the 25 screencasts met one or both of these criteria and were subjected to further analysis.

Screencasts Analyzed (collected)			Spring-Mass Exp 1	Spring-Mass Exp 2	Spring-Mass Simulation	Parachute Simulation	Indep. Exp
Group 1	3 males	2 juniors 1 senior	**2** (2)	**1** (1)	**1** (1)	**2** (2)	**1** (2)
Group 2	1 male 1 female	2 seniors	**1** (1)	**1** (2)	**1** (1)	(2)	(2)
Group 3	3 females	3 freshmen	**1** (2)	**1** (2)	(1)	(1)	**3** (3)

Table 1. There were often multiple class periods and screencasts for a single activity. Some screencasts were not useful for analysis, for instance, if there were no graphs and no indication of student thinking.

Group 1. This group of two junior males and one senior male, with Teacher A, varied widely on their focus depending on the day. They frequently joked and teased each other. On one of the days from which episodes were drawn, they did not appear to take the activity seriously, did not read instructions, and at first did not reason about their results. However, later in the period they became engaged in response to some unexpected results.

Group 2. This group of one senior female and one senior male with Teacher B exhibited some of the most advanced reasoning we observed. Although many of the groups had moments of reasoning intently when the graphs produced odd, irregular-looking results, this was one of the few groups who reasoned deeply because their graph, though not irregular, did not reflect the relationship between variables they had anticipated.

Group 3. The three freshmen females in this group, with Teacher B, were good students according to the teacher. They exhibited some deeper reasoning but also missed some on-screen details, and were observed trying to search the Internet with a smartphone for a correct "prediction" for their experiment.

Video Analysis

The purpose of this study is exploratory and descriptive; we wish to identify classroom activity that may not have been described previously. In a book chapter that was the result of an NSF panel on methodology, Clement (2000) made the case for the importance of generative/interpretive analyses that focus on constructing new observation concepts. These studies can range from Exploratory Studies that involve the open interpretation of large videotaped episodes by a single analyst, to Model Construction studies in which investigators begin to separate theoretical concepts from observations so that observations can be compared across different subjects and episodes. Continuing along a spectrum from more generative to more convergent studies are Explicit Analysis Studies, and finally, Independent Coder studies that can establish frequencies of well-defined categories of observations. Our study is properly described as near the exploratory end of Clement's spectrum. We wished to conduct data-driven rather than theory-driven analysis (Boyatzis, 1998), and did not attempt independent coding. Rather, our goal was agreement between the two authors, and differences in our categories were debated and clarified through discussion, as described in Harry, Sturges, & Klingner (2005).

When developing the codes ultimately used to describe episodes in the videotapes, we utilized a construct development cycle (Miles & Huberman, 1994; Glaser & Strauss, 1967) leading to the progressive refinement of observation categories. This involves the analysts in a cycle of: making observations from each segment, formulating a hypothesized construct for, or classification of, the student reasoning or activity, returning to the data to look for more confirming or disconfirming observations, comparing the classification of the statement to other instances, criticizing and modifying or extending the hypothesized category to be consistent with or differentiated from other instances, returning to the data again, and so on; followed by a stage of joining constructs that are too similar and rechecking against instances in the data. Sets of examples of each kind of activity are assembled into a descriptive study of each. Our goal for this first stage was the assembly of descriptive studies of the most interesting kinds of activity identified. When student activity that appeared to offer insight into our research questions was observed, we attempted to generate a rich description of the activity and its context.

First, the 15 screencasts were watched by one or both of the authors. We flagged all episodes where students were either a) working with the graphs and tables or b) responding to unexpected data patterns. We also looked for any episodes where students were engaged in metacognition; in other words, where they were evaluating their own thinking or their own experimental process. Codes such as “Data collecting,” “Reasoning about real world causes,” or “Describing pattern in p-space graph” were then assigned to these episodes. These observation categories were developed iteratively as additional episodes of student reasoning were identified and examined in light of previous episodes (Miles & Huberman, 1994; Glaser & Strauss, 1967). Once the observation categories stabilized, all relevant portions were re-coded. We then asked what implications each observation category had for our research questions and grouped similar categories into broader themes, in a process of inductive and deductive thinking that drew from Strauss and Corbin’s (1998) method of Axial coding.

This paper focuses on the three broader themes that emerged from this process, and constitutes a descriptive study of each, using student episodes as exemplars.

Results

The first two themes speak directly to the two research questions. The third theme was an emergent one that arose during this exploratory study.

Theme 1: A Strong Role for Graphical Anomalies

Research Question 1 asked what elements in the software environment appeared to facilitate student reasoning about patterns in the data. Videotape analysis suggested that it might not have been specific elements, but that something about the IS environment as a whole seemed to result in an unexpectedly strong role for graphical anomalies in motivating students to reason about data patterns.

In the present context, when we refer to an *anomaly* in a graph, we mean a data pattern that appears “odd” to the students. These anomalies could be unexpected outliers or irregularly shaped curves. During classroom observations of students working during the IS project, we noticed several strong reactions to unexpected shapes or patterns in the graphs, even when students did not appear to have any clear expectations concerning an outcome. At other times we observed students ignoring anomalies. Of the 15 screencasts analyzed, nine of them had episodes when anomalies were present and students were responding to them. However, this was likely not representative of the classes as a whole. Our goal was not to establish a representative frequency of episodes, but to find instances of phenomena and describe them. Therefore, we used purposeful sampling, which cannot indicate frequency or typicality of phenomena.

Reactions to these anomalies varied, and ranged from ignoring certain data points to redoing an experiment. We wanted to look more deeply at such episodes to see what was occurring. For instance, we wanted to know whether students were simply responding to what they perceived as “ugly graphs” or whether they were interpreting the meaning of the anomalies and taking appropriate actions in response. Three episodes can serve as exemplars to illustrate several different ways students responded to anomalies within the IS software environment.

Episode 1: Students Recognize Anomaly as Bad-Data Issue

The students in Group 1 had just completed an experimental run for a physical experiment in the first Spring-Mass activity, in which they had set in motion a mass on a spring. The position sensor data appeared fairly smooth but the time-series graph had an odd discontinuity at the end.

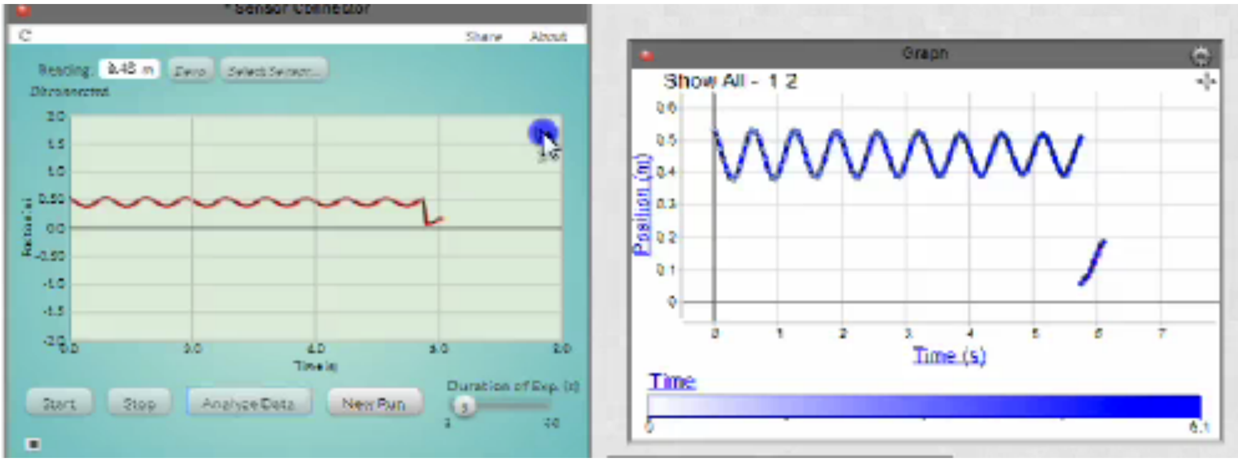


Fig. 2 “Run-time” position sensor data (left) and time-series graph (right).

S1: What’s with that?

S2: That’s when he moved it.

S1: Ah.

S1: Ooh, two data points.

S2: Yeah, you need to destroy that data.

(S3 agrees. The students delete the data from this run and re-do the run.)

Here S1 was surprised at an aspect of the time-series graph, the odd mark at the end of the run. S2 connected this mark to an action of S3, who had apparently moved the sensor before data collection was done. Although the students did not discuss the specifics of the graphical discontinuity, they did recognize that it was an anomaly and they mapped it to a real-world cause. They decided quickly that this was an issue of bad data (though they did not use that term) and deleted the data.

In this instance a student was able to reason quickly about data from a graph and make appropriate decisions concerning those data even though, as can be seen in Episode 5 below, he did not appear to have a particularly robust understanding of graphs. When the anomaly appeared, all three of these often-disengaged students responded immediately. They were able to move quickly to redo their experimental run and to collect data needed to address the question posed by the activity. Within another three minutes, they had two runs they liked and had generated a P-space graph.

Episode 2: Students Decide to Treat Anomaly in P-Space Graph as Noise

The same group, a few days later, was working with the Parachute computer model, varying the mass of the jumper for each run (each parachute drop) and determining the terminal velocity for that mass and parachute. Because they were not paying close attention, they exported one run incorrectly, resulting in the wrong value of mass being registered in the data table. The result was a beautiful time-series graph of five runs, but the data in the table looked odd to the students. They interpreted the anomalous pattern in the data table to mean that they had made two duplicate runs. They deleted one of the "duplicates," which, however, had been a correctly registered run. Not noticing that there was still an error in their data table, the students generated their P-space graph. Next they needed to place a "fit line" to figure out the relationship between the mass of the jumper and the terminal velocity.

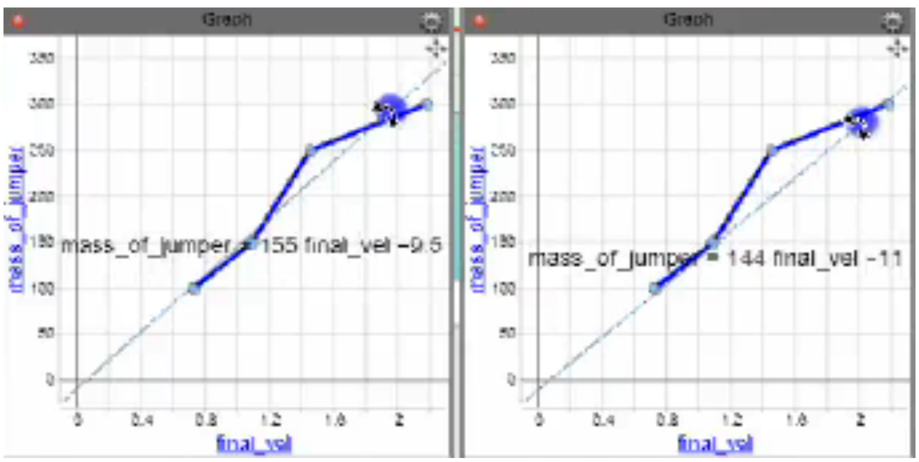


Fig. 3 Fit line feature fit to all four points (left), and fit to three points only with fourth point treated as outlier (right).

At this point the students noticed that their results were not in a straight line. S1 wanted to split the difference with the fit line (Figure 3a), while S2 noticed that three of the points did lie in a straight line and wanted to fit the line through them (Figure 3b). They discussed this for about a minute before S2 agreed that splitting the difference was probably all right. Even though they did not take the time to track down the source of the anomaly, the lack of symmetry in the graph and the requirement to fit a line to the data did catch these students' attention and appeared to elicit some awareness that they might want to ignore one of the points.

Although certain aspects of the graph were constructed for them, crucial aspects, such as determining the terminal velocity from data, choosing which runs to include, choosing variables for the axes of the P-space graph, and deciding where the fit line should go, were left up to them. This is an instance where students who initially appeared disengaged nonetheless were forced to make a decision as to whether or not to include a specific data point in their answer. As they became aware of two possible ways to deal with the wayward point, even though they did not make the best choice, we suggest they were coming to an intuitive and operational sense concerning properties of data, including that an outlier could be considered either acceptable variation or “bad data.” We will see in a later episode that this group was provoked by another outlier to engage in a considerably more intensive investigation.

Episode 3: Students Identify Cause of Anomaly and Interpret Its Meaning

Group 2, with two students, generated the P-space graph in Figure 4 to illustrate the results of a spring-mass experiment to investigate the relationship between period and amplitude.

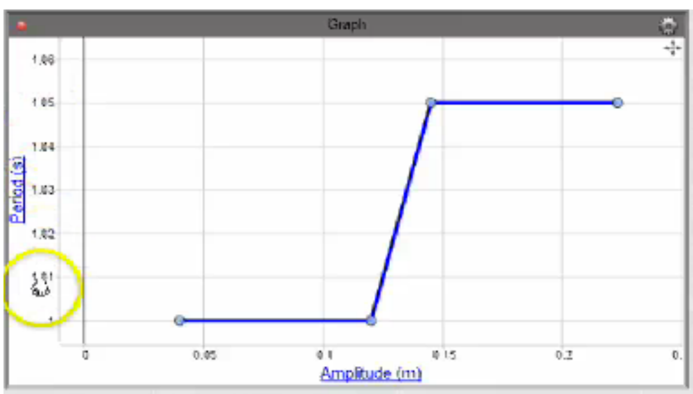


Fig. 4 P-space graph of spring amplitude versus period.

S4 *(referring to P-space graph)*: Oh that’s awkward because we have those two trials there. (*laughs*) So it should be a straight line, but instead it looks like a step function.

(They humorously speculate whether or not the step function is integrable, then click “show connecting lines.”)

S4: That's so unfortunate! (*laughing*)

S5: Yeah. But that’s because the scale is really small. *(Moves cursor along y-axis.)*

S4: I know, the scale is tiny.

S5: So, you know what --

S4: It's OK.

They decided that the graph would be a straight line if the y-axis were more "zoomed out." (One student joked that "if you squint so much that your eyes are closed, it would look better.")

In general, Group 2 responded to anomalies not only by mapping them to a physical action, but also by digging into the cause behind the specific graphical appearance. To them, this graph didn't just look "odd"; the shape meant something and they made sure that both members of the group understood this meaning before they moved on. In this case, they decided that the anomalous shape was not important and that they could ignore it. Although the graph did not meet their expectations, they were quickly able to determine correctly that what they were seeing in the graph reflected acceptable variation rather than bad data.

This was a group that had prior knowledge of graphs; nonetheless, features of the software environment prompted them to think both quickly and deeply about details of the graphical representation in order to interpret it in light of their experiment. In a later excerpt, we will see that their prior knowledge of graphs did not lessen the power of unexpected graphical results to trigger a strong response and impressive reasoning.

These three episodes illustrate something we saw repeatedly; many students, of varying levels of ability and prior knowledge, appeared to respond to the appearance of anomalous features of graphs before they knew how to interpret those anomalies. We suggest that the near-immediate display of experimental data in conjunction with a view of the computer model or physical experiment facilitated this response. We were impressed with the role graphical anomalies appeared to play in the present context, to provoke student reasoning and to trigger engagement. We hypothesize that students had a certain proprietary interest in these anomalies because of the close, and visible, association with their own actions.

Theme 2: Coordinated Use of Data Representations and Features to Make Sense of Puzzling Data

Research Question 2 asked whether and how students were able to make use of the different kinds of data representations in IS; these include data tables and several kinds of graphs (Figure 1). We wanted to know whether students were able to make connections between the representations, whether they appeared to experience cognitive overload (Sweller, 1994) from the multiple representations (Hegarty, Kriz, & Cate, 2003), whether they chose one kind of representation to focus upon, or whether they responded in some other way. We found that students did use the multiple representations in a coordinated way, especially when they were attempting to make sense of puzzling data.

Uniquely to IS, the data representations are linked; a change in one representation will produce an immediate change in the linked representations. The IS activities were designed to focus students on making correlations between variables rather than on the mechanics of using data points to create graphs on paper. This could help reduce cognitive load that might otherwise be created by the process of graph production and allow students to focus on more conceptually demanding aspects of the tasks. Because IS allows students to conduct multiple runs quickly, we believed students would, as a result, more easily see patterns and trends in their data.

During the classroom trials, it was clear that many students could follow the visual connections from one representation to another. Beyond this, we were surprised at the facility with which we saw students moving from physical set-ups to data tables, from data tables to time-series graphs, and from time-series graphs to P-space graphs and back again. This quick shift between data and graphical representations happened most often when students were trying to make sense of puzzling data. Thus, this theme builds on our first theme about the role of graphical anomalies.

The next two episodes serve as exemplars to show ways in which students could use multiple coordinated features to interpret anomalies and take appropriate action in response. At times this meant that they were able to distinguish good from bad data more quickly, freeing them up to consider the physical meaning of graphs in the presence of noise. At other times their investigations led them spontaneously into explorations of details of the graphical representations of their data, including how the graphs were constructed, though this was not the emphasis of the lessons. (Again, the purpose of this descriptive study is not to attempt to establish frequency data; we make no claim of typicality, but only that we observed multiple instances of this kind of behavior during our observations.)

Episode 4: Students Use Multiple Features in the IS Environment to Deal Quickly with Variation in Their Data

For their independent project, S6 and S7 (Group 3) had chosen to conduct experiments with a fan cart, a small cart propelled by a battery-operated fan. After a struggle, they finally produced a run with the cart that the motion sensor could register in the way they wanted. They selected the portion of the run-time graph that they felt represented good data (Figure 5, shaded portion of graph on the left) and exported those data to the data table and to two different time-series graphs, a distance-time and a speed-time graph. Although the students had exported an apparently smooth section of data, the speed-time graph that was produced was not completely smooth and had an anomaly at the end (most likely when the fan cart was no longer moving in proper alignment with the motion sensor). The students recognized the anomaly, laughing and pointing to it. They could not easily exclude these data once their time series graph was constructed, but they could have rejected the entire run. Instead, they turned on the fit line and positioned it in a way that dealt with the noise and ignored the anomalous portion of data. They then used the slope of the fit line to calculate the acceleration of the fan cart (Figure 6).

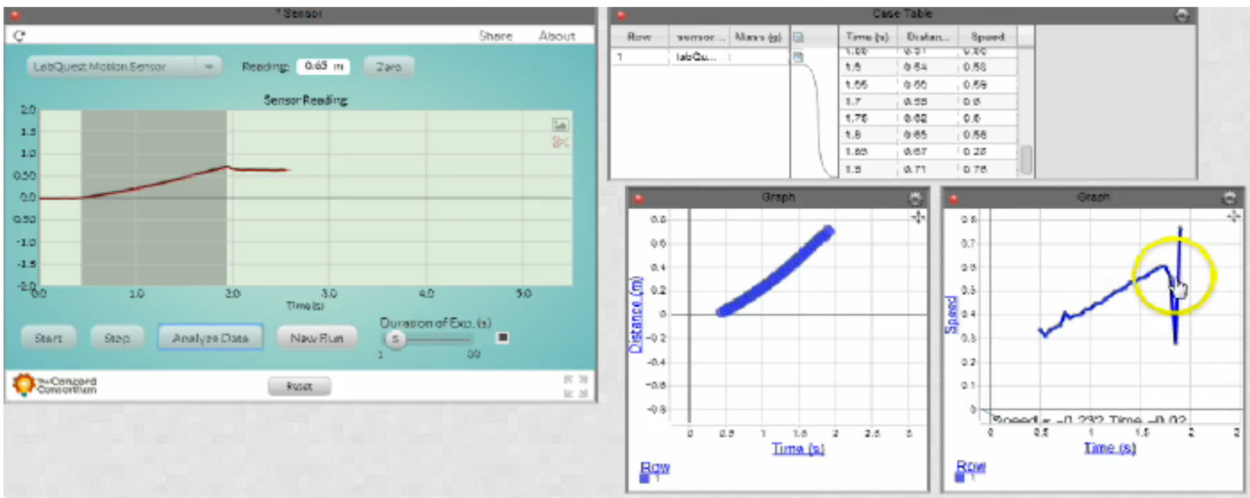


Fig. 5 Shaded area of the run-time graph is the portion of data the students have chosen to export for analysis. The time-series graphs on the right reflect only this portion of data.

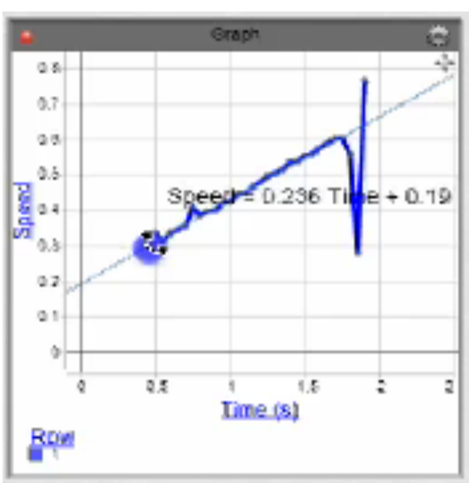


Fig. 6 Time-series graph with a fit line that treats the graphical anomaly as ‘bad data.’

Speed-time or velocity-time graphs tend to be very difficult conceptually for students to interpret correctly (Beichner, 1994; Clement 1989; McDermott, Rosenquist, & van Zee 1987). Height tends to be more salient for students than slope, resulting in Slope/Height confusion (Clement, 1989; Bell & Janvier, 1981; Leinhardt, Zaslavsky, & Stein, 1990; Planinic, Milin-Sipus, Katic, Susac, & Ivanjek, 2012) and students have considerable difficulty determining slopes (Beichner, 1994). We believe the CODAP fit line feature provided a scaffold that made slope a more salient visual aspect for the students. In addition, students did not have to calculate the average slope from the complicated graphs produced by these experiments, which typically contained both noise and anomalous data. Rather, they could match the fit line in a way that focused on the important variation in data for that experiment. The equation for the line automatically appeared along with the fit line, giving students easy access to the numerical value of the slope. These students were able quickly to identify the valid data in each of the graphical representations and employ CODAP's fit line feature to calculate an accurate answer.

Episode 5: Disengaged Students Use Multiple Representations to Figure Out a Puzzling Graph

Due to a lack of attention, Group 1 had intermingled data in their data table from two different parachute experiments, both conducted with the Parachute computer model, one in which they had varied the mass and one in which they had varied the parachute size. In addition, they had inadvertently completed two runs with identical values, one in each experiment, leading to nine data rows but only eight points on the P-space graph. (See Figure 7.)

Just before the following episode began, the students had expressed surprise at the odd shape of their P-space graph (Figure 7g). In this excerpt, as they worked to figure out the meaning of their data, they moved their cursor back and forth between the data table and the P-space graph.

S2: Wait a second! Hold on. Hold on. This straight line, that's the first half of the data. *(Indicating four data points in a horizontal line at the top of the P-space graph.)*

S1: Yeah, but it doesn't mean anything because it's all different kinds --

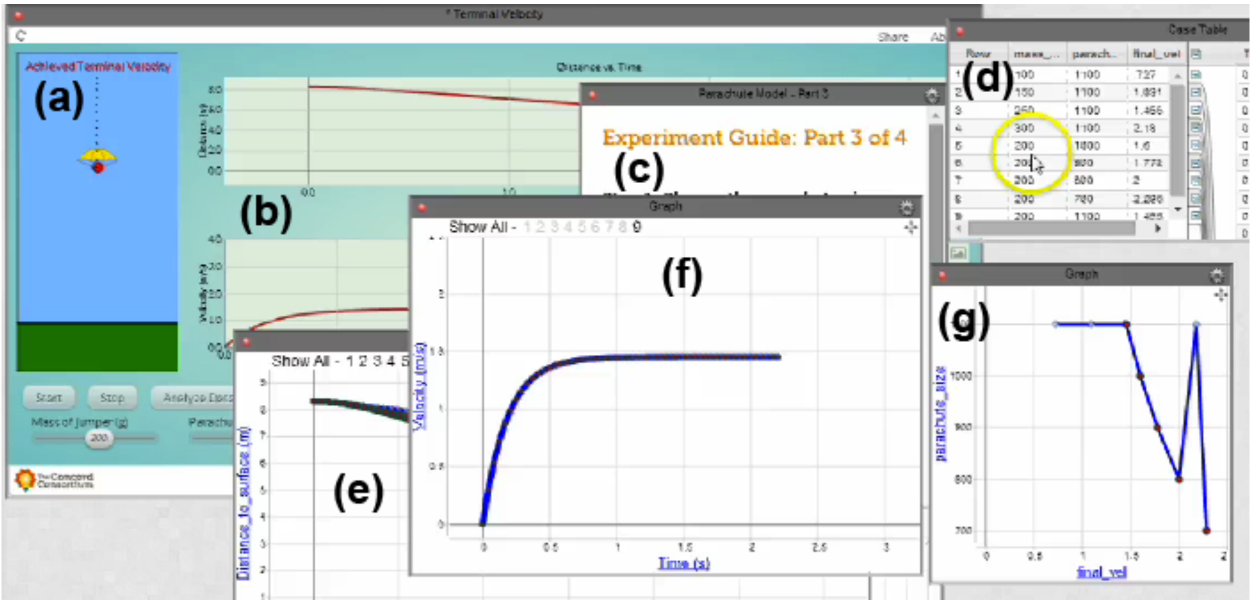


Fig. 7 (a) Parachute model, (b) two run-time graphs, (c) instructions, (d) data table, (e) time-series graph from their first experiment (mostly hidden), (f) time-series graph from their second experiment with only Run 9 showing, (g) P-space graph with anomaly.

S2: Right. So technically, the graph starts like that. *(Presumably gestures over the screen.)* Now I don't know why there's this big sp- Ohh. *(There is a spike in the graph.)* I know why. *(Scrolling through the data table.)* I know why it does that. Because we did that one last. So it goes in order of what we did. I don't think we can just rearrange it.

S2's last statement was incorrect; unlike the data table, the graph did not reflect the order in which the students conducted their runs. The spike appeared because the software couldn't make sense out of the two superposed graphs. In spite of S2's lack of clarity about how the graph was constructed (which persisted until the following day), in this episode he was able to visualize how the first graph would have looked had it not been superposed (it would have been a horizontal line). A moment after the above episode, he also offered a description of how the second graph would look, "a diagonal line going down." Although he did not mention that the "diagonal line" would be curved, his description and gestures indicated the points that properly belonged to the second experiment.

These students' attentions were arrested by a P-space graph that looked odd to them and they used additional data representations as they tried to figure out why. The P-space graph revealed a pattern (in fact, two superposed patterns) not easy to see in the data table, while the data table contained information about mass not in the graph. By moving back and forth between the two representations, S2 was able to connect individual experimental runs in the data table to the corresponding dots in the P-space graph and to see the pattern for each experiment. He was likely aided in this by an embedded CODAP feature: when S1 highlighted all the rows of Experiment 2 in the data table, the corresponding points in the P-space graph turned red (the darker points in Figure 7g), providing a visual cue to the results for that experiment.

The next day these students were faced with a similar situation; after they had separated their data into two P-space graphs, one of the graphs again had an outlier. This time S1 clicked on the outlier and observed which row in the data table lit up and determined that the point actually belonged to the other experiment. The two students then carefully clicked on each and every point in each of their P-space graphs to activate the visual connections between those points and the data table to insure that their data were now divided correctly between the two graphs.

By the end, these students were making connections between the nine runs of their model-based experiments, two time-series graphs, a conflated P-space graph, and the data table. Being able to connect separate elements across the data representations instantly, by highlighting data points across the representations, allowed them to connect the graphs to their own interactions[4] with the model, to increase their understanding of the different kinds of graphs, and to begin to investigate the physical relationships being represented. These students clearly had a strong sense that the lack of symmetry in the graph indicated a problem. We were struck again by how a graphical anomaly within the integrated IS environment had the power to lead relatively unengaged students to focus deeply on specific features of the graph, and to stay focused until they understood those features.

This and the previous episode illustrate how students could use multiple coordinated features within the environment to analyze their data and to troubleshoot unexpected issues. In both episodes, students were reasoning from multiple kinds of data representations, one of which showed an anomaly. They used additional features from the data analysis platform (fit line, coordinated highlighting of data points) to help them determine which patterns in their data were meaningful. Rather than producing cognitive overload, the visually coordinated features appeared to scaffold the students in becoming comfortable with, and adept at interpreting, unexpected data patterns that arose from their actions.

[4] e.g., systematically changing parameters

In general, we observed students moving fluidly between graphs, models, and data tables. Though the IS environment and embedded software were all designed with this in mind, the IS team (and the software teams) continue to be surprised with the ease with which these students were able to do this.

Theme 3: Moving Beyond Viewing Graphs as Task to Viewing Them as Tool

The third theme emerged during our analysis and moves beyond the research questions that motivated the study. During our analysis, we identified what could be termed an *epistemology of graphs*. Students appeared to frame the tasks in a variety of ways, including: “just going for the answer”; trying to understand graphing conventions and aspects of graphical representations; using graphs to try to figure out what happened in their experiment, including what constituted bad data or noise and how to treat these; using information in graphs to try to reason about real-world causes of phenomena.

These frames appeared to fall into two larger categories, depending on whether the graphs were viewed

- as a product, a task to be completed or
- as a resource that could provide information about their experiment.

When students were creating a graph primarily because the instructions said to do so, understanding the basics of the graph was important to them only inasmuch as it helped them produce an acceptable product. Early in the activity sequence, at least one group was observed not even looking at the P-space graph they had produced, satisfied merely with having produced one. Other groups focused on determining whether their graphs were acceptable. As the groups progressed through the activity sequence, however, we observed students beginning to view the graphs as a resource, a tool they could use to help them understand something about their experiment or about a scientific concept.

The points at which students began applying information from their graphs to help make decisions about their experiments appeared to occur at different times for different kinds of graphs. In general, students seemed to analyze run-time and time-series graphs for quality of data almost immediately. S2, for instance, was able to explain the discontinuity in the time-series graph of Episode 1 in relation to an action in the experiment (“That’s when he moved it”) and to use the graphical information to determine that those data were not valid.

The shift beyond viewing P-space graphs only as the successful completion of a task appeared to take longer. However, as the activities progressed, we noticed that students more often employed these graphs as a source of information about the validity of their data and the outcomes of their experiment. In Episode 5 above, S2 struggled to interpret the meaning of the P-space graph, but in spite of his confusion about the ordering of runs along the x-axis, he realized immediately that the anomaly indicated a problem with the group's data and he was motivated to figure out what was going on. In the next episode, we infer that S4 and S5 had, by this time, also begun to see the P-space graph as a source of information—in this case, information about the real world.

Episode 6: Using a Pattern in a P-Space Graph to Reason About the Underlying Physical Principle

This episode illustrates some of the most sophisticated reasoning we observed.

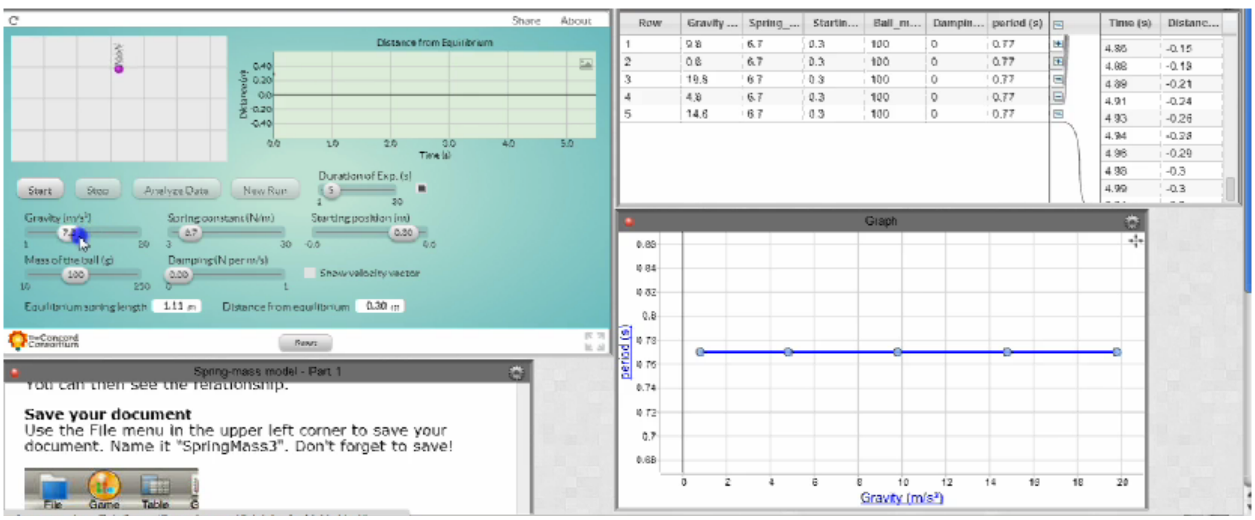


Fig. 8 Spring-Mass Model at rest, run-time graph cleared of data, data table with five runs, and P-space graph showing no relationship between acceleration of gravity and period. Student is manipulating the gravity slider in the model window (fuzzy blue dot below the model).

S4 and S5 (Group 2) had used the Spring-Mass Model to investigate how the period might be influenced by changes in gravity. (See Figure 8.) In the excerpts below, they were trying to answer a question in the lesson, "How would you explain this result?" They knew how to interpret the P-space graph—they could see that the period did not change when they changed the acceleration of gravity—but this result did not make sense to them. As the episode began, S4 compared these results to the results of their earlier hands-on experiment in which the period *had* varied with different masses. Their hands-on experimental set-up—a mass hanging from a spring—was still nearby. (The transcript is supplemented with descriptions of student actions from the observer notes. Ellipses indicate where the transcript has been lightly edited to reduce repetition.)

S4: I don't know, the other one *(experiment)* made sense to me, but this one -- does not. *(Handles mass on spring, pulls it down.)* A spring on Jupiter. A spring on the moon. *((Mutters too softly to hear.))* You'd think it would take longer to *((move through a period))*. *(pause)* I guess when gravity -- when gravity is less powerful, then the tension in the spring is less powerful, cause it's being -- no, it's being stretched out the same amount because the amplitude is the same. Why doesn't the amplitude change?!

(They spend a minute and a half searching the web for an answer but give up.)

S4: If gravity is more powerful, it will pull it down -- I just don't understand! Because the tension in the spring ... isn't affected by gravity. Right? ... So that means it should go further and take longer for the tension to pull it back up.

S5 *(softly)*: That didn't happen.

S4: My hypothesis is that the simulation *(model)* is wrong! There is no way to explain this with science! Uhhh -

(They laugh and look back at the model.)

S4: Here. What if we, let's vary like, *two* things. So this is --

S5: No! Because --

S4: No no, I just want to see what happens. Because I need gravity to affect *something*!

(S4 moves the gravity slider slowly back and forth. The simulated mass moves up and down slightly. S5 notices that the reading for equilibrium spring length is changing as S4 moves the slider. At first they sound shocked, but then they decide that with less gravity acting on it, the mass will hang higher.)

S4: Oh wait, so if it's higher up, when you start, like at low gravity – (…) and then you pull it down to here, there's going to be less tension in the spring than if it were regular gravity and you pull it down to here *(demonstrating with the nearby spring-mass set-up)*. There's more tension here -- Ooh! OK. I got it. Alright. So, there's more tension, alright.

S5: Yeah. Explain it to me?

S4 *(using the spring-mass set-up to demonstrate)*: So at low gravity, it's, the equilibrium length is like here, right? *(Holds mass up higher than its actual equilibrium point.)* So then the distance would be here *(lowers the mass)*. But from our gravity, *that's* the starting position. (...) Whereas here *(holds mass up higher than its equilibrium point),* the gravity is weaker, but you have less tension when you pull it, so it can still have the same motion. *(pause)* OK? Does that make --

S5: More or less.

(They start typing their answer, discussing how to word it in terms of the equilibrium spring length and the distance from equilibrium.)

These students responded to an unexpected graph as an indication that something about their thinking wasn't making sense. They viewed their graph as a source of information about their physical and virtual experiments as well as a foundation to reason more broadly about physical concepts.

This was not the only group that attempted to do an Internet search to answer one of the "why" questions and then gave up and reasoned out an answer on their own; Group 3 was observed doing this on another day. The fact that Internet searches did not help either group is an indication that the kinds of reasoning the students were being asked to engage in was of a higher order[5] than could be satisfied by quoting from authority (in this case, the Internet as "authority") or doing "plug-and-chug" with memorized equations. S4 and S5 did not spend a lot of time plotting graph points or going back to their distance-time graphs to figure out where they might have made a graphing mistake; the software had largely handled these technical aspects. Instead, the students focused on making sense of their results by using the experimental set-up and features of the computer model to come to a new understanding of the spring-mass system. They concluded by distinguishing between *equilibrium length* and *amplitude* in the case of a mass in harmonic oscillation. This is a distinction that can prove difficult even for college students who have studied harmonic motion (Frank, Kanim, & Gomez, 2008).

---

[5] Learning taxonomies in education reform tend to list thinking that involves evaluating, analyzing, and creating as higher order than learning from memorization or citing from authority. See Anderson, L., & Krathwohl, D., Eds. (2001)*, A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives,* Allyn and Bacon ISBN 978-0-8013-1903-7; and National Research Council (1987), *Education and learning to think*, Washington, DC: National Academy Press. Also see Roth, W. M., & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching*, *30*(2), 127-152 for higher order skills in the context of scientific inquiry.

Two Additional Episodes That Exemplify Theme 3

Groups 1 and 3 were also observed using graphs as a source of information about their experiments. During their final project with the fan carts, Group 3 spent 20 minutes over two days (ending shortly before Episode 4) responding to unexpected patterns in their time-series graphs. They used information in these graphs to help them redesign their physical set-up and to identify issues in the equation they were using to calculate speed, taking their experiment through multiple revisions before they were satisfied. Group 1 spent approximately 13 minutes over two days (Episode 5 is part of it) using their graphs to help them disentangle data from two experiments with the Parachute Model. Although the three groups varied widely in the depth of the reasoning they were able to achieve, each group spent substantial time and effort using their graphs as tools to help them think about their experiments and to revise either their experiments or their thinking.

Discussion

We have identified three major themes that we believe offer insights into our research questions asking whether and how elements in the IS software environment appeared to facilitate student reasoning about data patterns and how, if ever, students used the various kinds of data representations in that environment. We observed: 1) a strong response to graphical anomalies; 2) a coordinated use of data representations and features to make sense of puzzling data; 3) a movement beyond viewing graphs as task to viewing them as tool.

Visual anomalies in the graphs appeared to produce reactions from students and to provoke increased engagement and reasoning by many students. Research Question 1 asked what elements in the software environment appeared to facilitate student reasoning about patterns in the data. Our impression from the videotape analysis was that, rather than certain individual elements being of special importance, the whole environment seemed to work together in a way that resulted in a strong role for graphical anomalies in motivating students to reason about data patterns and how they related to the experimental results. We hypothesize that this was because there were strong associations for the students between the anomalous data represented in the graphs and their own actions that had produced the data. In the case of the model-based experiments, the associations were visual; in the case of the laboratory experiments and motion sensors, the associations also arose from physical interactions (as in Episode 6). Even when students appeared not to have strong prior conceptions about an outcome, if the symmetry of the resulting graphs did not match their intuitive sense of the actions they had taken, they responded to these representations immediately. These visual anomalies could be produced by errors in executing an experimental run, noise in the data, or mistakes in calculating an outcome, and appeared to lead students into reasoning about validity and reliability of their data. Most of the reactions we saw to anomalous data were appropriate to the situation, as groups decided to reject, redo, or ignore their data. Some groups moved beyond these practical decisions to reason deeply, to reconsider and revise their thinking. These findings agree with those of Brasell (1987) and Mokros & Tinker (1987) who found that use of sensors could increase comprehension of graphs. They also agree with Spence & Lewandowsky (1990), who pointed out that graphs can be used as tools for the detection of unusual features in data. These students certainly used graphs in that way, and for some students, the unusual features appeared to serve as a springboard to deeper reasoning.

Research Question 2 asked whether students were able to make use of the different kinds of data representations and we found that they did, and appeared especially motivated to do so when their data were puzzling. Even when there were no visual anomalies, if a graph did not make sense to students, a *coordinated use of features in the integrated IS environment* (such as connecting their actions with a physical set-up to the resulting graphs, or using scaffolds such as the fit line) appeared to enable students to focus on reasoning about the data rather than to focus on the mechanics of producing the graphs. Our results agree with Rubin, Hammerman, & Konold (2006) and Masnick, Klahr, & Morris (2007) that domain novice learners can reason about variation in data.

More specifically, we observed students using the *multiple kinds of data representations* in a coordinated way, fluidly following the visual connections back and forth between the model and the representations, and among the different representations, especially when they were trying to make sense of unexpected data. If a student clicked on an outlier in a P-space graph, for instance, the corresponding data table row and the corresponding run in the time-series graph would highlight (a feature of the CODAP representations), encouraging students to reason from the perspective of these different representations. The visual connections between corresponding data in different representations are intended to help students organize their perceptions, consistent with many of the principles described in Vekiri (2002) and Ainsworth (2006).

This does not mean that the complex visual environment was not a challenge for students; it was, especially in the beginning of the lesson sequence. However, much of this challenge appeared to be in making efficient use of the onscreen real estate in the relatively small display area on the available computers so that they could see all of their representations at once. As the lesson sequence progressed, we observed students becoming facile with moving the different representations around so that they could observe the connections between them. An example is in Figure 7, where the students were trying to map the connection between a single point in 7g to a data row in 7d to a single run in 7f—which they managed to do while keeping the IS instructional scaffolds in 7c and the computational model in 7a and 7b visible enough for quick reference.

The third theme was an emergent one that arose during this exploratory study, and goes beyond the initial research questions that motivated the study. In investigating how students made use of data representations in the software, we have identified what appeared to be a shift in many students' *epistemology of graphs*, a shift in how students were framing their activities with the graphs they were constructing. In the beginning, many students appeared to be approaching the graphs as tasks to be completed to an acceptable level of accuracy (or just completed, period) and handed in. Reminiscent of Shah & Hoeffner's (2002) findings for slightly younger students, many of these high school students did not appear initially to view graphs as a tool for thinking about data. However, as the activity sequence progressed, we began to observe most students approaching the run-time and time-series graphs as resources that could provide information about their experiment. Some groups also used the P-space graphs in this way, utilizing information in these graphs to help decide whether to accept or reject runs. At other times, the P-space graphs served as an impetus for students to reason about underlying physical relationships and to re-examine their own prior assumptions, as in Episode 6.

Further Discussion

Going beyond our data, we offer hypotheses that can explain our results.

The intention of the InquirySpace project was not to teach students to draw graphs but to have them make sense of graphical representations of their data and to use their graphs in the process of data analysis. With this integrated software environment, the loop is very short between students' data analysis and revising/repeating their experiment; when students suspect that something is not right about their results, they can redo the experiment in a matter of seconds and see the differences represented immediately. Because the students can quickly perform and analyze many experimental runs, their attention can become focused on patterns *across* runs and this, we believe, leads to moments of deep thinking.

Everything is connected in the IS environment, including the physical experiments, computational models, graphical representations, and other data representations. Because the visual analytical tools connect so well with the hands-on experiments that students are running, we believe sense-making becomes front and center, happening frequently and in real time about real-world phenomena in the labs. In our classroom observations and in the screencasts, we observed students engage in such *higher order reasoning* as:

- looking for trends;
- developing their own questions even when not directed to (e.g., why does that graph look that way);
- diagnosing and repairing (even when not directed to).

We believe students have the opportunity to focus on these kinds of higher order reasoning because they spend little—if any—time recalling rote knowledge and because they are not bogged down in the less important details of creating the graphs and data tables. Instead, they come to new understandings about data analysis through informal inference (Rubin, Hammerman, & Konold, 2006). We saw them *reasoning about properties of aggregate data*:

- signal and noise (Episodes 1-4)
- variability due to errors of measurement (Episode 2)
- variability due to multiple causes (Episode 5)
- sample-to-sample variability (Episodes 2 and 3).

We note that students' new understandings about noise and quality of data were largely implicit, as was any shift in how they viewed graphs. We suggest that these students had reached a point where they could have benefited from having their new understandings made explicit. Future research could investigate what kinds of curricular supports could help with this. A more challenging question is what kinds of supports could help teachers a) recognize these shifts in understanding when they occur and b) help make them visible to the students. One possibility is to provide more scaffolds to teachers to help them seed whole class discussions.

Significance

These results suggest that there are other ways to introduce data analysis and graphical representations than to focus first on the mechanics of graph-making. If students have puzzling data that they have some investment in, such as from experiments they themselves have designed, it may be that their existing intuitions about what they think the variation should be could lead them quite naturally into an exploration of their results, the meanings of the representations of those results, and the mechanics of how those representations were constructed. Anomalous features in graphs may be a positive instructional feature.

Therefore, introducing graphs as a tool that can aid investigations, with many of the mechanics supported by software, offers the potential to be a more powerful introduction to graphs, graphing, and the concept of variation in data than the pedagogically methods customarily employed. This suggests that a process of data collection, exploration, visualization, data modeling, and the interpretation of multi-level data may be a more effective introduction than tackling these many aspects piecemeal. We believe this has important implications for teachers, lesson design, and the design of pedagogical materials.

References

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and instruction*, *16*(3), 183-198.

Ainsworth, S., Bibb, P., Wood, D. (2002). Examining the effects of different multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences, 11*(1), 25-61.

Beichner, R. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics, 62*(8), 750-762.

Bell, A., & Janvier, C. (1981). The interpretation of graphs representing situations. *For the learning of mathematics*, *2*(1), 34-42.

Ben-Zvi D. (2004). Reasoning about data analysis. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 121-145). Dordrecht, the Netherlands: Springer Netherlands.

Ben-Zvi, D. (2006). Scaffolding students' informal inference and argumentation. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics.* [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Sage.

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching, 24*(4), 385-395.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*(1), 1-49.

Clement, J. (1989). The concept of variation and misconceptions in Cartesian graphing. *Focus on Learning Problems in Mathematics*, *11*, 77-87.

Clement, J. (2000) Analysis of clinical interviews: Foundations and model viability. In A. Kelly and R. Lesh (Eds). *Handbook of research design in mathematics and science education* (pp. 547-590). Mahwah, NJ: Lawrence Erlbaum.

Clement, J., Mokros, J., & Schultz, K., (1986). *Adolescents' graphing skills: A descriptive analysis*. Paper presented at the annual meeting of the American Educational Research Association (AERA), San Francisco, CA.

Finkelstein, N., Adams, W., Keller, C., Kohl, P., Perkins, K., Podolefsky, N., et al. (2005). When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics-Physics Education Research*, *1*(1), 010103.

Frank, B., Kanim, S., & Gomez, L. (2008). Accounting for variability in student responses to motion questions. *Physical Review Special Topics-Physics Education Research*, *4*(2), 020102.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education*, 124-158.

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory. *Strategies for qualitative research. London: Weidenfeld and Nicolson*.

Glaser, B. G., & Strauss, A. L. (1970). Theoretical sampling. In N. K. Denzin (Ed.), *Sociological methods: A sourcebook* (pp. 105-114). Chicago: Aldine.

Harry, B., Sturges, K. M., & Klingner, J. K. (2005). Mapping the process: An exemplar of process and challenge in grounded theory analysis. *Educational researcher*, *34*(2), 3-13.

Hegarty, M., Kriz, S., & Cate, C. (2003). The roles of mental animations and external animations in understanding mechanical systems. *Cognition & Instruction, 21*, 325–360.

Klahr, D., Triona, L., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science teaching*, *44*(1), 183-203.

Kozhevnikov, M., Hegarty, M., & Mayer, R. (2002). Revising the visualizer—verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction, 20*(1), 47-77.

Lee, H.-S., Linn, M. C., Varma, K., & Liu, O. L. (2010). How does technology-enhanced inquiry instruction with visualizations impact classroom learning? *Journal of Research in Science Teaching, 47*(1), 71-90.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of educational research*,*60*(1), 1-64.

Maries, A., & Singh, C. (2013). Exploring one aspect of pedagogical content knowledge of teaching assistants using the test of understanding graphs in kinematics. *Physical Review Special Topics-Physics Education Research*, *9*(2), 020120.

Masnick, A., Klahr, D., & Morris, B. (2007). Separating signal from noise: Children's understanding of error and variability in experimental outcomes. In M. Lovett, & P. Shah (Eds.), *Thinking with data* (pp. 3-26). Mahwah, NJ: Erlbaum.

McDermott, L., Rosenquist, M., & van Zee, E. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics. 55*(6), 503-513.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.

Mokros, J., & Tinker, R. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching, 24*(4), 369-383.

Paparistodemou, E., & Meletiou-Mavrotheris, M. (2008). Developing young students' informal inference skills in data analysis. *Statistics Education Research Journal*, *7*(2), 83-106.

Planinic, M., Milin-Sipus, Z., Katic, H., Susac, A., & Ivanjek, L. (2012). Comparison of student understanding of line graph slope in physics and mathematics. *International journal of science and mathematics education*, *10*(6), 1393-1414.

Roth, W. M., & Roychoudhury, A. (1993). The development of science process skills in authentic contexts. *Journal of Research in Science Teaching*, *30*(2), 127-152.

Rubin, A., & Hammerman, J., & Konold, C. (2006). Exploring informal inference with interactive visualization software. In *Proceedings of the Sixth International Conference on Teaching Statistics.* Cape Town, South Africa: International Association for Statistics Education. https://www.stat.auckland.ac.nz/~iase/publications/17/2D3_RUBI.pdf. Accessed 9 March 2015.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 14*(1), 47–69.

Shah, P., Mayer, R., & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, *91*(4), 690.

Spence, I., & Lewandowsky, S. (1990). Graphical perception. *Modern methods of data analysis*, *94*, 13.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Thousand Oaks, CA: Sage.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, *4*(4), 295-312.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction 12*(3), pp. 185-233.

Thornton, R. & Sokoloff, D. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. *American Journal of Physics 58*(9), pp. 858-867.

Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14*(3), 261-312.

Appendix A

Appendix A: The Ramp Game

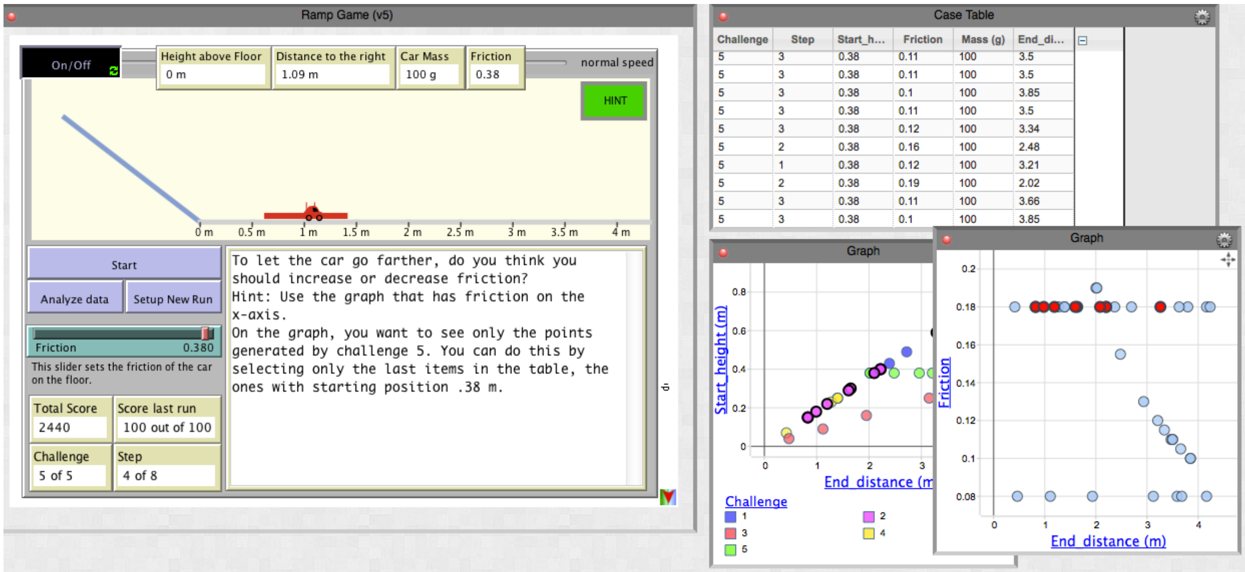


Figure A-1. The Ramp Game

In a series of activities, students could vary the car starting position on the ramp, the car mass, or the friction. The objective was to get the car to stop within the target zone, which changed position with each run. As the challenges increased in difficulty, the target zone became smaller.

Each run, whether successful or unsuccessful, produced data for a given independent variable that could be extrapolated by the student to estimate the starting conditions needed in order to hit each new target zone.

Appendix B: The Spring-Mass Activities

The three spring-mass activities used the set-up of Figure 1. The activity sheet for the first activity begins on the next page. Students were allowed some leeway to go at their own pace, so not all the small groups in a given class were on the same activity on the same day.

Appendix B

Name ______________________________________ Class________________

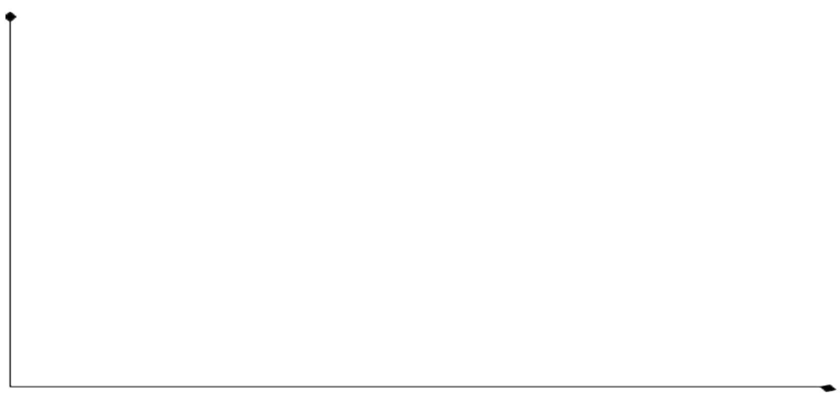
Team login______________________________

InquirySpace Mechanics Exploration

A. Introductory investigation

QUESTION:
How does the mass affect the period (time for one cycle) of a weight hanging on a spring?

Draw the graph you think shows this relationship. (You haven't even taken the data yet? How can you know? Do your best!)

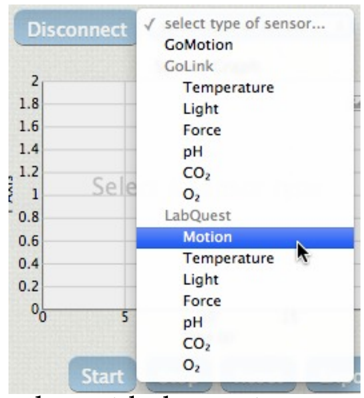


PLAN:
Investigate this question experimentally, using the equipment and sensors provided. Try different masses and measure the period in each case, using the motion sensor to make a record of the up-and-down motion of the mass. Then figure out the period of that motion.

EXPLORE AND PRACTICE
Go to the reference site (http://islinks.concord.org). All links are listed here.
Watch the video "Spring-mass experiment setup" at islinks.concord.org showing your experimental equipment.
Set up the mass-spring system. Start with one mass and one spring.
Watch the video "Collecting sensor data in DG" at islinks.concord.org showing how to collect and analyze data using the motion sensor.

Appendix B

Set up the data collection equipment: the LabQuest interface should be connected to your computer with a USB cable, and the motion sensor connected to the LabQuest (DIG1) with its special cable.
At islinks.concord.org, click the top link, Sensor-data-collector in Data Games.
ALWAYS LOG ON with the team login assigned to you by your teacher. Otherwise your data will not be saved. The username and the password are the same. A box called InquirySpace Importer appears.



Under select type of sensor, under LabQuest, choose Motion.

With the spring setup motionless, press the Zero button to zero the motion sensor.

Try

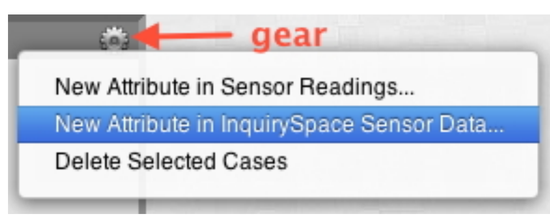
collecting data with the motion sensor.

After you press Stop, press the Autoscale button to make your graph expand.

Practice until you can collect a smooth run of at least 10 cycles.

EXPERIMENT

1. *Watch the video* "Analyzing sensor data in DG" at islinks.concord.org.
2. When you have a good run, click Export Data to send it to DG ("DataGames")
3. Make sure you have a table, and make a graph that shows the sensor reading as a function of time. Do this by dragging the time and reading headings into the X and Y axes of the graph.
4. Make new columns for mass and period in the "runs" (left-hand) section of the table. To make a new column:
    1. In the gear menu in the table, choose "New Attribute in InquirySpace sensor data".
    2. Enter the name of the new column (e.g., period).
    3. Do *not* enter anything in the "formula" box below the name!



5. Using the graph, measure the *period* for this run. Be sure everyone in the group understands how you decided to measure the period.

6. Enter values for mass and period in the table. Assume the cup weighs very little and that each 2-ounce mass is 60 grams. *Do not enter units*!

Now you have data for one complete run of your experiment. You need more runs!

7. Before each run, press "Reset" to clear the data in the data-collection box.
8. Record at least three runs. Record at least 10 cycles for each run. Be sure to export the data into DG. Enter the mass and period for each run.

ANALYZE DATA

Watch the video "analyzing sensor data in DG" at islinks.concord.org.
Create a new run-level graph in Data Games. Plot period against mass for each case. That is, drag mass (the column label) to the *x*-axis and period to the *y*-axis.
Use the graph to answer the question: "How does the mass affect the period (time for one cycle) of a weight hanging on a spring?"

SUMMARIZE WITH A SCREENCAST

Don't be shy! These screencasts are meant to provide a pleasant and expressive way for you to describe your work. This doesn't need to be a perfect movie and it won't go out on the web.

Make a *two-minute-max* screencast using *Jing* which includes the following:

- Make sure your graphs are visible on the recorded screen. Use the graphs to explain what patterns you found in the data.
- Identify the members of your team.
- State your question.
- Explain your procedure for collecting the data.
- Identify the variables and describe how you measured them.
- Describe the pattern you found between mass and period.
- Propose why you think this pattern exists.
- Describe any problems you had collecting data and how you overcame them.
- Explain how confident you are of your results and what you would do next to make them stronger.

Name and save your screencast in your team folder on the desktop. Use a name that includes the date and the team login.

Also save your screencast using the "FTP" button. This uploads it to the private Concord Consortium server.

Appendix C

Appendix C: The Parachute Activities

The Parachute model activity used the set-up in Figure 7 (Episode 5). In a sequence of activities, students varied the mass of the jumper, size of the parachute, and acceleration of gravity to see the effect on the terminal velocity of the jumper. A hands-on parachute activity was also attempted by many groups, using coffee filters as parachutes and lead weights as jumpers, but practical issues with the materials made it difficult to get to the data collection stage.

Appendix D: The Independent Experiment

The activity sheet for the Independent experiment begins on the next page.

E. Explore your own question

Now it's time to think of your own questions and explore them. You can go through several cycles of experimenting and exploration if you have time.

Plan of work

QUESTION

Decide with your team what you want to explore. To get you started, there is a separate handout with a list of possible topics.

Write down a **testable question** or questions that you will begin with. Remember that you can use a physical setup or simulations or both.

Your question should be testable with an experimental setup using models or sensors. The sensors you have available are force and motion.

Your question should be as specific as possible.

Your question should be as clear as possible.

Your testable questions(s):

____
____
____
____
____

What do you predict about the answer to your question?

____
____
____
____
____

Explain how you came up with your prediction.

____
____
____
____
____

PLAN

As a team, make a <u>rough</u> plan of how you will proceed.
<u>Make a short screencast</u> summarizing your rough plan.

- State your testable question.
- Describe your experimental setup.
- Describe the variables under consideration and how you are going to measure them.
- State what you expect to find.

Appendix D

MESS AROUND

Set up your experiment and try things. Treat this time as "practice" to figure out what behavior and relationships seem interesting, what you want to measure, and how to get good measurements.

REVISED QUESTION AND PLAN

Make a more careful plan of how you will proceed.
Make a short screencast summarizing your revised plan.

- State your testable question.
- Describe your experimental setup more precisely.
- Describe the variables under consideration and how you are going to measure them.
- State what you expect to find.

EXPERIMENT

Take measurements and send them to the DataGames chart for analysis, as before.

ANALYZE

Use the DataGames graphing and calculations to analyze your results.

SUMMARIZE

Make a short screencast which includes the following:

- Make your graphs visible on the recorded screen. Explain what they show about the variables you have chosen to explore.
- Identify your team.
- State your question.
- Explain your procedure for collecting the data.
- Identify the variables and describe how you measured them.
- Describe the patterns or relationships you found.
- Propose why you think these patterns exist.
- Describe any problems you had collecting data and how you overcame them.
- Explain how confident you are of your results and what you would do next to make them stronger.
- Describe something that was interesting or surprising that you would like to explore further.
- If your experiment has led to new questions or puzzles, state them.

Name and save your screencast as before.