# EXPANDING OUR UNDERSTANDING OF STUDENTS' USE OF GRAPHS FOR LEARNING PHYSICS

By

James T. Laverty

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics – Doctor of Philosophy

2013

# ABSTRACT

## EXPANDING OUR UNDERSTANDING OF STUDENTS' USE OF GRAPHS FOR LEARNING PHYSICS

By

James T. Laverty

It is generally agreed that the ability to visualize functional dependencies or physical relationships as graphs is an important step in modeling and learning. However, several studies in Physics Education Research (PER) have shown that many students in fact do not master this form of representation and even have misconceptions about the meaning of graphs that impede learning physics concepts. Working with graphs in classroom settings has been shown to improve student abilities with graphs, particularly when the students can interact with them.

We introduce a novel problem type in an online homework system, which requires students to construct the graphs themselves in free form, and requires no hand-grading by instructors. A study of pre/post-test data using the Test of Understanding Graphs in Kinematics (TUG-K) over several semesters indicates that students learn significantly more from these graph construction problems than from the usual graph interpretation problems, and that graph interpretation alone may not have any significant effect.

The interpretation of graphs, as well as the representation translation between textual, mathematical, and graphical representations of physics scenarios, are frequently listed among the higher order thinking skills we wish to convey in an undergraduate course. But to what degree do we succeed? Do students indeed employ higher order thinking skills when working through graphing exercises? We investigate students working through a variety of graph problems, and, using a think-aloud protocol, aim to reconstruct the cognitive processes that

the students go through. We find that to a certain degree, these problems become commoditized and do not trigger the desired higher order thinking processes; simply translating "textbook-like" problems into the graphical realm will not achieve any additional educational goals. Whether the students have to interpret or construct a graph makes very little difference in the methods used by the students.

We will also look at the results of using graph problems in an online learning environment. We will show evidence that construction problems lead to a higher degree of difficulty and degree of discrimination than other graph problems and discuss the influence the course has on these variables.

To my father, my hero, though I've never been able to say it to his face; And to my mother, the eternal educator, who raised me to be the person I am today.

## ACKNOWLEDGMENTS

I would like to thank my advisor, Gerd Kortemeyer for his guidance and for putting up with me for the past few years.

I would like to thank my group of friends (past and present) who supported me during the graduate school years. Despite all attempts of graduate school to take my sanity from me, I still have a few shreds of it left because of them.

I would like to thank Bob Geier for finding the time to help me with my dissertation. As for the rest of the CREATE for STEM Institute, I would like to thank them all for helping me bridge the gaps between the Physics and the Education communities and for providing me a second home.

I would like to thank my graduate committee for helping me through the graduate school process.

Lastly, I would like to thank Bhanu Mahanti, without whom I may never have made it to graduate school in the first place.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# Chapter 1

# Introduction

## 1.1 Overview

In this dissertation, I will be discussing how students use and learn from graphs in introductory physics classes. As always, we begin with a review of the current literature on the subject. After this, we will look at four separate but related studies on the topic that I have worked on. In Chapter 2, I will discuss the development of the Function Plot Response, a new problem type in the LON-CAPA homework system that allows students to construct graphs for themselves to submit as answers to homework problems. Chapter 3 discusses the usage of this system in real life courses. In Chapter 4, we will look at how students solve graph homework problems in physics, focusing on the strategies they use and what they might learn from them. Chapter 5 takes a step back and discusses what we might learn about graph problems from the large data sets that already exist from students trying to solve them over the years. And, finally, we will conclude in Chapter 6 by summarizing the results and discussing implications for future work.

## 1.2 Review of Literature

The ability to understand graphs is an invaluable tool in science, and life in general. A 2002 study found that the average number of graphs appearing in scientific journals almost doubled between 1984 and 1994, while the number of graphs found in newspapers more than doubled in the same timeframe[1]. The science education standards treat graphs as a base component of science learning[2] and graph literacy has been described as one of the most important abilities for developing scientific literacy[3, 4].

Two essential skills for any scientist or engineer are understanding data sets and visualizing the dependency of variables, and graphs are an essential tool for accomplishing either of these. In physics, the ability to work with graphs is important to these tasks because graphical representations allow larger trends to be more easily found and understood while keeping smaller details visible. However, instructors seem to be failing students when it comes to educating them on the use and understanding of graphs. Students frequently lack understanding of the subject matter behind them, fail to understand the connections between graphs and the real world, and have difficulties reading and interpreting graphs.[5, 6, 7, 8, 9]

Many introductory physics students do not attempt to gain a deep understanding of physics. In early physics courses, students often find that using expert-like methods (those necessary to actually understand the physics and do more complex problems) are more time-consuming than trivial methods (finding a formula and plugging in numbers); since most students are trying to get a good grade in the class in the least amount of time, they do not bother to learn expert-like methods.[10] The result is students who are capable of solving problems but don't understand the physics that the problem is trying to help them understand.[5]

When it comes to graphing, student strategies are no different. Before graphs become truly useful tools, students have to overcome difficulties translating between the graph, the real world, and formulas,[7] and using graphs appears to be more time-consuming (and confusing) than using strategies that, in the long run, are less effective. Thus, unfortunately, both students and instructors frequently attempt to navigate around these challenges. Students are rarely asked to construct graphs that have meaning. Instead, they are often asked to interpret features of graphs or to plot points to make a graph. This is, of course, because such problems are easier to write and can be multiple choice, which makes grading them much easier and less time consuming. Such algorithmic procedures allow students to get the right answer (which makes them believe they know what they are doing) without understanding what the graph represents. Asking students to construct graphs from other information will hopefully impose a deeper understanding of the material on the students. As Leinhardt, Zaslavsky, and Stein noted, "[W]hereas interpretation does not require any construction, construction often builds on some kind of interpretation."[11] Instead of constructing graphs, "[students] are usually given a formula or asked to select the appropriate formula from a well-defined (and very short) list and then to manipulate it using techniques from algebra or calculus."[12] In other words, students in introductory physics classes do not need to have any understanding of the physical world to solve most of their homework problems. Clearly, this is not what we (as educators) actually want. And even if students do understand the physics, many attempt to answer questions about graphing problems independently of the graphs involved.[13]

Student difficulties with graphing have been studied and many common misconceptions have been identified.[6] These include

- discriminating between the slope of a graph and its height.
- discriminating between changes in height and changes in slope.
- relating different graphs to each other.
- matching narrative information with relevant features.

- interpreting the area under a graph.
- connecting graphs to the real world, such as
  - representing continuous motions by continuous lines.
  - separating the shape of the graph from the (physical) path of the motion.
  - representing negative velocities on a velocity vs. time graph.
  - representing constant acceleration on an acceleration vs. time graph.
  - distinguishing position vs. time, velocity vs. time, and acceleration vs. time graphs from each other.

All of these common challenges need to be addressed in order to help all students effectively use graphs. In addition, Dykstra and Sweet noted that many students develop a "snapshot" view when it comes to changes in motion. They often refer to motions as being fast in the beginning and then slow at a later time, without noting the continuous change in the object's speed. Forty of the 99 students in the study drew a velocity vs. time graph as a series of step functions. Dykstra and Sweet further conclude that an understanding of changes in velocity (acceleration) is a necessity to understanding Newtonian forces.[14] This indicates that using step functions for velocity graphs may be misleading or even detrimental to students.

The term "graphicacy" has come to mean the ability to work with graphs, in much the same way that literacy is the ability to work with text. Bertin [15] divided the questions that graphs can answer into three categories: Elementary, Intermediate, and Comprehensive. These categories have been refined over the years;[16, 17, 18, 19] see Friel et al. [20] for a review. Elementary questions involve a simple extraction of data; Intermediate questions involve identifying trends; Comprehensive questions ask students to compare whole structures of the graph. For example, regarding a position vs. time graph, an Elementary question could be “What is the position of the car at $t = 3$ seconds?” while an Intermediate question could be “During what time interval was the car moving backwards?” and a Comprehensive question could be “When did the car have the highest speed?” Wagner [21] found that elementary school students had more difficulty understanding graphs than they did pie charts, bar charts, or tables. He noted that graphs may not be as useful for answering Elementary questions, but are more useful for Intermediate, and Comprehensive questions.

For decades, fostering higher order thinking among learners of science has been a continuing and ever-present theme in educational research, educational standards, and curriculum development. Most any compilation of skills associated with higher order thinking lists the understanding and employment of graphical representations (e.g. [22]), presumably because graphs are frequently used as a tool to analyze and evaluate data, as well as a tool to create and communicate new insights. In an effort to partway move this “hidden curriculum” [23] of fostering higher order and expertlike thinking to the foreground, we as educators sometimes allow ourselves the reverse conclusion: by assigning graph problems, we hope to foster and *instill* higher order thinking (e.g. [24]). Even students who can correctly read points off graphs and correctly plot graphs often lack higher order graphing skills, such as recognizing trends (e.g., [25, 26]). In addition, how students deal with graphing depends on the resources available at the time and even on factors such as attitude [27].

Given these results, it is questionable that by using graphs, we as educators achieve the desired instructional goals (i.e., foster higher order thinking). In fact, there are many indications that we do not: all too often, introductory physics students attempt to avoid higher order thinking processes, as they frequently find that using expertlike methods (those necessary to truly understand the physics) are more time-consuming and risky than trivial methods (e.g., "plug and chug") [28].

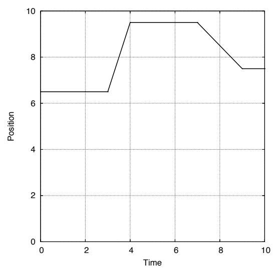
# Chapter 2

# Having Students Construct Graphs

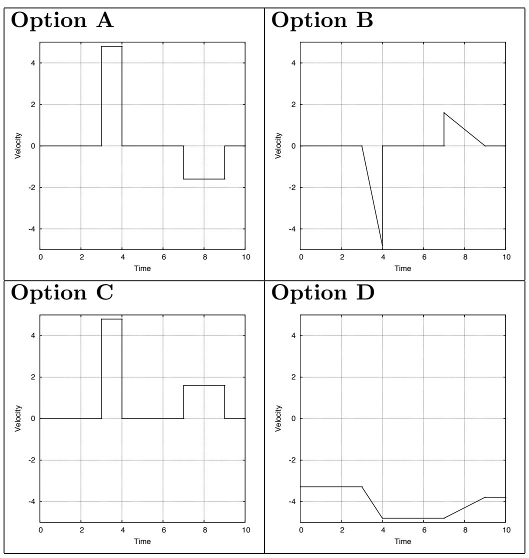
We believe that these "graphicacy" categories, designed to describe graph interpretation tasks, are mirrored in graph construction. An Elementary task would be to graph a set of value pairs or a particular function (which most students again accomplish by first calculating value pairs). An Intermediate task would, for example, be one in which a student has to draw a graph of position versus time for a car moving backwards. Finally, a Comprehensive task might be to draw the acceleration graph for a car that reaches maximum speed at a certain time. The Function Plot Response, introduced in this chapter, allows for graph construction questions from any of these three categories.

The particular strengths of Function Plot Response are that the scenarios can be randomized (different students get different versions of the scenario), that there can be more than one correct answer, and that by integrating it into the LON-CAPA learning content management system, the problems can be shared across courses and institutions.

However, before discussing the Function Plot Response, we will look at other systems that allow students to graph construction.

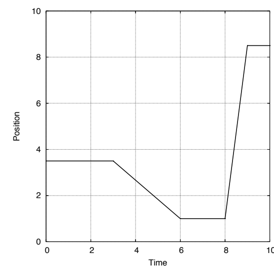



The position of a car moving in one dimension is shown as a function of time. The following are different predictions for the velocity of the car versus time:

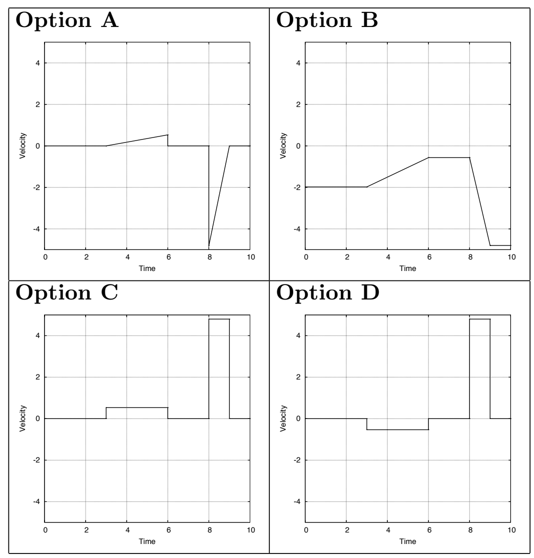



Which of these options could be true?

A. Option A
B. Option B
C. Option C
D. Option D




The position of a car moving in one dimension is shown as a function of time. The following are different predictions for the velocity of the car versus time:




Which of these options could be true?

A. Option A
B. Option B
C. Option C
D. Option D

Figure 2.1: A problem where the student must select the correct graph from a given set. Two versions are shown to demonstrate how each student receives a slightly different version of the problem. The text in this Figure is not meant to be readable but is for visual reference only.

Table 2.1: Comparison of different online graph sketching programs (Part 1).

| System | Function Plot Response | GraphPAD | MapleTA |
|---|---|---|---|
| Course System | LON-CAPA | PADs | MapleTA |
| Construction Method | Moving control points. | Creating and moving control points that snap-to-grid. | Creating and moving control points. |
| Function Type | Cubic Hermite Splines | Piece-wise Linear | Cubic Hermite Splines |
| Answer Evaluation | Server-side | Client-side | Server-side |
| Evaluation Method | Rules check values, derivatives, and/or integrals over specified or dynamic intervals, or comparing two points. | Can check values of control points, or coefficients of terms in underlying equations. | Checks points, slopes of specific points, concavity on specified intervals, and/or compares values of multiple points. |
| Feedback | Hint corresponding to first broken rule. | Concatenates feedback for all broken rules. | Adaptive feedback is possible, but technically difficult. |
| Problem Creation | Anyone with Author permission in LON-CAPA. | Must sign up for an authoring account. | Anyone with Instructor access to MapleTA. |
| Randomizable? | Yes | Yes | Yes |
| Background Plot? | Yes | Image | No |
| Multiple Tries? | At instructor's discretion | Yes | At instructor's discretion |
| Graded? | Yes | No | Yes |
| Program Required | Java™ (eventually Javascript/HTML5) | Java™ | Java™ |
| Mobile Devices? | No (but will after Javascript/HTML5 change) | No | No |
| Cost | Free, but requires server to host | Free | Subscription service |
| Topics | Any | Physics | Any |

Table 2.2: Comparison of different online graph sketching programs (Part 2).

| System | MasteringPhysics | SocraticGraphs |
|---|---|---|
| Course System | MasteringPhysics | BeSocratic |
| Construction Method | Create control points that snap-to-grid. Only one per $x$-interval. | Freehand draw, manipulate Control Points. The freehand draw has several options for interpretation. |
| Function Type | Piece-wise Linear | Cubic Bezier Splines (or Line Segments) |
| Answer Evaluation | Server-side | Client-side |
| Evaluation Method | Searches over specified ranges for values, linearity, concavity, minima, and maxima. | Rules based on evaluating minima, maxima, area under the curve, points, point pairs, slope, curve, segment, curve shape, and number of curves. |
| Feedback | Hint corresponding to first broken rule. | Concatenates feedback for all broken rules. |
| Problem Creation | Anyone with Instructor Access to Mastering Physics. | Anyone |
| Randomizable? | No | No |
| Background Plot? | Image | No |
| Multiple Tries? | Yes, reduced score for each incorrect answer. | Yes |
| Graded? | Yes | No |
| Program Required | Flash™ | Silverlight™ |
| Mobile Devices? | No | iTunes app currently in development |
| Cost | Subscription service | Free |
| Topics | Physics | Any |

## 2.1 Previous Work on Teaching Kinematic Graphs with Computers

There is considerable research to suggest that working with computers while studying kinematics is useful. Many people have studied the effects of using Microcomputer Based Labs (MBLs) to teach kinematics and shown that teaching with MBLs is better than traditional instruction [29, 30, 31, 32, 33, 34, 35]. Some evidence exists suggesting that having the graph drawn in realtime while the object is still in motion contributes most of these gains [36], but the benefits of the realtime view have not been generally confirmed [37, 38]). Regardless of the reasons why MBLs improve student learning, the graphs in these activities are generally created by the computer and not by the students. One study found that traditional lab instruction is better than using MBLs for teaching students graph construction and indicated that having students construct graphs by hand is worth the effort and should be pursued [39].

Mitnik et al. noted that "[s]everal computational tools have risen to improve the students' understanding of kinematical graphs; however, these approaches fail to develop graph construction skills" [40]. Recently, a number of online systems have been developed that allow students to do just that. In alphabetical order, they are GraphPAD [41], MapleTA$^{\text{TM}}$[42], MasteringPhysics$^{\text{TM}}$[43], and Socratic Graphs [44]. A description of each follows, but Tables 2.1 & 2.2 give a quick comparison between these and the Function Plot Response. It should be noted that this is intended to be a list of online graph construction programs for physics, and is not intended to be an exhaustive list of online programs that allow students to construct graphs. For instance, GraphPad in WebAssign$^{\text{TM}}$[45] (not to be confused with GraphPAD, in the first list) and MyMathLab$^{\text{TM}}$[46] have graph construction capabilities, but they are designed primarily for math education. Students can create lines (not line segments), circles, and parabolas, which are important in algebra and geometry, but do not allow for a very diverse set of problems in physics.

### 2.1.1 Physics Applets for Drawing

At Western Kentucky University, Bonham has created (using Java$^{TM}$) Physics Applets for Drawing (PADs) [47]. While PADs have a number of functions, only the graphing applet, GraphPAD, which allows students to construct a graph, will be discussed here. A blank grid (unless a background image is used) is given to the student who then can click on grid intersections (or partway between intersections) to establish line segments from one point to another. This creates a piece-wise linear graph, similar in appearance to the graphs in Fig. 2.1. Depending on the authoring of the problem, GraphPad can also use the student's control points to create an $n$th order polynomial (instead of a piece-wise linear graph), or even an exponential graph. However, graphs cannot go off to infinity. If students click in the wrong place, they are able to move or delete the points they created until they are satisfied with the result. Once a student has the graph they want, they can check to see if their answer is correct. This evaluation is done client-side, using the rules listed in one of the parameters of the Java$^{TM}$ applet. It should be noted that evaluating answers on the client's computer makes it much easier to "hack" these problems. However, since GraphPADs are not graded, the concern is minimal. These rules are capable of checking the value of the control points, or the coefficients of the underlying equations of any given segment of the graph. The system also gives immediate feedback, which can be individualized for each rule violated. If multiple rules are violated, the relevant feedback concatenates into one response for the student. Students can make as many attempts at a problem as they like. Problems written with GraphPAD can be randomized and are free to use, but are not part of any course management system, and do not work on most mobile devices. Authoring of problems can be done by anyone who has been given an account to do so. Signing up for an account takes less than a minute, but access is not immediately granted.

### 2.1.2 MapleTA

MapleTA™[42] is a subscription-based course management system and its graph construction tool was written in Java™. Students are given a blank set of axes on which they are allowed to click, creating control points. After creating two control points, the program fits cubic hermite splines to them (a line currently) and allows the student to create more control points (making a parabola, etc.), or move the control points they have already created to get the shape of the graph they desire. Once a student has obtained the graph they want, they can submit their answer to the server, where it is evaluated based on a set of rules written by the problem's author. The rules available to the problem's author allow the server to check a given $x$-value for its $y$-value, slope, or concavity; or to compare the $y$-value or slope of any two $x$-values. After the server evaluates the graph, it returns the result of its evaluation to the student, specifically whether or not their graph is correct. It is possible that this feedback could be more specific than just 'correct' or 'incorrect', but it is technically advanced for the author to implement such individualized feedback. Depending on the instructor's choices when they assign the problem, students may be given more than one try to get the answer correct. These types of problems can be created by anybody with instructor access in MapleTA and can be randomized so that each student receives a slightly different version of the problem. Some drawbacks to this system are that it does not work on mobile devices such as iPads or phones, and that students cannot create graphs that go off to infinity, such as $y = 1/x$ (this last feature is not very useful for kinematic graphs, but would be important to graph electric fields of point charges, for instance).

### 2.1.3 MasteringPhysics

MasteringPhysics$^{TM}$ is also a subscription-based homework system, which is usually coupled to textbooks by Pearson$^{TM}$ publishers. In it, one of the problem types allows students to construct graphs using a Flash$^{TM}$ applet. Students are given a set of axes (possibly with a background image) and asked to create a graph or graphs on it. Students can click on the graph to create control points, which automatically snap to the underlying (integer) grid. The applet uses the control points to create a piece-wise linear function. Only one control point may exist for any $x$-interval, e.g., if the graph goes from 0 to 6, there can only be 7 control points, but students can move them freely until they have the graph they want. Unlike many other systems, if a student does not know where to start a problem, hints can be "purchased" at the cost of the point value earned if they get the answer right. Once a student has the graph they want, he or she can submit it to the server. If the student's answer is incorrect, but matches an incorrect answer programmed into the problem, feedback specific to that mistake will be returned to to the student. Otherwise, the system responds with "Try Again". The program evaluates a student's answer by checking values, linearity, concavity, minima, and maxima over specified ranges. These problems can be written by anyone with instructor access to MasteringPhysics, but can only be edited using Internet Explorer on a Windows machine. The problems are not randomized, do not work on mobile devices, and cannot handle infinities.

### 2.1.4 SocraticGraphs

SocraticGraphs is the graphing element of the BeSocratic [**?**] system. It is still currently in development, and like the rest of BeSocratic, runs in Silverlight$^{TM}$. As such, this section will only be able to describe its current state. Students use their cursor to freehand draw a graph on the axes given. The program then interprets this drawing and turns it into a linear function, a piece-wise linear function, or a set of Cubic Bezier Splines. The interpretation is chosen by the problem's author. Once the graph has been drawn, students can either erase it, add another segment, or go into 'adjust' mode which allows them to move the control points of the splines. When the student is satisfied with their attempt, they can submit their answer. The attempt is graded client-side, but internal to the applet. The graph is evaluated by a set of rules which can test the following elements of the graph: minima, maxima, area under the curve, points, point pairs, slope, curve, segment, curve shape, and number of curves. While infinities could be drawn in this system, there is currently no way to evaluate them in the rules. After evaluating the graph, the program returns the 'correct' or 'incorrect' feedback written by the problem's author. In the 'incorrect' case, the feedback is specific to the first rule that was broken. Since BeSocratic is a free tutorial system, students have an unlimited number of tries to solve the graph problem and receive no benefit or penalty for their answers (except learning!). Currently, anyone who registers for instructor access has the ability to create problems in the BeSocratic system. The SocraticGraphs problems are not randomizable and do not work on mobile devices, but an iTunes application is also in development.

## 2.2 Components

Function Plot Response was created by integrating the Java applet GeoGebra into LON-CAPA.

### 2.2.1 GeoGebra

GeoGebra [48] is an open-source toolset initially developed for teaching mathematics in schools. Its functionality includes geometry, algebra, spreadsheets, and (most importantly for this project) graphing. GeoGebra has a clean, easy-to-use, intuitive interface that educators can customize to fit a particular exercise or activity. GeoGebra's effectiveness in teaching and learning has been studied in various settings,[49] and the project has received a number of educational software awards.[50] The GeoGebra collaboration is currently working on GeoGebraWeb (formerly GeoGebra Mobile [51]), and when this project finishes, the Function Plot Response will also be available on mobile devices such as phones and tablets.

### 2.2.2 LON-CAPA

LON-CAPA [52] (LearningOnline Network with Computer Assisted Personalized Approach) is a course management and homework system that is open-source (GNU General Public License) freeware, with no licensing costs associated. Both aspects were important for the success of this project: The open-source nature of the system allows researchers to modify and adapt the system in order to address research needs, and the freeware character allows easier dissemination of results, in particular, adaptation and implementation at other universities. The system started in 1992 as a tool to deliver personalized homework to students. "Personalized" meaning that each student sees a different version of the same computer-generated problem: different numbers, choices, graphs, images, simulation parameters, etc.[53, 54]

Over the years, LON-CAPA has been expanded with content management and standard course management features, such as communications, gradebook, etc., similar to those in commercial course management systems. In addition to standard features, the LON-CAPA delivery and course management layer is designed around STEM education, for example: It supports mathematical typesetting throughout (LaTeX inside of XML) (formulas are rendered on-the-fly, and can be algorithmically modified through the use of variables inside formulas); it evaluates multi-dimensional symbolic math answers (using sampling or the integrated Maxima and R symbolic math systems); and it fully supports physical units.

LON-CAPA has developed into a content sharing network of over 65 institutions of higher education including community colleges and four-year institutions, as well as about the same number of middle and high schools[55], and serves approximately 150,000 students every year. The shared content pool currently contains approximately 440,000 learning resources[56], including almost 200,000 randomizing homework problems. A number of studies have been carried out regarding the educational effectiveness of LON-CAPA.[54] It was found that the system is particularly helpful for female learners, as they take more advantage of the rich peer-to-peer interaction afforded by the problem randomization.[57]

## 2.3 Previous Graph Problems in LON-CAPA

Homework problems that involve graphs were already implemented in LON-CAPA prior to the introduction of Function Plot Response. Specifically, LON-CAPA had integrated GNU-plot support, which allowed the rendering of randomized graphs on-the-fly, and supported additional layered labeling of graphs and images. These problems, however, do not give students the chance to create the graph for themselves. Instead, they generally fall into one of two categories: multiple-choice or identify-a-feature.

In the first category, students are generally given a description and a set of possible graphs from which to choose the right answer, essentially a multiple-choice problem with graphical answer choices. This type of representation-translation problem would generally be classified Intermediate.[15] Since students generally attempt to solve problems with the least amount of effort, understanding the graph may not always be a high priority. Instead, students might try to find an individual feature that excludes one or more of the graphs as a possible answer. As an example, if a student can eliminate two of the four graphs as possible answers, and they have 2 chances to get it right, the problem is solved in their mind. This is a common thing to do in multiple-choice problems and is, in fact, seen as "effective test taking strategy." While it does require some understanding of the graph to do this, the student does not need to have a deep understanding of the graph or where it came from to get the answer right.

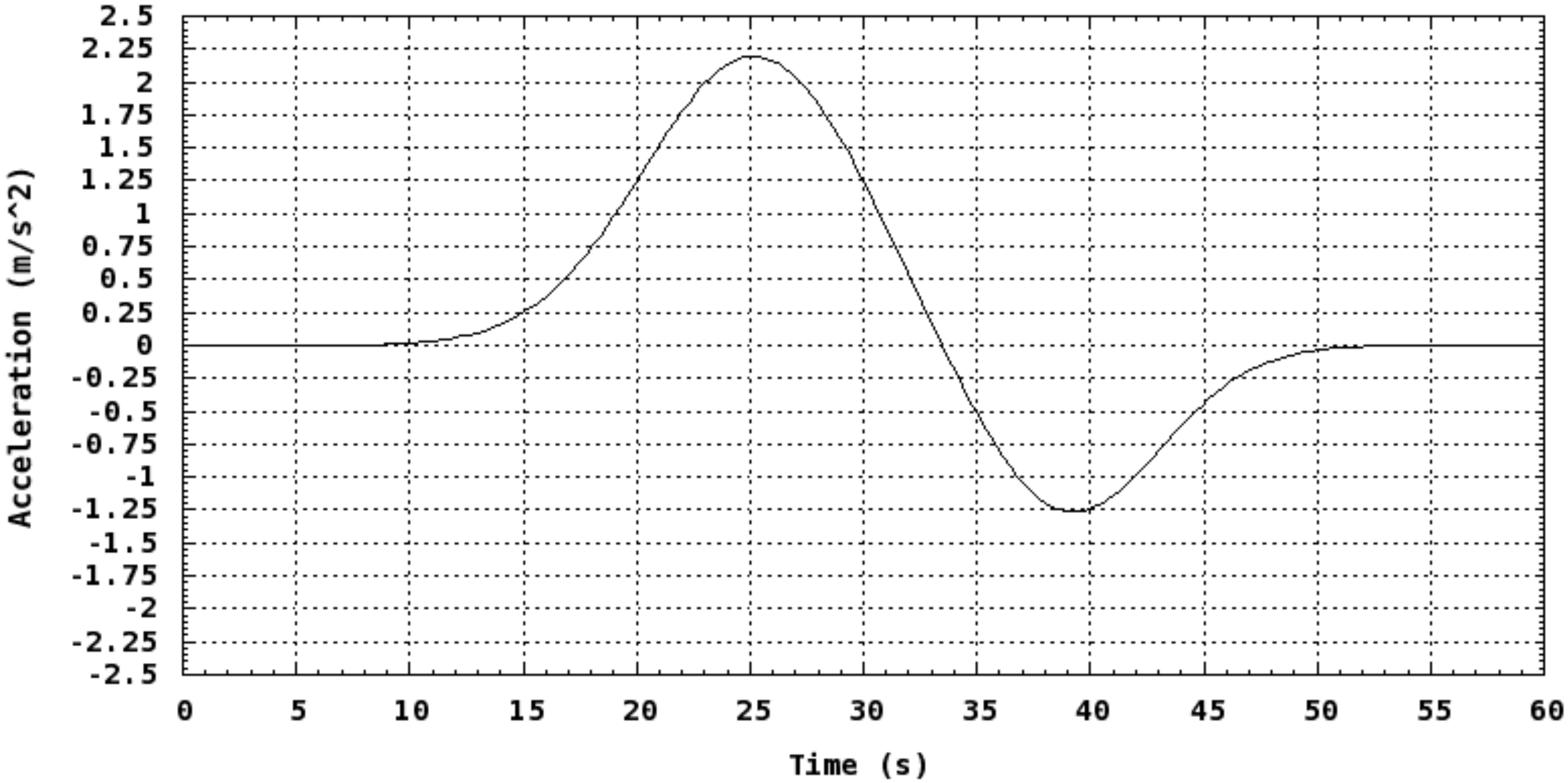



Figure 2.2: A problem where the student must extract information about a feature of the graph. The first part of the problem would be classified as Intermediate, while the remaining parts would be classified as Comprehensive. This Figure is continued in the next one.

At $t$=0, a car drives with a velocity of 26 m/s. Its acceleration on a straight road is shown over the next minute. Which one of the following statements is true?

- The car first slows down, then speeds up, and then slows down again.
- The car first speeds up and then slows down.
- The car first starts driving backwards and then forwards.
- The car first drives forwards and then backwards.
- None of the above

Submit Answer Tries 0/2

At what time does it have maximum velocity?

Submit Answer Tries 0/99

What is the maximum velocity?

Submit Answer Tries 0/99

What is the final velocity?

Submit Answer Tries 0/99

Figure 2.3: See the previous Figure for an explanation of this one. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

In the second category of graphing problems, students are given a graph and asked questions about its features. Such problems can span the whole range from Elementary to Comprehensive problems,[15] but again, students are likely to gain an understanding about some aspect of the graph, but generally not about the graph as a whole. An example of this type of graph problem is shown in Fig. 2.2; different randomizations of this problem may have the car first slowing down and then speeding up, and the car may in the end be slower or faster than before, depending on the integral of the acceleration.

## 2.4 Function Plot Response

To expand the range of graph problem types in LON-CAPA, we developed Function Plot Response. It allows students to construct their own graph and submit their answer to the server, which immediately grades the submission and returns relevant feedback to the student. While originally designed for back-of-the-envelope graph problems in physics, this problem type may be equally usable for any other subject that uses graphs.

The server grades the problem based on a set of rules defined by the problem's author, and thus it requires no hand-grading by the instructor. This makes it especially useful for large lecture classes often found in university settings, where hand-grading is particularly time-consuming.

Graphs are represented by one or more splines, which allows for discontinuities. However, a guiding principle was that realistic graphs are generally smooth and continuous, not piece-wise linear. If graphs have discontinuities, those need to be justified by other physics assumptions—for example, in the case of the electric potential of a point charge, the fact that we assumed a *point* charge.

An important design feature, shared with some of the systems discussed in Section 2.1, is the evaluation based on rules rather than value pairs. For example, rather than checking whether (within tolerances) the graph has value $y = 5$ at $x = 3$, we might check if the function is positive over an interval, or if its second derivative is smaller than 9.81. This allows for many correct solutions to a certain scenario, not just one particular graph. Through the built-in randomization, requirements and rules can vary from student to student. Interval boundaries can be flexible: An author can define checkpoints (internally called “labels”) that are determined based on the student’s answer, and which can be used in subsequent rules. For example, a scenario may state that an object is supposed to first accelerate and then move with constant velocity for a while. In this case, a checkpoint would be when the second derivative of the position is not positive anymore, and the next interval where the second derivative should be approximately zero starts at that rule-defined checkpoint.

### 2.4.1 Student Interface

The student interface of Function Plot Response was created in GeoGebra. While GeoGebra has a multitude of tools and uses, we have restricted most of these so that the student only has control over a set of cubic Hermite splines, which appear when the problem is first loaded. The student adjusts these splines in order to synthesize the desired graph.

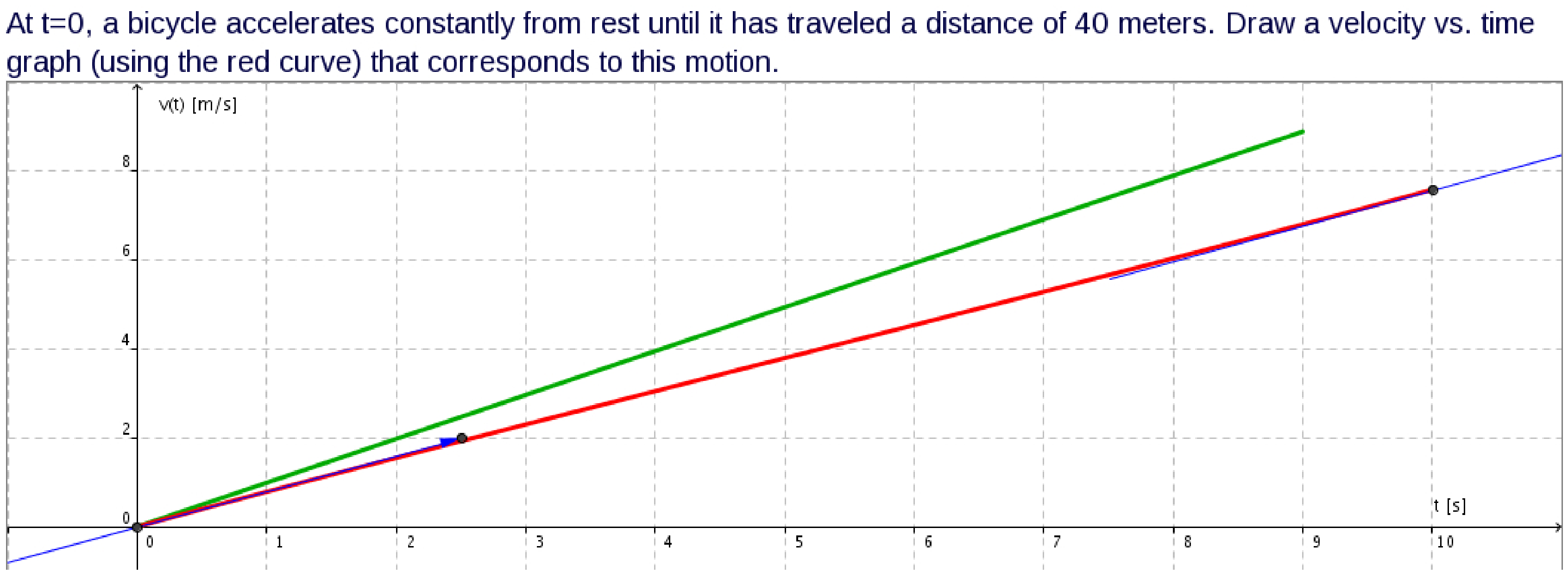



Figure 2.4: A simple problem using Function Plot Response. The lower line is the student input including the control points, the upper line is the sample answer given by the author. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

Figure 2.4 shows such a GeoGebra-based problem, which would be classified as an Intermediate problem[15]. Different students would get different total distances and different scales on the graphs, which is a rather simple randomization. The velocity needs to increase linearly with time, and the integral of the velocity over time must equal the distance covered. There are infinitely many correct answers, and the figure shows a correct student answer (lower line) as well as the author's sample answer (upper line) programmed into the problem.

Figure 2.5 shows another problem, which would also be classified as Intermediate.[15] Any answer that begins and ends on the $x$ axis, and is greater than or equal to zero everywhere in between, is considered correct. The top panel shows the freshly loaded problem, before the student has tried to answer it. This particular problem has one spline on it and the spline is defined by the six points on the graph. Adjusting any of the three points currently on the $x$ axis will move the position of the spline, while moving any of the three points currently not on the $x$ axis will adjust the slope at the relevant point. It is possible to have more than one spline (which one would need to graph the infinity in $1/x$), but the problem shown does not use this feature.

After students have manipulated the spline to where they want it, they submit the answer to the server (just like any other LON-CAPA problem). If the answer is wrong (as in the middle panel of Fig. 2.5), the server will give back the incorrect graph, as well as an author-provided hint. For instance, if a student gets all but one of the rules correct, a specialized error message can be returned to them to indicate where they went wrong. Only one hint can be given at a time. Thus, if a student gets more than one rule wrong, only the hint from the first violated rule will be returned to the student. This feature is intended to help the student focus on a particular aspect of the graph they are misunderstanding before

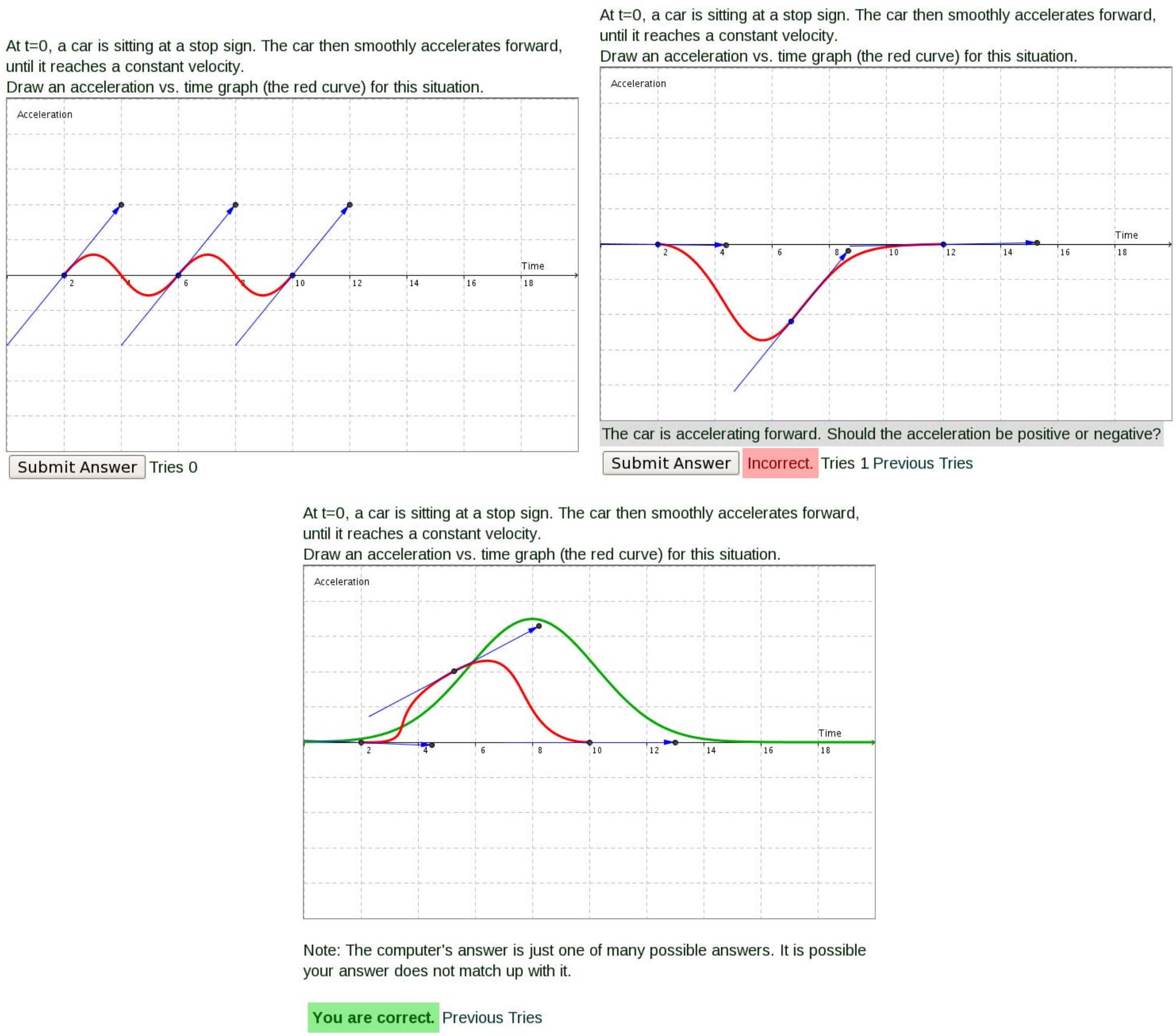



Figure 2.5: First panel: How the problem appears the first time a student opens it. Second panel: A wrong answer was submitted, and the server returned a customized hint to the student. Third panel: A correct answer was submitted, and the author's answer to the problem is also shown, which reaches its maximum at a later point in time. The text in this Figure is not meant to be readable but is for visual reference only.

continuing.

If the student's submitted answer is correct, he or she gets LON-CAPA's standard "green box" stating "You are correct." In addition, a new green curve (in this case, the Gaussian curve) is added to the graph area which shows the author's answer to the problem. In some cases it will be the only correct answer, but in other cases (such as this one), it can be just one of an infinite number of correct answers. As the bottom panel of Fig. 2.5 shows, the author's answer and the student's answer appear quite different. However, both answers are correct and Function Plot Response accommodates this.

### 2.4.2 Author Interface

Since homework problems for LON-CAPA are created by its users, it is important to allow any author in LON-CAPA to create Function Plot Response problems. The commonly nicknamed "colorful editor" is an author-friendly way of creating problems in LON-CAPA, see Fig. 2.6 for an example. The colors representing the underlying XML structures were initially randomly chosen, and the goal of implementing a more pleasing color scheme has not yet been addressed, thus its nickname. This editor allows authors to easily create the necessary XML code for their problems. Figs. 2.8 & 2.9 show the corresponding XML code. Both correspond to the problem shown in Fig. 2.5.

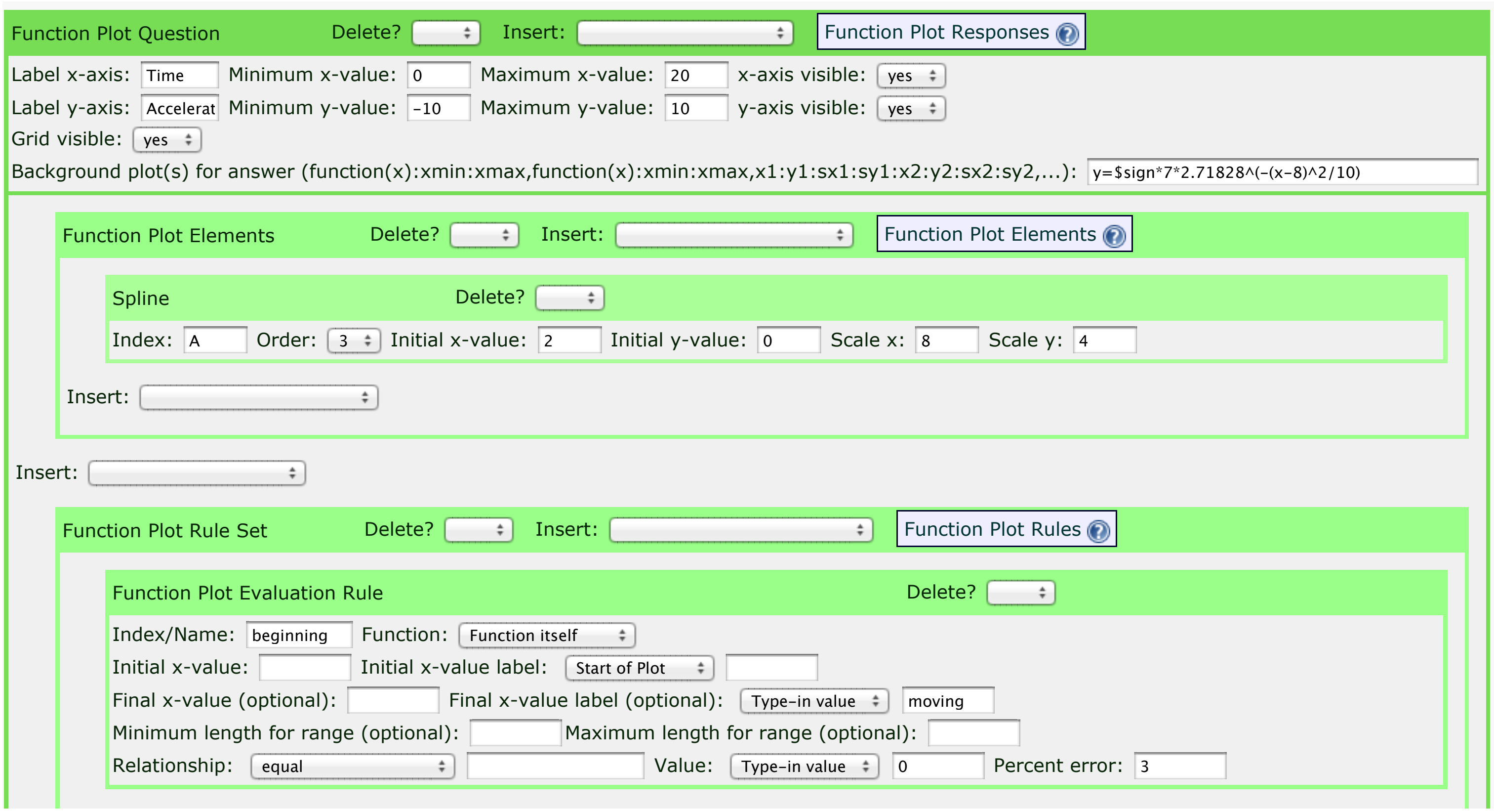



Figure 2.6: "Colorful editor" view for the Function Plot Response part of a graph problem. The first two rows of entered values in the "Function Plot Question" box determine the $x$ and $y$ axes for the problem. The next line determines whether or not the grid is visible and the following line determines the answer plot that will be displayed after the student gets the problem correct (the variable "$sign" is defined earlier and is 1 if the car is moving forward or $-1$ if the car is moving backward). The "Function Plot Elements" box contains the information about the splines and the background plot (not shown). The "Function Plot Rules" box contains the rules the server uses to determine if a submitted response is correct or not (the variable "$relation" in the second rule is defined earlier and is $\geq$ if the car is moving forward and $\leq$ if the car is moving backward . The entries shown correspond to the problem shown in Fig. 2.5. This Figure is continued in the next one. If the text contained in this Figure is unreadable, please see the Electronic version.

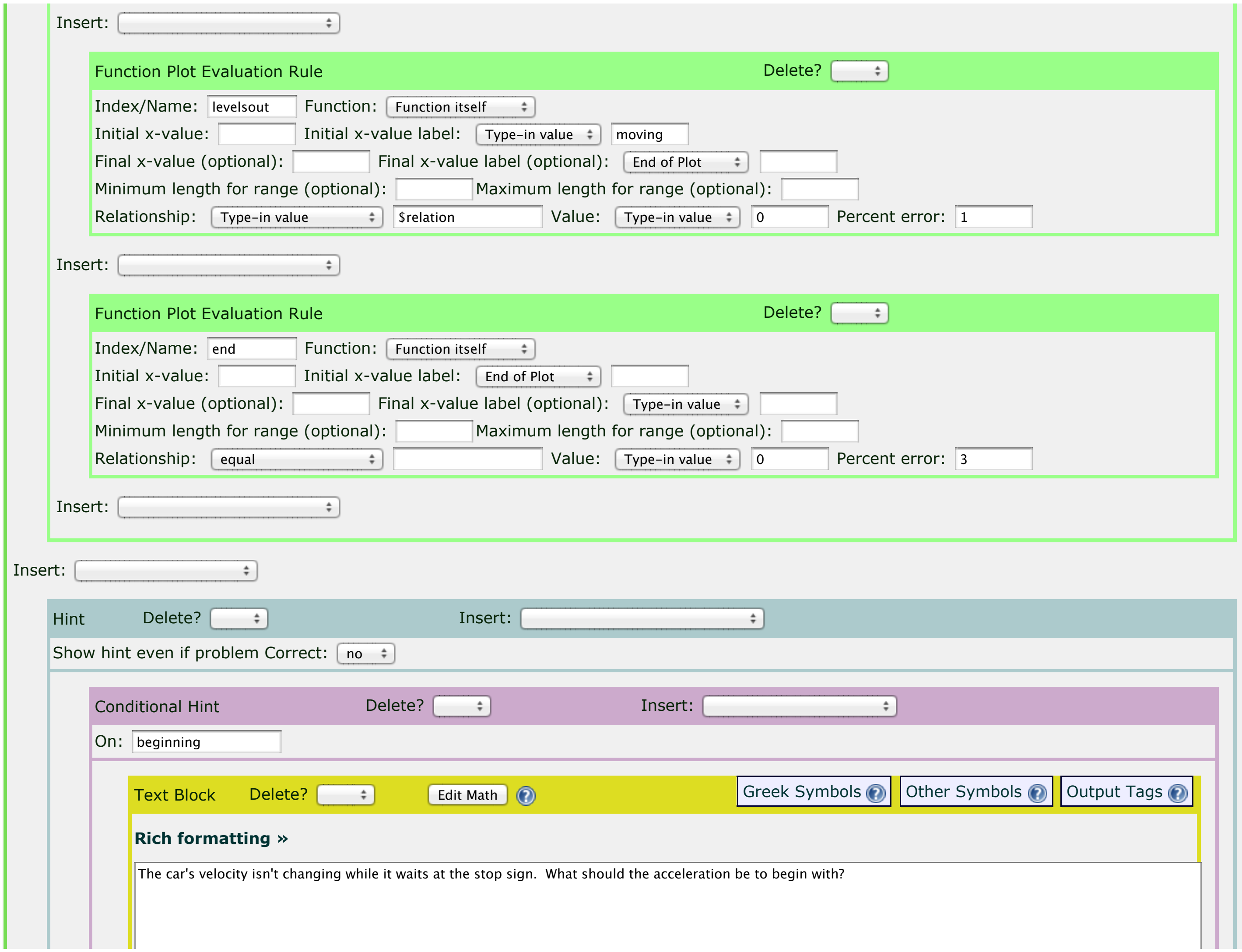



Figure 2.7: This Figure is a continuation of the previous one. The "Hint" box shows the customized hint that the student would see if the submission fails the first rule. The entries shown correspond to the problem shown in Fig. 2.5. If the text contained in this Figure is unreadable, please see the Electronic version.

Any instructor in the LON-CAPA system can be granted the author role, and these authors have a lot of control over the system. They have the ability to adjust the axes, turn the grid on or off, add labels to the axes, add a background plot, and add an arbitrary number of splines for the students to work with. Each spline can be controlled by $2n$ points, where $n$ is the 'order' of the spline, an integer between 2 and 8 (the upper limit of 8 is somewhat arbitrary, but the manipulation of the graph becomes increasingly cumbersome with more control points).

Authors also control the set of rules the server uses to grade the students' responses. As discussed, such rules include checking the value of the function, first derivative, second derivative, and/or integral on any given interval, or comparing any of these values (except the integral) between any two points. As an example, Table 2.3 shows the rules for the simple problem in Fig. 2.5, which are reflected in Fig. 2.6 and Figs. 2.8 & 2.9. The first rule, called "beginning" by the author, checks that the function is approximately equal to 0 to start out, and follows the graph until the value of the function is no longer close enough to 0. The rule then labels this point "moving". The next rule, called "levelsout" by the author, starts at this same label and extends to the end of the plot. It is controlled by the randomization of the problem; for some students the car accelerates forward while for other students it accelerates backward (implemented using the normal LON-CAPA randomization); the variable "relation" is set to "greater than or equal" or "less than or equal" accordingly: the function itself should be less than zero for a backward-accelerating car, and greater than zero for a forward accelerating car. This rule will fail if the label "moving" is not defined, so a flat line is not a correct answer. The final rule, called "end" by the author, insures that in the end the car is not accelerating anymore. For any violated rule, specific feedback can be given. For example, the hint "The car's velocity isn't changing while it waits at the stop sign." is given if the rule called “beginning” is not fulfilled.

```
<problem>
<script type="loncapa/perl">
$number=&random(1,2,1);
$direction=&choose($number,'forwards','backwards');
$relation=&choose($number,'ge','le');
$sign=&choose($number,1,-1);
</script>

<startouttext />
At t=0, a car is sitting at a stop sign.  The car then smoothly accelerates
$direction, until it reaches a constant velocity. <br />
Draw an acceleration vs. time graph (the red curve) for this situation.
<br />
<endouttext />

<functionplotresponse answerdisplay="y=$sign*7*2.71828^(-(x-8)^2/10)"
xmin="0" xaxisvisible="yes" ymax="10" xlabel="Time" ymin="-10"
gridvisible="yes" ylabel="Acceleration" id="11" xmax="20"
yaxisvisible="yes">

<functionplotelements>
<spline scalex="8" initx="2" inity="0" index="A" order="3" scaley="4" />
</functionplotelements>

<functionplotruleset>
<functionplotrule derivativeorder="0" value="0" relationship="eq"
percenterror="3" index="beginning" xinitiallabel="start"
xfinallabel="moving" />
<functionplotrule derivativeorder="0" value="0" relationship="$relation"
percenterror="1" index="levelsout" xinitiallabel="moving"
xfinallabel="end" />
<functionplotrule derivativeorder="0" relationship="eq" value="0"
percenterror="3" index="end" xinitiallabel="end" />
</functionplotruleset>
```

Figure 2.8: First part of the XML source code for the Function Plot Response problem in Fig. 2.5. While it may look like HTML, this code is never sent to the browser. Instead, this is the code which LON-CAPA evaluates server-side when rendering or grading the problem. While it can be edited directly by the author, most authors prefer to use the "colorful editor" shown in Fig. 2.6 when constructing more complex problem types like Function Plot Response. For part two of the XML source code, see Fig. 2.9.

```
<hintgroup showoncorrect="no">
<hintpart on="beginning">
    <startouttext />The car's velocity isn't changing while it waits at
    the stop sign.What should the acceleration be to begin with?
   <endouttext />
</hintpart>
<hintpart on="end">
    <startouttext />Once it has reached a constant velocity, what should
    the acceleration be?<endouttext />
</hintpart>
<hintpart on="levelsout">
    <startouttext />The car is accelerating $direction.  Should the
    acceleration be positive or negative?<endouttext />
</hintpart>
</hintgroup>

</functionplotresponse>

<solved>
<startouttext /><br />
<br />
Note: The computer's answer is just one of many possible answers.
It is possible your answer does not match up with it.
<br /><br /><endouttext />
</solved>

</problem>
```

Figure 2.9: Part two of the XML source code for the Function Plot Response problem in Fig. 2.5. For part one of the XML source code, see Fig. 2.8.

Table 2.3: Ruleset for the Function Plot Response problem in Fig. 2.5.

| Rule Name | Derivative | Initial $x$-value | Final $x$-value | Min. Length | Max. Length | Relationship | Value | %-Error |
|---|---|---|---|---|---|---|---|---|
| *beginning* | Function itself | Start of Plot | Label: *moving* | | | equal | 0 | 3 |
| *levelsout* | Function itself | Label: *moving* | End of Plot | | | Variable: *relation* (set to 'greater than' or equal' or 'less than or equal') | 0 | 1 |
| *end* | Function itself | End of Plot | | | | equal | 0 | 1 |

Beyond the type of rules, as discussed, the author can also include individualized feedback for specific rules being missed. In the example in Fig. 2.6, the hint “The car’s velocity isn’t changing while it waits at the stop light.” is given if the rule called “beginning” is not fulfilled, which checks if the initial value of the acceleration is zero.

The system also incorporates the type of problem individualization the rest of LON-CAPA is built around. For instance, the values used in the rules can be dynamically generated and thus be different from one student to the next. In the example shown in Figs. 2.5 and 2.6, half of the students have the car moving forward, and half have the car moving backward. Since the Function Plot Response is just another problem type in LON-CAPA, it can be matched with any other resources allowed by the system:

- Students can be given a dynamically generated graph and asked to draw a related graph;
- A video of an object moving may be shown and the students asked to graph the motion;
- Due to the coupling of LON-CAPA to the MAXIMA computer algebra system and the R statistics package, symbolic terms can be evaluated alongside the graphs;
- Using the built-in grading queues, written student responses can be graded alongside the graphs.

In short, most computer-based resources can be used in conjunction with the Function Plot Response in a problem, which is the advantage of implementing the problem type using the same XML-structure as the remainder of the system.

# Chapter 3

# Usage of Function Plot Response in Classes

## 3.1 Results from In-Class Usage

As part of the ongoing process of refinement, a test group was given some Function Plot Response problems and asked to give feedback during Spring 2011. After a few updates, three of these problems were pilot-tested in a Michigan State University physics course during Summer 2011, and eventually several problems were used in a different class during Fall 2011. Between the oral and written feedback, discussion boards, word of mouth, and server logs, a few common themes have emerged.

### 3.1.1 User Experience

There is an abrupt learning curve when students first encounter this type of problem. Similar to a student's first interaction with any online homework system, it is useful to have the first problem or two help students get accustomed to it. Specifically, we have found it useful to include a practice problem describing how to create a graph using Function Plot Response, which does not yet involve any physics and is ungraded.

On the technical side, as is often the case with Java, we encountered some compatibility problems with older student computers. Also, as opposed to the rest of LON-CAPA, Function Plot Response does not run on most mobile devices. Overall, though, GeoGebra is remarkably compatible, and we hope to eliminate any remaining problems once the switch to GeoGebraWeb is accomplished.

Writing open-ended graph problems is a difficult task. Authors must conceptualize the rules that define a correct graph and then be sure that no correct graph can be made that doesn't fit these rules, and vice versa. Beyond this, the wording of the question is very important. For instance, one problem stated "A car is traveling in a straight line at constant velocity. It covers 56.7 m between 2.7 sec and 9 sec." Many students drew a line that started and ended at the given points in time, and the server often returned "incorrect" because the function ended up being undefined at the specified times themselves. The new version of the problem states "A car travels in a straight line at constant velocity, beginning at $t$=0 sec. It covers 56.7 m between 2.7 sec and 9 sec," which helps clarify that the car travels at a constant velocity over the whole width of the graph.

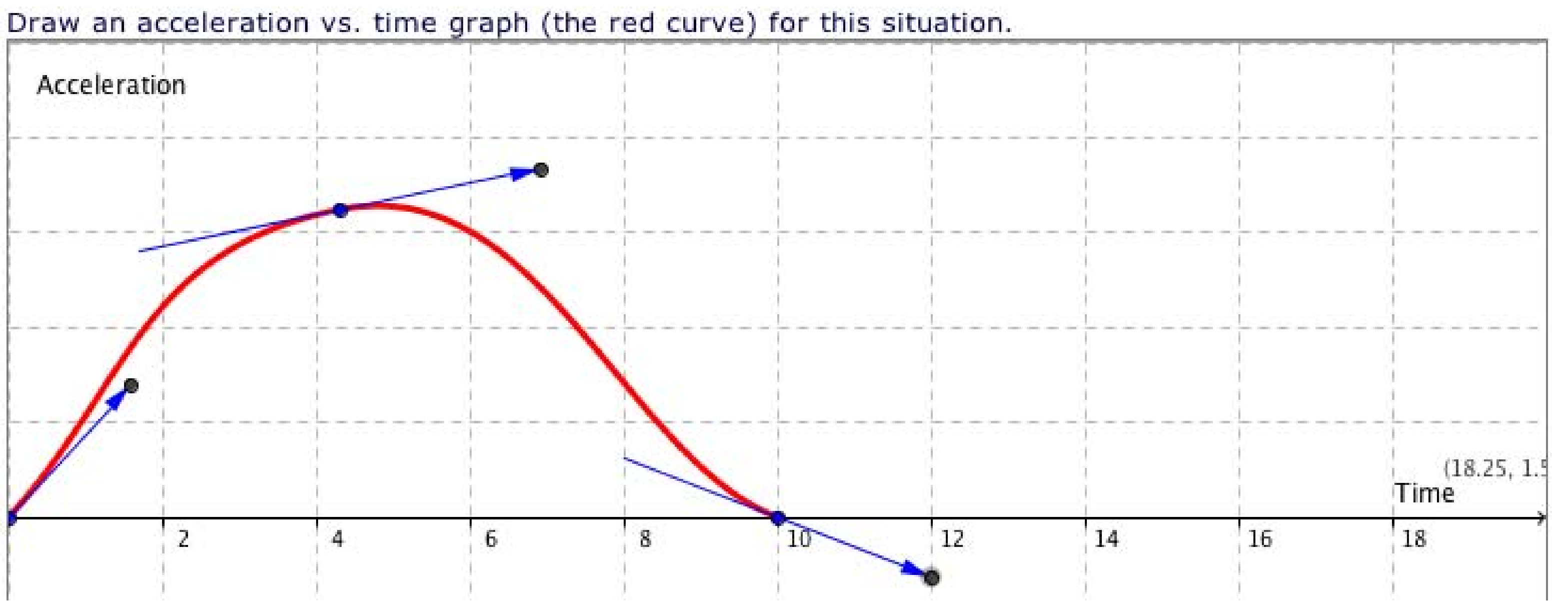



Figure 3.1: Sample student solution for the problem in Fig. 2.5. If the apparent non-smoothness at $t = 0$ is taken literally, the car starts from rest with a discontinuous jerk. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

As with all homework questions, some judgement calls must be made by the problem's author. For example, the problem shown in Fig. 2.5 has gone through several iterations. Originally, the word "smoothly" was meant to denote that the derivative of the $a$ vs. $t$ graph had to be zero at the end points. Many test subjects complained that this was not effectively communicated to them, and so the rules that checked the derivative at the end points were removed. This means that it is now possible to make graphs that are not actually smooth but are still considered "correct" for this problem (see Fig. 3.1). We found it interesting that even advanced students and instructors often wanted to build graphs that were not smooth in response to this problem; see Fig. 3.2 for three examples.

Although the sample size is somewhat small, the discussion boards for these problems show similar patterns to other problem types in LON-CAPA. While some students make emotional statements ("someone please help with a more specific answer?? i understand what needs to be graphed but for some reason what i'm doing will NOT work"), others ask about or discuss the graphs ("my distance is 40 and i have tried every possible combination of heights and base lengths that could give me an area of 40 and nothing works."), while still others discuss the physics behind the graphs ("If a car is not moving it has a zero acceleration which would be along y=0"). However, the graphical problems resulted in more of these discussions: In our sample, while purely traditional problems on average had 1.7 contributions in their associated discussion boards, problems involving graph interpretation had an average of 3.5 contributions, and problems involving graph construction an average of 6.5 contributions. These types of discussions are well documented [58] and are a good sign that this problem type is useful for students.

In individual feedback forms, a notable number of students made it clear that they did not feel as comfortable with this new problem type, saying, for example, "I prefer numerical

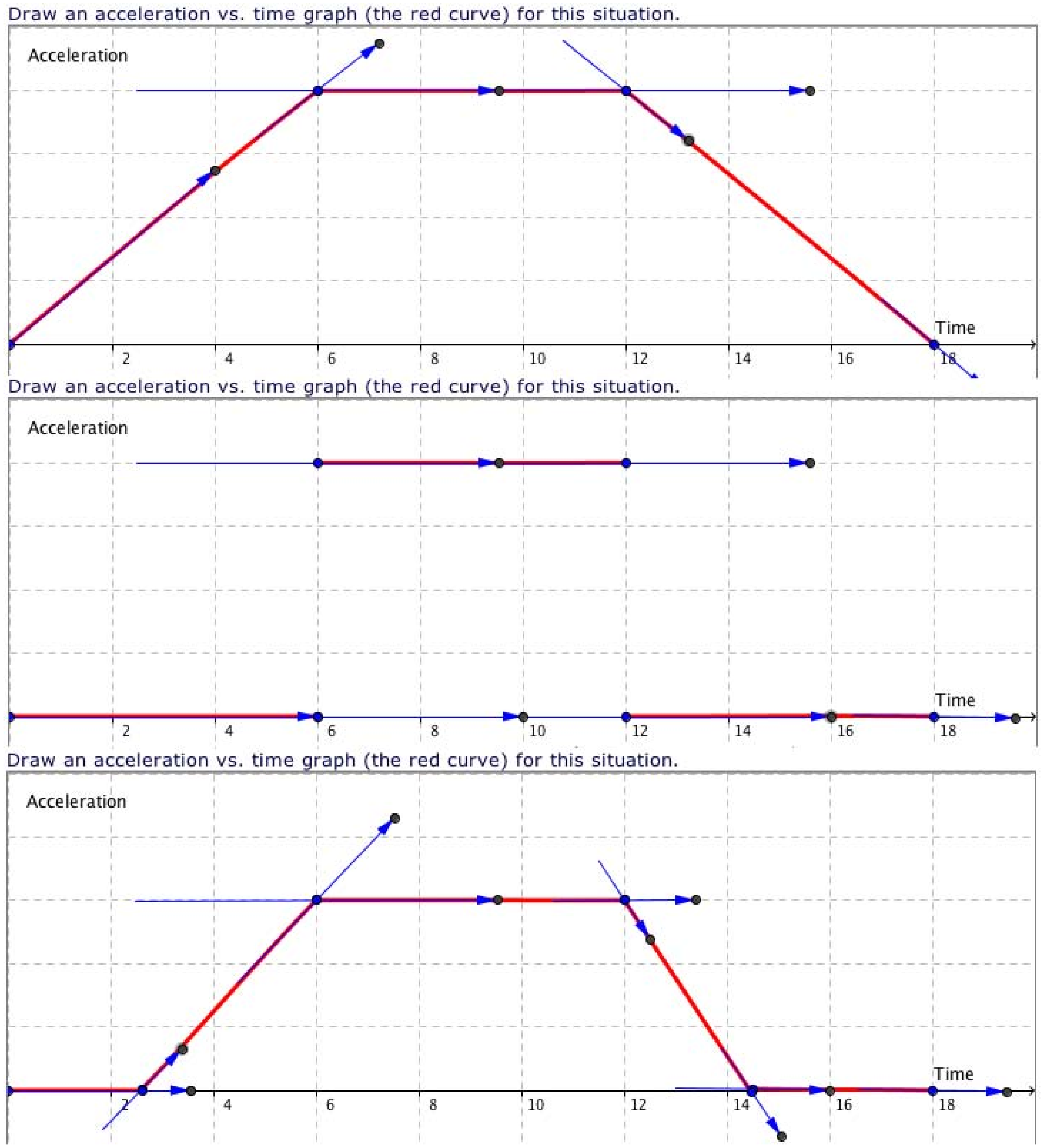



Figure 3.2: Discontinuous solutions that both introductory students and instructors expected to be acceptable for the problem shown in Fig. 2.5. If the apparent non-smoothness is taken literally, the first and the third graph would require a discontinuous jerk, while the second would require an infinite jerk. The text in this Figure is not meant to be readable but is for visual reference only.

answers as opposed to having to graph like this." This uncomfortable feeling seems to be related to the non-definite answers to many of the problems. More precisely, many of the questions have an infinite number of possible answer graphs that satisfy the question, which is very different from the solve-for-this-number, end-of-chapter type of problem to which students are accustomed. Questions with multiple answers have been studied in mathematics education with regard to creativity in problem solving, a valued skill in physics. For a quick overview, read Silver's paper [59] about creativity in math education. One student, in particular, made it very clear that they did not like the indefiniteness of the answer, claiming "the questions lacked adequate detail and depth of information to complete accurate graphical representations. . . . There are better ways to teach graphs than with this program. For instance providing multiple choice options for such problems would be more beneficial."

One possible drawback of these problems is that they do not lend themselves to error analysis, an important metacognitive skill for physicists. For most problems, when students get the feedback from the system that their answer is incorrect, they often go back and search their math for where they may have gone wrong, often discovering that their equation was incorrect, or that they mistyped something in their calculator. For Function Plot Response problems, since students often don't have anything on paper to go back to and look at, they seem more likely just to try making a small adjustment on the graph they already tried. If this adjustment works, the students often claim that the software is "picky" or "touchy," instead of trying to understand the difference between their submissions, and why one of them is right, but the other is not. In other words, students often fail to distinguish between changes to the graphs that bring them to within tolerance of the rules ("pickiness" of the grading algorithm) and changes that actually represent different physics (where the adjustment changes the described scenario). This trial-and-error approach is similar to the finding that online homework systems may “turn thinkers into guessers.”[60] One possible way to deal with this concern is to be careful to set tolerances in such a way that “pickiness” can hardly be an excuse. Another way may be to expand Function Plot Response to also plot the most recent wrong answer as a background plot, so students can better compare their correct answer to their most recent wrong answer, and hopefully pick out the salient “make or break” differences.

### 3.1.2 Problem Characteristics

Server logs were evaluated for the fall 2011 class, which had 80 students. This is a calculus-based introductory physics course with mostly pre-medical students. Eight graph construction (using Function Plot Response) and three graph interpretation problems were embedded into the 61 online homework problems for the chapters on linear dynamics, rotational dynamics, and energy.

Based on user feedback, one would expect that Function Plot Response problems are much more time-consuming than other types of problems. We compared traditional problems (multiple-choice and mostly numerical problems), graph interpretation problems (see Section 2.3), and graph construction (Function Plot Response) problems. We expected graphical problems to show more total time-on-task and longer intervals between attempts, but in fact they turned out to have moderate to low average times between subsequent submissions—probably the result of the instructor allowing 99 tries to get them correct. Students spent a little more than average total time-on-task on the graph construction problems, but the difference is not significant.

Analyzing subsequent submissions that occur within two minutes provides more insights, as students will most likely have been on-task during such short intervals. As Fig. 3.3

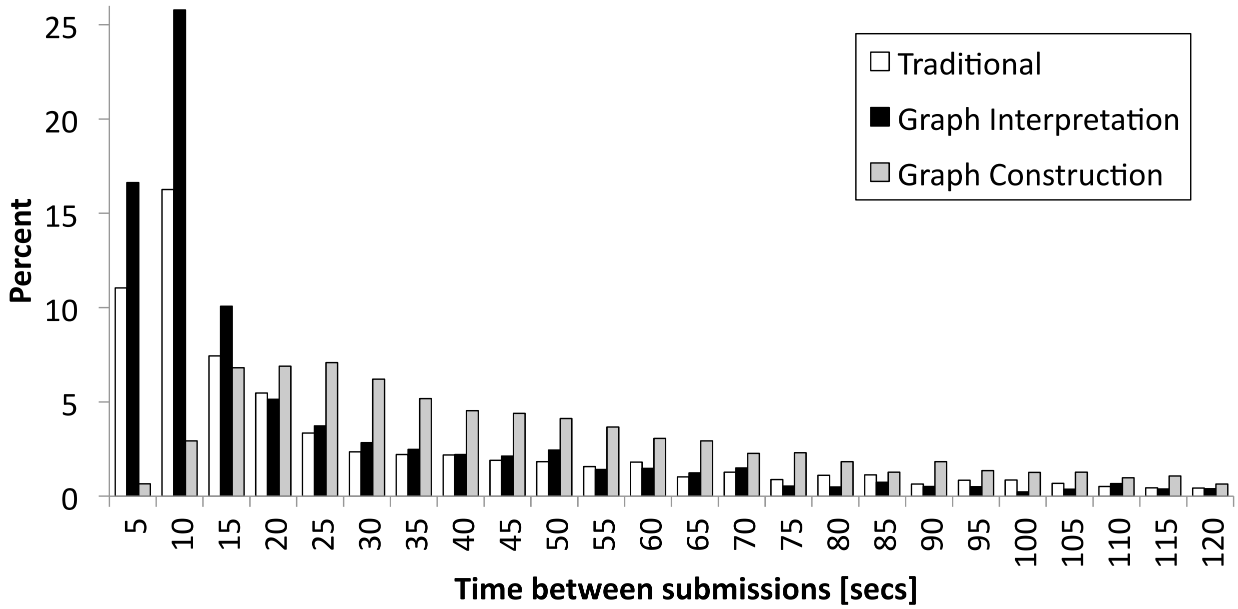



Figure 3.3: Time intervals between subsequent submission to a problem after a failed attempt for traditional, non-graphical problems (white), graph interpretation (black), and graph construction problems (gray). For both traditional and graph interpretation problems, a subsequent answer submission occurs between 5 and 10 seconds later, while for graph construction problems, the most frequent interval is 20 to 25 seconds later.

shows, for both traditional and graph interpretation problems, most subsequent answer submissions occur between 5 and 10 seconds later, while for graph construction problems, most frequently another graph is submitted 20 to 25 seconds later. However, loading a traditional problem on a good internet connection takes about two seconds, while loading a graph construction problem takes about nine seconds due to the applet initialization. Thus, most of the shift of the maximum can be attributed to technical reasons. In either case, the graph indicates a disappointingly high percentage of guessing and trial-and-error. While the distribution for the graph construction problems has a longer tail, this may be for trivial reasons: Manipulating the graph control points is more cumbersome than clicking on multiple-choice fields or entering numbers, and thus it takes longer to enter the next guess.

While we found no (non-trivial) differences in the problem timing characteristics, we did find that graphical and traditional problems have different characteristics with respect to student performance measures. We analyzed two performance characteristics: degree of difficulty and degree of discrimination.

The degree of difficulty in LON-CAPA is defined as

$$\text{DoDiff} = 1 - \frac{\text{Total number of correct solutions}}{\text{Total number of tries}}, \tag{3.1}$$

which is always between 0 and 1. DoDiff = 0 would indicate that all students get the problem right on the first attempt, while DoDiff = 1 would indicate that no student got it correct, and DoDiff = 0.5 would indicate that on the average students got the problem correct on the second attempt.

The degree of discrimination in LON-CAPA is defined as the percentage of students from the top quartile on the problem set getting the problem correct, minus the percentage of students from the bottom quartile on the problem set getting the problem correct:

$$\begin{aligned} \text{DoDisc} &= \frac{\#\text{ correct by top 25\% on set}}{\#\text{ students in top 25\%}} - \frac{\#\text{ correct by bottom 25\% on set}}{\#\text{ students in bottom 25\%}} \\ &\approx 4 \cdot \frac{(\#\text{ correct by top 25\% on set}) - (\#\text{ correct by bottom 25\% on set})}{(\#\text{ students working on set})} \end{aligned} \tag{3.2}$$

This quantity is always between $-1$ and 1. The DoDisc considers an individual problem in the context of the complete problem set ("assignment") of which it is part. If only the top quartile of students get the problem correct, then DoDisc = 1, and the problem would be highly discriminating. If, on the other hand, only the students in the bottom quartile get the problem correct, then DoDisc = $-1$, and something is very likely wrong with the problem. When DoDisc = 0, it means that performance on the problem is not correlated with overall performance on the problem set.

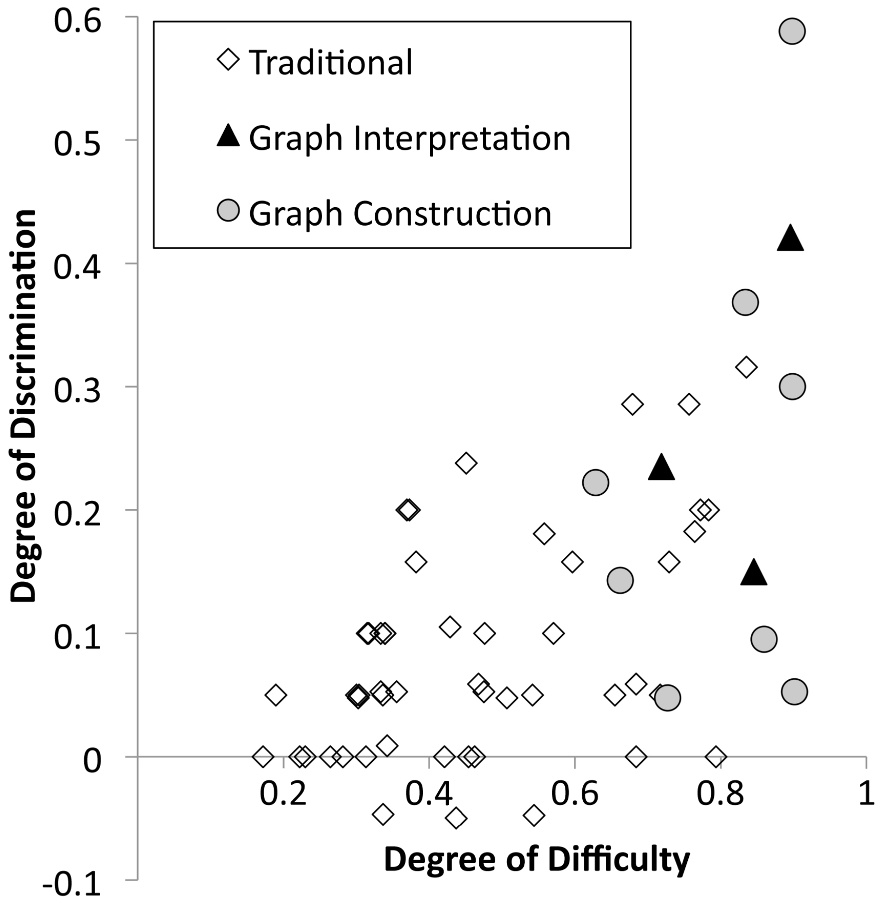



Figure 3.4: Degree of discrimination versus degree of difficulty for traditional, non-graphical problems (white), graph interpretation (black), and graph construction problems (gray).

As Fig. 3.4 shows, graphical problems turned out to be more difficult, but also more discriminating, than traditional problems. Students who generally do better in a given topic also do better on the related graph problems, and vice versa. Thus, performance on graph construction problems may be a better indicator of student learning than most other types of online homework problems.

The problem with the highest degrees of difficulty and discrimination was one in which students were asked to graph the turning angle over time of a wheel that is subject to a constant given torque—a problem that would be classified as Comprehensive.[15] The problem with the lowest degree of discrimination was one that asked students to construct a $v$ vs. $t$ graph for a car moving with constant velocity (i.e., just a straight line), placed at the beginning of the linear kinematics chapter. For comparison, the problem in Fig. 2.5 had a degree of difficulty of 0.66 and a rather low degree of discrimination of 0.14.

## 3.2 Do Graph Construction Problems Improve Learning?

In order to better analyze and characterize difficulties students have with graphs, a method to measure students' understanding of graphs in physics was needed. In response, the Test of Understanding Graphs-Kinematics (TUG-K) was created[61]. Beichner found that the level of instruction had little to no effect on how much students understood about graphs. The author further concluded that "[S]tudents must be given (1) the opportunity to consider their own ideas about kinematics graphs and then (2) encouragement to help them modify those ideas when necessary." Specifically, he suggested making students predict the shapes of graphs.[61]

One obvious way to try to improve student understanding of graphs is with homework. Here we will divide graph related homework problems into two categories: interpretation and construction. For the purposes of this study, graph interpretation problems will be defined as problems where students are given a graph (or graphs) up front and either asked to select the correct graph from a set or asked questions about a particular graph. Graph construction problems are defined as questions that require students to draw a graph for themselves based on the information given. The overlapping case where a student is asked to draw a graph based on a given graph is also classified as construction (though no such problems are involved in this study.)

All of the homework problems in the relevant classes were distributed using the LON-CAPA course management system. Until recently, only graph interpretation problems could be easily assigned using LON-CAPA. However, the recent development of the Function Plot Response problem type now allows for questions where students are required to construct graphs themselves. Fig. 3.5 shows a graph construction question using the Function Plot Response [62].

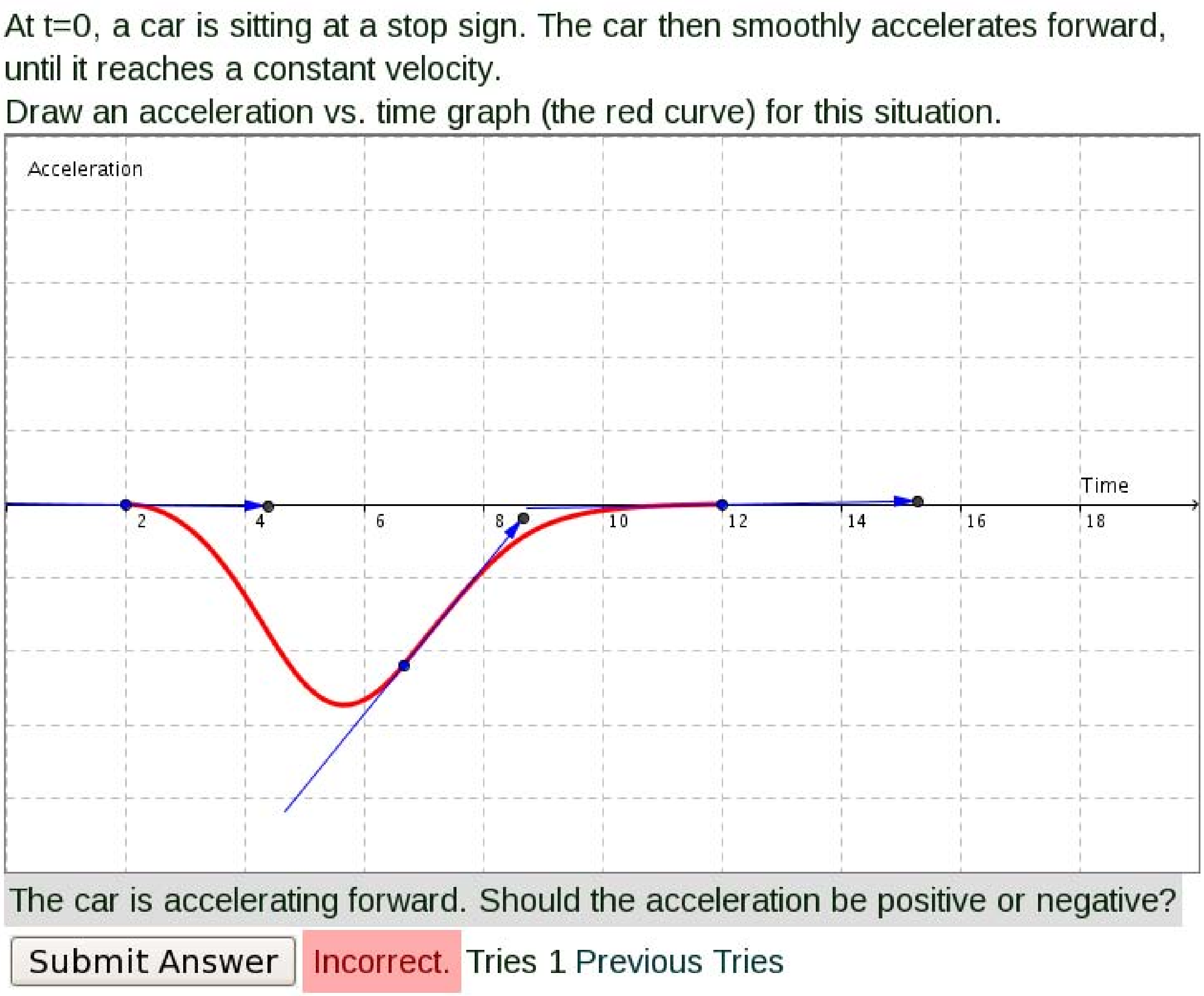



Figure 3.5: An example of a graph construction problem using Function Plot Response. In this case, a wrong answer has been submitted to the server. The student has multiple chances to get the answer right, and has been given a hint as to what is wrong with their graph. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

The primary intervention we are hoping to examine is whether or not including graph construction problems in students' homework improves their understanding of kinematic graphs (as measured by the TUG-K) more than having only graph interpretation problems, or no graph problems.

The data used in this study comes from two different courses (labeled A & B here) and was collected over several semesters at Michigan State University. Both are calculus-based, first semester physics courses, that had two sections. Course A is primarily populated by students majoring in engineering, while the students in Course B are predominantly premedical. In total, the data analyzed in this paper comes from six different classes, where we define a "class" to be the set of students in a course during a semester. Each class was given the TUG-K during the first week of the semester (before discussing kinematics) and again in the last week of the semester. Since not all students took both exams, we list both the number of students in the class ($N$) and the number of students who took the pretest and posttest ($n$) in Table 3.2. The analysis presented here is only on the students who took both tests.

In Classes 1 and 2 (both in Course A), students had no homework problems that involved kinematic graphs. Class 3 (Course A) and Class 5 (Course B) had homework problems that involved interpreting graphs, but no problems that involved constructing graphs. Class 4 (Course A) and Class 6 (Course B) had problems involving graph interpretation and graph construction. The problems were part of the regular weekly homework, alongside regular homework for kinematics, so we were only able to deploy a limited number of these problems.

It should also be noted that Classes 1, 3, 5, & 6 were taught in the Fall, while Classes 2 & 4 were taught in the Spring; Class 1 & 3 also had the same instructor (all other Classes had different instructors); and in Class 4, only one section participated in the study, thus the significantly lower $n$ value.

## 3.3 Analysis of TUG-K Results

Table 3.2 shows the relevant macroscopic values for each of the classes.

Table 3.1: Comparison of the TUG-K data for the six different classes. The first columns show the course, class, semester, number of students in the course ($N$), number of students who took both the pretest and posttest ($n$), as well as the number of graph interpretation (No. Interp.) and graph construction (No. Constr.) problems. The following columns show the average pretest and posttest scores, as well as the average gain and normalized gain.

| Class | Course | Sem | $N$ | $n$ | No. Interp. | No. Constr. | Ave. Pre | Ave. Post | Ave. Gain | Norm. Gain $g$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | A | F | 464 | 259 | 0 | 0 | 12.6 | 14.1 | 1.5 | .178 |
| 2 | A | S | 452 | 189 | 0 | 0 | 13.1 | 14.5 | 1.4 | .177 |
| 3 | A | F | 452 | 165 | 6 | 0 | 13.2 | 14.5 | 1.3 | .172 |
| 4 | A | S | 485 | 118 | 8 | 5 | 13.3 | 15.2 | 1.9 | .242 |
| 5 | B | F | 124 | 84 | 4 | 0 | 11.1 | 14.3 | 3.1 | .317 |
| 6 | B | F | 77 | 73 | 1 | 6 | 12.5 | 16.6 | 4.1 | .485 |

### 3.3.1 Gain and Normalized Gain

The gain for an individual is simply the difference between the posttest and pretest scores.

$$G = \text{Posttest Score} - \text{Pretest Score}$$

The normalized gain is defined as

$$g = \frac{\text{Average Posttest Score} - \text{Average Pretest Score}}{\text{Maximum Score} - \text{Average Pretest Score}}$$

The normalized gain became popular when Richard Hake ran his analysis of 62 introductory physics courses, which ultimately demonstrated that interactive engagement classes had better normalized gains than traditional classes[63]. The primary reason for using the normalized gain is that it does not correlate with pretest scores. This made it a reasonable method to compare classes at different instructional levels from different institutions.

We note that the introduction of graph interpretation problems in Course A had little effect on the gain and the normalized gain, while there is an increase in gain and normalized gain in both Courses A and B with the introduction of graph construction problems. This is remarkable given the small number of problems. However, we have too few classes to make a useful comparison of the normalized gains, and comparing the gains would be dangerous because they correlate with pretest scores. An Analysis of Variance (ANOVA) on the pretest scores indicates that they are statistically significantly different so we have chosen a more suitable analysis method.

Table 3.2: ANCOVA Results for the two courses. The null hypothesis is that the graph construction problems made no difference in TUG-K outcomes.

| Comparison | Source of Variation | SS | df | MS | F | P-value |
|---|---|---|---|---|---|---|
| Course A | Adjusted Means | 32 | 3 | 10.93 | 1.03 | 0.379 |
| | Adjusted Error | 7681 | 726 | 10.58 | | |
| | Adjusted Total | 7714 | 729 | | | |
| Course B | Adjusted Means | 88 | 1 | 88.42 | 9.60 | 0.002 |
| | Adjusted Error | 1418 | 154 | 9.21 | | |
| | Adjusted Total | 1506 | 155 | | | |

### 3.3.2 ANCOVA using Pretest as a Covariate

Analysis of Covariance (ANCOVA) is a combination of ANOVA and a linear regression model. Since a large portion of the variance in posttest scores can be explained by the variance of the pretest scores (48% in this case), we can remove some of the within-group variance from the posttest scores. It should be noted that an ANCOVA analysis is primarily for groups that were randomly selected. While we did not randomize a single set of students into the different semesters, we argue that the students who enroll in one semester versus another are effectively random.

Running an ANCOVA on the posttests, using the pretest as a covariate, answers the question, "if the students in these groups started at the same pretest level, would their posttest scores be different?" Given the difference in class size between Course A and Course B, and the difference in student population, we have chosen to run the analysis for each course separately. The results of the ANCOVA $F$-test are displayed in Table 3.2, which shows the effect of introducing the graph construction problems. Within Course A (with engineering students), the result is not statistically significant, while in Course B (with the premedical student), the handful of graph construction problems made a significant difference.

Given that the only direct intervention was to add or change a small number of homework problems during the course of the entire semester, it is surprising to see a statically significant effect in the course for premedical students. What is equally surprising, and much more confusing, is that the results are not significantly different for the engineering students. In fact, Course A has all of the advantages for finding a statistically significant result (if the null hypothesis should be rejected.) The $n$-value is notably higher, and the total number of homework problems related to graphs in Class 4 was significantly larger than for any other class. Obviously, understanding the reasons behind this discrepancy requires further investigation. We will pose a few hypotheses here.

The difference in student populations between the two courses is known to be significant. One obvious possibility is that these differences in student population are the reason behind our conflicting results. Perhaps engineering students are already comfortable with graphs and only needed to learn the kinematics, while premedical students needed to learn both, or perhaps the students in Course B are more dedicated and thus were able to get more out of the graph construction problems.

It may also be the case that the lack of difference is the result of it being the ‘off-semester’ class. It may be the case that students who take their first physics class in the Fall tend to do better than those who take it in the Spring, meaning enrollment in one semester versus another may not be random after all (in Course A). It’s possible that if Class 4 had been in the Fall instead of the Spring, the posttest scores would have been statistically significantly different.

Another possible reason for the discrepancy is that in Classes $1-5$, kinematics was covered in the second and third weeks of class, whereas in Class 6, kinematics was covered in the sixth and seventh week of the semester. An argument can be made that the temporal proximity to the posttest led to higher scores for Class 6. However, the opposite argument can also be made: being taught kinematics just a week after taking the pretest could have primed their learning. In other words, since the students knew what they would be tested on, they paid more attention to the relevant details.

# Chapter 4

# How Students Solve Graph Problems

In this study, we examine the thought processes involved in solving graph problems in physics. To that end, we employed a think-aloud protocol [64] as students worked through a variety of graph problems, which were chosen to mimic problems routinely assigned as homework. We then looked for evidence of higher order thinking, and if indeed graph problems foster these desirable strategies.

Along the way, we were interested in possible factors that may influence cognitive processes in graph problem solving:

- The effect of requiring students to construct graphs rather than interpret them. As Leinhardt, Zaslavsky, and Stein noted, "[W]hereas interpretation does not require any construction, construction often builds on some kind of interpretation"[11]. Instead of constructing graphs, "[students] are usually given a formula or asked to select the appropriate formula from a well-defined (and very short) list and then to manipulate it using techniques from algebra or calculus [12]." Will graph construction force students toward higher order thinking?

- The order of problems. If construction problems indeed lead to higher order thinking, do construction problems prime students to approach subsequent interpretation problems differently, or vice versa?

- The gender of the subjects. At least at younger ages, gender has been found to have

Table 4.1: Self-reported background information. The "Sem." category lists whether the subject was taking Physics I or II at the time of the interview. The "Grade" category lists the subject's grade in the first semester course, either self-reported or "self-predicted" if they were still in the class (Jodie did not venture to predict her grade).

| Name | Sem. | Grade | Other Graph Exp. | Other Physics |
|---|---|---|---|---|
| Calvin | I | 4 | | High School |
| Jodie | I | - | Calculus | High School |
| Isaac | I | 3.5 | | High School |
| Abbie | II | 4 | Calculus | High School |
| Erica | I | 4 | Calculus | |
| Andrew | I | 3.5 | Calculus | Physics 101 |
| Cindy | II | 4 | Calculus | High School |
| Gideon | I | 3 | Calculus | High School |

significant influence on graph problem solving strategies (e.g. [65]).

- The effect of the problem medium. Would the delivery method, on paper or electronic, influence their thinking process, given that the electronic online medium can be in the way of employing higher order thinking skills ("turning thinkers into guessers" when coupled with multiple possible attempts and instant feedback [60]).

## 4.1 Population and Methodology

Subjects for this study were recruited as volunteers from introductory physics courses. At the beginning of the interview, each subject was asked to fill out a short questionnaire on their physics and graphing background (Table 4.1). Of the eight subjects in the study, some had completed their first semester of Physics, while others had not, but all had completed the material relevant to the problems in the interviews.

The interviews were conducted using a think-aloud protocol, [64] and lasted from 30-60 minutes. They were video-recorded for later transcription. In each interview, subjects were given five homework-style graph problems to complete, one at a time. During the interview, the subjects had access to scrap paper, a calculator, and an introductory physics textbook. The subjects were responsible for determining when to move on to the next problem (either after finishing or giving up on a problem), but they could not return to a previous problem after they had moved on. While working on the problems, intervention by the interviewer was limited to reminding the subject to keep thinking aloud. After completing the problems, there was a brief follow-up interview that was unique to each subject. The questions asked during the follow-up interview focused on the methods the subject used to solve the problem.

The problems used in this study were developed by the authors and were intended to be similar to (and somewhat more difficult than) the problems found in many standard textbooks or lab manuals. Bertin [15] divided the problems that graphs can answer into three categories: elementary, intermediate, and comprehensive. These categories have been refined over the years [16, 17, 18, 19], see Friel et al. [20] for a review. Elementary questions involve a simple extraction of data; intermediate questions involve identifying trends; comprehensive questions ask students to compare whole structures of the graph. [18] Using these categories, the problems in our study would mostly be characterized as "elementary" and "intermediate." Since many textbooks do not include graph construction problems, the ones used here were designed to be similar in format to graph interpretation problems (e.g. having an answer that could be given in the back of the book).

Four subjects were given graph problems on paper, while the other four were given their graph problems using the LON-CAPA homework system [54]. Before the interview, each of the subjects using the LON-CAPA homework system was given an opportunity to familiarize himself or herself with the input method for the graph construction problems [62]. Each interview consisted of

- an introductory "baseline" problem, where subjects were given an equation of the form $y = mx + b$ and asked to graph it,
- two graph interpretation problems, and
- two graph construction problems.

It is important to note that one major difference between the two delivery mechanisms is that the problems in the LON-CAPA system allowed subjects multiple attempts, offering the subject immediate feedback on whether their answer was correct or not (possibly combined with a hint).

We will use a shorthand to identify these problems in the remainder of this paper, where the first letter designates electronic (E) or paper-based (P), and the second letter designates baseline problems (L), interpretation (I), and construction (C) problems. Each problem appears in one of the Figures in this paper. For example, EC2 is the second electronically administered construction problem and is shown in Fig. 4.3.

For the purposes of this study, graph construction problems are defined as any problem where the subjects must construct a graph for themselves. Graph interpretation problems are defined as any problem where a graph is given and one does not need to be constructed. Problems where graphs appear in both the question and answer are somewhat less straightforward. For the purposes of this study, if a graph is given and the subject must choose the related graph from a given set (as in EI1), it is classified as graph interpretation. However, if a graph is given and the subject must use that graph to construct a related graph (as in PC2), it is classified as graph construction.

Within the categories of electronic vs. paper based problems, two of the subjects completed the graph interpretation problems first, while the other two completed the graph

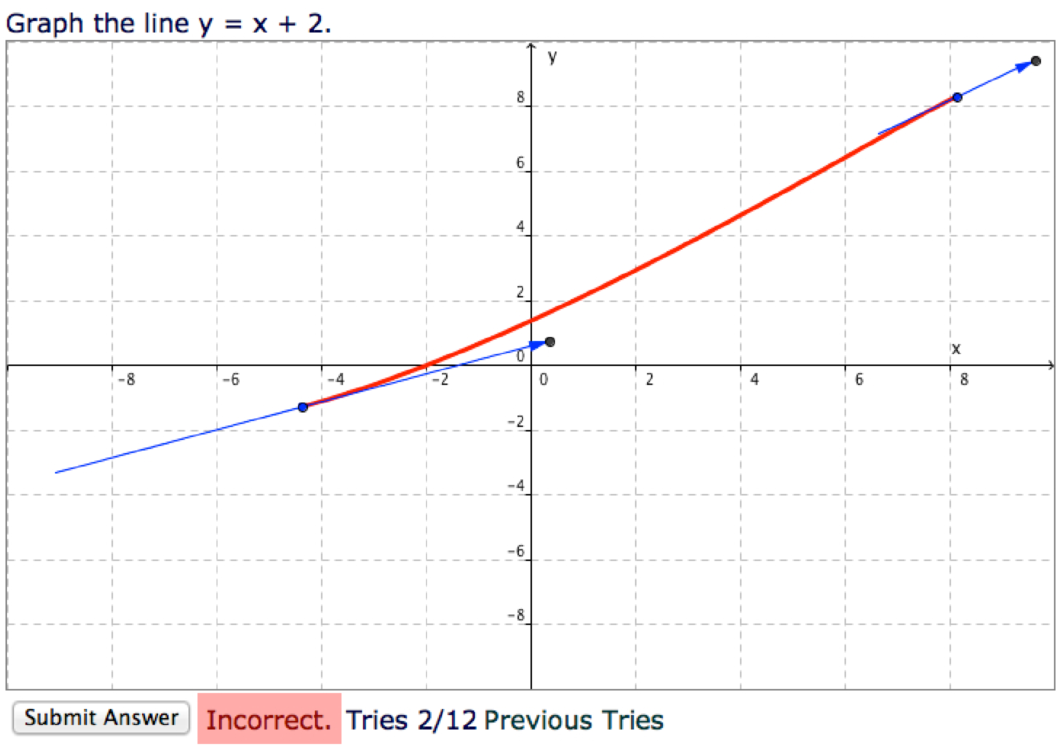



Figure 4.1: Problem EL. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

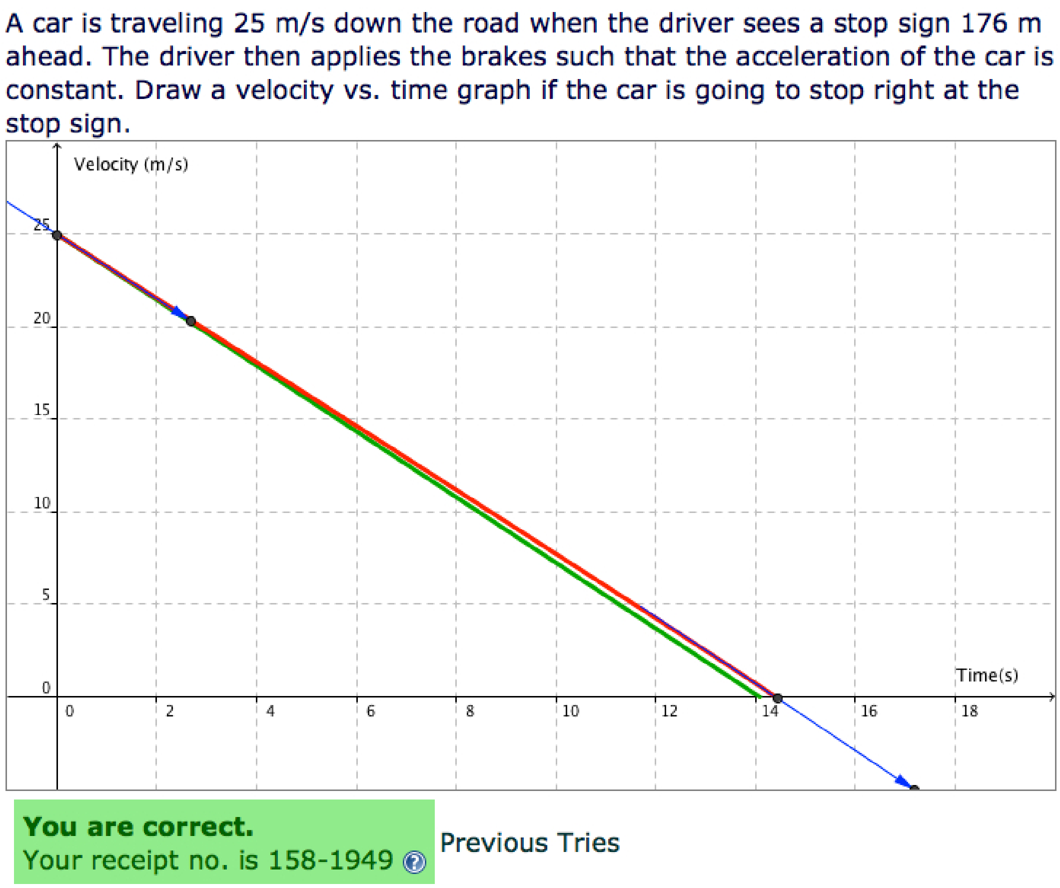
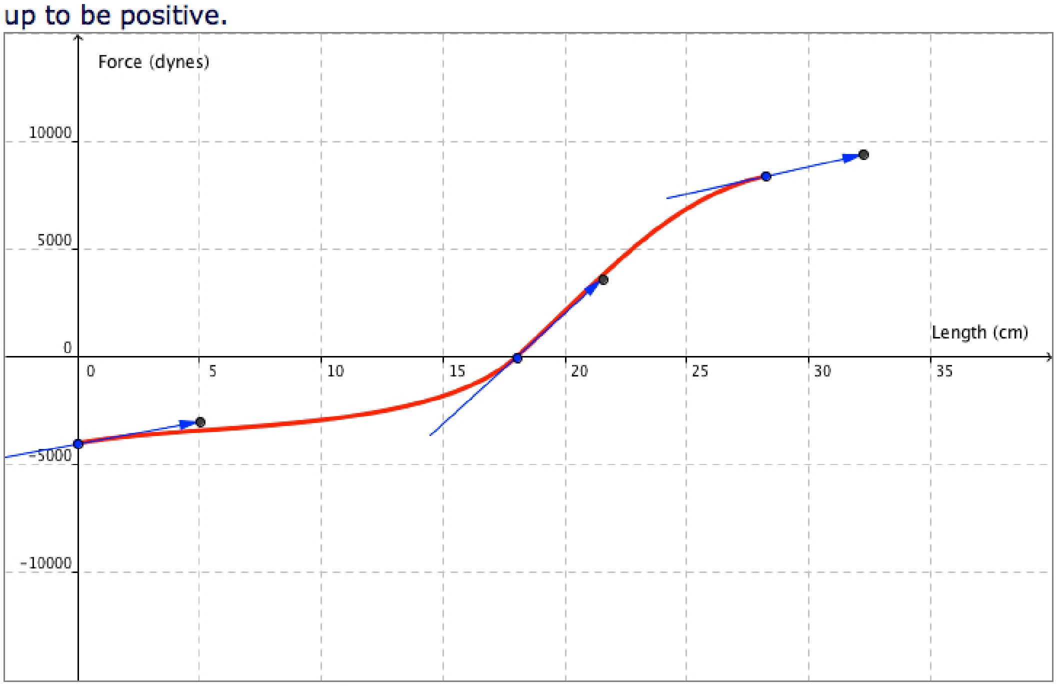



Figure 4.2: Problem EC1. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

A spring with spring constant 610 dynes/cm is hanging vertically from a hook attached to the ceiling. A mass of 50 g is hanging motionless on the end of the spring, which is now 20 cm long. The spring is then compressed and stretched to various lengths, and the net force acting on the mass is measured each time. Make a graph of this net force as a function of the length of the spring. Assume up to be positive.




Submit Answer Incorrect. Tries 3/12 Previous Tries

Figure 4.3: Problem EC2. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

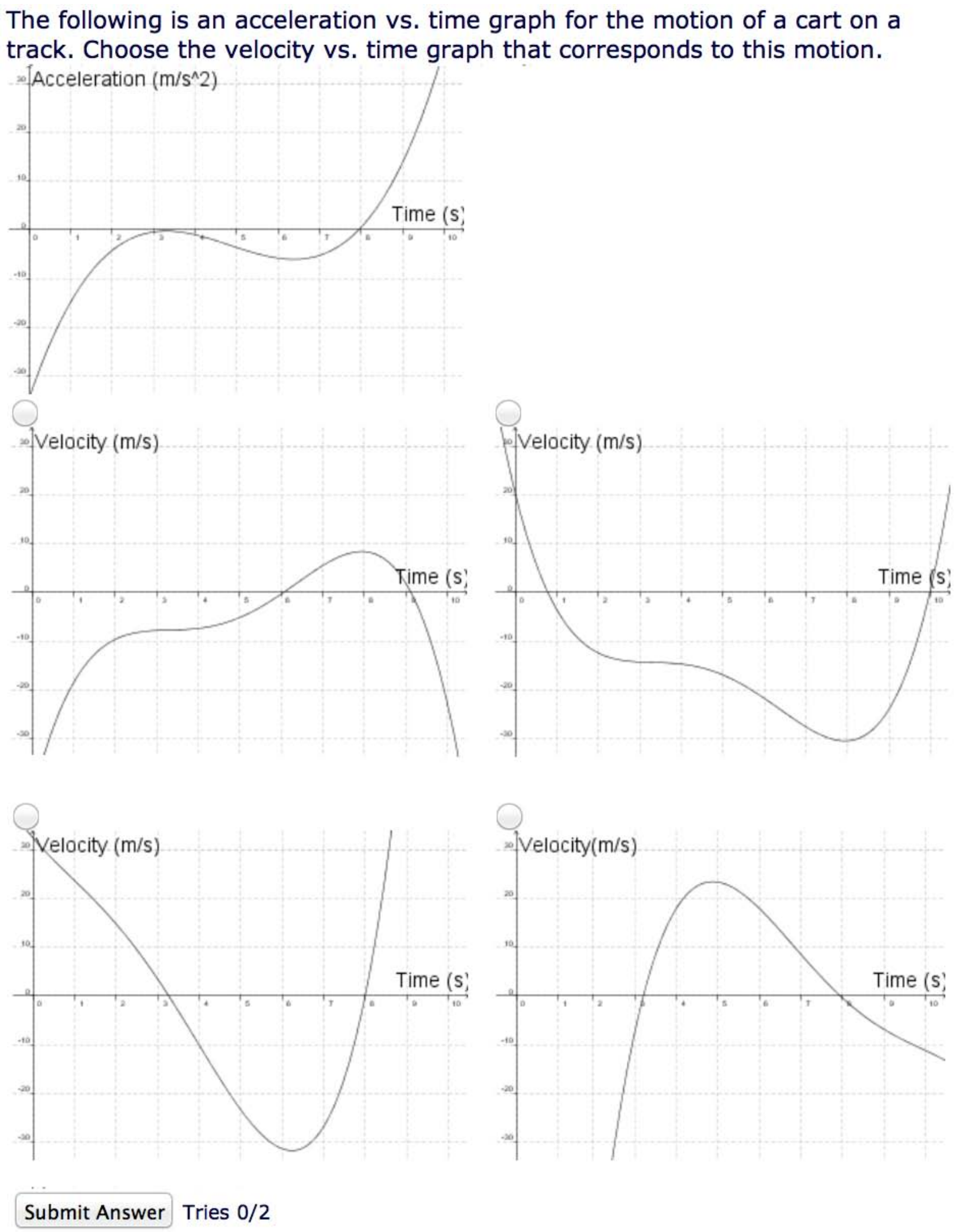



Figure 4.4: Problem EI1. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

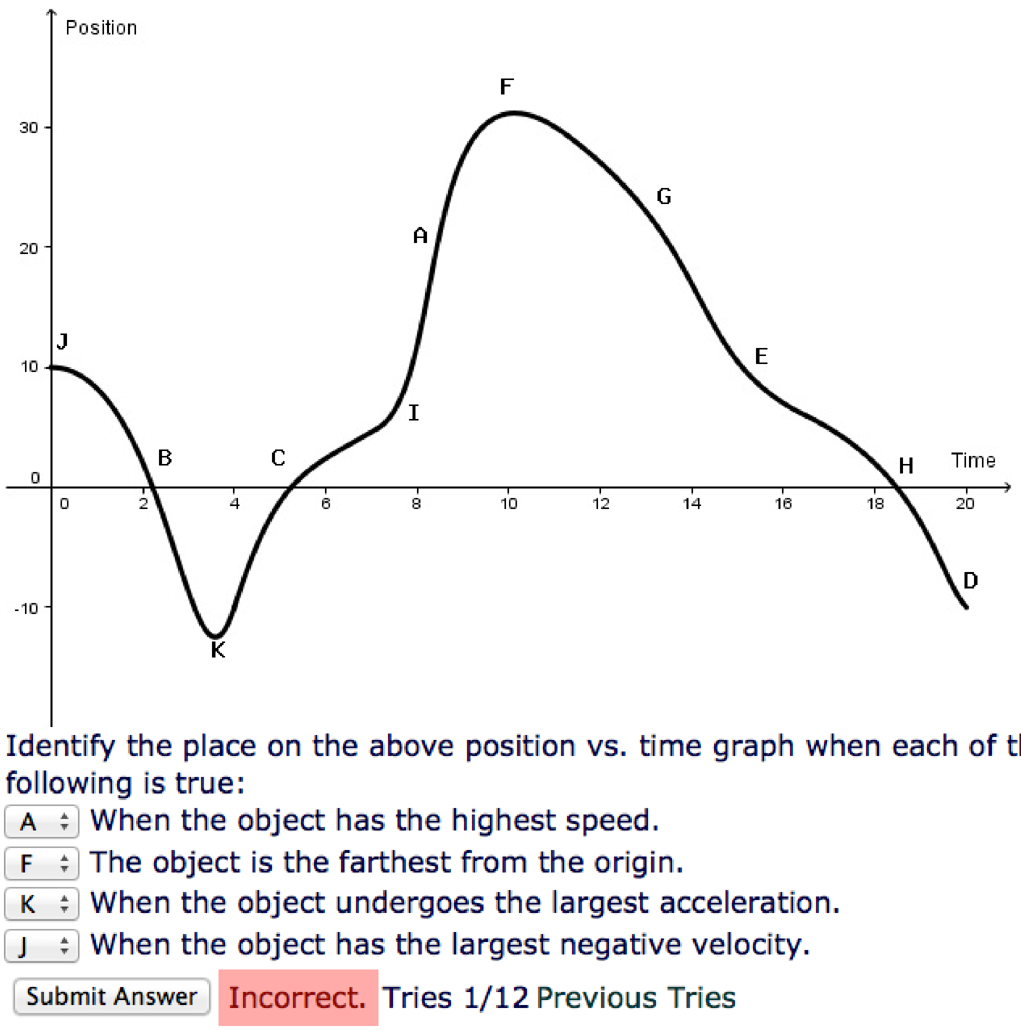



Figure 4.5: Problem EI2. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

PL: On the Grid below, graph the line y = x+3.

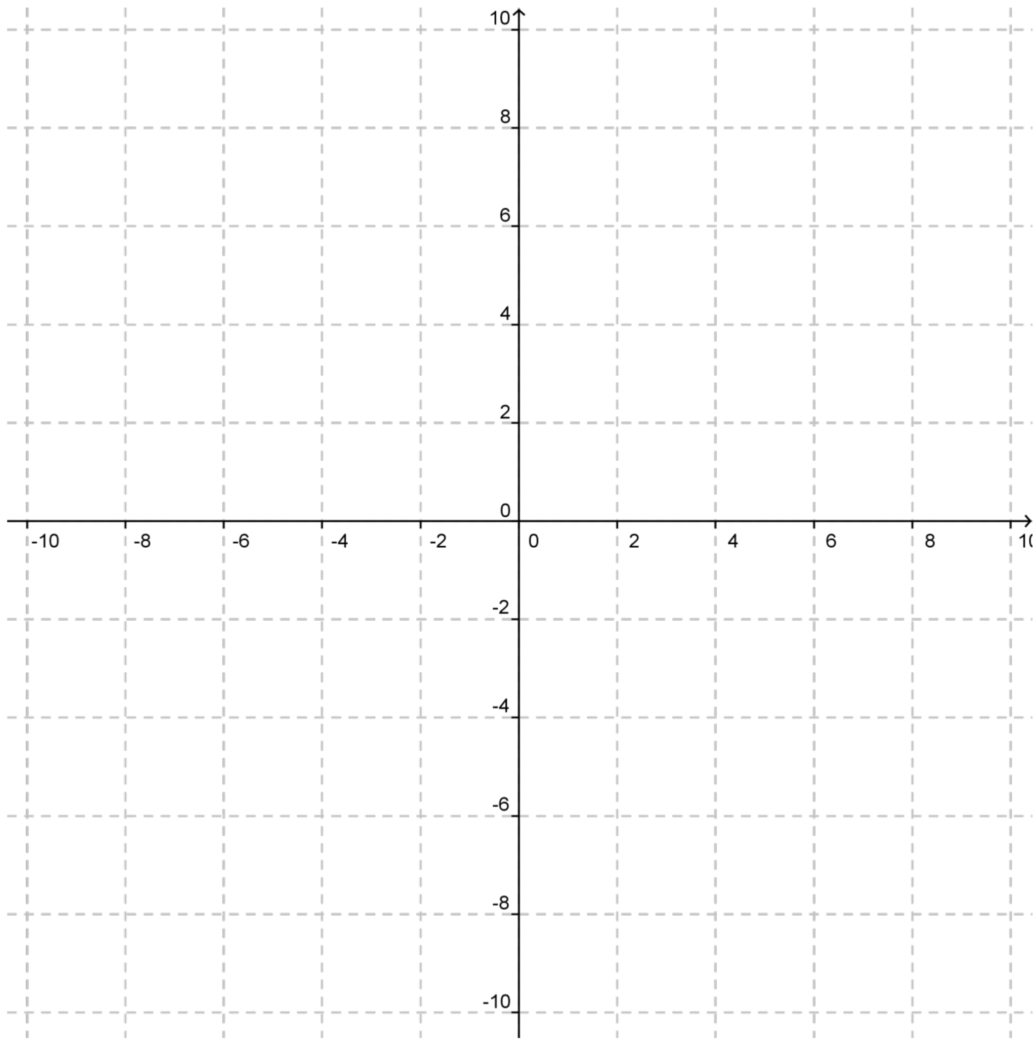



Figure 4.6: Problem PL. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

PC1: A rock is thrown from a height of 2 meters above the ground. Its initial velocity is 13 m/s at an angle of 60 deg. above the horizontal. Sketch the path (y-value vs. x-value) the rock travels until it hits the ground.

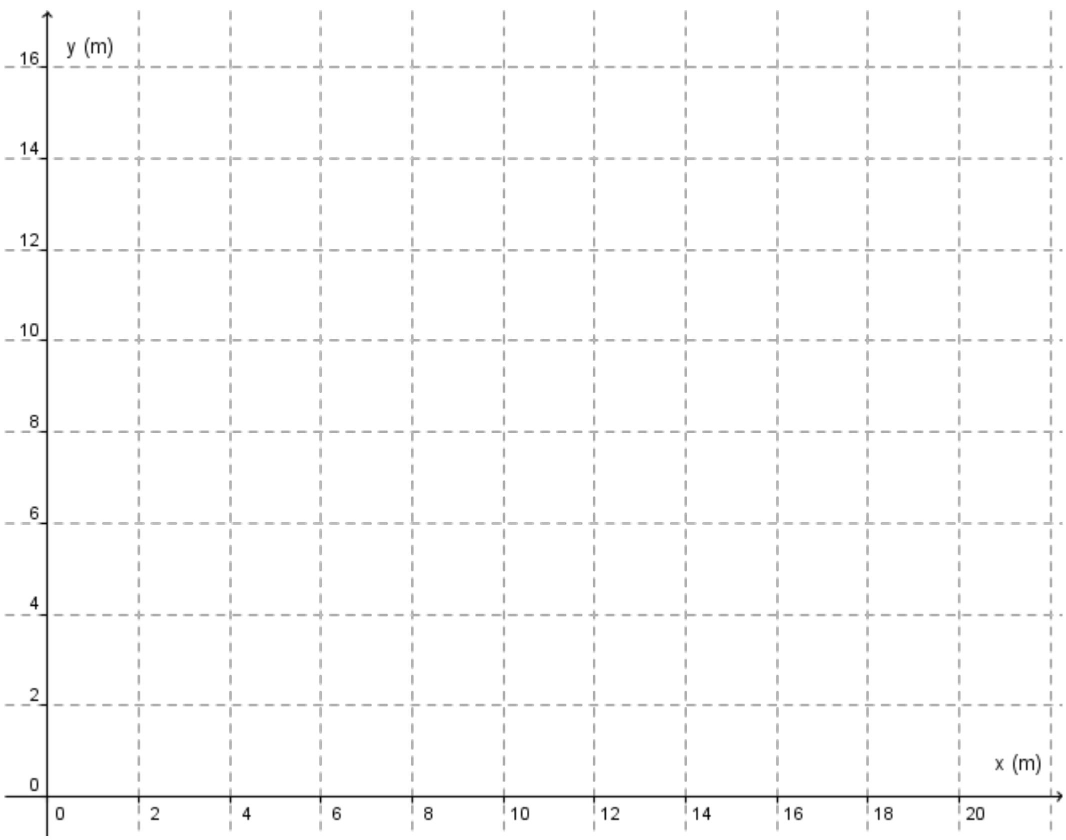
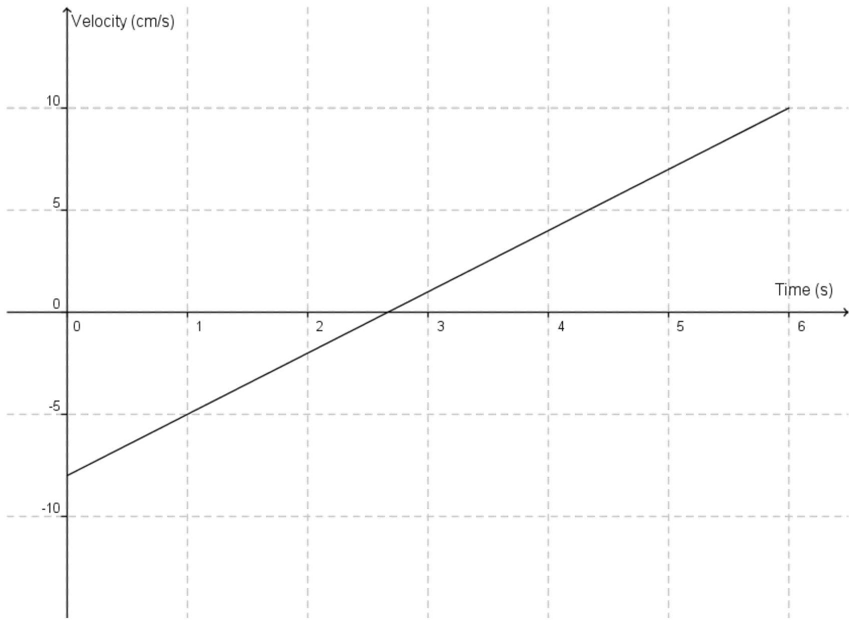



Figure 4.7: Problem PC1. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

PC2: The following graph shows the velocity of a ball on a track as a function of time.




Use the v vs. t graph above to create a position vs. time graph below. Assume that the ball is at x=0 at t=0.

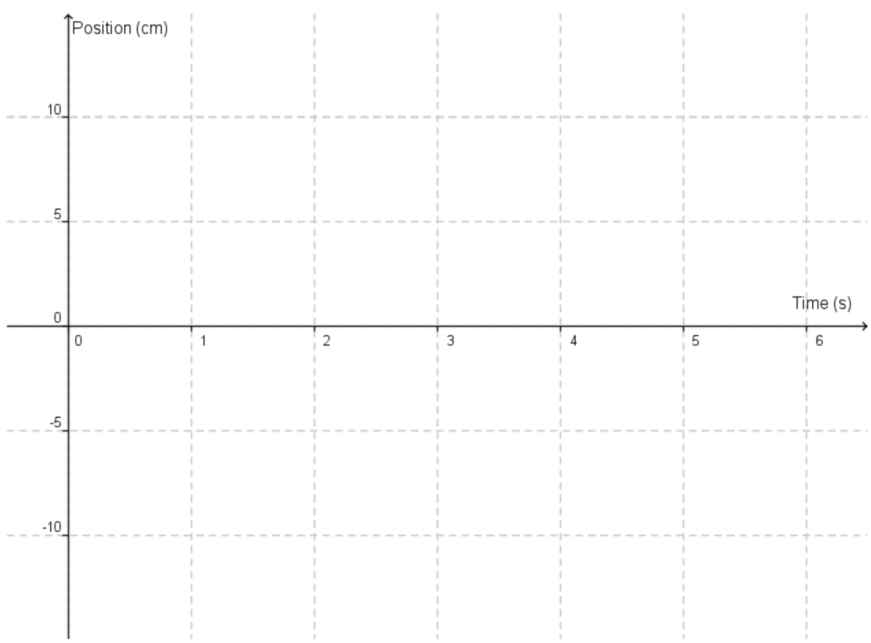
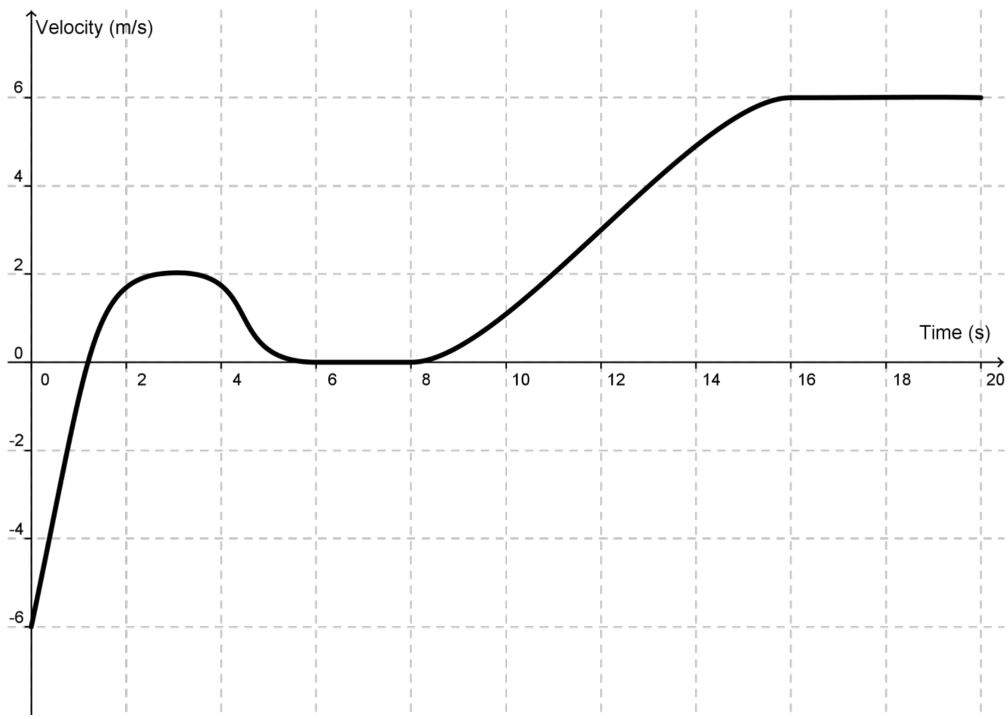



Figure 4.8: Problem PC2. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

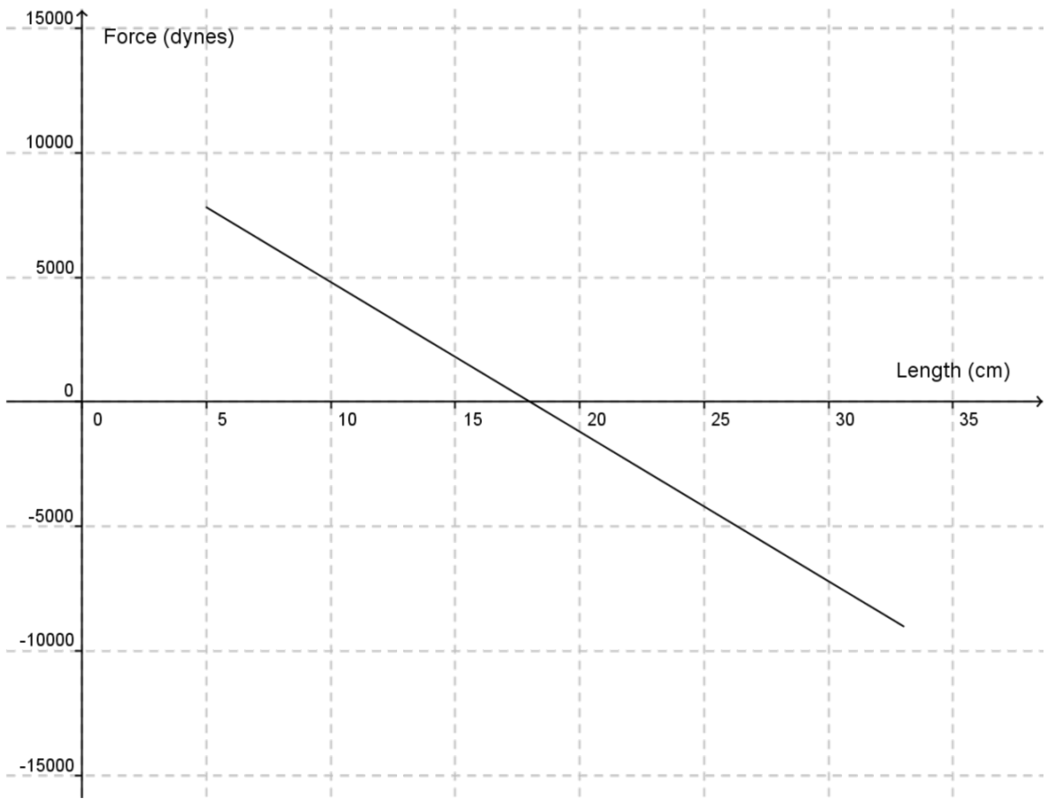
PI1: The graph below shows the velocity of a car as a function of time. Use it to answer the following questions.




1. When is the car the farthest behind where it started?
2. What is the car's final velocity?
3. During what time interval(s) is the car's acceleration negative?

Figure 4.9: Problem PI1. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

PI2: The graph below shows the force exerted by a spring as a function of the length of the spring. Use it to answer the following questions.




1. What is the spring constant, k, of the spring?
2. What is the length of the spring when it is in equilibrium?

Figure 4.10: Problem PI2. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

Table 4.2: Stratification of subjects into different interview arrangements.

| | Construction First | Interpretation First |
|---|---|---|
| Electronic | Calvin, Jodie | Isaac, Abbie |
| Paper | Erica, Andrew | Cindy, Gideon |

Table 4.3: Correct Answers — Electronic. An 'X' indicates a correct solution as determined by LON-CAPA.

| | EL | EC1 | EC2 | EI1 | EI2 | % Correct |
|---|---|---|---|---|---|---|
| Calvin | X | | | X | X | 60 |
| Jodie | X | | | | X | 40 |
| Isaac | X | X | | X | X | 80 |
| Abbie | X | X | X | X | X | 100 |
| % Correct | 100 | 50 | 25 | 75 | 100 | 70 |

construction problems first. The subjects were intentionally stratified so that one male and one female subject was in each of these four groups (see Table 4.2). Overall, the subjects performed well on the warm-up problems EL and PL (see Tables 4.3 and 4.4), so in spite of their different backgrounds (Table 4.1), all subjects demonstrated a basic knowledge of graphing. Overall, the subjects struggled more with the construction than with the interpretation problems, independent of medium.

Table 4.4: Correct Answers — Paper. An 'X' indicates a correct solution. Problems PI1 and PI2 had multiple parts, where in the averages, each part was counted as a separate problem. Problems were graded by the interviewer. No partial credit given.

| | PL | PC1 | PC2 | PI1a | PI1b | PI1c | PI2a | PI2b | % Correct |
|---|---|---|---|---|---|---|---|---|---|
| Erica | X | | | | X | | | X | 37.5 |
| Andrew | | X | | | X | X | X | X | 62.5 |
| Cindy | X | | X | X | X | X | | X | 75 |
| Gideon | X | | | | X | X | | X | 50 |
| % Correct | 75 | 25 | 25 | 25 | 100 | 75 | 25 | 100 | 56.3 |
| % Correct (tot.) | | | | | 66.7 | | 62.5 | | |

## 4.2 Analysis

Transcripts of the interviews were analyzed for problem solving strategies and evidence of higher order thinking, i.e., evidence of analysis, synthesis, and evaluation. Instead, we mostly found evidence for strategies associated with lower-level educational goals. In addition, we found ample evidence for what can only be characterized as "guessing."

### 4.2.1 Transcripts

The video recordings of the interviews were transcribed with a focus on the strategies the subjects used to solve the problems. The transcripts were not verbatim, but kept track of the details relevant to how the subject solved the problem. Tables 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11, and 4.12 in turn show synopses of these transcripts, giving an overview of how the subjects went about solving the electronic and paper-based problems, respectively. One of the concerns of the think-aloud protocol is that it requires some amount of working memory to voice thoughts. If the task given is taxing to the subject, they may have a difficult time voicing their thoughts due to the load on their working memory. In the interviews, there definitely seemed to be a correlation between how much a subject was silent and how poor their performance was, and at times subjects had to be repeatedly reminded to "think aloud."

Table 4.5: Transcript synopsis for Calvin. Instances of higher order thinking processes are italicized.

| Name | EL (Fig. 4.1) | EC1 (Fig. 4.2) | EC2 (Fig. 4.3) | EI1 (Fig. 4.4) | EI2 (Fig. 4.5) |
|---|---|---|---|---|---|
| Calvin | Makes a data table, calculates points on graph. Correct. | Writes down $x$-equation from memory, as well as $\Delta v/\Delta t = a$. Realizes that $v$-graph should be linear with negative slope. Sketches $v$-graph on paper and states that area under $v$-line should be "distance curve" and correctly draws general shape of distance graph. Eventually writes $\Delta v \cdot \Delta t = D$ and $a \cdot \Delta t = \Delta v$. Gives up. | Draws diagram of spring hanging from ceiling. Draws sine graph suggesting the answer will be similar. Moves middle point to location of equilibrium point. States $F$ will be negative when spring is compressed. Aligns remaining points to form a monotonic, increasing graph. Writes down $k\Delta x^2 = F$ from memory. Calculates $F$ using this equation and obtains values not on the graph. Considers problem might be a units issue. Notes graph should have symmetry. Gives up. | Describes shape of given graph. *Attempts to understand physical situation.* Notes $v$ is not decreasing in the beginning for two of the options. Notes that the rate of decrease of $v$ should be lower near the middle. Correct. | Chooses "greatest position". States steepest slope will give largest $a$. States $v$ is largest when largest change in $x$ over smallest change in time. *Revisits part two because it should not be same answer as part three.* Chooses point at 8 second mark. Guesstimates change in $x$ and change in $t$ for various points with negative slope, tries to identify steepest. Incorrect. *Makes an educated guess for point with largest $a$.* Correct. |

Table 4.6: Transcript synopsis for Jodie. Instances of higher order thinking processes are italicized.

| Name | EL (Fig. 4.1) | EC1 (Fig. 4.2) | EC2 (Fig. 4.3) | EI1 (Fig. 4.4) | EI2 (Fig. 4.5) |
|---|---|---|---|---|---|
| Jodie | Manipulates graph into line $y = x$. Uses a graphing calculator to draw function and manipulates graph to match. Correct. | Correctly finds initial point, states she is going to draw a straight line, "but it doesn't say how long." States that "acceleration is concavity," draws curved line. States that $a$ is constant, thus the "graph will level out." Tries curved graphs. Never uses the given distance. Gives up. | States graph will be similar to sine or cosine graph. Tries to remember equations for period and "whatnot". Makes a roughly sinusoidal graph with a period equal to the width of the graph. Incorrect. Knows $k$ is $F$ over change in $x$. *Considers the graph may be a straight line, but rejects it because the spring changes length.* Searches book. Gives up. | States incorrect relationship between the given graph and the options. Chooses two options based on "obvious" points. Realizes she is out of attempts. Gives up. | States she wants steepest slope of tangent line for highest speed. States largest negative $v$ is when "the slope of the tangent line is decreasing at the steepest". "Acceleration is concavity". Chooses point farthest from origin. Correct. |

Table 4.7: Transcript synopsis for Isaac. Instances of higher order thinking processes are italicized.

| Name | EL (Fig. 4.1) | EC1 (Fig. 4.2) | EC2 (Fig. 4.3) | EI1 (Fig. 4.4) | EI2 (Fig. 4.5) |
|---|---|---|---|---|---|
| Isaac | Mentally calculates points on graph. Draws line. Correct. | Notes initial $v$. Writes down $x$-equation from memory. Looks in book for another equation. Finds $v_f^2 = v_i^2 + 2ax$. Solves for $a$. Solves $v$-equation for time. Starts to graph a parabola, but realizes it should be a line because $a$ is constant. Correct. | Solves for $F$ of gravity on mass ($mg$) and force on spring ($kx$). Attempts to create a conversion from Newtons to dynes with this, but realizes he doesn't know $x$. Obtains correct conversion factor. Is confused by the results. Gives up. | Notes that $a$ is negative until 8 seconds. *Eliminates options that don't agree with this. Repeats this to double-check.* Correct. | Copies graph, *evaluates some elements,* then reads questions. States largest negative $v$ is when slope is steepest and decreasing. Chooses point that is nearly vertical for highest speed. *For largest $a$, draws a rough $v$ vs. $t$ graph.* Realizes answer after drawing it. Reads off point farthest from origin. Correct. |

Table 4.8: Transcript synopsis for Abbie. Instances of higher order thinking processes are italicized.

| Name | EL (Fig. 4.1) | EC1 (Fig. 4.2) | EC2 (Fig. 4.3) | EI1 (Fig. 4.4) | EI2 (Fig. 4.5) |
|---|---|---|---|---|---|
| Abbie | Realizes that equation is of form "$y = mx + b$" Notes intercept and slope. After failed attempts, examines other specific points. Adjusts graph to match. Correct. | Notes that constant $a$ means graph should be a line. Identifies initial value. Calculates final time dividing distance by initial $v$. Incorrect. Pulls position equation from book. Identifies that there are two unknowns. Finds and solves $v_f^2 = v_i^2 + 2ad$ for $a$. Solves $v$-equation for time. Correct. | Converts dynes to Newtons. Calculates $mg$. Adds $kx$ to $mg$ to find $F$ at length given, obtains a number not on graph. *Considers what happens as spring length goes to zero; realizes that given length is the equilibrium point, and that the distance in "$kx$" is from equilibrium.* Calculates a few points on graph. *Decides shape should be line based on the equation $F = -kx$.* Guesses that her graph is off by a negative sign. Correct. | States integral of $a$ is $v$ and derivative of $v$ is $a$. *Looks at the four graphs and eliminates two because they don't have negative slopes in the beginning. Identifies correct graph by where the slope is zero first.* Correct. | Determines highest speed is when slope is greatest. Chooses point farthest from axis for farthest from origin. Notes largest negative $v$ is steepest negative slope. *Identifies that trough has an abrupt change (educated guess).* Correct. |

Table 4.9: Transcript synopsis for Erica. Instances of higher order thinking processes are italicized.

| Name | PL (Fig. 4.6) | PC1 (Fig. 4.7) | PC2 (Fig. 4.8) | PI1 (Fig. 4.9) | PI2 (Fig. 4.10) |
|---|---|---|---|---|---|
| Erica | Types equation into her calculator. Based on equation, identifies that slope is 1 and $y$-intercept is 3. | Identifies all given values while reading the question. Sketches a rough version of what she thinks the graph will look like. Looks in book for something similar. Finds and solves max height equation. Finds range equation but does not think it will be helpful. | States she needs the area under the $v$-graph to get position. Finds areas and uses them to create points on graph. Connects those points, but doesn't like her answer. Starts problem over by finding the equation of the line in the given graph and integrating it. Graphs the integrated function in her calculator. Copies to paper. | Identifies position as the integral of $v$ and shades in the area under graph. Identifies final $v$ as a point she can just read off. Recognizes that she must differentiate and chooses "the only place where, uh, the slope is negative". | Finds equation $F = -kx$ in book. Solves for $k$, writing down $k = -\Delta F/\Delta x$. Identifies that this is the slope of the graph. Notes that equilibrium is when $F = 0$. |

Table 4.10: Transcript synopses for Andrew. Instances of higher order thinking processes are italicized.

| Name | PL (Fig. 4.6) | PC1 (Fig. 4.7) | PC2 (Fig. 4.8) | PI1 (Fig. 4.9) | PI2 (Fig. 4.10) |
|---|---|---|---|---|---|
| Andrew | Identifies that "$y$ equals $x$ is just a horizontal line." Draws graph of $y = 3$ for answer. | Identifies initial values, drawing the initial position and its $v$-vector on the graph. Finds equations for range and max. height. Solves them in calculator. Draws the max. height as halfway between initial point and max. range. | *Attempts to find out what negative $v$ looks like on an $x$-graph by thinking about a physical situation. Hypothesizes a 'U' shaped track.* Notes that $x$-intercept is when $v$ changes from negative to positive. Attempts to calculate an area. Connects the point he calculated using area to initial point with a line. Justifies shape claiming that "since it's a constant velocity, it'll move away at a constant rate." | Identifies answer to part two because he can "just look at the end of the graph." Notes that the acceleration is negative when $v$ is decreasing. Guesses for the first question. "I'm just gonna go with time twenty cause that's when it's farthest away." | Identifies that equilibrium is when $F$ is zero. Finds $F = -k\Delta x$ in book and reads a point off the graph, plugs in values of $F$ and $\Delta x$ to find $k$. |

Table 4.11: Transcript synopsis for Cindy. Instances of higher order thinking processes are italicized.

| Name | PL (Fig. 4.6) | PC1 (Fig. 4.7) | PC2 (Fig. 4.8) | PI1 (Fig. 4.9) | PI2 (Fig. 4.10) |
|---|---|---|---|---|---|
| Cindy | States equation is of form $y = mx+b$. Finds and draws intercept. Uses slope to draw another point and connects them with a line. | States answer is "upside-down looking parabola" and draws sketch of it off to the side. Identifies initial position and that $v_x$ is constant, while $v_y$ changes because of gravity. Finds top of graph using $v_f^2 = v_i^2 + 2ad$. Uses resulting position to find time and uses this for peak of her graph, mistakenly thinking it is $y$ vs. $t$. States that graph is symmetric around this point and draws parabola accordingly. | Notes that $v$ shows slope for the $x$-graph and that $a$ is constant. States that since $v$-graph is a straight line, the $x$-graph is a parabola. Notes that ball will return to its initial location around 5.5 seconds because the $v$-graph is "just a mirror over here." Finds area under $v$-graph up to $x$-intercept and uses this to place the trough of her $x$-graph. *Notes that slope of velocity graph has positive slope and so her answer makes sense for being concave up.* | States that "farthest behind" happens when integral is most negative. States that final $v$ is just the value at the end. States $a$ is negative when slope of $v$ is negative. | Immediately mentions $F = k\Delta x$ and that $k$ is slope of given graph. Calculates slope from two points on graph. States that equilibrium is when there is no force on it. |

Table 4.12: Transcript synopsis for Gideon. Instances of higher order thinking processes are italicized.

| Name | PL (Fig. 4.6) | PC1 (Fig. 4.7) | PC2 (Fig. 4.8) | PI1 (Fig. 4.9) | PI2 (Fig. 4.10) |
|---|---|---|---|---|---|
| Gideon | Finds $y$-intercept and draws it. Draws another point using the slope from the original point and makes a line. | Draws the initial position and $v$-vector on graph. Searches book for an example and finds one. Extends the $v$-vector he drew into a line. Ultimately guesses height and range and just draws a parabola that's "gonna fall just under this line". | Draws point at origin. *Describes what $v$-graph shows in words. Attempts to describe the motion of the object's path.* Ultimately seems to guess a point and draws a line connecting it to origin. | *Uses graph to describe car's motion.* Chooses initial point for part one, presumably because it is the most negative. Attempts to use fact that $v$ is increasing to calculate final $v$. Reads off the value of $v$ for the end point. Decides $a$ is negative because car was going from higher $v$ to lower $v$ (not slope). | Finds $F = -kx$ and uses it at each end of graph "to see if the constant remains the same." Due to miscalculation, obtains two different values for $k$ and decides to average them to find value of $k$ in the middle. It is unclear why he chose the correct answer to the second question. |

### 4.2.2 Bloom Levels

The cognitive domain of Bloom's Taxonomy [66] forms a useful framework for the characterization of educational goals. Bloom and his colleagues defined six fundamental levels of educational goals: knowledge, comprehension, application, analysis, synthesis, and evaluation. However, the latter three goals are frequently treated as the combined goal of "higher order thinking," as it has been suggested that they are not truly hierarchical, but are three aspects at the same level of difficulty. [67]

We will use the cognitive domain of Bloom's taxonomy to discuss and identify subjects' strategies solving graph problems. Whenever this framework is employed, some interpretation is required in order to apply it to the specific instructional scenario under investigation. In our case, we needed to answer the question of how each of these levels manifests itself in solving graph problems. A summary of the levels and examples of each from this study can be found in Tables 4.13, 4.14, and 4.15. For the most part in this study, we do not concern ourselves with whether or not the subjects obtain the correct answer with their strategies, only that they are employing certain strategies.

Table 4.13: Part one of Bloom's Taxonomy in this study. While it is not a learning goal, we have added Guessing to the list of cognitive levels for the purposes of analysis in this study. More elaborate descriptions of how the subjects solved these problems can be found in Tables 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11, and 4.12

| Category | Level | Definition | Example |
|---|---|---|---|
| Lower Order | Guessing | Attempting to solve without any underlying reason | While a truly "Educated Guess" can in fact be evidence of higher order thinking, most guessing that we found is not a desired educational outcome. Pure guessing can come in many forms:<br>• If an answer seems right, *maybe* it is just off by a minus sign, so try the negative solution (e.g., Abbie solving problem EC2).<br>• If the problem asks where something particular happens on a graph, *maybe* it happens at the point that has the most striking feature (e.g., Abbie and Calvin, Problem EI2).<br>• If all else fails, just give it a try, *maybe* it is correct (e.g., Gideon, Problems PC1, PC2, and possibly PI2). |
| | Knowledge | Pulling facts, relationships, or equations from memory | In our study, not surprisingly, we found that the subjects, whenever possible, fall back to this level. The most common method for determining the shape of the graph was to recognize a certain relationship that the subject had learned. Examples are:<br>• A problem with a spring and a mass triggering "oscillator" and associated sinusoidal graphs, even if the problem does not address oscillation (for example Calvin and initially Jodie on Problem EC2).<br>• A problem asking for a line immediately triggering the memorized equation $y = mx + b$ (Abbie and Cindy on Problems PL and EL, respectively). |

Table 4.14: Part two of Bloom's Taxonomy in this study. While it is not a learning goal, we have added Guessing to the list of cognitive levels for the purposes of analysis in this study. More elaborate descriptions of how the subjects solved these problems can be found in Tables 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11, and 4.12

| Category | Level | Definition | Example |
|---|---|---|---|
| Lower Order | Comprehension | Understanding relationships and being able to switch between representations | • "Automatic" representation translations, for example going from $y = mx + b$ to slope and $y$-intercept in a graph (Cindy and Gideon, Problem PL.<br>• Understanding the meaning of axes on graphs, for example, Jodie and Isaac on Problem EI2 translating "far away" into "toward the end of the $x$-axis." |
| | Application | Applying acquired understandings to new situations | Subjects might not have seen a particular graph, but they do know how to apply certain rules to unknown graphs:<br>• Finding the integral as the area "under the curve" (e.g., Erica on Problems PC2 and PI1).<br>• Finding the derivative from the slope of a curve (e.g., Jodie on Problem EI2).<br>• Finding or attempting to find similar problems in memory or in a book, and adjusting them to the new situation (e.g., Jodie on Problem EC2, or Gideon on Problem PC1).<br>• Using kinematic equations (e.g. Isaac on Problem EC1 (Table 4.6)) or Hooke's Law (e.g., Abbie on Problem EC2). |

Table 4.15: Part three of Bloom's Taxonomy in this study. While it is not a learning goal, we have added Guessing to the list of cognitive levels for the purposes of analysis in this study. More elaborate descriptions of how the subjects solved these problems can be found in Tables 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11, and 4.12

| Category | Level | Definition | Example |
|---|---|---|---|
| Higher Order | Analysis | Make inferences or generalizations | • Problem EI1 (Fig. 4.4) was predominantly solved by process of elimination. Subjects identified and analyzed salient features in the graphs to eliminate incorrect answer options (Calvin, Isaac, and Abbie on Problem EI1). |
| | Synthesis | Combining elements in new ways | • Attempting (even unsuccessfully) to translate a graphical situation into a real physical scenario (for example Andrew on Problem PC2). |
| | Evaluation | Metacognitive strategies to check one's work | • Recovering from failed attempts. For example, on Problem EC2, Abbie recovers from misinterpreting $x$ in $F = -mx$ by considering a limiting case of spring length going to zero. <br> • Rejection of possible solutions. For example, on Problem EC2, Jodie rejects a solution she is considering based on the physical scenario. |

Instructors frequently hope that assigning graph problems in their homework will trigger or require higher order thinking processes. In this study, several patterns emerged for solving these problems, but more often than not, they employed lower order cognitive levels. In Tables 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, 4.11, and 4.12, we have highlighted the instances of higher-order thinking that we found in our study. In the following subsections, we describe the overarching patterns which emerged, as well as their associated Bloom-level building blocks.

### 4.2.3 Lower Order Thinking

The majority of strategies used by the subjects while working on these graph problems come from the three lower levels of Bloom's taxonomy or from our added category of "guessing". Almost all of the lower order attempts to solve the graph problems fall into the following categories.

- Determination of Points - Identifying or calculating points on the graph without looking at larger trends.
- Memorized Relationships - Relationships between concepts that the student has committed to memory.
- Formula Reliance - Using formulaic representations to solve the problem.
- Algorithmic Solving - Using algorithms ("programs") to solve problems that are similar to what they have seen before.

#### 4.2.3.1 Determination of Points

Particularly when it comes to graph construction, the hope is that students would sketch graphs from a holistic point of view, taking into account features and trends. Instead, the subjects frequently calculated points on the graph (or let their graphing calculator do that job from beginning to end). Students are trained to plot functions, often using value tables, and not really well trained to sketch "back of the envelope" graphs [25, 26]. Frankly, most graph problems do very little to force students into another mode, as the majority of points were obtained because they were either given in the problem statement, or could be evaluated from the values given in the problem statement:

> "So at time zero, the velocity should be twenty one meters per second." - Calvin (EC1)

> "I'm looking for the, uh, the equations that would tell you how far it would go or how high it would go." - Andrew (PC1)

In some of the problems, obtaining these points was a little bit more involved, as it required reading the values from another graph or, alternatively, noting symmetries:

> (Indicating what will become the minimum of her graph) "About here it's gonna be at some value. What value is that? It's gonna be the integral of this (shades in the area under the curve up to the $x$-intercept) the area of a triangle so I don't have to do the integration." - Cindy (PC2)

> (Indicating the right side of the given graph) "That's just a mirror over here so at like six seconds it's finally gonna come back up to the original starting position." - Cindy (PC2)

When subjects did not find any way to nail down their points, they sometimes reverted to guessing their location:

> "I guess the only thing I'm confused is is for how long cause it doesn't really say." (she then moves the end point to a new location and submits her answer)
> - Jodie (EC1)

It is to be expected that not all thoughts will be voiced in an interview using a think-aloud protocol, but it remains the best window we have on the thought processes of subjects. That being said, it is interesting to note that during the graph construction problems, none of the subjects ever voiced any thoughts on why they chose to evaluate certain points. For instance, in problem PC1, which asks the subject to graph the trajectory of a rock, each of the subjects expressed wanting to find the maximum height and/or range of the rock. However, there were no thoughts expressed as to why they chose to evaluate these points. This is in contrast to finding the shape of the graph, where the subjects often commented on what shape the graph would be and why.

There are indications, of course, that these choices are being made. During Erica's attempt to solve PC1, she finds the range equation in the book, "but the one here is, I believe that it starts from height is equal to zero and that's not what I'm looking for." Ultimately, she seems to have decided that the point she wants to find is somewhere on the $x$-axis, and since the range equation would not give this to her, she ignores it and eventually gives up on the problem. Of course, the range equation would have allowed her to find the point on the graph that intersects with the line $y = 2$, and ultimately get the problem correct.

Another indication of a choice being made to search for a specific point occurs while Calvin is working on problem EC1. Just after he identifies that the graph will be a line with negative slope, he notes, "But then you have to solve for the time. So at what time does the velocity equal zero?" Again, a clear indication that he wants to find the point where the line crosses the $x$-axis, but no explanation as to why he wants to find this point.

The implication is that these decisions are happening at a more sub-conscious level and it may be worth future research to investigate how individuals identify such points. It is also worth noting that in almost all cases, the points that the subjects thought were important were useful for solving the problems they were given.

#### 4.2.3.2 Memorized Relationships

Many of the subjects had a good number of relationships memorized when they were interviewed. There were many instances of the subjects mentioning one or more of these during the course of a problem and predominantly, the relationships they mentioned were correct. This is not surprising, as typically, students do not have much difficulty with graph problems that only require these kinds of memorized relationships [6].

The subjects were also very confident about these relationships: while some of the time, the subjects questioned the relationship they thought was true, at no point did any of them attempt to verify or contradict themselves by looking it up. Of course, most problems require interpretation or construction well beyond these simple relationships.

These memorized relationships are clearly associated with lower order thinking processes, they are recipes that are mostly at the application level, though some were simply mentioned (knowledge) as the subject attempted to work a problem in the hopes that it would help them.

In graph interpretation, the subjects frequently invoked relationships such as acceleration being the slope of velocity or position being the integral of velocity.

> “Ok, so this is a position versus time graph so we want slope of the tangent line where it is the steepest.” - Jodie (EI2)

> “And then, position is the integral of velocity so it’s not necessarily where it’s the most negative, but when the integral is most negative.” - Cindy (PI1)

> “Acceleration is concavity, I believe.” - Jodie (EI2)

Alternatively, the relationships could be between the concepts of kinematics, such as velocity.

> “[...] so it’s going from a higher velocity to a lower velocity, and so accelerating downward.” - Gideon (during the follow-up interview, regarding PI1)

Meanwhile, for graph construction, these memorized relationships could be used to determine the shape of a graph, or to solve for points on the graph. In some cases, they involved general calculus principles:

> “Alright, so the acceleration’s constant. So, like that (draws horizontal line on scratch paper). So, then the velocity is going to decrease linearly (draws a line with negative slope on scratch paper).” - Calvin (EC1)

> “If this (pointing to the given graph) is a straight line that means that this (pointing to the area where she must construct a graph) is gonna be a parabola, because the graph of the derivative of a parabola is a straight line.” - Cindy (PC2)

> “So the velocity is, shows the slope for my position graph... and then the slope of the velocity, um, determines the concavity for the position graph because acceleration is the (inaudible) second derivative and second derivatives determine concavity.” - Cindy (PC2)

#### 4.2.3.3 Formula Reliance

We found some evidence of a behavior we call "formula reliance," again at the application level. This behavior is characterized by the use of mathematical equations, often to the exclusion of other strategies for trying to solve the problem. Erica gives us many examples of this, as she made no attempt to solve any of the problems using anything but mathematical formulas and memorized relationships.

Upon receiving the first problem (PL), which asks her to graph the line $y = x + 3$, she notes, "I can easily just do it by calculator without even thinking and it would be much easier." After typing the equation into her calculator, Erica notes the slope and $y$-intercept of the equation and draws the answer on her paper.

The second problem Erica encounters (PC1, also mentioned above) asks her to graph the trajectory of a rock which initially has a velocity that points above the horizontal. While she is able to immediately draw the general shape of the graph, she is ultimately unable to solve the problem because, "I wanted to find the range and I couldn't really find, like, something straightforward in the book." She did, however, find and solve the equation for the maximum height along the way.

The third problem Erica works on (PC2) gives a velocity vs. time graph for a ball on a track and asks her to produce a position vs. time graph. Erica's initial attempt to solve the problem involves finding the area underlying the curve, which for many students is a cumbersome procedure. [6] However, the areas she finds or attempts to use are not very useful — she remembers that the way from velocity to position involves "area under the curve" (knowledge), does the automatic representation translation to looking for areas under the curve (comprehension), and relies on these puzzle pieces of mathematics, but fails to apply both of these correctly in the given situation (application), likely because of the "negative area" in the first part of the graph [6]. After graphing her answer using this method, she decides to start over and try a different approach.

In her second attempt, Erica elects to find the equation of the line in the velocity vs. time graph, however she uses the wrong sign for the $y$-intercept. After integrating the equation she obtained, she types the equation into her calculator and obtains an incorrect answer. She seems confused by this answer, but ultimately decides to trust it over the answer she already has and erases her initial answer, replacing it with this new one. Clearly, Erica foregoes any potential benefits from higher order thinking, in this case evaluation, and instead relies on her graphing calculator. She then announces that she is done with the problem stating "I know it's not really that."

In the follow-up interview, the interviewer asks her why she chose her second answer over the first, she claims, "I have no idea." When the interviewer presses further, she claims "the first one is the correct one." Figure 4.11 shows an overview of this process.

Also in the follow-up interview, when the interviewer goes to ask her about the next problem, she exclaims, "Here I was feeling a lot more comfortable." When asked if there was any particular reason, she says, "I don't know, I like those graphs. They're really straightforward." Apparently, Erica felt more comfortable working on the graph interpretation problems than on the graph construction problems.

Having arrived at the interpretation problems, Erica seems much more at home. Working on PI1, she quickly notes how to find the answer to all three parts from using memorized relationships.

For the last problem, Erica goes straight to the book "to double check that what I have in my head is right." It isn't, but she ultimately finds the correct equation for the force of

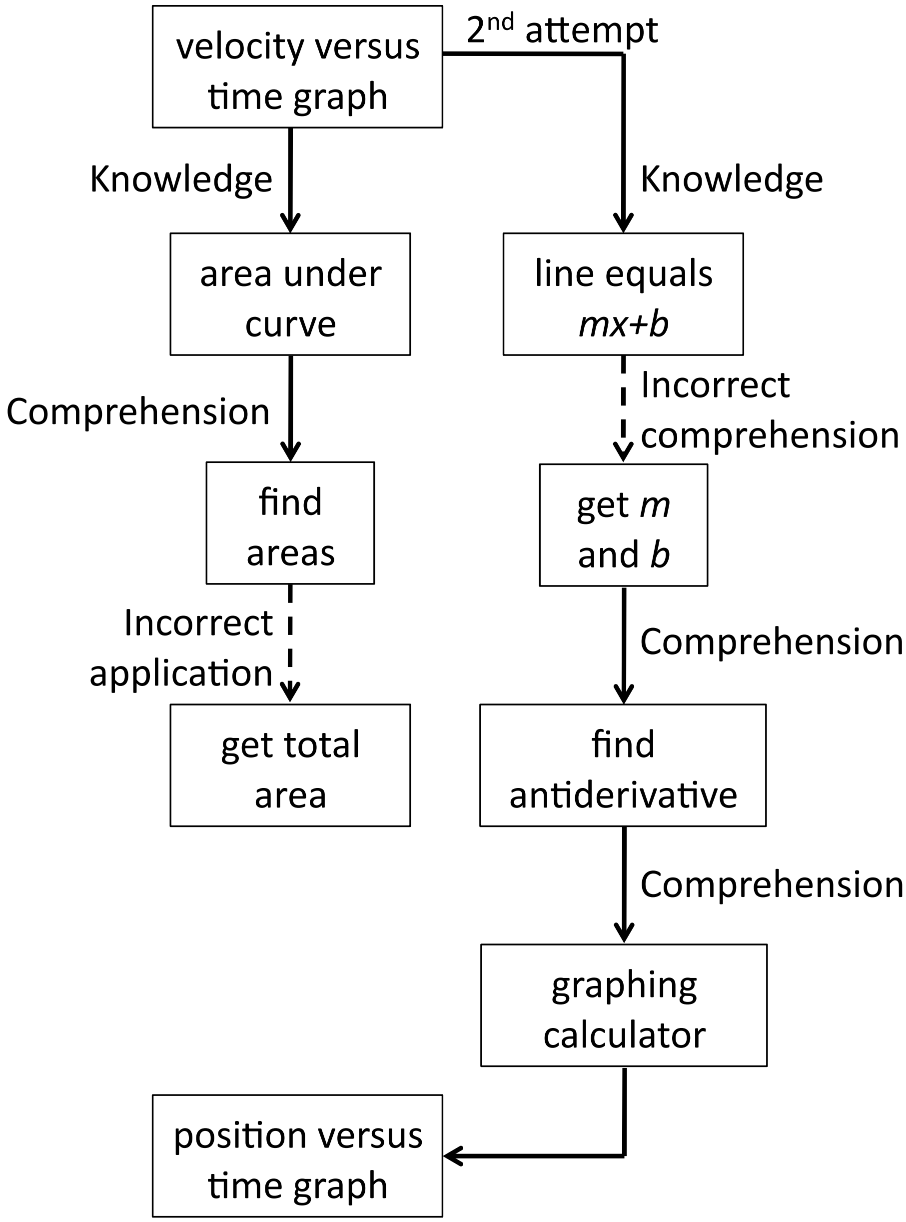



Figure 4.11: Erica's solution of PC2. She made two independent attempts at solving this problem, ignoring any form of higher order thinking.

a spring. On her paper, she rearranges the equation to obtain $k = -F/x$. For reasons that are unclear, she then changes the equation to $k = -\Delta F/\Delta x$ and notes that this is just the slope of the line. When asked in the follow-up interview, she indicates that she added the $\Delta$'s because it's a graph, so $F$ and $x$ are not constants.

While Erica got the incorrect answer for most of the problems, many of her errors were simply computational mistakes (missed a negative sign in PC2, or mis-labeled a point on PI2.) Given that she was obviously nervous about being video recorded, she very well could have solved most (all?) of the graph problems using no higher order thinking whatsoever. This clearly is cause for concern if we are trying to teach our students to use graphs to learn about the physical world.

Beyond Erica's personification of solving graph problems using as many formulas (and as little else) as possible, other subjects used similar arguments during parts of their interviews.

> "So velocity's basically change in $x$ over change in time (writes down $\Delta x/\Delta t$.) So where does it have the biggest change in x in the smallest amount of time." - Calvin (EI2)

> "I actually think this is gonna be a straight *line* based on the equation I'm using." - Abbie (EC2)

#### 4.2.3.4 Algorithmic Solving

The interview of Jodie was interesting primarily because of how poorly she was able to do with three of the problems, while adeptly and rapidly solving the other two. We refer to the method of solving graph problems she employed as "algorithmic" (or "programmatic"). Jodie's attempts to solve a given problem were analogous to finding an applicable computer program which could be run to find the solution. If the applicable subroutine exists to solve a given problem, it will run; But if no subroutine exists, the attempt to solve the problem fails almost outright.

For the first problem (EL), Jodie readily constructs the line and submits her answer. The server responds that she is incorrect, so she graphs the equation in her calculator. After seeing the calculator's graph, she moves her line up slightly and resubmits to get the right answer. This is the epitome of operation at the application level.

For her second problem (EC1), Jodie began by noting where the graph should start and that it should be a line. She then made and submitted to the server a series of graphs that were not linear and used wild guesses for the location of the end of the graph. She ultimately gave up on the problem without having checked the book, written anything down, or used her calculator.

Problem three (EC2) also left Jodie very confused. While working on the problem, she notes that "We never really did, a problem like this, um, where you have to look at the graph." In the follow-up interview, she states that she expected the answer to be sinusoidal because those are the graphs that she had been doing in class. She also explicitly states that she has never seen an example problem like this one.

Jodie attempts to solve the fourth problem (EI1) by looking at the "obvious points." It is not entirely clear what made certain points obvious, except that in at least one case, it was the global maximum of the graph.

Finally, after stumbling through the last three problems, Jodie arrives at the fifth problem (EI2). After reading each part, she immediately identifies what she needs to look at on the graph. For the highest speed, she wants the slope of the tangent line where it's the steepest. For the largest negative velocity, she wants to find where "the slope of the tangent line is decreasing at the steepest." For the largest acceleration, she wants the concavity. And finally, for the point farthest from the origin, she wants the point that is farthest from the starting point. While this last one is technically incorrect, the point farthest from the starting point is also farthest from the origin. This problem ultimately caused her no trouble and she was able to complete it quickly.

In summary, Jodie had no difficulty solving the first and last problems she was given, but struggled intensely on the other three, citing a lack of experience with problems of that type multiple times. Perhaps more concerning, is that she made no attempt to learn anything from the problems. Once she decided that she did not know how to solve a problem, she gave up instead of trying to work through it.

While Jodie is a somewhat extreme example, searching for similar or analogous problems is very common — and while it is low on Bloom's taxonomy, it is not necessarily non-expertlike, as experts often work from a wide base of similar problems and situations to solve a new problem. But the method can backfire if no similar problems are found, or wrong similarities employed. For instance, when Cindy attempts to solve problem PI2, she immediately knows what to do and begins writing down and solving equations. Unfortunately, her equations and attempts are actually solving the problem as if it asked for a $y$-position vs. $t$ graph.

### 4.2.4 Higher Order Thinking

Any attempt to solve a problem in this study that does not fall into one of the above categories could be categorized as higher order thinking. As such, the higher order strategies the subjects employed to solve their graph problems were much more diverse than the lower order ones. Instead of trying to categorize them all, we will instead focus on the strategies that occurred repeatedly.

- Process of Elimination - Eliminating options on the multiple choice question.
- Interpreting the Physical Situation - Identifying what is happening in the real world.
- Error Checking - Evaluation of their answer.

#### 4.2.4.1 Process of Elimination

In one of the problems assigned to the subjects (EI1), an acceleration vs. time graph is given to the subjects and they must choose the correct velocity vs. time graph from the four options given.

Of the four subjects to try this problem, three of them got it correct. In each of these instances, the subject solved the problem in much the same way that you would expect them to solve any multiple-choice problem: process of elimination. While this is not too surprising, what is interesting is that all three subjects eliminated the three incorrect options using one (and only one) criterion and that the criterion was the same for all three of them. Specifically, the subjects all noted that the velocity should be decreasing from the beginning to about about the $t = 8$ second mark.

The other subject (Jodie) attempted to solve this problem by matching the "obvious" points from the given graph. For all of the other electronic problems, the subject was given 12 tries to get the problems correct, but for this problem they were only given two. Jodie did not realize this until after she had already (rapidly) used both tries.

Even though it is not the behavior desired by the instructor, who may have hoped for a more complete understanding of and process for solving the problem, the multiple choice triggered a well-studied higher order thinking process in the subjects, who grew up with multiple choice problems and have been trained in this situational mode of thinking. It is important to note, however, that this higher order thinking is with respect to answering homework problems, not to the physics or the graphs. On the other hand, it *is* an expertlike skill to zoom in on salient features and neglect spurious information.

#### 4.2.4.2 Interpreting the Physical Situation

During the course of the interviews, the subjects often did not mention what was happening in the physical situation that the problems were describing. Given that each graph used in this study is meant to represent a real situation, it is somewhat disappointing that so few attempts were made to think about those situations. Somewhat surprising and disturbing is the fact that when the subjects did attempt to think about the physical situation, they either very often were wrong, ignored it, or both.

For instance, both Andrew and Gideon attempted to determine the shape of the track the ball was following in problem PC2.

> "Well it seems here that since the, it starts down here that the velocity is... negative and then increases to positive. But, I'm thinking of how to draw a negative velocity, so it probably means it's going *down*. So it's like falling and then maybe like a, it's like a loop track or something. (Draws a 'U' shape in the air)" - Andrew (PC2)

> "At six seconds it's going to be at it's highest 'cause it was going faster than it was initially, so the track's gonna, slant down at that point. So I guess the track is going to be. . . probably pretty close to the line that the velocity versus time, 'cause it looks like it's at negative and then increasing." - Gideon (PC2)

While Gideon's description may be right (it's not entirely clear if he meant the track is just a ramp or something else), both of them ultimately backed away from this strategy and tried to solve the problem by another method. When asked about it during the follow-up interview, Andrew implies that he thought it would be 'U' shaped because a negative velocity implies that it's going down, while a positive velocity would be going up. Ultimately, he states, "I don't know, it just, it didn't really seem to make sense [. . . ] so I just decided to just forget about that." During Gideon's follow-up interview, the interviewer asks Gideon directly what he thinks the track looks like. Gideon responds with several answers, ranging from 'U' shaped to the shape of his position vs. time graph, which looks like a ramp (a known common misconception).[6]

But those aren't the only instances where the subjects consider the physical situation. While reading problem EI1, Calvin wonders, "What type of track?" He then proceeds to describe the motion based on the acceleration graph, describing the object's motion as "slowing down from whatever its speed was," "it goes a little bit in the opposite direction more," and "finally, it has positive acceleration so it's going away from the axis." Ultimately, he ends up solving the problem by process of elimination in the manner discussed earlier, ignoring his physical interpretation.

Additionally, before Gideon reads any of the problem parts below the graph in PI1, he accurately describes the physical motion represented by the graph.

> "So, the car is accelerating and then as it's leveling off it's holding speed. Then it looks like it decelerates back down to zero, stops for a period of time, and then accelerates again and continues on at that speed." - Gideon (PI1)

Even though he has correctly identified the motion, he does not return to thinking about the physical motion for the remainder of the problem.

Both Calvin and Abbie determine the sign of the Force in problem EC2 by thinking about the physical situation, getting it correct and incorrect, respectively. Calvin's use of the physical situation here is the only case where interpreting the physical situation helped solve the problem.

The fact that in our study, trying to link the graph with a physical situation appears to be associated with failure to correctly solve the problem may be one of the reasons that students might get discouraged from using this higher order thinking skill. A similar effect has been found by Wemyss and van Kampen [68]: students were much more successful answering otherwise identical graph problems when they were not given in a physics context, suggesting that "the school experience with the distance-time graphs negatively impacted on the students' problem solving strategy and their likeliness to engage with the question." [68]

#### 4.2.4.3 Error Checking

There are many instances where the subjects attempted to evaluate their answer, or their steps along the way to answering the problems. Generally, this is considered an important metacognitive skill. However, most of these instances are trivial, in the sense that the subject repeated the same steps to see if they obtained the same answer (recheck).

For instance, Isaac double checked his answers for three of the problems during the course of the interview. For the graph a line problem (EL), he mentally calculated two points to draw a line. After setting up his answer, but before submitting it, he validated his answer by calculating a third point to make sure that it was on the line. For EI1, he used the criteria listed earlier in the paper to select his answer (velocity is decreasing before $t = 8$ sec). Before submitting it, though, he started the problem over and, using the same criteria, made sure that he arrived at the same answer. For EC2, upon noticing that one of the values he calculated was not on the graph, Isaac typed the values into his calculator again, receiving the same answer.

Similar to Isaac, Abbie recalculated values during her interview, as well as questioning negative signs she obtained on both problems EC1 and EC2.

There was only one instance of a non-trivial evaluation of an answer by one of the subjects (crosscheck). While working on problem PC2, Cindy notes the symmetry in the problem to identify when the object will cross the x-axis again, and uses the area under the curve to determine the value of the minimum. Having drawn her (correct) answer, she notes that "it makes sense since acceleration is always positive slope so it's gonna concave up which is what it's doing the entire time."

As discussed earlier, Erica made two separate attempts to solve problem PC2, as shown in Fig. 4.11. However, she never attempted to reconcile her two distinct answers, foregoing the opportunity for higher order thinking.

### 4.2.5 Graph Construction Strategy

In our study, we were particularly interested in the question whether or not graph construction problems more frequently lead to higher order thinking processes than interpretation problems. We could not find evidence for this hypothesis, most likely due to another pattern that emerged. Specifically, the subjects (seemingly subconsciously) divided solving the construction problems into two concerns.

1. Identifying the shape of the graph (line, parabola, etc.)
2. Identifying the locations of points needed to draw that shape (two for a line, three for a parabola, etc.)

Without exception, the subjects answered every graph construction problem in this manner. Not always, but most of the time, they tackled these two concerns in the order listed. The methods used by the subjects varied as described earlier in this section, but using this two-step process, they were able to circumvent many higher order processes. In fact, upon identifying the shape of the graph, the subjects were able to treat the problem as though the graph did not exist, instead working through it in a manner similar to solving any other quantitative physics problem.

> "I just don't know what the shape of the graph is gonna look like so that could be a problem." - Abbie (EC2)

> "So the velocity's gonna decrease constantly until it's at a certain time, so it's gonna be linear, but then we have to solve for the time." - Calvin (EC1)

This particular strategy is confirming earlier findings that students tend to avoid the complications of graphing if at all possible. [13, 12].

## 4.3 Discussion and Outlook

In our study, we identified some general patterns of solving graph problems which were associated with lower and higher order thinking processes. There were three areas that the authors would not have been surprised to see a difference in these problem solving strategies, namely gender, paper vs. electronic problems, and order of problems (e.g., construction before interpretation). The transcripts were analyzed looking for uncommon themes between each of these groups and none were found. Further exploration may find differences between these groups, but there was not sufficient evidence in these interviews.

The problems used in this study were rather difficult and involved physics topics (e.g., kinematics) along the way to be solved. The problems were also frequently numerical/calculational in nature. These characteristics likely triggered the same non-expertlike problem solving strategies that students use for other type problems, and thus kept them from considering alternative, more expertlike approaches — they never left their comfort zone. This finding is similar to what Wemyss and van Kampen found when confronting their students with a kinematics graph problem that asked for numerical values: "when asked to determine a numerical value for the speed, the wording appears to have cued many students to apply a formula, and they ceased to engage with the graphical representation or the question they had just answered. It appears that in the context of this question, the graph no longer represented information about the nature of the motion, but just served to provide data necessary to find an answer through an equation." [68]

The graph problems used in this study were intentionally designed to be "textbook-like." As a result, it is important to note that all of the graph construction problems had a unique, polynomial answer. It is not immediately obvious what the subjects would have done with a graph construction problem that could not be solved using the "shape-points" strategy. For instance, Fig. 4.12 shows a graph construction problem that has an infinite number of correct answers. While a subject might be able to identify that both end-points must be on the $x$-axis, it is unclear how they might deal with the portion of the graph in between the two endpoints, since there are no points that can be solved for and while a polynomial could be used for the shape, it may not be obvious which polynomial to use.

One possibility, as suggested by Gideon's attempt at problem PC1, is trying to copy an example from the book. Ultimately, Gideon looked up a problem that was similar to what was in front of him, and tried to mimic that graph as his answer. While he got the general shape and idea correct, he lacked having the correct maximum height and range. Effectively, he was able to answer the open ended question of "draw a trajectory", but not able to solve

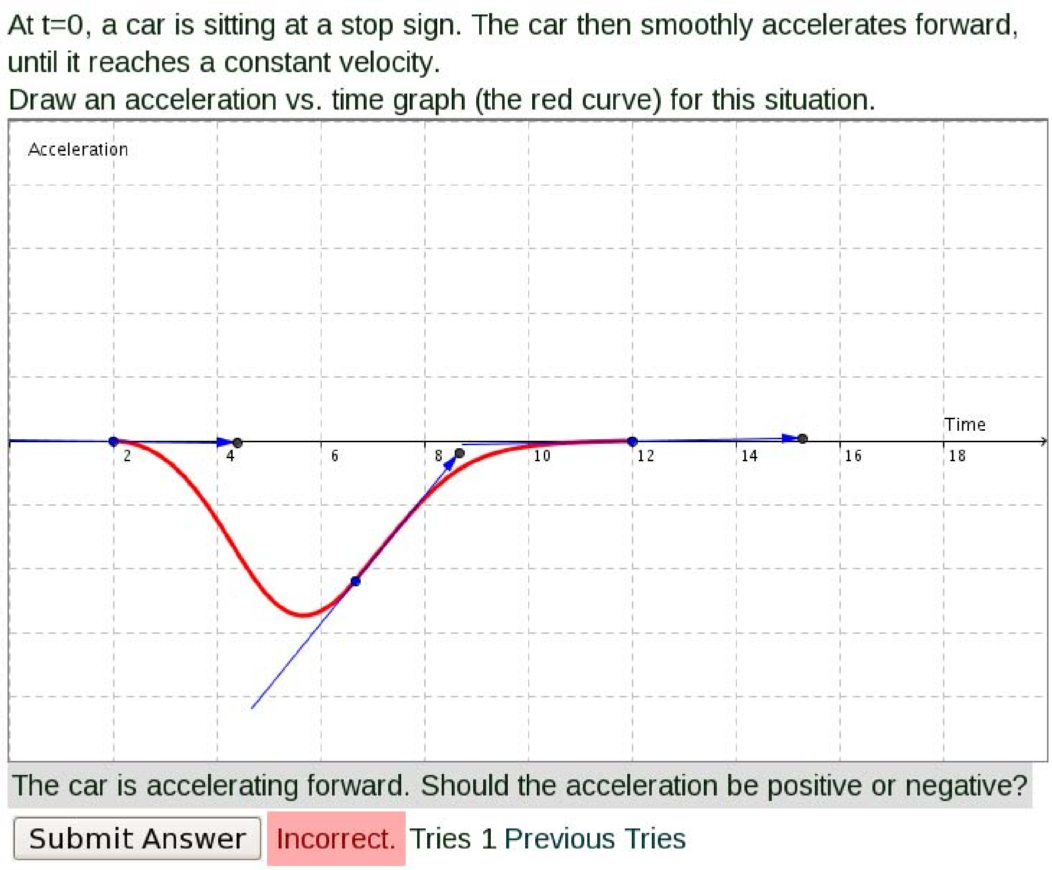



Figure 4.12: An open-ended problem. The text along the axes in this Figure is not meant to be readable but is for visual reference only.

a specific trajectory. Accordingly, writing problems that are not corrolaries from the book would prevent this.

The think-aloud protocol is somewhat similar to "listening in" on student discussions. Many of the "textbook-like" problems were rather difficult, involving multiple steps and possibly calculations. In an earlier analysis of online homework discussion, we had found that problems with a high degree of difficulty lead to procedural, non-expertlike discussion, while a "sweet-spot" for fostering conceptual discussions are medium difficulty problems: not too easy, so discussions do not become trivial, but also not too hard, so students do not feel that they can understand the problem on a conceptual level and do not immediately resort to apparently safe procedures. [58, 69] We may see a parallel in this study.

Also in graph problems, it is not enough to present students with scenarios where experts would use expertlike strategies, but not to leave the door open for employing the strategies outlined in Subsections 4.2.3 and 4.2.5. Problems need to be designed such that they "focus on making adoption of expertlike views truly useful and relevant for students." [28] Future studies should thus involve non-calculational, open-ended problems, along the lines of giving an example of a graph for acceleration versus time for a particular textually described situation (as in Fig. 4.12, for example). They may also need to be truly "comprehensive" [15] in order to require higher order thinking processes.

# Chapter 5

# Analyzing Graph Problems in LON-CAPA

Recently, there has been substantial growth in the number of massively open, online courses (MOOCs) offered by such websites as Coursera, Udacity, and EdX. These types of courses are generally seen as a way for researchers to obtain large amounts of data for studying student learning, but there are already systems with large amounts of data that we can be studying. Such studies can inform decisions on the design of MOOCs and course management systems in general, as well as what type of data should be collected for researchers.

For instance, students have more and better engagement on discussion boards for formative assessment problems with certain characteristics. Specifically, problems with a medium degree of difficulty around 0.5 correlated with discussions more focused on the physics involved in the problem and less emotional responses or entries purely focused on the solution. In turn, participants in these types of discussions were more likely to succeed in the class than their counterparts[58]. Descriptive statistics are, by their nature, computed after the fact. Is it possible that these characteristics can be predicted? One of the goals of this study is to investigate the possibility that we may be able to select or design problems to have a certain degree of difficulty to encourage these discussions. This will allow instructors to select or design better problems for their courses based on recommendations by the system, instead of the usual “guess and check” method, which is the standard now.

In fact, this idea has been investigated in the context of summative assessments. Mesic and Muratovic used a linear regression approach to predict the item difficulty of physics problems to improve the test design process[70]. Ultimately, they were able to explain 61% of the variance of item difficulty with the factors they included in their model. These factors generally reflected the “automaticity, complexity, and modality of the knowledge structure” needed to solve the problems. Another desirable characteristic of a problem is a high degree of discrimination, i.e. how well the problem distinguishes between students who understand the broader underlying concepts and those who do not.

One of the most cited papers in the history of Physics Education Research is the work done by Chi on the categorization of physics problems by experts and novices[71]. Although the results of this paper have been difficult to replicate[72, 73], the general idea that novices categorize physics sproblems by surface features while experts focus on deep structure is still useful and suggested by other works[74]. In this study, the questions used have been characterized by features which may be considered surface features. Since the questions in LON-CAPA are generally directed at novices, this study will also be investigating the possibility that the difficulty of questions can be explained by surface features, instead of deep structure or the steps required to solve them, as other studies have looked at[75]. In the realm of online problems, these surface features could be gathered by automated harvesting processes, which would make such mechanisms attractive for large scale learning content management. This would be particularly useful when it comes to open educational resources (OERs) where traditional deep structure metadata is incomplete or missing[76].

## 5.1 The Data

This study will focus on a subset of problems inside the LON-CAPA[52] course management system that include kinematic graphs. Each data point corresponds to the use of a problem part (referred to as "question", starting now) in a particular class. Questions used in the context of an exam were removed from the data sets, as well as questions that were attempted by less than 20 students.

The primary data set comes from classes from various institutions across the LON-CAPA network, compiled from over 138 million homework transactions. The information gathered is at the class level (not individuals) and was deidentified before collection. The institutions span a range from high schools, to two-year and four-year universities. There are 572 data points in this set, using 32 distinct graph questions. Due to the nature of the harvesting process, problems with multiple parts could not be included in this data set.

The secondary data set comes from data logs taken from thirty introductory Physics classes at Michigan State University. These classes come from five different courses, across multiple years and semesters. In total, these thirty classes used a total of 56 distinct graph questions, most of which are used in multiple classes and multiple semesters, resulting in a total of 342 data points.

The two data sets share no data points in common, but they do share 16 of the same questions.

While these two data sets could be combined into one, it is common practice to divide data sets in order to verify the results of the model. Since these two sets were obtained by different methods, it was convenient to use them separately in our analysis.

### 5.1.1 Dependent Variables

For each of the data points in both studies, we will consider two outcome variables: degree of difficulty and degree of discrimination. These variables are given per question per class, and no information about individual student's responses are known.

#### 5.1.1.1 Degree of Difficulty

The Degree of Difficulty (DoDiff) in LON-CAPA is a Classical Test Theory inspired measure. The value is given by the equation

$$\text{DoDiff} = 1 - \frac{\text{Number of Correct per Student}}{\text{Number of Tries per Student}} \tag{5.1}$$

It is worth noting that if LON-CAPA did not allow students multiple tries, this would be the same value as Classical Test Theory and would map monotonically to the Degree of Difficulty used in Item Response Theory. However, since students have multiple attempts to solve most questions, the exact meaning of DoDiff in LON-CAPA becomes cloudier.

For instance, if a question is given in which 60% of the students get the answer correct on the first (and only) try, this would result in a DoDiff of 0.4. However, if multiple tries are allowed, and on the second attempt 60% of the remaining students got the question correct, this would also result in a DoDiff of 0.4. This same logic can be repeated for all subsequent attempts, still arriving at a DoDiff of 0.4. Thus, it is not straightforward to interpret what the DoDiff means, as the same value can represent multiple situations which are not the same. However, this does not render it meaningless, particularly since it allows comparisons with previous literature.

#### 5.1.1.2 Degree of Discrimination

The Degree of Discrimination (DoDisc) is also inspired by Classical Test Theory. The organization of most courses in LON-CAPA follows the same basic model as most textbooks: a system of folders that can contain files and other folders, similar to chapters and assignements. Often, each assignment in a semester is given its own folder. To calculate the DoDisc, the folder that a question resides in is treated like a test. The students are divided into quartiles based on how well they did on all of the questions in the folder. Then the top and bottom quartiles are compared for the individual question. The equation used is

$$\begin{aligned} \text{DoDisc} &= \frac{\#\text{ correct by top 25\% on set}}{\#\text{ students in top 25\%}} - \frac{\#\text{ correct by bottom 25\% on set}}{\#\text{ students in bottom 25\%}} \\ &\approx 4 \cdot \frac{(\#\text{ correct by top 25\% on set}) - (\#\text{ correct by bottom 25\% on set})}{(\#\text{ students working on set})} \end{aligned} \tag{5.2}$$

Since most students eventually get most questions correct in their homework (as opposed to exams), often the difference between members of the top and bottom quartile is one or two correct answers. This may result in a fair amount of volatility in the measured DoDisc from class to class.

### 5.1.2 Independent Variables

The independent variables originally proposed in this study (and found in Table 5.1) were specifically chosen to be surface features for three reasons. First, research has suggested that students distinguish between physics problems based on surface features, whereas experts distinguish them by "deep structure". The second reason is that these features should be much more readily agreed upon by raters, without having to create an elaborate rubric for categorization. We will see in a later section that the inter-rater reliability is quite high considering there were very few elaborate instructions on what the variables meant. This would allow for the distribution of such assessments to be distributed across a network of people, instead of relying on a select few individuals who have been trained. The third and most important reason behind the independent variables used here is the possibility that human input might not even be needed to categorize such problems. Variables such as "Question Type" and "Graph Type" can be identified by the homework system itself. Furthermore, items such as "Area" might also be identifiable by computers using lexical analysis software[77].

### 5.1.3 Inter-rater Reliability

A total of three people analyzed eight randomly chosen questions to categorize on the nineteen characteristics listed in Table 5.1. The information about the categories given to the two reviewers other than the author was similar to the descriptions listed in the Table.

Of the 162 values, the raters disagreed on a total of ten of them and the resulting Fleiss' Kappa value was $\kappa = .91$, which is generally categorized as high agreement.

### 5.1.4 Exploratory Factor Analysis

As certain question characteristics may be strongly associated with one another, an exploratory factor analysis (EFA) was run to search for characteristics that often or almost never occurred together. The results of the EFA suggested that seven of the original characteristics should be reduced to just two to avoid collinearity concerns.

Table 5.1: The graph questions used in this study were initially characterized by 19 different elements. With the exception of "Items", each characteristic is a dichotomous variable where a "1" indicates that the characteristic is included in the question.

| Category | Characteristic | Description |
| --- | --- | --- |
| Graph Usage | Graph in Question | Is a kinematic graph given in the statement of the question? |
| | Make a Graph | Does the student need to make a graph in order to solve the problem? |
| | Graph as Answer | Is the answer to the question a kinematic graph? |
| Question Type | Multiple Choice | Several options are given for answers. |
| | Drop Down Boxes | Several Statements are given. Students must choose the correct answer for each statement from a list. |
| | Graph Construction | The student must submit a graph that they created. |
| | Numerical | Students must submit a numerical answer that they calculated. |
| Graph Types | Position | The question or answer involves a Position vs. Time graph. |
| | Velocity | The question or answer involves a Velocity vs. Time graph. |
| | Acceleration | The question or answer involves an Acceleration vs. Time graph. |
| Procedures | Area | Identifying or Calculating the area under the graph is needed to solve the question. |
| | Point | Identifying or Calculating the values of a point on the graph is needed to solve the question. |
| | Slope | Identifying or Calculating the slope is needed to solve the question. |
| | Concavity | Identifying or Calculating the concavity is necessary to solve the question. |
| Misc | Smooth | Is the graph differentiable everywhere? |
| | Calculations | Does the student need to calculate specific values? |
| | Multi-part | Is the question part of a multi-part problem? |
| | Items | How many separate answers are necessary to get correct for this question. |
| | Options | Are there an infinite number of options for the correct answer? |

Table 5.2: An exploratory factor analysis suggested reducing seven of the original characteristics into two, due to significant correlations between them. The newly created characteristics are the average of the values from the original characteristics, after swapping the values for “Graph in Question” (ie. now 1 means no and 0 means yes.)

| New Characteristic | Original Characteristics |
|---|---|
| Construction | Graph Construction<br>Graph in Question<br>Make a Graph<br>Graph as Answer |
| Numerical Response | Numerical<br>Options<br>Calculations |

After the coding on one of the original characteristics (“Graph in Question”) was reversed (zeros becames ones and vice-versa), the two resulting characteristics were defined to be the average of the contributing original characteristics. This change in characteristics is summarized in Table 5.2.

The resulting list of fourteen categories was used for the analyses found in the rest of this study. You can find them in the left most column of Table 5.4.

### 5.1.5 Weighting

Certain questions show up significantly more often than others in these data sets. In order not to give undue preference to questions that appeared often, the data points were weighted in such a way that all of the individual questions were on equal footing. Also, since classes with more students should have more reliable values for the DoDiff and DoDisc, the data points were weighted accordingly.

$$\text{Weight} = \frac{\text{Number of Students in this Class Who Solved this Problem}}{\text{Number of Students in all Classes Who Solved this Problem}} \tag{5.3}$$

In the simplest terms, the numerator weighs within questions while the denominator weighs across questions.

## 5.2 Multiple Linear Regression Analysis

Using SPSS 19, two separate multiple linear regression analyses were run for each data set (one for DoDiff as the dependent variable, and one for DoDisc), using the question characteristics as the predictor variables.

A graph question without any characteristics is trivial. Such a question would effectively correspond to one in which clicking the submission button would automatically be a right answer. Thus, such a question should have a DoDiff and DoDisc of 0. For this reason, these regression analyses were run without an additive constant (ie, the fit is forced to go through the origin).

### 5.2.1 Primary Data Set

Surprisingly, despite the low level of sophistication of the characteristics, the models fit especially well for the primary data set, as shown in Tables 5.3.

Table 5.4 shows the resulting coefficients for the two models for the primary data set. The results suggest that the type of question is the most influential in determining both the DoDiff and DoDisc, with numerical response and graph construction problems having the two largest contributions to both DoDiff and DoDisc. Beyond that, the information the students must read off the graph and the graph type have the next largest contribution to problem difficulty, all having similar values for their coefficients. The data also suggests that including a position vs. time, velocity vs. time, and/or acceleration vs. time graph reduces

Table 5.3: Model Summaries for the Primary Data Set

| | $R$ | $R^2$ | Adjusted $R^2$ | Std. Error of the Estimate |
|---|---|---|---|---|
| DoDiff | .964 | .929 | .927 | .153 |
| DoDisc | .922 | .850 | .847 | .222 |

Table 5.4: Model Results for the Primary Data Set. None of the questions in the primary data set were part of multi-part problems, thus the coefficients could not be calculated for it.

| | Degree of Difficulty | | | Degree of Discrimination | | |
|---|---|---|---|---|---|---|
| Characteristic | Coeff | Std Err | $p$-value | Coeff | Std Err | $p$-value |
| Construction | .528 | .078 | *** | .983 | .113 | *** |
| Numerical Response | .565 | .026 | *** | .602 | .038 | *** |
| Multiple Choice | .070 | .036 | .054 | .562 | .053 | *** |
| Drop Down Boxes | .283 | .035 | *** | .144 | .051 | .005 |
| Position | -.093 | .023 | *** | -.122 | .033 | *** |
| Velocity | -.243 | .026 | *** | -.286 | .038 | *** |
| Acceleration | -.202 | .026 | *** | -.081 | .038 | .032 |
| Area | .272 | .027 | *** | .057 | .039 | .142 |
| Point | .116 | .020 | *** | -.130 | .029 | *** |
| Slope | .184 | .026 | *** | -.215 | .038 | *** |
| Concavity | .295 | .032 | *** | .106 | .046 | .023 |
| Smooth | -.082 | .019 | *** | -.020 | .028 | .483 |
| Multipart | N/A | N/A | N/A | N/A | N/A | N/A |
| Items | .028 | .010 | .004 | .131 | .014 | *** |

the DoDiff. We suspect this is because these are the types of graphs students are frequently exposed to, as opposed to trajectories, or, say, a position vs. velocity graph.

### 5.2.2 Secondary Data Set

The DoDiff model fit for the secondary data set is similar to the primary data set, while the model fit for the DoDisc is notably lower. However, both are still a reasonably good fit for the data. Table 5.5 shows these model fit results.

When we look at the values of the coefficients from these two models, shown in Table 5.6,

Table 5.5: Model Summaries for the Secondary Data Set

| | $R$ | $R^2$ | Adjusted $R^2$ | Std. Error of the Estimate |
|---|---|---|---|---|
| DoDiff | .965 | .931 | .929 | .065 |
| DoDisc | .825 | .680 | .666 | .050 |

Table 5.6: Model Results for the Secondary Data Set

| | Degree of Difficulty | | | Degree of Discrimination | | |
|---|---|---|---|---|---|---|
| Characteristic | Coeff | Std Err | $p$-value | Coeff | Std Err | $p$-value |
| Construction | .470 | .044 | *** | .154 | .033 | *** |
| Numerical Response | .356 | .048 | *** | -.009 | .037 | .800 |
| Multiple Choice | .055 | .051 | .277 | -.108 | .039 | .006 |
| Drop Down Boxes | .238 | .102 | .020 | -.090 | .078 | .247 |
| Position | .039 | .039 | .325 | .025 | .030 | .403 |
| Velocity | -.092 | .039 | .017 | -.021 | .029 | .470 |
| Acceleration | .022 | .032 | .497 | -.006 | .024 | .803 |
| Area | .202 | .031 | *** | .104 | .023 | *** |
| Point | -.030 | .029 | .301 | -.040 | .022 | .071 |
| Slope | .084 | .029 | .004 | .049 | .022 | .030 |
| Concavity | .149 | .045 | .001 | .010 | .034 | .764 |
| Smooth | .013 | .023 | .560 | -.033 | .018 | .060 |
| Multipart | .020 | .021 | .336 | .097 | .016 | *** |
| Items | .041 | .021 | .057 | .040 | .016 | .013 |

we discover that while the values do not agree well with the results of the primary data set, they do generally show the same trend. That is, the question type has the largest influence, while the rest of the characteristics tend to be lower and around the same level. Table 5.7 lists the characteristics by importance in each of the models.

Comparing the results from the two data sets suggests that there are likely to be population differences that contribute to the differences in the models. After all, the primary data set includes students from a much broader range of abilities than the secondary data set, which is more homogeneous by comparison.

Additionally, it is somewhat naïve to suggest that the DoDiff and the DoDisc are inde-

Table 5.7: List of independent variables in order of decreasing absolute value of the coefficient for the two data sets. Variables with $p > .05$ or where the absolute value of the coefficient is $< .1$ are ignored.

| Degree of Difficulty | | Degree of Discrimination | |
|---|---|---|---|
| Primary Data Set | Secondary Data Set | Primary Data Set | Secondary Data Set |
| Numerical Response | Construction | Construction | Construction |
| Construction | Numerical Response | Numerical Response | Multiple Choice |
| Concavity | Drop Down Boxes | Multiple Choice | Area |
| Drop Down Boxes | Area | Velocity | |
| Area | Concavity | Slope | |
| Velocity | | Drop Down Boxes | |
| Acceleration | | Items | |
| Slope | | Point | |
| Point | | Position | |
| | | Concavity | |

pendent of the students/courses that the problem is used in. Taking this argument to an extreme case, the DoDiff for any of the questions used in this study would likely be very close to 1 if it were assigned to an elementary school classroom, while it would hopefully be much closer to 0 if given to a class of physics instructors.

## 5.3 Structural Equation Model

The data in the primary data set comes from a significantly more diverse set of courses. Due to the deidentified nature of the data, we have no background information about the courses involved. However, the secondary data set comes from only five distinct courses. Given the likelihood of effects from the courses involved, we used structural equation modeling to try to determine the effect of 'Course' on the DoDiff and DoDisc using the secondary data set. This also allowed us the opportunity to investigate another element of the system: whether or not 'Course' has a significant effect on what kinds of questions are assigned.

The model was constructed using the AMOS add-on to SPSS19 and its overall structure is shown in Fig. 5.1. Given the large number of paths in the model, we will only bother to report the values that were significant at the $p < .05$ level. These results can be found in Tables 5.8, 5.9, & 5.10, looking at the effect the courses have on the DoDiff and DoDisc, the effect of the characteristics on the DoDiff and DoDisc, and the course effects on choosing questions with certain characteristics, respectively. These results are summarized in Fig. 5.2.

The first thing to note is the lack of considerable effects on the DoDiff and DoDisc due to the different courses. While these effects should, in principle, exist, it seems that the difference in students from one course to another in the secondary data set does not have much of an effect on these variables. This suggests that at least for courses within the same institution, there is not much variation in ability to solve graph problems between the different courses' populations, regardless of whether the class uses calculus, or is taught traditionally or in a more reformed manner.

The results in Table 5.9 should reflect the same general results as the multiple linear regression analysis, and they do. Because the structural equation model has many more parameters, we lose some of our statistical power, but we still have graph construction problems leading the way in their contributions to DoDiff and DoDisc with the other variables trailing behind around the same, lower value.

A fairly interesting result is reflected in the density of significant results from this model. A majority of the significant results were returned in connection with the course's influence on question characteristic selection. Presumably, this is a result of instructor choice, past or present, as there were many instances of question sets being repeated in the same course. In other words, the strong dependence of question characteristics on 'Course' suggests that the selection of question strongly depends on the preferences, priorities, and tastes of the

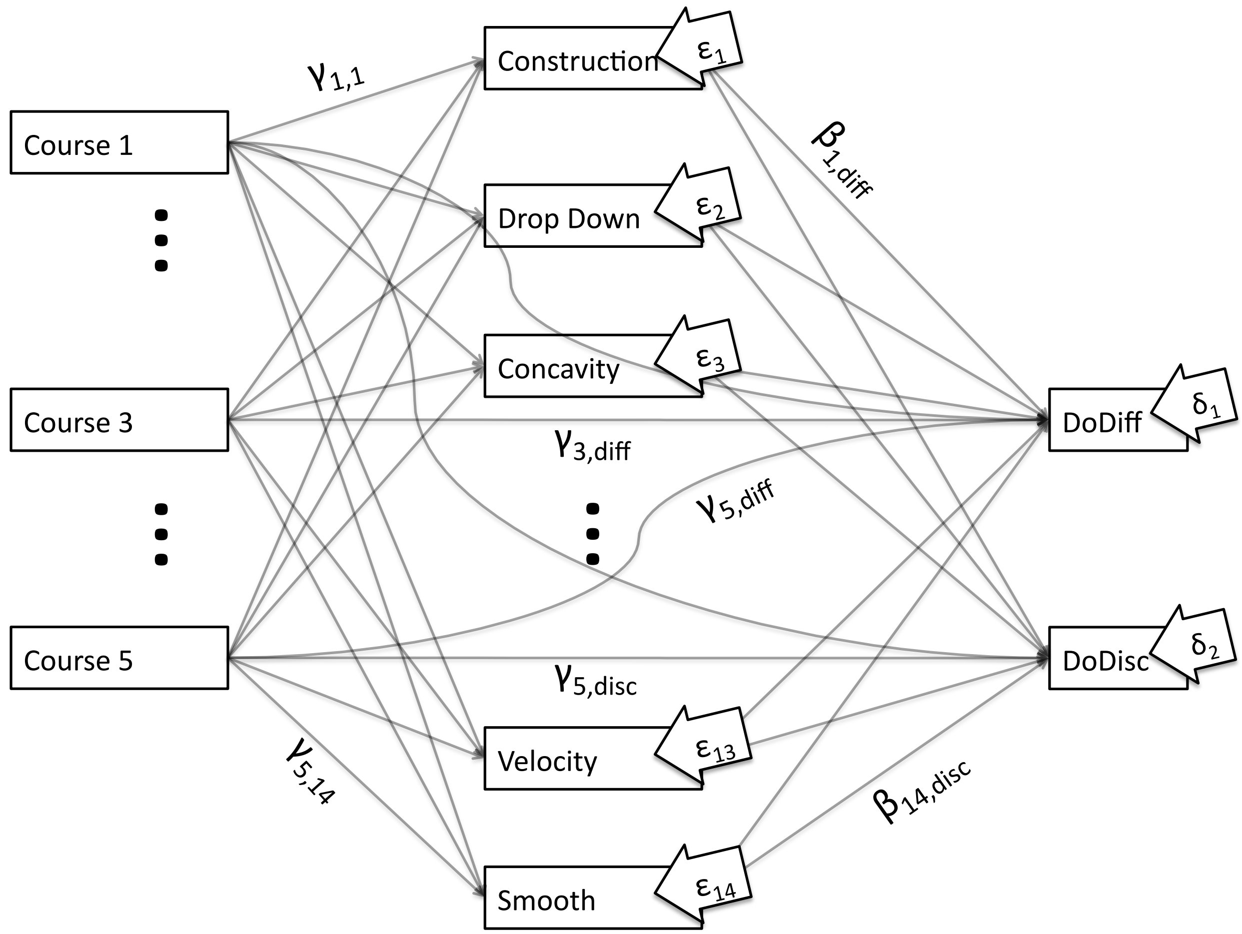



Figure 5.1: The structural equation model used for the secondary data set.

instructors. It is our hope that through pre-facto mechanisms like the one described in this study, as well as post-facto analytics, question choices could become less subjective and increasingly data-driven.

Table 5.8: Influence of Courses on Degree of Difficulty and Degree of Discrimination in the Structural Equation Model. Results are shown only for coefficients whose absolute value is $> .1$ and whose results have $p < .05$.

| Course ($n$) | Variable ($m$) | Coeff ($\gamma_{n,m}$) | Std Err | $p$-value |
|---|---|---|---|---|
| 1 | DoDiff | .111 | .018 | *** |
| 3 | DoDiff | -.110 | .049 | .025 |
| 3 | DoDisc | .253 | .038 | *** |
| 4 | DoDiff | .110 | .020 | *** |
| 5 | DoDiff | -.110 | .049 | .024 |

Table 5.9: Influence of Question Characteristics on Degree of Difficulty and Degree of Discrimination in the Structural Equation Model. Results are shown only for coefficients whose absolute value is $> .1$ and whose results have $p < .05$.

| Characteristic ($n$) | Variable ($m$) | Coeff ($\beta_{n,m}$) | Std Err | $p$-value |
|---|---|---|---|---|
| Construction | DoDiff | .392 | .025 | *** |
| Construction | DoDisc | .157 | .019 | *** |
| Area | DoDiff | .188 | .016 | *** |
| Area | DoDisc | .118 | .012 | *** |
| Concavity | DoDiff | .166 | .022 | *** |
| Numerical Response | DoDiff | .173 | .018 | *** |
| Velocity | DoDiff | -.109 | .016 | *** |

Table 5.10: Influence of Course on Question Characteristics in the Structural Equation Model. Results are shown only for coefficients whose absolute value is > .1 and whose results have $p < .05$.

| Course ($n$) | Characteristic ($m$) | Coeff ($\gamma_{n,m}$) | Std Err | $p$-value |
|---|---|---|---|---|
| 1 | Drop Down Boxes | .101 | .038 | .008 |
| 1 | Construction | .138 | .036 | *** |
| 1 | Concavity | .117 | .039 | .003 |
| 1 | Slope | .184 | .054 | *** |
| 1 | Items | .486 | .178 | .006 |
| 1 | Numerical Response | .110 | .050 | .027 |
| 1 | PositionVs.Time | .212 | .054 | *** |
| 1 | Velocity | .146 | .053 | .006 |
| 2 | MultipleChoice | .202 | .042 | *** |
| 2 | Area | .162 | .060 | .007 |
| 2 | Numerical Response | .153 | .053 | .004 |
| 2 | Multipart | .234 | .059 | *** |
| 2 | Velocity | .235 | .057 | *** |
| 3 | Acceleration | .454 | .151 | .003 |
| 3 | Area | .389 | .163 | .017 |
| 3 | Numerical Response | .332 | .145 | .022 |
| 3 | Multipart | .460 | .160 | .004 |
| 4 | Acceleration | .256 | .052 | *** |
| 4 | Drop Down Boxes | .197 | .039 | *** |
| 4 | Concavity | .144 | .040 | *** |
| 4 | Mustidentifypoints | .162 | .047 | *** |
| 4 | Slope | .122 | .054 | .025 |
| 4 | Items | .880 | .178 | *** |
| 4 | PositionVs.Time | .139 | .055 | .011 |
| 4 | Smooth | .399 | .051 | *** |
| 5 | Area | .453 | .159 | .004 |
| 5 | Numerical Response | .31 | .141 | .028 |
| 5 | Smooth | -.307 | .143 | .031 |
| 5 | Velocity | .681 | .151 | *** |

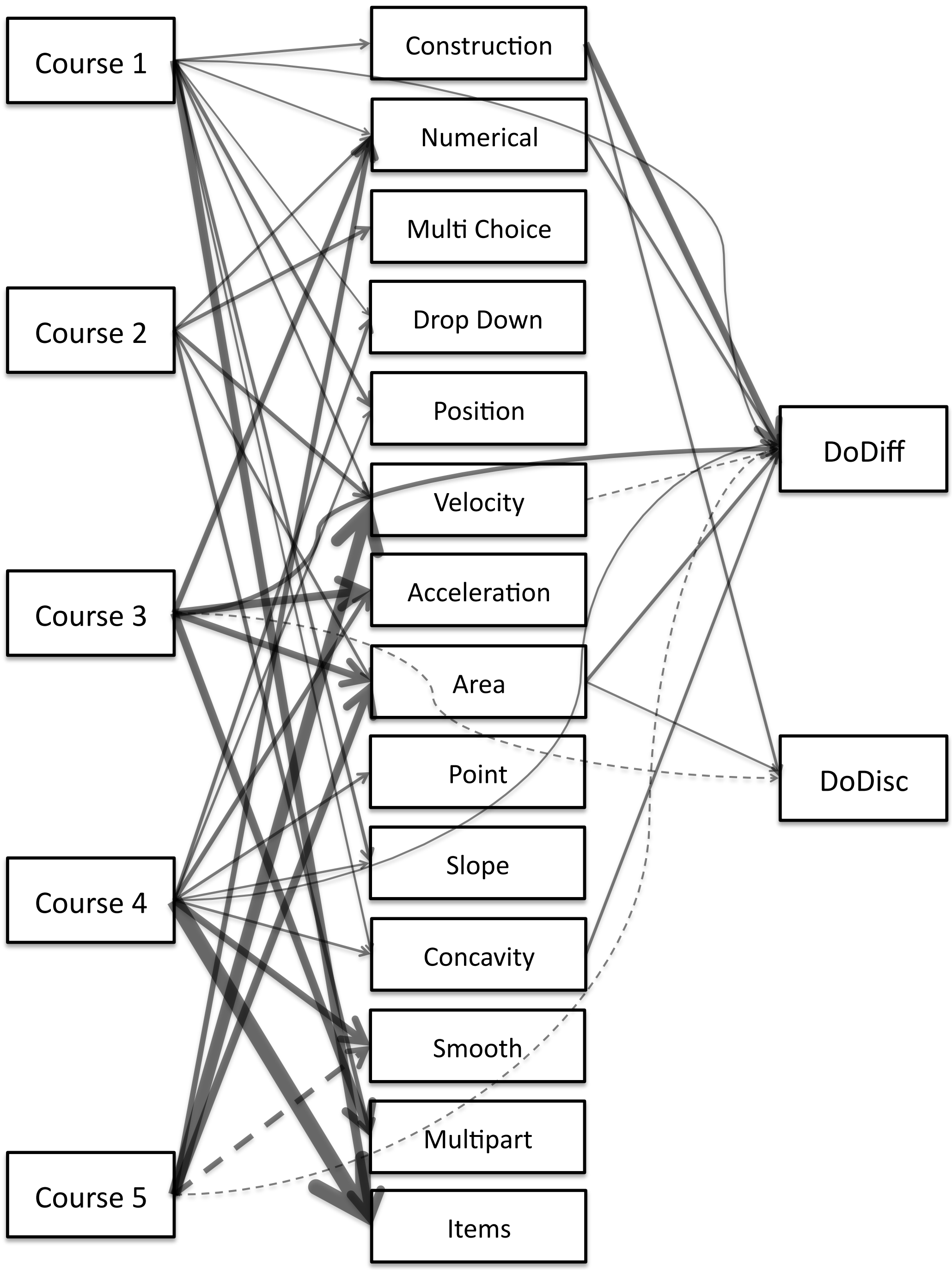



Figure 5.2: The full model showing only the paths whose results were significant. Positive/Negative coefficients are denoted by solid/dashed line, and the thickness of the lines denotes the magnitude of the coefficient.

# Chapter 6

# Conclusions

## 6.1 Summary of Results

We developed a new problem type in LON-CAPA that allows students to construct graphs instead of, more traditionally, interpret graphs. Special emphasis was put on not plotting some particular function, but instead to provide functionality to automatically evaluate and provide feedback on “back-of-the-envelope” graph sketches. Authoring such problems can be difficult, not so much from a technical point of view, but due to the fact that authors must anticipate a large range of answers (both correct and incorrect) and learner expectations. Using these problems in introductory physics courses, we found that they can in fact run counter to student expectations, and there was some resistance asking for more traditional graph-related problems. While initially perceived as cumbersome, in the long run, these graph construction problems did not turn out to be much more work-intensive than other problems, but did lean more toward the difficult end of the homework spectrum.

We then explored the effect of introducing graph interpretation and graph construction problems to two physics courses, one for engineers and one for premedical students. The results suggest that graph interpretation problems have no effect on students’ ability to interpret graphs as measured by the TUG-K. The introduction of graph construction problems in students’ homework resulted in significant differences in posttest scores for the course with premedical students. However, for the course primarily made up of engineers, the increase in posttest scores was not significant.

Interviewing students revealed little evidence of higher order thinking in the problem solving strategies they employed on graph problems, regardless of whether they required construction or interpretation. Instead of analysis, synthesis, and evaluation, we mostly found evidence of strategies associated with knowledge, the lowest of the levels, and some evidence of comprehension and application. The graph construction problems, if they have unique answers, can and will be reduced to mostly non-graph related problems by learners; in general, the only time the subject needed to think about the graph in these problems was to decide on its general shape.

What is worse is that when students do employ higher order thinking while solving graph problems, it generally proves unfruitful. Methods such as interpreting the physical situation were often abandoned as the subject did not know what to do with it or how to use it. Even when it was not abandoned, the subjects often obtained incorrect information from their attempt. From the student's point of view, it actually makes more sense to solve these problems using lower order thinking skills. From an instructor's point of view, we are effectively disincentivizing the students to solve graph problems by actually thinking about the graphs and what they represent.

As a result, simply translating "textbook-like" problems into the graphical realm will not achieve any additional educational goals. As with other problem types, educators need to leave the realm of calculational problems with unique solutions and move toward open-ended conceptual problems in order to move their students toward the next level.

An investigation of the results of using graph problems in LON-CAPA suggests that construction problems have a higher degree of difficulty and degree of discrimination than other graph problems. Beyond this, we investigated the effect courses have on these values and discovered that courses have a significant effect on the problem characteristics used in their classes. This underscores the need for more effective, data-driven suggestions in course management systems in general. Additionally, at least within a single institution, the course generally does not have a large direct effect on the degree of difficulty or degree of discrimination.

## 6.2 Implications for Instruction

As a result of this research, there are some clear statements that can be made with regard to teaching graphs to students in introductory physics courses.

First, giving students multiple choice graph problems is ultimately not very useful. Students will fall back on a single piece of understanding they have to solve the problem. The same students have been given multiple choice tests for most of their lives and have learned a number of meta strategies to solve them. Ultimately, these questions do not require the students to understand much about the graphs involved.

Graph construction problems may prove to be more useful for helping students learn, but this is certainly not automatic. Graph construction problems that mirror those found in standard textbooks allow students to circumvent deep engagement with the graphs. In order to prevent this, we (as instructors) should develop and use problems that can only be solved by engaging with the graph on more than surface level. Assigning problems with an infinite number of correct answers will hopefully force students to actually think about the general shapes and trends of graphs instead of trying to assign polynomials as answers.

Graph problems that do not explicity require connections to the real world are effectively math problems. While the math the students need to use may come from physical principles (kinematics, or $F = ma$ for example), asking them to read the necessary values off a graph or plot the points they calculated using those physical principles is not really using graphs for physics. Instead, students either read a graph and solve a standard physics problem, or solve a standard physics problem and then make the graph in the same manner they would if they were taking a math class.

Since most introductory physics courses are taught to large classes, it is important to develop these types of problems in ways that can easily scale to larger numbers, such as the Function Plot Response. It is not feasible to expect instructors to assign graph construction questions to hundreds of students and then grade them all by hand.

## 6.3 Implications for Future Research

In the preceding subsection, I suggested that graph construction problems with an infinite number of correct answers would hopefully useful. However, this is something that still needs to be investigated directly. It would be interesting to see how students approach these types of problems and how those approaches might be different from the strategies discussed in Chapter 4. The obvious extension of that work would be to interview students using the same methods, but new problems that reflect these suggestions.

Another interesting question that remains to be answered is whether or not teaching students graphs in introductory physics helps them deal with graphs in the rest of their lives. In other words, does the ability to answer questions involving graphs in physics correlate with the ability to understand and interpret graphs in general? If so, then it makes sense to teach these to all introductory students. However, if it turns out that teaching students graphs in physics does not improve their abilities with graphs in general, then perhaps it is a topic we should only address in courses with physics and engineering majors.

In Chapter 5, we investigated a model for predicting problem statistics using a fairly unsophisticated method. This vein of research should be extended in two ways. First, there should be an attempt to predict values for new problems and see if the results agree with the model's predictions. Second, there needs to be work that attempts to generate similar predictions for physics problems in general, not just graph problems.

## 6.4 Final Thoughts

During the course of this dissertation, two events happened in my life that made it clear to me the importance of this research. The first was a phone call from my mother, who was considering whether or not to go through chemotherapy again. The decision she ultimately made came down to the interpretation of a graph, which we discussed at length. I am happy to say that she is still here to see me finish this. The second was a conversation with my doctor about the results of a test performed on myself. Essentially, I ended up having to explain to my doctor how to interpret the graphs on which the data was represented. It bothers me to think about how those conversations go with other patients.

When I initially began this research, I generally assumed that giving students problems involving graph construction instead of graph interpretation would be more beneficial to learning. As a result of the research described in this dissertation, it has become clear to me that it is not as simple as I had hoped. Simply replacing graph interpretation with graph construction does not result in considerable increases in student learning, nor does it push students to consider the graphs more thoroughly or deeply.

However, there is still hope. If we can shift away from problems that require procedural understandings of reading graphs, and instead move toward questions that require comprehensive connections to the physical world, we will be doing a great service to our students.

# BIBLIOGRAPHY

# BIBLIOGRAPHY


[1] J. Zacks, E. Levy, B. Tversky, and D. Schiano, *Graphs in Print*, ch. Diagrammatic Representation and Reasoning, pp. 187–207. 2002.

[2] *National Science Education Standards*. Washington, DC: National Academy Press, 1996.

[3] *Project 2061: Science for all Americans*. New York: Oxford University Press, 1989.

[4] A. A. for the Advancement of Science (AAAS), *Benchmarks for Science Literacy*. Oxford University Press, 1993.

[5] A. V. Heuvelen, "Learning to think like a physicist: A review of research-based instructional strategies," *Am. J. Physics*, vol. 59, pp. 891–897, 1991.

[6] L. McDermott, M. L. Rosenquist, and E. H. van Zee, "Student difficulties in connecting graphs and physics: examples from kinematics," *Am. J. Phys.*, vol. 55, pp. 503–513, 1987.

[7] D. E. Meltzer, "Relation between students' problem-solving performance and representational format," *Am. J. Phys.*, vol. 73, pp. 463–478, 2005.

[8] D.-H. Nguyen and N. S. Rebello, "Students' understanding and application of the area under the curve concept in physics problems," *Phys. Rev. ST Phys. Educ. Res.*, vol. 7, p. 010112, Jun 2011.

[9] W. M. Christensen and J. R. Thompson, "Investigating graphical representations of slope and derivative without a physics context," *Phys. Rev. ST Phys. Educ. Res.*, vol. 8, p. 023101, Jul 2012.

[10] H. Lin, "Learning physics vs. passing courses," *Phys. Teach.*, vol. 20, pp. 151–157, 1982.

[11] G. Leinhardt, O. Zaslavsky, and M. K. Stein, "Functions, graphs, and graphing: Tasks, learning, and teaching," *Review of Educational Research*, vol. 60, pp. 1–64, 1990.

[12] J. Clement, J. Lochhead, and G. S. Monk, "Translation difficulties in learning mathematics," *The American Mathematical Monthly*, vol. 88, pp. 286–290, 1981.

[13] R. N. Steinberg, M. C. Wittmann, and E. F. Redish, "Mathematical tutorials in introductory physics," in *The Changing Role Of Physics Departments In Modern University, AIP Conference Proceedings*, pp. 1075–1092, 1997.

[14] J. Dewey I. Dykstra and D. R. Sweet, "Conceptual development about motion and force in elementary and middle school students," *Am. J. Physics*, vol. 77, pp. 468–476, 2009.

[15] J. Bertin, "La semiologie graphique," *The Hague, Netherlands: Mouton-Gautier*, 1973.

[16] F. R. Curcio, "Comprehension of mathematical relationships expressed in graphs," *Journal for Research in Mathematics Education*, vol. 18, pp. 382–385, 1987.

[17] C. C. McKnight, "Critical evaluation of quantitative arguments," in *Assessing higher order thinking in mathematics* (G. Kulm, ed.), pp. 169–185, American Association for the Advancement of Science, Washington DC, 1990.

[18] H. Wainer, "Understanding graphs and tables," *Educational Researcher*, vol. 21, pp. 14–23, 1992.

[19] C. M. Carswell, "Choosing specifiers: An evaluation of the basic tasks model of graphical perception," *Human Factors*, vol. 34, pp. 535–554, 1992.

[20] S. N. Friel, F. R. Curcio, and G. W. Bright, "Making sense of graphs: Critical factors influencing comprehension and instructional implications," *Journal for Research in Mathematics Education*, vol. 32, pp. 124–158, 2001.

[21] H. Wainer, "A test of graphicacy in children," *Applied Psychological Measurement*, vol. 4, pp. 331–340, 1980.

[22] W. E. Brock, ed., *What work requires of schools: A SCANS report for America 2000*. Secretarys Commission on Achieving Necessary Skills, U.S. Department of Labor, 1991.

[23] E. F. Redish, R. N. Steinberg, and J. M. Saul, "Student expectations in introductory physics," *Am. J. Phys.*, vol. 66, pp. 212–224, 1998.

[24] K. F. Connery, "Graphing predictions," *The Science Teacher*, vol. 74, pp. 42–46, 2007.

[25] M. Padilla, D. McKenzie, and E. Shaw, "An examination of the line graphing ability of students in grades seven through twelve," *School Science and Mathematics*, vol. 86, pp. 20–26, 1986.

[26] H. Brasell and M. B. Rowe, "Graphing skills among high school physics students," *School Science and Mathematics*, vol. 93, pp. 63–70, 1993.

[27] S. Szyjka, *Cognitive and attitudinal predictors related to graphing achievement among pre-service elementary teachers.* PhD thesis, Southern Illinois University, Carbondale, 2009.

[28] K. E. Gray, W. K. Adams, C. E. Wieman, and K. K. Perkins, "Students know what physicists believe, but they don't agree: A study using the class survey," *Phys. Rev. ST Phys. Educ. Res.*, vol. 4, p. 020106, Nov 2008.

[29] I. S. Araujo, E. A. Veit, and M. A. Moreira, "Physics students' performance using computational modelling activities to improves kinematics graphs interpretation," *Computers & Education*, vol. 50, pp. 1128–1140, 2008.

[30] P. Hale, "Kinematics and graphs: Students' difficulties and cbls," *The Mathematics Teacher*, vol. 93, pp. 414–417, 2000.

[31] J. R. Mokros and R. F. Tinker, "The impact of microcomputer-based labs on children's ability to interpret graphs," *Journal of Research in Science Teaching*, vol. 24, pp. 369–383, 1987.

[32] R. K. Thornton and D. R. Sokoloff, "Learning motion concepts using real-time microcomputer-based laboratory tools," *Am. J. Phys.*, vol. 58, pp. 858–867, 1989.

[33] M. Svec, "Improving graphing interpretation skills and understanding of motion using micro-computer based laboratories," *Electronic Journal of Science Education*, vol. 3, 1999.

[34] I. Testa, G. Monroy, and E. Sassi, "Students' reading images in kinematics: the case of real-time graphs," *International Journal of Science Education*, vol. 24, pp. 235–256, 2002.

[35] D. W. Russell, K. B. Lucas, and C. J. McRobbie, "The role of the microcomputer-based laboratory display in supporting the construction of new understandings in kinematics," *Research in Science Education*, vol. 33, pp. 217–243, 2003.

[36] H. Brasell, "The effect of real-time laboratory graphing on learning graphic representatiosn of distance and velocity," *Journal of Research in Science Teaching*, vol. 24, pp. 385–395, 1987.

[37] R. J. Beichner, "The effect of simultaneous motion presentation and graph generation in a kinematics lab," *Journal of Research in Science Teaching*, vol. 27, pp. 803–815, 1990.

[38] J. B. Brungardt and D. Zollman, "Influence of interactive videodisc instruction using simultaneous-time analysis on kinematics graphing skills of high school physics students," *Journal of Research in Science Teaching*, vol. 32, pp. 855–869, 1995.

[39] D. D. Adams and J. W. Shrum, "The effects of microcomputer-based laboratory exercises on the acquisition of line graph construction and interpretation skills by high school biology students," *Journal of Research in Science Teaching*, vol. 27, pp. 777–787, 1990.

[40] R. Mitnik, M. Recabarren, M. Nussbaum, and A. Soto, "Collaborative robotic instruction: A graph teaching experience," *Computers & Education*, vol. 53, pp. 330–342, 2009.

[41] "Pads." `http://www.wku.edu/pads/`.

[42] Maplesoft, "Maple ta." `http://www.maplesoft.com/products/mapleta/`.

[43] Pearson, "Mastering physics." `http://www.masteringphysics.com/`.

[44] M. Klymkowsky, M. Cooper, and S. Bryfcynski, "Besocratic." `http://besocratic.clemson.edu/`.

[45] "Webassign." `http://www.webassign.net/`.

[46] Pearson, "Mymathlab." `http://www.mymathlab.com/`.

[47] S. Bonham, "Graphical response exercises for teaching physics," *The Physics Teacher*, vol. 45, pp. 482–486, 2007.

[48] "Geogebra homepage." `http://www.geogebra.org/`.

[49] M. Hohenwarter, M. Hohenwarter, and Z. Lavicza, "Introducing dynamic mathematics software to secondary school teachers: The case of geogebra," *Journal of Computers in Mathematics and Science Teaching*, vol. 28, pp. 135–146, 2009.

[50] "Geogebra awards." `http://www.geogebra.org/cms/en/info`, 2010.

[51] "Geogebra mobile." `http://www.geogebra.org/trac/wiki/GeoGebraMobile`.

[52] "Lon-capa homepage." `http://www.lon-capa.org/`.

[53] D. A. Kashy, G. Albertelli, G. Ashkenazi, E. Kashy, H.-K. Ng, and M. Thoennessen, "Individualized interactive exercises: a promising role for network technology," in *Proc. Frontiers in Education*, vol. 31, pp. 1073–1078, 2001.

[54] G. Kortemeyer, E. Kashy, W. Benenson, and W. Bauer, "Experiences using the open-source learning content management and assessment system lon-capa in introductory physics courses," *Am. J. Phys*, vol. 76, pp. 438–444, 2008.

[55] "Lon-capa user institutions listing." `http://www.lon-capa.org/institutions.html`.

[56] "Lon-capa shared content repository statistics." `http://www.lon-capa.org/sharedpool.html`.

[57] G. Kortemeyer, "Gender differences in the use of an online homework system in an introductory physics course," *Phys. Rev. ST-PER*, vol. 5, p. 010107, 2009.

[58] G. Kortemeyer, "An analysis of asynchronous online homework discussions in introductory physics courses," *Am. J. Phys.*, vol. 74, pp. 526–536, 2005.

[59] E. Silver, "Fostering creativity through instruction rich in mathematical problem solving and problem posing," *ZDM*, vol. 29, pp. 75–80, 1997.

[60] A. Pascarella, "The influence of web-based homework on quantitative problem-solving in a university physics class," in *Proc. NARST Annual Meeting*, 2004.

[61] R. J. Beichner, "Testing student interpretation of kinematics graphs," *Am. J. Phys.*, vol. 62, pp. 750–762, 1994.

[62] J. T. Laverty and G. Kortemeyer, "Function plot response: A scalable system for teaching kinematics graphs," *Am. J. Phys.*, vol. 80, pp. 724–733, 2012.

[63] R. R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *Am. J. Phys.*, vol. 66, pp. 64–74, 1998.

[64] K. A. Ericsson and H. A. Simon, "Verbal reports as data," *Psychological Review*, vol. 87, pp. 215–251, 1980.

[65] T. Lowrie and C. M. Diezmann, "Solving graphics tasks: Gender differences in middle-school students," *Learning and Instruction*, vol. 21, pp. 109–125, 2011.

[66] B. S. Bloom, *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain.* David McKay Co Inc. (New York), 1956.

[67] L. Anderson and D. A. Krathwohl, *Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives.* Longman, 2001.

[68] T. Wemyss and P. van Kampen, "Categorization of first-year university students' interpretations of numerical linear distance-time graphs," *Phys. Rev. ST Phys. Educ. Res.*, vol. 9, p. 010107, Feb 2013.

[69] G. Kortemeyer, "Correlations between student discussion behavior, attitudes, and learning," *Phys. Rev. ST-PER*, vol. 3, p. 010101, 2007.

[70] V. Mesic and H. Muratovic, "Identifying predictors of physics item difficulty: A linear regression approach," *Phys. Rev. ST Phys. Educ. Res.*, vol. 7, 2011.

[71] M. T. H. Chi, P. J. Feltovich, and R. Glaser, "Categorization and representation of physics problems by experts and novices," *Cognitive Science*, vol. 5, pp. 121–152, 1981.

[72] G. H. Veldhuis, "The use of cluster analysis in categorization of physics problems," *Science Education*, vol. 74, pp. 105–118, 1990.

[73] S. F. Wolf, D. P. Dougherty, and G. Kortemeyer, "Empirical approach to interpreting card-sorting data," *Phys. Rev. ST Phys. Educ. Res.*, vol. 8, p. 010124, May 2012.

[74] A. Feil and J. Mestre, "Change blindness as a means of studying expertise in physics," *J. of the Learning Sciences*, vol. 19, pp. 480–505, 2010.

[75] R. E. Teodorescu. PhD thesis, George Washington University, Washington, DC, 2009.

[76] G. Kortemeyer, "Ten years later: Why open educational resources have not noticeably affected higher education, and why we should care," *EDUCAUSE Review*, 2013.

[77] K. C. Haudek, L. B. Prevost, R. A. Moscarella, J. Merrill, and M. Urban-Lurain, "What are they thinking? automated analysis of student writing about acid-base chemistry in introductory biology," *CBE - Life Sciences Education*, vol. 11, pp. 283–293, 2012.