THE DEVELOPMENT OF THE PROBLEM DECOMPOSITION DIAGNOSTIC

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By
David J. Van Domelen

*********

The Ohio State University

2000

Dissertation Committee:
Professor David Andereck, co-chair
Professor Alan Van Heuvelen, co-chair
Dr. James Stith

Approved by
Co-chair

Physics Graduate Program
ABSTRACT

The Problem Decomposition Diagnostic is a testing instrument designed to help determine student ability in two areas of problem solving. As the name implies, all versions of the test have been intended to gauge student ability to decompose, or break up, complex physics problems into simpler sub-problems. Additionally, later versions of the PDD sought to also examine the strategy the student would use to solve such sub-problems.

This thesis is organized in a roughly chronological manner, following the development of the Problem Decomposition Diagnostic from early open response tests and interview tasks to the current multiple-choice format instrument. Both qualitative and quantitative methods were used in developing this instrument and measuring its validity. Extensive appendices include every version of the test, plus a complete laboratory course which helped inspire this work and informed a number of the aspects of the PDD.
Dedicated to the student who, during a tutoring session in 1990, forced me to consider the possibility that students might not be able to decompose complex problems.
ACKNOWLEDGMENTS

I wish to thank my advisor, Alan Van Heuvelen, for always supporting me in what I chose to do, giving me ideas and encouragement when I needed them, and perhaps most importantly, giving me the freedom to pursue my own path.

I thank Dr. David Andereck for agreeing on somewhat short notice to be my thesis committee co-chair.

Thanks also to Dr. James Stith, Dr. Greg Lafyatis and Dr. Ben Patton for serving on my committees throughout the years.

I am grateful to all those who I bounced ideas off over the course of developing this test, including but not limited to: Kathy Andre, Leith Allen, Tom Foster, David May, Dr. Edward F. Redish, Seth Rosenberg and Xueli Zou. And thanks to Dr. Edward Adelson for being patient with me when I wanted to come into the Physics 131 courses and take up time giving my tests.

This research was supported in part by NSF grants DUE 9396205 and GER 9553460, plus the Graduate Research Trainee grant DGE-9553460.
VITA

July 18, 1970.................................Born - Waukesha, WI, USA

1992.............................................B.S. Physics, Northeast Missouri State University, summa cum laude

1992-1995....................................Graduate Teaching Assistant

The Ohio State University

1995-present...............................Graduate Research Trainee

The Ohio State University

PUBLICATIONS

Peer-Reviewed Publication


FIELD OF STUDY

Major Field: Physics
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedication</td>
<td>.................................................................................................</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>...........................................................................................</td>
<td>iv</td>
</tr>
<tr>
<td>Vita</td>
<td>.................................................................................................</td>
<td>v</td>
</tr>
<tr>
<td>List of Tables</td>
<td>............................................................................................</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>.............................................................................................</td>
<td>xii</td>
</tr>
</tbody>
</table>

Chapters:

1. Introduction...........................................................................................................1
   1.1 The issue of problem decomposition............................................................2
      1.1.1 The importance of problem decomposition skills.......................3
      1.1.2 Why we want to assess proficiency in problem decomposition........6
      1.1.3 Why a new instrument?..............................................................7
   1.2 Other work in the content area of physics............................................9
   1.3 Summary of PDD development...............................................................10
   1.4 Reliability and Validity.........................................................................12

2. The Zone Test.......................................................................................................14
   2.1 Administration of the Zone Test............................................................14
   2.2 Results.....................................................................................................16

3. The Problem Decomposition Diagnostic version zero..................................19
8. PDD v2.2 interviews .................................................................98
  8.1 Interview protocol .............................................................99
  8.2 Interview summaries ......................................................100
  8.3 Discussion of interviews ..................................................114

9. Large-scale administration of the PDD v2.2 ..........................116
  9.1 The experimental group ....................................................116
  9.2 Goals of this administration ............................................117
  9.3 Pretest comparisons with other measures ......................118
  9.4 Posttest comparisons with other measures .....................122
  9.5 Comparison of pretest PDD to posttest PDD results ..........127

10. Future directions, Conclusions ...........................................139
   10.1 The Problem Decomposition Diagnostic v2.3 .................140
   10.2 New mechanics-content items .......................................142
   10.3 Additional content areas ..............................................146
   10.4 Administration at other institutions ...............................147
   10.5 Conclusions .................................................................148

Appendices:
A: Toys In Motion laboratory manual version 2.0 ....................150
B: The Zone Tests .................................................................198
C: The Problem Decomposition Diagnostic version 0.5 ..........223
D: ActivPhysics interview results ..........................................231
E: Card-sorting task ..............................................................240
F: The Problem Decomposition Diagnostic version 1.0 ............253
G: The Problem Decomposition Diagnostic version 1.1 ..........260
H: The Problem Decomposition Diagnostic version 2.0 ..........269
I: The Problem Decomposition Diagnostic version 2.1 ..........277
J: The Problem Decomposition Diagnostic version 2.2 .............................291
K: The Problem Decomposition Diagnostic version 2.3 ..........................304
L: Problem Decomposition Diagnostic expansion items.
   open response.................................................................317
M: Problem Decomposition Diagnostic expansion items.
   multiple choice............................................................325

Bibliography.........................................................................................331
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Summary of PDD v1.1 scores</td>
<td>53</td>
</tr>
<tr>
<td>5.2</td>
<td>Correlations and means tests for 131. a) Correlations, b) Difference</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>in means for students with scorable and unscorable tests</td>
<td></td>
</tr>
<tr>
<td>5.3</td>
<td>Correlations and means tests for 131E. a) Correlations, b) Difference</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>in means for students with scorable and unscorable tests</td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>Correlations and means tests for H131. a) Correlations, b) Difference</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>in means for students with scorable and unscorable tests</td>
<td></td>
</tr>
<tr>
<td>5.5</td>
<td>Percent of class correct by sub-problem. PDD v1.1</td>
<td>69</td>
</tr>
<tr>
<td>6.1</td>
<td>Results of the PDD v2.1 on items 1 through 3.</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>The letters across the top are responses on the PDD (correct responses are</td>
<td></td>
</tr>
<tr>
<td></td>
<td>underlined), compared against course grades.</td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>Results of the PDD v2.1 on items 4 and 5.</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>The letters across the top are responses on the PDD (correct responses are</td>
<td></td>
</tr>
<tr>
<td></td>
<td>underlined), compared against course grades.</td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>Results of the PDD v2.1 on items 6 through 8.</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Only four options are listed across the top, as this section had only four</td>
<td></td>
</tr>
<tr>
<td></td>
<td>options per item, and no “wrong” answers.</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Results of the PDD v2.1 on items 9 and 10.</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td>Only four options are listed across the top, as this section had only four</td>
<td></td>
</tr>
<tr>
<td></td>
<td>options per item, and no “wrong” answers.</td>
<td></td>
</tr>
</tbody>
</table>
6.5 Results of the PDD v2.2 on items 1 through 3. The letters across the top are responses on the PDD (correct responses are underlined), compared against course grades.................................................................81

6.6 Results of the PDD v2.2 on items 4 and 5. The letters across the top are responses on the PDD (correct responses are underlined), compared against course grades.................................................................82

6.7 Results of the PDD v2.2 on items 6 through 8. Only four options are listed across the top, as this section had only four options per item, with no “wrong” answers.................................................................83

6.8 Results of the PDD v2.2 on items 9 and 10. Only four options are listed across the top, as this section had only four options per item, with no “wrong” answers.................................................................84

9.1 Pretest/Posttest responses for Item 1 on the PDD v2.2.................................129

9.2 Pretest/Posttest responses for Item 2 on the PDD v2.2.................................130

9.3 Pretest/Posttest responses for Item 3 on the PDD v2.2.................................131

9.4 Pretest/Posttest responses for Item 4 on the PDD v2.2.................................132

9.5 Pretest/Posttest responses for Item 5 on the PDD v2.2.................................133

9.6 Pretest/Posttest responses for Item 6 on the PDD v2.2.................................134

9.7 Pretest/Posttest responses for Item 7 on the PDD v2.2.................................135

9.8 Pretest/Posttest responses for Item 8 on the PDD v2.2.................................136

9.9 Pretest/Posttest responses for Item 9 on the PDD v2.2.................................137

9.10 Pretest/Posttest responses for Item 10 on the PDD v2.2.........................138

G.1 Multiple choice options for the Problem Decomposition Diagnostic v1.1....263
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Wickelgren’s schematic, demonstrating the usefulness of sub-goals.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a) Decision tree for a complex problem. b) Reduced decision tree for a complex problem containing a sub-goal</td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>Sample worked item from Zone Test 1</td>
<td>15</td>
</tr>
<tr>
<td>5.1</td>
<td>PDD v1.1 scoring histograms for Physics 131 students. N=19</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>PDD v1.1 scoring histogram for Physics 131E. N=56</td>
<td>57</td>
</tr>
<tr>
<td>5.3</td>
<td>PDD v1.1 scoring histogram for Physics H131. N=22</td>
<td>58</td>
</tr>
<tr>
<td>A.1</td>
<td>Diagrams for possible setups for Lab 1. a) Flat surface. b) Sloped surface</td>
<td>157</td>
</tr>
<tr>
<td>A.2</td>
<td>Diagrams for Lab 3. Part A. a) A car rolling down a slope. b) A falling object</td>
<td>163</td>
</tr>
<tr>
<td>A.3</td>
<td>Diagrams for Lab 3. Part B. a) A car rolling on a horizontal surface. b) a car rolling down a slope</td>
<td>165</td>
</tr>
<tr>
<td>A.4</td>
<td>Diagrams for Lab 4. a) Missile launcher version. b) Hot Wheels™ version</td>
<td>169</td>
</tr>
<tr>
<td>A.5</td>
<td>Lab 5 diagrams. a) Truck pulling mass hung from a pulley. b) Truck pulling mass in cart tied behind it on slope</td>
<td>174</td>
</tr>
</tbody>
</table>
A.6 Diagrams for Lab 6. a) Setup for Lab 6. Initial positions are shown by dashed lines, final positions by solid lines. All velocities not explicitly indicated are zero. b) Sample velocity versus time graph for Lab 6. c) Sample position versus time graph for Lab 6.

A.7 Diagrams for Lab 7. a) Physical representation. b) Force diagram.

A.8 Setup for Lab 8, the Vertical Loop.

E.1 Card Sorting Task: Card 11

E.2 Card Sorting Task: Card 2

E.3 Card Sorting Task: Card 10

E.4 Card Sorting Task: Card 7

E.5 Card Sorting Task: Card 4

E.6 Card Sorting Task: Card 9

E.7 Card Sorting Task: Card 3

E.8 Card Sorting Task: Card 6

E.9 Card Sorting Task: Card 1

E.10 Card Sorting Task: Card 8

E.11 Card Sorting Task: Card 5

E.12 Card Sorting Task: Card 12

F.1 Problem Decomposition Diagnostic v1.0 item 1

F.2 Problem Decomposition Diagnostic v1.0 item 2

F.3 Problem Decomposition Diagnostic v1.0 item 3
K.5  Problem Decomposition Diagnostic v2.3 item 5.................................311
K.6  Problem Decomposition Diagnostic v2.3 item 6.................................312
K.7  Problem Decomposition Diagnostic v2.3 item 7.................................313
K.8  Problem Decomposition Diagnostic v2.3 item 8.................................314
K.9  Problem Decomposition Diagnostic v2.3 item 9.................................315
K.10 Problem Decomposition Diagnostic v2.3 item 10...............................316
L.1  Problem Decomposition Diagnostic expansion sample item ..................319
L.2  Problem Decomposition Diagnostic expansion item 1.........................320
L.3  Problem Decomposition Diagnostic expansion item 2.........................321
L.4  Problem Decomposition Diagnostic expansion item 3.........................322
L.5  Problem Decomposition Diagnostic expansion item 4.........................323
L.6  Problem Decomposition Diagnostic expansion item 5.........................324
M.1  Problem Decomposition Diagnostic expansion item 1: multiple choice....326
M.2  Problem Decomposition Diagnostic expansion item 2: multiple choice....327
M.3  Problem Decomposition Diagnostic expansion item 3: multiple choice....328
M.4  Problem Decomposition Diagnostic expansion item 4: multiple choice....329
M.5  Problem Decomposition Diagnostic expansion item 5: multiple choice....330
CHAPTER 1

INTRODUCTION

The following exchange is true, in the sense that it has happened in the past and will likely happen in the future. The details vary from instance to instance. of course, but the essential events remain about the same.

STUDENT: Could you help me with this problem? [Student presents Tutor with a complicated homework problem that can’t be solved by a single equation.]

TUTOR: Sure. The first thing you need to do is figure out what velocity you need at the end in terms of the energy at this point. [Tutor draws a diagram, labels the energy and velocity, then writes down an equation.]

S: Is that the answer?

T: No, not yet. Next you need to figure out the energy in terms of your givens, like the initial height. [More labels, another equation.]

S: Is that the answer?

T: No, now you need to put the pieces together. [Does so.]

S: Is that the answer?

T: Yes, that’s the answer.

S: [Pause] Isn’t there an equation to do this. so I can put it on my cheat sheet for the midterm?
1.1 The issue of problem decomposition

In short, the above student's strategy for problem-solving is similar to that suggested in many textbooks (Fishbane, 29; Gettys, 9-10; Giancoli, 64-65; Halliday, xxix; Hecht, 22; Van Heuvelen (1982), 19; Weidner, 206-207; Young, 41, 65-66, 100, et al): Read the problem, draw a diagram, identify an equation, solve the equation, check the results. When confronted with a more complex problem that cannot be solved by a single textbook equation, the student may be unable to identify the sub-problems into which the task can be decomposed. The student is unable to Analyze the problem.

Bloom's Taxonomy of understanding (Bloom) provides a commonly-used way to describe where a person is in their progress towards mastery of a subject. Analysis is one of the six levels in this taxonomy, which are:

1) Knowledge. The student has memorized facts and basic procedures. In the context of physics tasks, students who have reached this level are able to perform simple back-of-chapter homework problems that involve solving equations provided in the book or in the student's notes.

2) Comprehension. The student understands some of the reasoning and meaning behind the rote knowledge. In other words, the student can apply his understanding to basic qualitative reasoning about a situation.

3) Application. More specifically, application to an unfamiliar situation. Exam problems and harder textbook problems may require this level of understanding from the student, but not always.

4) Analysis. The student is able to look at a complicated situation and apply his or her understanding of physics in order to break it into simpler pieces that can then be solved. The pieces are then reassembled for solve the larger problem.
5) Synthesis. The student can put together aspects of different domains or sub-domains to approach a novel situation or design an experiment or process.

6) Evaluation. This is the ability to not only determine whether a situation meets established criteria, but also the ability to establish those criteria in a meaningful manner. A full understanding of the relevant domain is needed in order to do this well.

In the example that leads off this chapter, the student is attempting to master the Knowledge level, while the tutor is demonstrating Analysis skills. The problem was written with the intent of helping students acquire Analysis skills, but the student is convinced that only Knowledge will be required in order to pass the course.

This leads to three basic questions which will be answered in this chapter and lead into the remainder of the thesis.

1) Is it really important for students to be able to decompose complex problems into simpler sub-problems? In other words, is Analysis really a useful skill to teach in physics?

2) Why would a teacher or researcher want to know a student's level of proficiency in problem decomposition?

3) Is a new instrument truly necessary?

1.1.1 The importance of problem decomposition skills

This would seem to be almost self-evident, but the question does need to be asked and answered. Is problem decomposition a skill that students should possess? Given the number of single-step exercises in the average college physics text, it is entirely possible for a student to perform well in class without ever demonstrating problem decomposition skills, so the question is worth asking.
Given that life involves complicated problems, and that solving back-of-chapter problems is not something students will do once they leave class (physics majors will certainly move on to more complex problems), being able to break something into smaller, more easily solved pieces is certainly an important skill. One selling point in the classroom for problem decomposition is that students can take a difficult problem and break it into sub-problems that resemble the back-of-chapter homework exercises. Obviously, if one can reduce a task to elements you know how to solve, the task is more than half done. However, this is not the only benefit of problem decomposition. As stated in How To Solve Problems (Wickelgren. 91-94), simply by setting subgoals between the givens and the ultimate goal of a problem, it is possible to reduce the work required to reach a solution. Searching hit-or-miss for a solution path to a complex problem involves a large number of possible paths. However, if a subgoal is established, then one large search can be reduced to two much smaller searches (see Figure 1.1). Being able to evaluate progress at the subgoal reduces the number of blind alleys that the problem-solver may have otherwise follow. In the case of very complex problems, the difficulty can be reduced by an order of magnitude or more by appropriate choice of subgoals. Wickelgren demonstrates that this method is so powerful that, even if multiple guesses are required to set an appropriate subgoal, the work required of the problem-solver is still greatly reduced.

This does, however, raise an additional point. Choice of appropriate subgoals is going to be at least partly driven by student content knowledge and their ability to identify the deep structure of the subproblem that would meet the subgoal. After all, in the simple case of zero content knowledge, a problem-solver will have no idea what constitutes an appropriate subgoal. Therefore, in order for a student to acquire the desired skill in problem decomposition, the student will also need a grasp of the content area in order to identify subgoals and evaluate the results of subproblems.
Figure 1.1 - Wickelgren’s schematic, demonstrating the usefulness of sub-goals.

a) Decision tree for a complex problem  b) Reduced decision tree for a complex problem containing a sub-goal.
Beyond simply improving problem-solving skills, an emphasis on teaching problem decomposition can help students build to the Analysis level of understanding. There is a push, part due to the importance of problem-solving skills in the workplace, to bring students higher up Bloom’s Taxonomy than they are currently rising in physics (Van Heuvelen (1999)), and students who gain mastery of problem decomposition have acquired useful Analysis skills.

1.1.2 Why we want to assess proficiency in problem decomposition

There are two reasons to assess student proficiency in problem decomposition. The first is to determine if students need help in developing this skill, and the second is to determine if any help that is given is effective.

In the first case, an assessment instrument should not only inform the instructor or researcher that a problem exists, but also diagnose specific weaknesses in each individual student as well as in the class as a whole. Students who are perfectly able to decompose one variety of task may not recognize a viable subgoal in a different sort of problem. Ideally, the instrument should also measure the student’s inclination to decompose a problem, but it may not be feasible to cover both “will they?” and “can they?” in a single instrument. For the purposes of the instrument described in this thesis, the “can they?” question will be given greater weight.

Having established that students need instruction designed to improve their problem decomposition skills, it is vital that there be some way to determine if that instruction has had any effect. To give an example from the author’s work, the Toys In Motion laboratory course (Appendix A) was designed to, among other things, encourage students to break complex and vague assignments up into smaller tasks which could be easily performed.
Anecdotally this lab course seemed to have the desired effect, but it could not be determined whether students in the TIM labs acquired better problem decomposition skills than classmates in a laboratory course that had less of an emphasis on those skills. Clearly, in such a case where hundreds of students are studied at once, some method of assessing problem decomposition skills, other than in-depth interviews, is desirable.

1.1.3 Why a new instrument?

It is fair to ask why a teacher or researcher could not simply assign complex problems on an exam and determine problem decomposition ability from student responses. This approach can be useful in some cases, but it runs into certain difficulties.

First, it is time-consuming for both student and teacher. A single complex problem can require up to an hour for a single student to solve, and several would be required to elicit a reasonable range of problem types and potential pitfalls. Likewise, grading a hundred or more of these test papers takes a great deal of time, especially if partial credit is to be given. After all, if there are four valid subgoals to a problem, a student could correctly pick three of them and still write a test paper that is not only incorrect, but also very difficult to decipher.

Secondly, problem decomposition is often not done explicitly, or in an easily-interpreted manner. Student answers often wander about the page, with sections connected by circles and arrows and other vague indicators. Even worse, students' reasons for choosing their subgoals will often be left off the test sheet entirely. Did the student omit the collision because of an incorrect assumption about energy conservation, because the student assumed velocity would be conserved, et cetera? Students can be very results-motivated, and leave out important steps in their work.
Additionally, current work at Iowa State University and Southwest Missouri State University (Manivannan) has been probing the relationship between student problem-solving abilities and conceptual knowledge via a "problem dissection" protocol. This work suggests that problem-solving skills are not necessarily a predictor of conceptual understanding. Students who successfully solved moderately complex (if not multi-part) physics problems still did poorly on conceptual questions directly related to the problem they had just solved. This implies that students are approaching these problems from a Knowledge level of understanding, without having mastered any higher levels in Bloom's Taxonomy. It is therefore unlikely that these students have reached the point of being able to Analyze problems.

Exam problems, therefore, require a large investment in time for all concerned, and may not even require Analysis skills to successfully solve, despite any apparent complexity. The loss of information can be solved by working individually with students in an interview process in which the student solves complex problems while the interviewer guides the student via a protocol, but this raises the time investment to prohibitive levels for the average instructor. It is certainly not practical for gauging the ability level of an entire introductory mechanics class. An objective test focusing more specifically on problem decomposition would certainly be desirable.

Does an existing, widely-available objective test exist that measures a student’s ability to solve complex problems in general, or problem decomposition skills in specific? The Force Concept Inventory (Hestenes (1992a)) and Force and Motion Conceptual Exam (Thornton) do look at the underlying content knowledge that is important in problem decomposition, but they do not address complex problems. The Mechanics Baseline Test (Hestenes (1992b)) does measure a student’s problem-solving ability, but in the domain of single-part problems rather than complex multiple-part problems. Various attitudinal
surveys could be modified to measure a student’s inclination to see complex problems as being composed of multiple simpler pieces, but would say nothing of the student’s ability to perform the decomposition.

So, looking at the situation as it stands, there is not only a need for an instrument to diagnose student proficiency at problem decomposition, that need is not currently being met by existing tests and methods.

1.2 Other work in the content area of physics

The author would like to first issue a disclaimer for this section. Physics Education Research is still a young and rapidly-growing field, and there is much research done in relative isolation. Therefore, this section will not attempt to include all work being performed on physics problem-solving as pertains to problem decomposition. New work is being performed all the time, and in fact, one example came to the author’s attention during the writing of this thesis. Instead, this section will attempt to cover a few important pieces relevant to the Problem Decomposition Diagnostic’s development and purpose.

Some of the best work in problem decomposition has been performed by Frederick Reif. Over the course of numerous student interviews (Heller (1984), Frederick Reif and Joan Heller attempted to develop a problem-solving strategy which could be taught successfully to students. A balance was sought between a method that was too vague to be useful and too prescriptive to be widely-applicable. The results of this work can be found in the textbook Understanding Basic Mechanics (Reif. 52-61). Students are not only told to look for ways to break apart a complex problem, they are given advice on the sort of things to look for, such as events in a sequence or known relationships between elements
of the problem. A number of other research-based physics texts focusing on problem-solving have been produced, including *The Competent Problem Solver* (Heller (1992a)).

Any test of problem decomposition skills should be able to tell something about why a student decomposes problems in a particular way, and for this one should look at the extensive work done by M. Chi and her collaborators (Chi). Subjects, both instructors and students, were given a number of physics problems on index cards and asked to sort them by similarity. Experts tended to group the problems based on underlying physical concepts, while non-experts grouped the cards based on surface features. In other words, while an expert might have a pile of energy conservation problems, an non-expert would have a pile of “inclined plane” problems for one of his or her groupings.

It is important to consider a task along the lines of Chi’s card-sorting experiment, to determine whether domain-specific knowledge is vital to decomposition skills. If students demonstrate the ability to correctly decompose physics problems, yet consistently classify the resulting sub-problems in a non-expert fashion, that might suggest two things. One, problem decomposition skills can be learned at a strictly Knowledge level, and two, non-domain-specific problem-solving skills may have a stronger influence on student problem-solving skills than domain-specific understanding. Chi’s research is very robust, and at the least would provide a strong case for whether a subject has reached the Comprehension level in Bloom’s Taxonomy.

### 1.3 Summary of PDD Development

The following is a brief summary of what the remainder of this thesis will cover.

After experiencing a number of situations similar to the one that started this chapter, the author decided to devise a test of problem decomposition ability, called the *Zone Test*. 
Students were presented with the diagrammatic representation of a physics problem and asked to label significant points on the diagram, then explain what the "zones" between or at those points signified. While results were very inconclusive, certain consistent responses eventually led to further investigation during the later development of the PDD.

Some years later, as part of work on the Toys In Motion labs (Appendix A), the author attempted to design a decomposition test that would not depend on physics content knowledge, but this was quickly proven impractical. Physics content was necessary to reduce the ambiguity of the problems and the number of possible correct answers.

The next step involved a series of small-group interview tasks, including the use of computer animations of complex problems and a card-sorting exercise inspired by Chi's work. These led to the first version of the PDD, which was semi-open in nature. The first version of the PDD placed a 7 interval line below the diagram of a physics problem, similar to the successive "frames" in a computer animation. It also used a 26-choice menu of problem types in an attempt to simulate Chi's card-sorting exercise with a multiple choice test. The choices were based on student interviews, both before the PDD was written and during initial testing of version 1.0.

Large group implementation of the PDD version 1.1 revealed serious problems with the test design, in that it placed too high a cognitive load on students. For example, the 26-option multiple choice seemed to baffle some subjects. However, some very compelling data were found in these test results, and the student responses were used to create distractors for a fully multiple choice version of the PDD, version 2.0.

The new PDD was then tested on two small classes over a summer term and further refined into the PDD v2.2, which has been implemented with some success. Additionally, a small number of interviews have been performed to help establish the validity of the items
on this test, and new items have been created and administered in open-response format to help expand the test. The next section of this chapter will discuss the reliability and validity issues that future chapters will attempt to establish.

Finally, closer analysis of the data from the PDD v2.2 was used to correct a number of the deficiencies in that instrument.

1.4 Reliability and Validity

Since these concepts will be addressed repeatedly in this paper, statistical reliability and test validity should be discussed at this point.

Reliability is a measure of how consistent a testing instrument is: does it measure the same thing every time? There’s two basic ways to measure reliability. The first is test-retest, in which the same test (or two similar tests) is given twice to either the same group or to two demographically identical groups. If the performance is roughly the same on both iterations, then the test is considered to be reliable in that sense. Of course, in the case of testing the same group twice, it is important that no significant treatment be given to that group that would affect test scores.

The other basic way to look at a test’s consistency is single-test reliability. This is usually measured by splitting a test into two or more sections randomly, then comparing those sections as if they were separate tests. However, this requires that the original test be fairly large. If the original test is too small, the resulting half-tests could suffer from two major problems. One, the likelihood of the two half-tests covering different material becomes higher, since all the items involving one concept or task could easily cluster in one of the halves. Two, reliability increases with the number of items, and a small half-test will have little statistical reliability no matter how well the halves match.
Next there is validity, which tells how well an instrument or item measures what it is supposed to measure. Conventional wisdom says that one cannot have good validity if the reliability is poor, since if the test doesn’t measure the same thing every time, it can’t be measuring the right thing every time. In recent years, convincing arguments have been put forth that it is possible to have high validity without high reliability (Linn, 12), but most of these arguments do not apply to a multiple choice test such as the later forms of the PDD, at least, not in its entirety. Individual items may be tested for validity without establishing reliability, but that will be discussed more in a later chapter.

Validity can be measured in many ways. Three common ways are face validity, construct validity and predictive validity.

Face validity can be defined as asking an expert, “Does this look like it tests what I want it to test?” In other words, does the instrument look reasonable? It is possible for an instrument to be valid without having good validity “on the face of it,” as some counter-intuitive instruments do exist. But establishing face validity is usually a good start.

Construct validity is somewhat more rigorous. The researcher must have some underlying theoretical construct on which the instrument has been built. The validity of the theory therefore lends itself to the instrument. Chi’s work mentioned in the previous section is one such theory that part of the PDD is built on, although the connection is somewhat tenuous.

Finally, predictive validity could also be called comparative validity. The researcher compares performance on his new instrument to performance by the same subjects on other, already accepted measures of ability. High correlations between measures help support the idea that the new instrument is as valid as the older one. This is often useful when the accepted measure is difficult to obtain compared to the newer measure, which can then serve as a substitute.
CHAPTER 2

THE ZONE TEST

The author's first attempt at a test of problem decomposition ability was the Zone Test. It was a fairly open-ended instrument that presented students with a diagram of a possible problem and then asked them to pick "significant points" in the diagram and define "zones" where relevant physical processes took place. At the time that these tests were written and administered (1995-6), the author did not have any formal training in test construction, and the Zone Test had several serious flaws. Despite these flaws, the Zone Test did yield useful information that affected the eventual development of the PDD.

2.1 Administration of the Zone Test

The Zone Test was administered in two different academic quarters. Both times, it was administered to students in the calculus-based introductory physics course at the Ohio State University, Physics 131. The test was administered during the laboratory sections both quarters. All Zone Test items may be found in Appendix B.
1-0: Pendulum-Cart Bash

motion detector

starts here

D

rises to here after impact

C

B

\[ \ell \]

\[ \theta \]

\[ \mu \]

\[ x \]

Solve for: \( \mu \)

A: PENDULUM RELEASED, STARTS FALLING
A-B: PENDULUM FALLS UNDER INFLUENCE OF GRAVITY
B: IMPACT BETWEEN PENDULUM AND CART
B-C: PENDULUM RISES TO NEW MAXIMUM HEIGHT, WILL KEEP SWINGING BACK AND FORTH TO THIS HEIGHT
B-D: CART, STARTED MOVING AT IMPACT, ROLLS ALONG UNTIL IT SLOWS TO A STOP DUE TO FRICTION

Figure 2.1 - Sample worked item from Zone Test 1.
For the first quarter of administration, ten items were prepared and split into two tests, Zone Test 1 and Zone Test 2. The two versions had similar instructions, but different sample items. The author's intention was to determine if a more complex worked example would help students perform better on the Zone Test, but due to complications with scoring and analysis, this issue was never resolved. Figure 2.1 shows the worked example from Zone Test 1.

Zone Tests 1 and 2 were randomly mixed together before administration so that students would be assigned a version of the test solely by whatever packet was at the top of the pile when they took their test. This resulted in 71 students taking Zone Test 1 and 124 students taking Zone Test 2.

Results of Zone Tests 1 and 2 resulted in a redesigned Zone Test that was administered to students the next quarter. All students in the lab section were given Zone Test 3, resulting in a sample size of 226 students.

2.2 Results:

While the sample size was large, flaws in test design (including not determining scoring criteria in advance) resulted in data that was difficult to meaningfully analyze. Part of this problem lay in the fact that students not only would omit important subproblems in their analysis, they would often create divisions that did not exist. Additionally, a significant number of students had difficulty following the directions of the Zone Tests: some tried to fully solve the problems, others turned in answer sheets that were undecipherable, or blindly tried to apply the exact zones of the example item to each subsequent item. The redesign for Zone Test 3 did largely eliminate student tendencies to
try to solve the problem (largely because Zone Test 3 did not provide the students with variables to use), but more than one in ten students still did not follow the directions, either because they misunderstood the directions or for other reasons.

The general thrust of student responses on all three versions of the Zone Test was, however, that when asked to break an item into subproblems, they were generally able to do so. In cases where the crude scoring system indicated student deficiency, the item was examined more closely to determine what might have been the reason for the difficulty. There were four elements that seemed to account for most of the student difficulties: changing slopes, ballistic arcs, springs and collisions. It is important to keep in mind that these results were very nebulous due to the flaws in test design. Exact numbers will not be presented because the author is not convinced these numbers have much meaning.

Changing Slopes: In many items, an object’s path changes slope abruptly, such as from a flat surface to an incline. The student tendency was often to neglect this factor so long as nothing else changed at the change of slope. However, since the flat surface involved was usually frictionless, that stretch of the item formed a trivial subproblem that may have been “chunked” in with the inclined section by the student.

Ballistic Arcs: In some items with objects moving through the air under the influence of gravity, students broke the symmetric ballistic arc into two half-arcs. However, as such problems can be meaningfully broken up that way, this “error” was discounted. In any case, decomposition of ballistic arcs into two half-arcs did not have a large effect on student scores.

Springs: Many items used springs either to get objects moving, to stop the object at the end, or both. All items involving springs clearly marked the distance over which each spring extended or compressed. Despite this, students had a marked tendency to consider
the spring's effect as acting only at a single point. The author dubbed this the "pinball paradigm," after the rapid and point-like interaction of a pinball machine's plunger and the pinball itself.

Collisions: Because of the way the test was scored, a student who did not consider a collision to be a "zone" of its own was not penalized as seriously as one who missed a zone that was extended in space and time. As a result, collisions did not generally affect final scores enough to be noticed in the analysis. However, over the course of scoring the tests, the author noticed repeated omission of the collision as a "zone," even when the collision was clearly inelastic. This agreed with examination of lab reports for an exercise similar to the item in Figure 2.1. Out of 39 laboratory groups in Physics 131, roughly one third correctly included the conservation of momentum in their analysis. The remainder either claimed that the collision was fully elastic (which it wasn't) or simply ignored it entirely without explanation. This result was repeated, on a more anecdotal level (copies of lab reports were not analyzed), during other quarters where the "Pendulum Box Bash" lab was given. Between the lab results and the Zone Test responses, the students seemed to be demonstrating a sort of blindness for collisions when they were part of a larger problem.

The Zone Tests accomplished three purposes in the context of the overall development of the PDD. First and foremost, the experience stressed the importance of careful test design. Second, the Zone Tests highlighted the difficulties of using a completely open test design for such an endeavor. Finally, the results, while only suggestive, pointed towards several areas that could be focused on more clearly in any later test design: slope changes, springs and collisions.

The next chapter will discuss the initial attempt by the author to create an objective test of problem decomposition.
CHAPTER 3

THE PROBLEM DECOMPOSITION DIAGNOSTIC VERSION ZERO

Some years after the administration of the Zone Tests, the author had developed the Toys In Motion laboratory course (Appendix A) and was seeking ways to measure the effects of the TIM course. While also examining conceptual knowledge through the Force Concept Inventory (Hestenes (1992a)), attitudes towards physics using the Maryland Physics Expectations Survey (Redish) and overall class performance through exam grades and final grades, the author wished to specifically test student problem-solving ability. It was decided that a test of problem decomposition would be useful and cover territory not measured by the other assessment tools being used.

The first version of the PDD was intended to be usable as both a pretest and a posttest, so that improvement could be gauged and differences in treatment groups could be dealt with. Because students do not all take physics before college, it was decided to attempt a PDD that did not require physics content knowledge. In other words, the language would be as natural as possible, and the subproblems determinable without recourse to specific physical concepts such as conservation of momentum. The PDD "version zero" used several situations in the form of stories, followed by short descriptions of parts of the story. Students were asked to then respond to each short description, designating it as either a single "step" in the situation, more than one step lumped together,
or something that wasn't even a complete meaningful step in the procedure. The entire test at its point of last development can be found in Appendix C, but here is the first story and the four items following it:

(Items 1 - 4)

You’re watching a movie in which an unpleasant main character’s life is changed after an automobile accident, and the scene with the accident is coming up. The “hero” is driving down a straight stretch of country road while talking on his cellphone, and doesn’t see the stop sign ahead. A driver on the cross road sees him and hits the brakes, but the two cars collide anyway. The twisted mass skids off the road and up an embankment where it comes to rest a short ways up the slope. Assuming this stunt was done all in one take, it’s possible to analyze the situation using physics, once it’s broken down into manageable parts.

Possible pieces:

1) Motion of the main character’s car from the beginning of the stunt until just before the impact.
2) Motion of the other vehicle from the beginning of the stunt until just before the impact.
3) The crash as both cars hit and lock together.
4) Motion of the combined mass of cars from just after the impact until they come to a stop.
If you look at item 3, you will see an implied inclusion of conservation of momentum, to see if students really thought collisions were important or not. If they thought a collision wasn’t an important thing to consider ever, the hypothesis was that they would answer “not a significant part” to item.

In addition to extended story problems, an alternate type of item was considered, in which short descriptions would be compared pairwise, asking students which item, if either, had more pieces to it.

However, this version of the test was never actually administered to students. During discussions with colleagues, the author realized that in its own way, this test was even more ambiguous than the open Zone Tests. Without a specific physics context, and given the rather hazy use of terms like “part” or “piece,” numerous correct interpretations were possible for any given situation. The pairwise comparison items might have reduced this problem somewhat, but did not go far enough towards eliminating it.

To give a concrete example, consider item 4 above. If interpreted in physical terms, that segment of the action could be broken into two subproblems. In the first, an object slides across a flat surface with friction. In the second, an object slides up an incline with friction. However, without bringing in physics concepts, you could correctly claim that only one thing is happening...the wreck is slowing down. Additionally, depending on how one uses conservation of energy (Van Heuvelen (1982), 142-3 has an alternative to the work-energy equation), it could be solved in a single subproblem using energy principles.

The result is, that by attempting to eliminate physics formalism, the items become too ambiguous to be used in an objective test. They would certainly be useful in an interview situation where the interviewer could probe the subject’s reasoning for various
choices, but as the PDD was intended to be a multiple-choice test, the ambiguity turned out to be a fatal flaw. The physics had to be included in order to make the test work as intended.

In short, this version of the Problem Decomposition Diagnostic lacked even face validity. Experts who examined it found numerous ambiguities and places where the instrument simply could not be counted on to measure anything.

Another result of the failure of the PDD v 0.5 (the final iteration in the version zero series) was to convince the author to spend some time on student interviews to gather qualitative data that might suggest a better format for the test. The next chapter will discuss the results of these interviews and the development of the PDD v1.0.
CHAPTER 4

EARLY DEVELOPMENT - STUDENT INTERVIEWS

This chapter concerns the results of interviews conducted with students in the process of developing the PDD version 1. Various tasks were presented to the students in these interviews, but the protocol remained essentially the same throughout. This protocol was granted an exemption by the Human Subjects Institutional Review Board of the Ohio State University as Protocol Number 98e0167.

The protocol was as follows. First, students who volunteered to be interviewed were given a form that outlined the basics of the research and assured them that more detailed information would be available at the end of the interview period. After responding that they had read and understood it, the subjects signed and dated the form. The subjects were then given a task to perform that was related to the development of the PDD. They were encouraged to speak aloud and describe what they were doing and the reasons why they were doing it. Tape recorders captured this verbal record, and the interviewer took notes during the process. The tapes were only used to clarify points during analysis, and have been stored securely. Once the student completed the task, the interviewer fully explained the reasons behind the task so that the subject was fully informed on the nature of their work. In most cases, volunteers were financially compensated for their time, and receipts for monies received were signed at the end of the
interview session. Performance on the interview task did not impact the grades of the subjects in any way, and the names of the students were not attached to any reporting of the results, in keeping with the privacy concerns of the protocol.

This phase of the research included three interview tasks and one small-group trial run that did not involve tape recordings or payment. They were as follows:

4.1. ActivPhysics (Van Heuvelen (1997)) simulation decompositions
4.2. Card Sorting (Two groups)
4.3. PDD version 1.0 development trial
4.4. PDD version 1.1 trial

4.1. ActivPhysics Simulation Decompositions

On April 22-23 of 1998, four students volunteered for this interview task. These students were all taking Physics 132, calculus-based electricity and magnetism. Two had been enrolled in laboratory sections using the Toys In Motion labs (Appendix A) while the other two had been enrolled in the Constructing and Applying the Concepts of Physics (Van Heuvelen (1995)) laboratory sections while taking Physics 131. Three subjects were male, one female. All interviews took place during the afternoon, with one student and the interviewer present.

Three computer animations were chosen from the ActivPhysics CDROM: Skier Into Cart, Pendulum Box Bash and Pendulum Person Bowling. ActivPhysics’s ActivPad application was used to create specific local webpages that allowed students to examine these three simulations in isolation from other ActivPhysics instructional material. Each simulation consisted of a computer animation with playback controls. Students could watch the animation play out, stop it at any time, step it forwards or backwards, or move to
specific frames. A slide bar at the bottom of the simulation also showed which frame in the animation was being displayed. Each animation had at least 60 frames, starting at zero.

"Skier Into Cart" involved a skier on a frictionless flat surface colliding with a cart attached to a spring. The inelastic collision resulted in both skier and cart compressing the spring. The three sub-problems involved in this simulation were: motion at constant velocity, inelastic collision and transference of kinetic energy to spring potential energy.

"Pendulum Box Bash" was a variation of the problem seen in the worked example of Zone Test 1 (Appendix B), but with a box instead of a cart, and with the pendulum stopping after hitting the box. The three sub-problems were: motion of the pendulum, the collision, and motion of the box as friction slowed it.

"Pendulum Person Bowling" was slightly more complex. A person on the edge of a cliff was struck by a pendulum and followed a ballistic path through the air to land on a massive cart on the ground below the cliff. The pendulum stopped after striking the person, as in the previous simulation. The four sub-problems were: motion of the pendulum, the collision of the pendulum and the person, the ballistic motion of the person and the collision between the person and the cart.

Students were provided with a worksheet with short verbal descriptions of each simulation and an area broken into four columns beneath that. Students were instructed to break each problem into "parts," where the term "part" was defined only loosely and left to the student to determine more fully. The first column on the worksheet was labeled "Part," and students were to number or otherwise label the parts into which they broke the problem on this column. The second column was labeled "Frames," and students were instructed to list the animation frames that composed or were included in that part. The third column asked the student to describe what type of problem that part was, and the final column gave the student room to explain his or her choices.
For example, if the student decided that the first piece of a problem took place during frames zero through twenty and was a conservation of energy problem, he or she would write “1” in the first column, “0-20” in the second column, “conservation of energy” in the third, and an explanation in the fourth.

During the course of the tasks, the interviewer would encourage the subjects to speak aloud about whatever was going through their heads, and took notes on such verbalizations in addition to tape recording the subjects. These notes and recordings were then analyzed, the summaries for each subject can be found in Appendix D.

Analysis of all student responses led to the following conclusions and suggestions for further work:

1. In these exercises, students often described the type of problem by the goal of that problem. For instance, a sub-problem that involved finding the velocity of a skier would be classified as a “velocity problem.” This suggested an additional classification of problem types beyond the “surface features” and “deep structure” of Chi’s card sorting research. The author’s own card sorting interviews sought to further explore this idea.

2. Students frequently misused physics terms, saying one word when the context suggested they meant another. For example, talking about force when the context is momentum or energy conservation: “The force goes from the pendulum to the box.” This suggested that an instrument should try to avoid vocabulary-based errors unless specifically designed to test for them.

3. Computer simulations and animations, in addition to requiring expensive equipment for administration, seem to bring with them a number of problems, including but not limited to: the illusion that an object is not moving before the beginning of the animation or after the end of the animation even if it is supposed to be, graininess of the time steps at lower speeds leading to the impression that objects are moving in fits and
starts, and the suggestion that time is a discretely divided quantity at the macroscopic scale. Additionally, the ActivPhysics simulations had specific aspects that led to student confusion, including the lack of audio or deformation cues on inelastic collisions and the way in which the motion of the slider bar confused students.

The difficulty with the time intervals suggested a solution for a paper and pencil test. Rather than using a large number of discrete times (as on the ActivPhysics animations) or an unlabeled continuum (as with the Zone tests), the problem could be broken into a small number of discrete intervals with labels between the boundary points. This would emphasize that the labels for the intervals referred to extended sections rather than merely to points. The PDD would incorporate a 7-interval system, see any Appendix containing a PDD v1 or v2 for examples of this.

The matter of which laboratory course (TIM or CACP) the student was taking did not have a noticeable impact on the results, possibly due to the very small sample size.

4.2. Card Sorting Exercises

Given the suggestion in the previous interview set of a student classification scheme that focused on the goal of a problem, the author designed a card sorting exercise to see if this behavior would be repeated. Rather than use the problems from Chi's study, the author created twelve problems that specifically took into account different problem goals. These twelve problems can be found in Appendix E. The problems were designed so that there were fairly clear rough groupings by surface features, by deep structure and by goal features. Students were not expected to stay exclusively with one type of grouping, however, and this initial division was mostly intended to assure a good variety of potential similarity-based groups.
For this exercise, the students were asked to sort the cards into groups based on similarity, then explain why each group belonged together. It was stressed that the students would not actually have to solve the problems. Students were encouraged to speak their thoughts aloud into a tape recorder during this exercise, but the tapes did not prove to be very useful in either trial of this interview task. The exercise was performed twice, each time with three student volunteers from calculus-based physics courses.

The following are the “intended” groupings for the cards by surface, goal and deep structure organizational schemes.

Surface: Cards 1, 3, 5, 6, 8 and 12 involve a golf club hitting a ball and sending that ball on a ballistic path. Cards 2, 4, 7, 9, 10 and 11 all involve an object sliding down a circular ramp. Because minor details do differ within each group, a student focusing on surface details may break the cards into more than these two groups.

Goal: Cards 1, 2, 3, 4, 8 and 10 ask the student to find a velocity. Cards 5, 6, 7, 9, 11 and 12 ask for a force. Four of the cards ask for an average force, while one asks for weight and one for normal force. Students using a broad goal orientation might end up with two groups.

Deep Structure: Cards 1, 7, 8 and 10 involve either conservation of momentum or the relationship between impulse and momentum. Cards 4, 6, 9 and 12 involve conservation of energy principles. Cards 2 and 11 use Newton’s Second Law, while cards 3 and 5 use geometric kinematics. Cards 2, 3, 5 and 11 could be considered to be “kinematics/dynamics” problems as a broader group.

A special note about card 6. Of all the cards, this is the only one with a problem markedly unlike anything the students might have encountered in homework or on tests, as it asks them to make assumptions about the force of air resistance. As a result, the author expected to see it standing alone in its own group at least some of the time, simply because
the students might not be able to figure out how to classify it. This turned out to be the case, as half of the students did assign card 6 to its own group, but not all for the same reasons.

Subjects S2-1, S2-2 and S2-3 performed the interview task on the afternoon of May 28, 1998. They were students in Physics 131 during the Spring quarter. Subjects S2-4, S2-5 and S2-6 performed the interview task on the afternoon of July 7, 1998. These students were enrolled in the Summer quarter section of Physics 132, and had all taken Physics 131 within the previous six months. The following is a summary of the results of these two runs. The groups will be presented in the following manner: Group number, then in parentheses the cards assigned to that group, then a brief explanation of why those cards formed that group.

S2-1: Female, took 30 minutes to complete the task.
1 (2, 3, 4, 6, 9, 11, 12) - Application of Newton’s Laws, find the forces
2 (7, 10) - Center of mass problems with similar solutions
3 (5, 8) - “Find the average force” problems
4 (1) - Conservation of momentum problem

S2-2: Male, took 55 minutes.
1 (5, 8) - Impulse/Impact-related problems
2 (4, 7, 9) - “Find acceleration” problems
3 (6) - Kinematics and ballistics problem, in a group by itself because the air resistance can’t be ignored
4 (1, 10) - Mass and Δx are provided, so you have enough information to use conservation of momentum
5 (3, 12) - Parabolic motion problems

6 (2, 11) - Simple Harmonic Motion problems. can be treated as pendulums even though they aren’t.

S2-3: Male, took 45 minutes.

1 (9, 12) - Problems with no friction and unknown gravity

2 (5, 8) - “Find Average Force” problems

3 (2, 11) - “Use Sum Of Forces” problems

4 (7, 10) - Conservation of momentum problems

5 (1, 3, 4) - Problems with elastic collisions that want information about the system right after the collision

6 (6) - “Find air resistance” problem

S2-4: Female, took 40 minutes

1 (6) - Problem where you need to deal with air friction

2 (3, 4, 9, 12) - Problems involving distance or displacement

3 (2, 11) - Problems have the same variables. use a force diagram

4 (1, 5, 7, 8, 10) - “Complicated” problems. Explanation then split the group further without specifying which card was in which subgroup. Apparently the student felt that the only common bond was that the problems were hard. However, this group does include all the conservation of momentum problems, and the student’s rambling explanation included the phrase “conservation or collisions.”
S2-5: Female, took 25 minutes.
1 (9, 11) - Solution requires kinematic equations
2 (3, 12) - Solution requires “projectile motion” equations
3 (2, 4) - Solution requires circular motion equations
4 (5, 6) - Solution requires use of “changes in energy” and energy conservation
5 (1, 7, 8, 10) - Solution requires conservation of momentum (note: as seen above, this is the correct conservation of momentum group.)

S2-6: Female, took 25 minutes.
1 (5, 6, 8) - Find the average force
2 (7, 9, 11) - Need to find forces using $\sum F_x$ or $\sum F_y$
3 (1, 3, 12) - Find the velocity of a projectile
4 (2, 4, 10) - Find the velocity on a ramp or of a ramp

As expected, students rarely replied simply in one mode. Because of the limitations of the protocol, it’s difficult if not impossible to say for sure if any of these groupings truly reflect a deep understanding of the material. For instance, the correct grouping of conservation of momentum problems could also be simply due to the surface feature of a collision or the “right equation” involving momentum. However, it was fairly clear when a student was categorizing the problems based on the goal of the problem. In fact, S2-6 was very clearly in this mode, with a bit of surface features added to break the cards into four groups. This exercise did accomplish the desired result of demonstrating that students would indeed categorize problems based on the goal of the problem rather than only on surface features or the underlying physics.
However, an additional “flavor” of classification also appeared in this experiment. S2-1’s group 2, S2-3’s group 3, S2-5’s groups 1 and 3 seemed to be related to goal-oriented classification, but instead grouped the problems based on what equations would be used to solve the problems. It is commonly accepted that students solve problems via means-ends analysis (Gick), but here it seemed that they were using the means and the ends as ways to describe the problems as well. In retrospect, this makes a great deal of sense: if the student is inclined to use means-end analysis to solve problems, then he or she will also sort problems based on the type of means-end analysis the problems will require. This behavior is consistent with the ontological belief on the part of students that physics is all about the equations, not about the underlying organization (Hammer. 158). The distinction between “find this variable” and “use this equation” was not initially recognized by the author, however, and development of the PDD version 1 primarily focused on surface features, goal orientation and deep structure. The emphasis on the “equation hunt” would, however, later informed development of the PDD version 2.

4.3. The Problem Decomposition Diagnostic version 1.0

Based on the results of the previous two interview tasks, the author devised a first draft of the Problem Decomposition Diagnostic version 1. Some inspiration was taken from the Toys In Motion labs (Appendix A) and the Zone Test tasks (Appendix B) for the items created for this test. Appendix F has this version of the PDD in full.
4.3.1 Description of the PDD v1.0

For each item on the PDD v1.0, the student was given two tasks. The first was to break the problem into sub-problems. A set of seven intervals was placed below the diagram for each item. At the bottom left hand of the page was a table, with the interval numbers in one column, and blanks for student responses in the other column. The student’s goal would be to put a number in each blank for every sub-problem that took place partially or completely within that interval. Thus, if sub-problem 1 ended in interval 3 and sub-problem 2 began in interval 3, the student would write “1, 2” in the blank across from interval 3 on the sheet. The first item on the test demonstrated this decomposition for the students.

The second task involved classifying the type of problem that each sub-problem was. The student was provided with a multiple choice list of 22 options, including such choices as “conservation of energy” or “find a distance.” These options were generated using the results of both the card-sorting and the ActivPhysics interviews. Students were told to select as many as they thought fit the sub-problem, then circle the one they felt best described how they would solve the sub-problem. Students were also told that the list was most likely incomplete, and that they should add their own options to the multiple choice sheet if they felt something was missing.

The reason for the long multiple choice list was to present a limited set of responses that could be interpreted by the scorer, yet also not steer students towards the correct decomposition in the first task by presenting them with only the classification choices appropriate to each item on the test. The concern was that if only 5-10 choices were listed on each page, students with an entirely different way of breaking the problem up might decide their solution was incorrect. Hence, the author desired a multiple choice list
comprehensive enough to cover any potential solution on the test. However, given the
difficulty interpreting the earlier Zone test, it was decided that completely open responses
here would be counter-productive. The primary concern regarding the multiple choice list
was that this would confuse the students, confounding them with too many choices to
meaningfully evaluate.

A scoring system was devised at this time, but as it was not applied until the large
scale administration of version 1.1, discussed in the next chapter.

The primary purpose of this administration was to determine if the format of the test
was feasible for use in a classroom setting. In other words, would the students be able to
understand the directions and complete the test in a reasonable amount of time
(approximately half an hour or less)? Could students make effective use of the large
multiple choice list? A secondary goal was to determine the nature of any missing items on
the multiple choice classification list.

Because the primary goal was to determine the usefulness of the test in a standard
classroom or laboratory setting, this was not a think-aloud protocol, and students were not
tape-recorded. They were, however, encouraged to write any questions or comments on
the test sheet. Four student volunteers from the summer quarter Physics 131 course
attended a single session on August 25, 1998 for a maximum of one hour. All completed
the exercise in twenty minutes or less.

Subjects S3-1, S3-2 and S3-3 were female. Subject S3-4 was male. No additional
demographic information was collected.
4.3.2 Results of administration

S3-1 did not add any items to the multiple choice list. She consistently picked only one descriptor for each sub-problem.

Item 1: Part 1 was classified by surface feature, part 2 by an inappropriate deep structure.

Item 2: S3-1 added an extra sub-problem in interval 1, labeled it as conservation of energy. She correctly identified the sub-problem of the pendulum’s swing, but classified it by surface feature. S3-1 omitted the collision, included the slide of the block, but classified it as conservation of momentum.

Item 3: Added two extra sub-problems in interval 1, classified as conservation of energy and constant acceleration kinematics. She identified intervals 1-6 as a single part, “object on a spring,” suggesting she misunderstood the set-up. Finally, she picked interval 7 as a single part, classified as conservation of energy.

Item 4: She chose essentially correct sub-problems, but cut sub-problem 1 off at interval 5 and only considered the final sub-problem to take place in interval 7. It’s uncertain why she did this, as she had no qualms about overlapping sub-problems in the previous items. She also correctly classified the sub-problems based on the deep structure.

Overall: This subject seemed inconsistent and uncertain. It’s possible that the large volume of instructions caused her confusion.

S3-2 added three choices to the classifications: “Kinetic energy,” “Potential energy,” and “Find position or distance in y direction (y components).” All three were judged to be goal-oriented choices. She consistently picked multiple classifications for each sub-problem and then circled her preferred choice or choices.
Item 1: S3-2 listed three options for sub-problem 1, then circled conservation of energy. She listed eight options for sub-problem 2, then circled both constant acceleration kinematics and the surface feature of projectile motion.

Item 2: Subject correctly divided the problem into sub-problems, but ended the pendulum swing in interval 2 and had the collision take place over both interval 2 and interval 3, possibly a result of misreading the diagram. She picked two options for sub-problem 1, two for sub-problem 2 and eight for sub-problem 3, but circled the deep structure choices in all three cases (conservation of energy on sub-problems 1 and 3, conservation of momentum for the collision in sub-problem 2).

Item 3: S3-2 identified two sub-problems taking place in intervals 1 and 2 respectively. Sub-problem 1 was classified as conservation of momentum, while sub-problem 2 was classified as “kinetic energy,” suggesting that she split the extension of the spring into two pieces. She identified sub-problems 3 and 4 as both taking place during intervals 2 through 7. Both were classified primarily as constant acceleration kinematics, but one also listed “find distance (x)” and the other listed “find distance (y).” This suggests she treated the mathematical parts of solving simultaneous equations in x and y directions as being two separate, but parallel sub-problems. In all four sub-problems, at least three classifications were indicated, with one primary choice circled.

Item 4: Like S3-1, she correctly broke the problem into sub-problems, but did not have the sub-problems correctly overlap in interval 6. Sub-problem 1 was given five classifications, but the surface feature of “object moves with friction” was the circled choice. Sub-problems 2 and 3 were each given three choices, and the correct deep structure was circled.

Overall: This subject’s responses suggest two things. One, that the directions to circle only one choice needed to be more explicit, and two, that simultaneous sub-problems...
would have to be dealt with by the PDD. The author decided that in the event of simultaneous parallel sub-problems, they would be treated as a single sub-problem for scoring.

S3-3 did not add any options to the multiple choice list. She sometimes only picked one classification for a sub-problem, but always circled one main choice.

Item 1: S3-3 picked two classifications for sub-problem 1, circled conservation of energy. She picked only one classification for sub-problem 2, the surface feature of projectile motion.

Item 2: She correctly identified sub-problem 1, but extended the collision of sub-problem 2 into interval 4. Sub-problem 3 was listed as taking place over intervals 4 through 6, ignoring motion in the final interval. Sub-problem 1 was classified only as a pendulum problem, sub-problem 2 only as a conservation of momentum problem, and sub-problem 3 was classified both as constant acceleration kinematics and "find a velocity," with the former circled.

Item 3: Sub-problem 1 was correctly identified, sub-problem 2 was listed as only taking place in intervals 2 through 6, with the final position in interval 7 not being part of any sub-problem. Sub-problem 1 was identified as conservation of energy and "object on a spring," with the former circled. Sub-problem 2 was identified as constant acceleration kinematics (circled), "find a velocity" and "find an acceleration."

Item 4: Sub-problem 1 was listed as only taking place in intervals 1 through 5, ignoring motion in interval 6 prior to impact. Sub-problem 2 covered intervals 6 and 7. The collision was omitted. Sub-problem 1 was classified with two options, constant acceleration kinematics being circled. Sub-problem 2 was given four classifications, with conservation of energy circled and no suggestion of a collision.
Overall: This subject followed the directions correctly, and her test paper was interpretable. While the actual responses were not always correct, this subject demonstrated that the test was not beyond the ability of a student to take.

S3-4 did not add any options to the multiple choice set. He always picked at least three classifications for each sub-problem, and always circled one of them as his preferred choice.

Item 1: S3-4 circled the deep structure classifications in both sub-problems.

Item 2: He omitted the collision, but did have sub-problems 1 and 2 overlap in interval 3. In both cases, he circled “conservation of energy” as the primary classification. Also, he included conservation of momentum in his list for sub-problem 2.

Item 3: S3-4 identified sub-problem 1 as only taking place in interval 1, while sub-problem 2 covered intervals 2 through 7. He classified sub-problem 1 as energy conservation, suggesting that he misinterpreted the diagram or did not consider the extension of the spring in interval 2 as being relevant to the sub-problem. He gave primary classification of “projectile motion” to sub-problem 2, although “constant acceleration kinematics” was included in his list.

Item 4: He omitted the collision as a separate sub-problem, but correctly identified the other two overlapping sub-problems. S3-4 classified both sub-problems as conservation of energy, but also included conservation of momentum in his list for sub-problem 2. This suggests that the collision may have been “folded” into the spring compression, following the pattern of some responses in the ActivPhysics interviews.

Overall: In two cases, he had two adjacent sub-problems classified as conservation of energy. This could suggest that while he may be comfortable with conservation of energy equations, he has not yet accepted the concept of a state process. If he had, the two
conservation of energy sub-problems would have been merged into a single conservation of energy problem. Additionally, in both of these cases, the second sub-problem's list of classifications included conservation of momentum. This suggests that while the student was aware of the collision, he didn’t think it needed to be treated separately, and could be merged with an adjacent sub-problem. This matches some of the behavior observed in the ActivPhysics interviews.

General Conclusions: From this limited sample, it appeared that the test met minimum criteria of administrability. There was some confusion regarding the instructions, so the next version was made more explicit on certain points. Four more choices were added to the multiple choice list ("potential energy" was split into spring and gravitational potential energy). It was decided that with minor modifications the PDD would be ready to use in a large classroom setting.

4.4. Problem Decomposition Diagnostic version 1.1 trial run

The opportunity arose in November 1998 to administer the PDD version 1.1 to a small group of students at the Marion branch campus of the Ohio State University. This administration was used to test some of the changes to the instrument, as well as to get more data on the parts that did not change.

4.4.1 Description of the PDD v1.1

One significant addition to the PDD for this version was a single survey question at the end of the instructions. Students were asked to rate how confusing they found the large
multiple-choice list, on a five point Likert scale, from "very confusing" to "not at all confusing." The intent of this scale was to look for a correlation between self-reported confusion regarding the instructions and demonstrated ability to follow the instructions. For this trial, of course, no correlations would be sought due to small sample size. But if all subjects ranked it as "very confusing," that would suggest the need for a redesigned test.

The other significant addition was item five, an entirely new problem on the test. Because all subjects completed the test in twenty minutes or less in the previous administration, it was decided to make the test a little longer, as there seemed to be room to grow.

A minor alteration based less on research and more on concerns for clarity was to alter the labeling of the sub-problems. Rather than label them with numbers, the instructions told students to label them with capital letters. This way, the intervals, sub-problems and classifications all had their own set of labels, to avoid some potential confusion.

The revised instructions and complete version 1.1 of the PDD can be found in Appendix G.

Four subjects were given the PDD v1.1 during an afternoon section of Physics 131 at OSU-Marion. They were told to work in groups of two and focus on trying to find things on the test that they felt needed changing. The author considered the "gang up" factor to be important in finding faults. Subjects were not recorded, and were given up to 45 minutes to complete the test. One group took half an hour, the other spent the entire 45 minutes. Groups G4-1 and G4-2 both consisted of two males of traditional college age. Subjects were told to individually answer the Likert scale survey.
4.4.2 Results of administration

G4-1: Confusion ratings of 3 and 4 (5 is "very confusing"). They selected a varying number of classifications for each sub-problem, and usually circled the favored choice. The deep structure classifications were never picked.

Item 1: G4-1 selected surface feature classifications for both sub-problems.

Item 2: They omitted collision and did not merge it into one of the other sub-problems. G4-1 correctly identified the other two sub-problems. They classified the motion of the pendulum by surface features and the ballistic motion by goal features.

Item 3: G4-1 provided correct identification of the second sub-problem, but listed the first as only occurring in interval 1. They classified sub-problem 1 as "find the spring potential energy," and sub-problem 2 by surface features.

Item 4: They correctly decomposed the sub-problems. G4-1 did not pick between the two classification choices on sub-problem 1 (goal and surface classifications). They did not classify the collision as involving conservation of momentum at all, rather listing it as a "find the kinetic energy" problem. They classified sub-problem 3 by surface features.

Item 5: G4-1 ignored the motion of the ball up the ramp and declared this to be a single part problem involving the motion of the ball through the air, classified by surface feature. They redrew the ball as a motorcycle, suggesting they at least considered tying the problem to their physical intuitions.

Overall: While they expressed some confusion over the directions, this group was able to follow them quite well. They totally avoided conservation principles in their classifications. A possible explanation is that conservation of momentum had just been discussed that day in class, and these students were not yet comfortable enough with the concept to want to use it.
G4-2: Subjects self-assessed confusion levels at 3 and 4. However, they did not realize they needed to circle their preferred choice and were told to go back and do so before handing the test in. For each sub-problem a large number of classifications was listed. The group took great pains to include any classification that seemed even marginally relevant, in part explaining why this group spent 45 minutes.

Item 1: G4-2 circled surface feature classifications and did not even list the deep structure classifications.

Item 2: They correctly identified the three parts of the problem. G4-2 circled surface features for sub-problems 1 and 3, but included and circled conservation of momentum for sub-problem 2. However, it is difficult to determine if a student means conservation of momentum as a deep structure or a surface feature, since most conservation of momentum problems share the surface feature of collisions.

Item 3: G4-2 added an additional sub-problem in interval 1 and another in interval 2, but included the two correct sub-problems as well. They classified sub-problem 1 by the surface feature of the spring, suggesting a “pinball paradigm” of seeing springs acting only at one point. They did, however, include conservation of energy in their long list for sub-problem 1. They classified sub-problems 2 and 3 by surface features, but also included option c (kinematics) in their lists. Sub-problem 4 was classified as a conservation of momentum problem, which would be correct if the tray’s motion were of interest.

Item 4: G4-2 displayed correct decomposition of the item into sub-problems. As with item 1, the students classified the collision correctly, but classified the other sub-problems by their surface features. Deep structure options were included in all three lists.
Item 5: They correctly decomposed the region of interest. They also added a third sub-problem for motion of the ball after it had landed. All sub-problems were classified by surface features, but deep structure options were included in the lists.

Overall: This group took the directive to be thorough seriously, but still missed an important part of the instructions. Like G4-1, this group did not seem comfortable with conservation principles, but they at least listed such classifications in their comprehensive answers.

Conclusions: At least in this small sample, confusion did not seem to be an insurmountable problem, although the decision was made to have instructors administering this test verbally repeat parts of the instructions to help with clarity. Collisions do seem to present a difficulty in scoring, because it is difficult to tell if a student is picking "conservation of momentum" for deep structure reasons, or because of a belief that all impacts involve conservation of momentum, making it more of a surface feature.

It was decided, based on this sample, to proceed with administration of the PDD version 1.1 to a large sample at the end of the Autumn quarter at the Ohio State University. The next chapter discusses this administration and the analysis of student responses and scores.
CHAPTER 5

LARGE-SCALE ADMINISTRATION OF THE PDD VERSION 1.1

In the Autumn quarter of 1998 the Problem Decomposition Diagnostic was administered for the first time to a large number of students. This chapter will discuss the administration of the test and the results gained. It will be broken up into six sections:

5.1. Discussion of the method chosen for scoring the PDD v1.1
5.2. Description of the subject groups
5.3. Identification of other measurements made on the subject groups
5.4. Discussion of sources of error
5.5. Results of scoring, additional scoring methods devised after data collection
5.6. Suggestions for further development of the PDD

5.1. Scoring the PDD v1.1

See Appendix G for a copy of the PDD v1.1. Prior to administration of the PDD, a scoring system was devised, consisting of five scores: Decomposition score, Classification score, percent Deep Structure, percent Surface Features, and percent Goal Oriented. The first two numbers were intended for use as a rough score of student ability, with the remaining three numbers providing a bit more detail.
For the Decomposition score, the student was given 2 points for each correct sub-
problem listed on the left-hand side of their answer sheets. To be counted as correct, there
had to be a sub-problem listed with exactly the correct intervals listed. In this way, both
combining multiple sub-problems into one and breaking a single sub-problem into multiple
parts would result in loss of points. However, the penalty was greater for failing to
decompose (4 points lost for combining two sub-problems) than for excessive
decomposition (2 points lost for breaking a single sub-problem into multiple sub-
problems). This was considered acceptable by the author, as excessive reductionism could
more easily be dealt with in general problem-solving than insufficient decomposition.
However, another consequence of this scoring system was that a student who shifted the
boundaries of a sub-problem (for instance, thinking that the pendulum stopped swinging in
interval 2 on the pendulum-box-bash item) would lose points despite an essentially correct
decomposition. This would need to be considered in any discussion of error.

The Classification score looked only at fully correct sub-problems. If a student
picked a relevant “deep structure” answer as the primary classification of the sub-problem,
then 2 points were awarded. If the list of classifications contained the deep structure
answer, but a different classification was circled, then 1 point was awarded. And if an
appropriate deep structure option did not appear at all in the student’s classification, zero
points were given. As a result, the student’s Classification score would always be less
than or equal to his or her Decomposition score.

12 correct sub-problems could be found in the PDD v1.1, so the maximum
Decomposition and Classification scores were 24. The first item showed the correct
decomposition as an example, and 4 points were given to the student so that the
Classification score could not exceed the Decomposition score. Decomposition scores
could range between 4 and 24, while Classification scores could range between 0 and 24.
The remaining three scores were intended to give a slightly more detailed picture of student classification patterns, although the intent would always be to look at individual test responses for the complete picture. For these scores, no consideration was made regarding the correctness of decompositions. Instead, the student's preferred classifications for each sub-problem (whether or not the sub-problem was correct) were grouped into Deep Structure, Surface Features and Goal Oriented. Once all three groups were totaled, a percentage was found for each. The intent here was to get some feel for the student's overall preferences for problem classification. For instance, a student with 90% Deep Structure would be a qualitatively different problem-solver than one with 90% Surface Features. Percentages were used rather than raw scores in order to normalize results and meaningfully compare, for instance, students who over-decomposed with those who under-decomposed.

In summary, three tiers of scoring were envisioned. At the simplest level, an instructor could just look for the Decomposition and Classification scores. For example, 20/10, meaning 10 correct sub-problems which were not strongly classified by deep structure, but the student recognized some deep structure aspects. The next level would look at student tendencies towards using the three types of classification, such as 50%/20%/30%. Such a result would mean the student picked deep structure classifications about half the time, went for surface features one in five times, and otherwise focused on the goal of the sub-problem. Finally, instructors wishing more detailed information could look at the test sheets themselves to see what specific responses the student had made. The hope of the author was that the simpler scales would prove sufficiently valid and useful measures.
5.2. Subject groups

Three different groups of students were tested using the PDD at the end of Autumn quarter 1998 at The Ohio State University. These students came from three different versions of the introductory calculus-based physics class: regular Physics 131, honors Physics H131 and engineering honors Physics 131E.

Due to limitations in the time available to administer testing instruments to the regular Physics 131 class (henceforth to be identified simply as 131), it was not possible to administer both the PDD and the Force Concept Inventory to a given laboratory section of 131. Thus, three sections were given the PDD and the remainder (who are not counted in this study) were given the FCI. FCI scores were desired for reasons unrelated to the development of the PDD, so it was not possible to reduce the number of lab sections taking the FCI by much.

Students in regular 131 are typically engineering majors, although a number of other majors do require 131. Students in 131 were subjected to a relatively traditional mode of instruction, although their laboratory sections used the Constructing and Applying the Concepts of Physics (Van Heuvelen (1995)) lab manual. One of the lab exercises performed by 131 students was virtually identical to the pendulum-box-bash item on the PDD, and students had been graded on their lab reports for that exercise before taking the PDD. 52 students in 131 took the PDD, 19 followed the instructions to the point that full analysis of their scores was possible.

Students in honors Physics H131 (to be henceforth identified simply as H131) are primarily physics majors, but a number of computer science majors also take this course. Instruction was fairly traditional but intensive. H131 students used the Toys In Motion lab
course (Appendix A), and had several class periods devoted to the action of springs. 41 students in H131 took the PDD, 22 followed the instructions well enough that full analysis of their scores was possible.

The Physics 131E course was designed for freshman engineering honors students, who also took courses together in Calculus and Engineering Graphics as part of their honors program. Physics 131E instruction drew from a great number of Physics Education Research theories and results, including a special demonstration exercise involving elastic and inelastic collisions (Zou). 131E students also took the Toys In Motion lab course. 93 students in 131E took the PDD, and 56 followed instructions well enough that full analysis of scores was possible.

It should be noted that with only a handful of exceptions, all student responses could be analyzed at some level, especially the Springs/Collisions analysis which will be discussed in section 5.5. But the scoring system discussed in section 1 could only be applied to the smaller number of students listed for each group.

5.3. Other measures taken of the subjects

The students involved were tested in many ways, with the hopes of being able to find correlations between various scores on the PDD and these other measures of ability. The following measures were studied for potential correlations. The first two were directly involved with the PDD, but were not part of the scores explained in section 1.

1. Followed Instructions: As mentioned earlier, not all students followed the instructions well enough that their responses could be fully scored in terms of Decomposition Score and Classification Score. This variable was dichotomous, a "yes/no" on whether the student's responses were scorable.
2. Confusion Level: In Chapter 4, a Likert scale was introduced. Students self-reported their level of confusion regarding the instructions on a scale of 1 to 5. The author was interested in seeing if those reporting more confusion performed differently on the PDD than those who reported less confusion. Especially interesting would be whether students reporting more confusion were less likely to follow the instructions correctly.

3. Final Exam Grade. Each of the three groups had their own final exam. Only the 131E final exam involved any serious decomposition tasks, so this was taken more as a general measure of problem-solving skills, rather than of problem decomposition skills. Of course, as discussed in chapter 1, it's possible that exam scores do not necessarily reflect higher order problem-solving skills.

4. Final Course Grade. This was examined as a measure of the student's overall ability in physics, to see if there were any correlations with "general physics ability." Again, each course had its own grading methods.

5. The Force Concept Inventory (Hestenes (1992a)). The FCI is a commonly-used instrument to test conceptual understanding of Newtonian mechanics in introductory physics. The test was administered both as a pretest at the beginning of the term and as a posttest at the end. As mentioned earlier, the 131 students who took the PDD did not take the FCI. As the FCI is a measure of conceptual understanding and not problem-solving ability, a strong positive correlation with any PDD score was not necessarily expected. However, a strong negative correlation could indicate that students with strong comprehension (Bloom's level 2) might be scoring poorly on the PDD because their problem-solving skills were primarily based on knowledge and memorization of patterns.

Due to time constraints no further tests were administered to the students in this study.
5.4. Discussion of sources of error

The primary source for error in scoring the PDD involved the instructions themselves, and that will be discussed first. However, individual scores did have their own sources of error, and the tests in section 3 also contain some error. Due to the high level of variability possible and the fact that a specific model was not devised to describe incorrect answers, a quantitative error analysis will not be performed, nor was statistical reliability computed. However, measures for dealing with the uncertainties in scores will be discussed.

The instructions, a page of text followed by a page of multiple choice options, were judged to place a high cognitive load on students. The fact that only 54% of students overall followed the instructions well enough that their papers could be scored certainly indicates problems with the test. Additionally, it is entirely possible that some of that 54% probably misunderstood the instructions in smaller ways that still allowed their papers to be scored. This source of error is likely to result in a random noise in the Decomposition and Classification scores, as well as in the three classification percentages.

As mentioned earlier, the Decomposition score does not necessarily penalize students in a way that accurately reflects the severity of their error. This is especially important in the case of what the author calls "boundary problems," where the student either starts or ends a sub-problem in the wrong interval, but does indicate the correct sub-problem qualitatively. Therefore, it is possible for a student to have a good idea how to decompose all the items on the PDD, but still get a low score due to boundary problems.

Additionally, the Decomposition score is deliberately rough. It does not distinguish among the three main student mistakes: over-decomposition, under-decomposition and boundary problems. Thus, correlations might be masked by the roughness of the scales.
Likewise, the Classification score has some sources of uncertainty. Despite the research done in interviews (Chapter 4), some of the multiple choice options might be interpreted in more than one way. The primary example being conservation of momentum, which might be strongly linked with the surface feature of a collision and thus picked as Surface Feature but scored as Deep Structure. And, similarly to the Decomposition score, the roughness of this scale can mask particular student difficulties.

The classification percentages suffer as well from the first difficulty mentioned above for the Classification score. It is possible to have an inflated Deep Structure percentage if certain surface features get attached to the deep structure options.

The self-reported confusion scale is subject to a fair amount of "fuzziness," both due to being self-reported and due to the nature of Likert scales. One student's rating of 4 may actually represent less real confusion than another student's rating of 3, for instance. The data from this scale should be treated with a great deal of caution as a result.

All exams and final grades are subject to the normal variability of such measures. Additionally, the fact that each group has its own exams and its own grading procedures means that a correlation found in one group does not necessarily mean one should find a correlation in another group. Each test group must be considered separately, because they are separate tests.

Finally, the FCI is probably the most stable of the instruments used, having gone through many years of development. There will be a small amount of random noise from students who do not take the test seriously, but the error variance on the FCI will probably be less than on any other measure in this study.

In addition to all these instrument-specific sources of error, there is also the general error resulting from student attentiveness and commitment to the task. A number of test sheets did show signs of deliberate "doodling around" by students on the PDD, and as
mentioned, it is not impossible for the FCI to have similar problems. Because of the grades attached to them, however, exam scores and final scores are expected to be the result of students taking their work seriously.

The following are some of the steps taken in order to deal with these sources of error. The first, of course, is to take any results from this stage of the research as primarily advisory or suggestive rather than as proof of anything. The proverbial “grain of salt” is to be taken. Second, in cases where a correlation is found, no correlation of absolute magnitude less than 0.3 will be considered even suggestive. Even in cases where a t-test would mark as statistically significant a much lower correlation, this limit is hewn to. However, in some cases statistical significance requires a correlation with magnitude 0.4 or 0.5. When discussing results for such cases, the disclaimer will be brought up. Finally, when examining differing means, each case will be examined separately to determine if the difference is large enough to overcome the noise generated by the various sources of error. Hypothesis testing will also be performed, but will not be considered the sole criterion for acceptance of results as noteworthy.

More quantitative methods of error analysis, such as examination of the standard deviation of error, do exist. However, given the nature of this data and its sources of error, such rigorous error analysis methods may not offer any particular advantage over simpler ways of dealing with error, especially in light of the caution to take this data as more advisory than confirmatory.

5.5. Results

Table 5.1 summarizes the results of scoring the PDD v1.1.
<table>
<thead>
<tr>
<th>Quantity</th>
<th>131</th>
<th>131E</th>
<th>H131</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of tests taken</td>
<td>52</td>
<td>93</td>
<td>41</td>
</tr>
<tr>
<td>Number of tests scorable</td>
<td>19</td>
<td>56</td>
<td>22</td>
</tr>
<tr>
<td>Self-reported Confusion Mean</td>
<td>3.5</td>
<td>3.3</td>
<td>3.5</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.2</td>
<td>1.2</td>
<td>1.0</td>
</tr>
<tr>
<td>Decomposition Score Mean</td>
<td>17.5</td>
<td>19.2</td>
<td>18.7</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>3.6</td>
<td>3.7</td>
<td>2.3</td>
</tr>
<tr>
<td>Classification Score Mean</td>
<td>4.9</td>
<td>11.1</td>
<td>8.2</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>4.3</td>
<td>6.1</td>
<td>4.5</td>
</tr>
<tr>
<td>Deep Structure Percentage Mean</td>
<td>32%</td>
<td>53%</td>
<td>38%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>28%</td>
<td>30%</td>
<td>31%</td>
</tr>
<tr>
<td>Surface Feature Percentage Mean</td>
<td>47%</td>
<td>25%</td>
<td>41%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>30%</td>
<td>18%</td>
<td>31%</td>
</tr>
<tr>
<td>Goal Oriented Percentage Mean</td>
<td>21%</td>
<td>22%</td>
<td>21%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>23%</td>
<td>27%</td>
<td>22%</td>
</tr>
<tr>
<td>FCI Pretest Mean (all students)</td>
<td>n/a</td>
<td>62%</td>
<td>67%</td>
</tr>
<tr>
<td>FCI Posttest Mean (all students)</td>
<td>n/a</td>
<td>81%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Table 5.1 - Summary of PDD v1.1 scores
5.5.1 Group means

As can be seen from the scores in Table 5.1, there is little to no difference between groups when it comes to the Decomposition score. Given that this test should be able to discriminate between groups as different as these, that suggests that the method of scoring decomposition skills is not effective. Classification scores do show more variation, as do percentages of selection of deep structure or surface feature strategies. However, given that this part of the test was only scorable on 97 out of 186 test sheets, that strongly indicated that this too needed serious changes. Self-reported levels of confusion seemed to be roughly the same for all groups, with the values reported in Table 5.1 coming solely from students whose tests were scorable. Table 5.1 does not include exam grades or final course grades, because those are not equivalent measures across groups. The FCI scores listed are for all students, not just those whose tests were scorable. At the time the PDD was administered (concurrent with the FCI), 131E and H131 were comparable groups in terms of what the FCI measures.

Of course, as per the earlier caveats, none of these numbers should be trusted too heavily. In addition to considering means and standard deviations, it is a good idea to look at histogram plots of student scores. If the graphs all have a similar appearance, that suggests that the scoring method does not do a good job of distinguishing between the groups. However, should one group’s histogram look markedly different, then perhaps the scoring method is useful, and it is simply the use of mean scores that doesn’t suit the data.
5.5.2 Score distributions

Figures 5.1, 5.2 and 5.3 show these histograms for 131, 131E and H131 respectively. The top graph in each case is for the Decomposition score, and the bottom graph is for the Classification score. On the top graphs, note that part of the area is shaded. Because students were given 4 points for the example decomposition, no scores under 4 are possible on that score, and the shaded area emphasizes that point.

On examining these graphs, it is clear that in Decomposition there is very little qualitative difference between the groups. All histograms have a single mode and similar shape, although H131's is a bit more sharply peaked. The Classification graphs differ both in shape and where the peaks (if any) are located. These observations support the conclusions gained by examination of the means, that the Decomposition score is not useful, and the Classification score may be, difficulties with the instructions aside.
Figure 5.1 - PDD v1.1 scoring histograms for Physics 131 students. N=19
131E PDD Results: Decomposition

131E PDD Results: Classification

Figure 5.2 - PDD v1.1 scoring histogram for Physics 131E. N=56
Figure 5.3 - PDD v1.1 scoring histogram for Physics H131. N=22
5.5.3 Correlations

The next step in examining the scores on the PDD was to look for correlations with the other measures taken. It was possible that, for instance, the similar Decomposition scores were the result of the groups performing differently on measures that themselves had different correlations with the Decomposition score. Additionally, a strong correlation between any score on the PDD and scores in other areas would suggest that the score in question had some validity and should be retained in some form in the next version of the test. Correlation analysis was also performed to answer questions about the format of the test itself. Both self-reported confusion and whether the student followed the instructions were measures of the cognitive load placed on students by the construction of the PDD. A valid question could be asked as to whether there was any link between these measures and, for example, overall performance on the course. In other words, were students who were more successful in the course overall also more likely to understand the directions (low confusion) and follow them?

Tables 5.2, 5.3 and 5.4 list the correlations between measures for 131, 131E and H131 respectively and means tests for those groups. Underlined correlations are considered significant. Blank spaces represent correlations that were not measured, either because there was no interest or because the scores as defined were entangled with each other.
### a) Correlations

<table>
<thead>
<tr>
<th></th>
<th>Ex</th>
<th>Fi</th>
<th>Con</th>
<th>Dec</th>
<th>Class</th>
<th>D%</th>
<th>S%</th>
<th>G%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exams</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Grade</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion</td>
<td>0.31</td>
<td>0.30</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decomp.</td>
<td>-0.33</td>
<td>-0.24</td>
<td>-0.48</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classify</td>
<td>0.43</td>
<td>0.49</td>
<td>0.13</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep %</td>
<td>0.47</td>
<td>0.48</td>
<td>0.35</td>
<td>-0.23</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface %</td>
<td>-0.43</td>
<td>-0.44</td>
<td>-0.40</td>
<td>0.36</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal %</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.19</td>
<td>-0.20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### b) Means

<table>
<thead>
<tr>
<th></th>
<th>Ex</th>
<th>Fi</th>
<th>Con</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scorables</td>
<td>128</td>
<td>443</td>
<td>3.44</td>
</tr>
<tr>
<td>Unscorable</td>
<td>120</td>
<td>431</td>
<td>3.53</td>
</tr>
<tr>
<td>z-score</td>
<td>0.77</td>
<td>0.50</td>
<td>0.23</td>
</tr>
<tr>
<td>Significant?</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 5.2 - Correlations and means tests for 131. a) Correlations, b) Difference in means for students with scorable and unscorable tests
a) Correlations

<table>
<thead>
<tr>
<th></th>
<th>Ex</th>
<th>Fi</th>
<th>FCIp</th>
<th>Con</th>
<th>Dec</th>
<th>Class</th>
<th>D%</th>
<th>S%</th>
<th>G%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exams</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCI posttest</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion</td>
<td>-0.07</td>
<td>-0.05</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decomp.</td>
<td>-0.01</td>
<td>0.03</td>
<td>0.12</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classify</td>
<td>0.64</td>
<td>0.67</td>
<td>0.46</td>
<td>-0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep %</td>
<td>0.50</td>
<td>0.50</td>
<td>0.29</td>
<td>-0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface %</td>
<td>-0.31</td>
<td>-0.24</td>
<td>-0.14</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal %</td>
<td>-0.35</td>
<td>-0.40</td>
<td>-0.23</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) Means

<table>
<thead>
<tr>
<th></th>
<th>Ex</th>
<th>Fi</th>
<th>Con</th>
<th>FCIpost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scorable</td>
<td>118</td>
<td>85%</td>
<td>3.34</td>
<td>24.5</td>
</tr>
<tr>
<td>Unscorable</td>
<td>121</td>
<td>87%</td>
<td>3.37</td>
<td>24.2</td>
</tr>
<tr>
<td>z-score</td>
<td>-0.48</td>
<td>-0.65</td>
<td>0.27</td>
<td>-0.10</td>
</tr>
<tr>
<td>Significant?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 5.3 - Correlations and means tests for 131E. a) Correlations. b) Difference in means for students with scorable and unscorable tests
### a) Correlations

<table>
<thead>
<tr>
<th></th>
<th>Ex</th>
<th>Fi</th>
<th>FCIp</th>
<th>Con</th>
<th>Dec</th>
<th>Class</th>
<th>D%</th>
<th>S%</th>
<th>G%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exams</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Grade</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCI posttest</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confusion</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.26</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decomp.</td>
<td>-0.02</td>
<td>-0.14</td>
<td>-0.16</td>
<td>-0.09</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classify</td>
<td>0.22</td>
<td>0.13</td>
<td>-0.31</td>
<td>-0.20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deep %</td>
<td>0.36</td>
<td>0.29</td>
<td>-0.19</td>
<td>-0.28</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface %</td>
<td>-0.15</td>
<td>-0.23</td>
<td>0.15</td>
<td>0.15</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goal %</td>
<td>-0.28</td>
<td>-0.08</td>
<td>0.08</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### b) Means

<table>
<thead>
<tr>
<th></th>
<th>Ex</th>
<th>Fi</th>
<th>Con</th>
<th>FCIpost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scorable</td>
<td>67%</td>
<td>73%</td>
<td>3.45</td>
<td>25.5</td>
</tr>
<tr>
<td>Unscorable</td>
<td>60%</td>
<td>65%</td>
<td>3.72</td>
<td>23.2</td>
</tr>
<tr>
<td>z-score</td>
<td>1.33</td>
<td>1.73</td>
<td>3.00</td>
<td>1.54</td>
</tr>
<tr>
<td>Significant?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 5.4 - Correlations and means tests for H131. a) Correlations, b) Difference in means for students with scorable and unscorable tests

62
As the above data demonstrates, while there were certainly some noteworthy correlations between scores, they were mostly confined to the classification measures. Self-reported confusion did not generally correlate with anything strongly, nor did whether a student followed directions seem to imply anything about their performance on non-PDD measures. For the next subsection, the results for each class will be examined a little more closely.

131: This group had the smallest sample size, making it more difficult to consider correlations as statistically significant (and giving low post-hoc statistical power). The four correlations marked in Table 5.2 as significant (at the P<0.05 level) are barely above the minimum values required for statistical significance. Given the sources of error mentioned earlier, these correlations must therefore also be considered with a grain of salt. Additionally, they did not take the FCI, reducing the data pool.

Most of the higher correlations for 131 involved the choice of deep structure classification (represented by both the Classification score and the Deep Structure Percentage) correlating with the final exam score (out of 150 points) and overall grade (out of 600 points). This suggests that in 131 students who were more comfortable with the “expert-like” knowledge structures of physics, as evidenced by their higher Classification scores and Deep Structure percentages, were also more likely to succeed on the final exam and in the course overall. However, it must be kept in mind that on version 1.1 of the PDD, it is not stressed that the student’s personal answer is sought rather than the “right” (professor’s) answer, and successful students may simply have been more adept at divining what sort of answer was expected of them.

The other moderately strong correlation was between Decomposition score and Confusion rating. Students who were more confused by the instructions were less likely to
do well on the decomposition task. Oddly, they were not significantly more or less likely to do well on the classification task, which would seem to be where most of the difficulty of the instructions lay. On the other hand, this data came solely from students whose test sheets were scorable, so they had apparently managed to decipher the instructions well enough to follow them. Therefore, another mechanism may be at work, or the correlation may simply be an unlikely result of random chance.

Considering the division of the class into those able to follow the instructions and those whose test sheets were unscorable, there was no significant difference between these groups in exam scores, final grades or self-reported confusion. This suggests that for this class, trouble with the instructions is not particularly tied to their facility with the subject matter or their own perception of their ability to comprehend directions.

131E: This group had the largest sample size and was more closely examined as a result of being the subject of extensive treatment by the Physics Education Research Group at The Ohio State University. Statistical significance for correlations is strictly obtained at levels lower than the 0.3 standard the author imposed on the data.

Self-reported confusion did not correlate with any other measure in this group, with the strongest correlation being $r = 0.14$, neither statistically significant nor large enough to worry about. Decomposition scores likewise had very low and insignificant correlations, although the overall high mean Decomposition score for 131E (19.2 out of a possible 24) implies a lack of variance with which to covary.

The Classification score correlated extremely strongly with other measures for this group. As with 131, this could simply indicate that the successful student was better able to determine the desired "opinion". However, the 131E exams placed a large emphasis on higher order skills within Bloom's Taxonomy, so it is also possible that the successful
student better understood the material and hence was also more likely to pick a Deep Structure option on the PDD. Successful students were also less likely to pick Surface Features or Goal Oriented options, suggesting that as these students improved they abandoned both non-expert strategies in favor of a more expert-like one, rather than shifting between the non-expert strategies. Additionally, one of the final exam items was an explicit problem decomposition task, with students graded on breaking the situation into sub-problems and correctly identifying a strategy for solving each sub-problem. This would help explain some of the correlations, as students were given similar tests.

However, once again there was no real difference between the students who followed instructions and those who did not. Final exam scores (out of 150), cumulative grade percentages, self-reported confusion ratings and FCI posttest scores had virtually identical means for both groups.

**H131:** This was another smaller sample, and in this case none of the correlations rose high enough to be considered statistically significant. In fact, only a handful reached even the $r = 0.3$ level needed to be considered large enough for any consideration. It is possible that this group was fairly homogeneous in performance, although its lower average exam and overall course scores suggest more room for variation than there was in 131E.

However, in further examining the course materials and exams, it also appears that problem decomposition skills were very weakly covered, if at all, and that appropriate deep structures were tightly tied to particular surface features or equation cues. For instance, while each item on the final exam had at least a half dozen sub-problems, each sub-problem was clearly labeled and presented separately. With one exception, each lettered item
involved picking a single appropriate equation, solving it for the desired quantity and then putting in known quantities. The one exception required solving a system of two linked equations.

However, in contrast to the lower correlations, the H131 students were more clearly broken into two groups by the question of whether the PDD test sheets were scorabled. The absolute differences between groups were all larger than in the other two classes, and one even attained statistical significance. Students with unscorable sheets were significantly more confused by the instructions, according to their own reporting.

5.5.4 Examination of alternatives

The above results were certainly interesting, and continued to suggest that it was worthwhile to incorporate some sort of “card sorting task” variation in the PDD, but they clearly demonstrate that the Decomposition score does not do a very good job of measuring student ability to decompose complex problems (low validity). Its primary weakness is that it does not discriminate well. Three very different groups of students had mean Decomposition scores within about a two point spread...the difference of one sub-problem. And this lack of spread meant that correlations would be difficult to find. After all, data from a restricted range rarely correlates well with data from other sources: you need not know a person’s score on other instruments to be able to predict roughly where it will lie on the Decomposition score.

Additionally, it was fairly clear that the PDD v1.1 needed serious revision to avoid the serious problem of unscorable tests. That scorability did not seem to relate to any other
measure suggested that the data obtained from scorable tests was reasonably representative of the larger sample of all tests, but a nearly fifty percent unscorable rate would not be acceptable on a test intended for wide use.

Therefore, the raw data was reexamined with two goals in mind: 1) find some way to better discriminate on student decomposition ability, and 2) reduce the cognitive load of the instructions. These goals led the author back to suggestions from the earlier Zone Tests: look at collisions and springs. If considering just those sub-problems led to better discrimination, then a test could be designed to specifically check for those situations and others like them. Without the need to examine student decomposition of the entire item, a more focused and less confusing test could be designed.

5.5.5 Springs and collisions:

All test sheets were reexamined at this stage in the research to determine student success at identifying sub-problems having to do with springs and with collisions. It was decided that \textit{false mastery} errors were preferable to \textit{false non-mastery} errors in this case, so scoring was made as lenient as possible in order to give students the benefit of the doubt. The following criteria were used.

Springs: Student had to indicate a sub-problem covering both of the intervals in which the spring extends or compresses, but without also indicating a sub-problem taking place in just one interval that is clearly identified as being a spring-related problem. In the case of doubt, score it correct. Example: in the Spring Launcher item, the spring extends over intervals 1 and 2. Students who indicated sub-problem A in interval 1 and sub-problem B in intervals 1 and 2 would only be marked incorrect if the classification of sub-problem A included clear indications that it was meant to refer to the spring extending (with
the implication that B was just the ball moving up the ramp after the spring had “hit” it). This is because A could also have been a “find the initial conditions” pseudo-sub-problem.

Collisions: Students had to indicate a sub-problem occurring in just the interval containing the collision, and not classify it in such a way that they couldn’t have intended the sub-problem to be a collision. Such a classification was not observed, however.

Because the above criteria were somewhat looser, the author was able to apply them to otherwise unscorable test sheets, raising the sample sizes considerably. It must be kept in mind that the percentages that will now be reported represent an estimate of the upper bounds of correct responses, and include a fair number of students who most likely did not understand the subject but who were ambiguous enough to be given credit.

The items considered here are as follows:

2C: The partially inelastic collision between the pendulum and the box in item 2.
3S: The extension of the spring as it pushed the object up a ramp in item 3.
4C: The totally inelastic collision between blocks in item 4.
4S: The compression of the spring in item 4.

The 131 students performed a lab essentially like item 2. The 131E students received special instruction in collisions. The H131 students spent extra time on the system of a spring pushing a block up an incline. These three differences in treatment should be kept in mind when evaluating the mean scores of the three groups on 2C, 3S, 4C and 4S. The percentages listed on Table 5.5 represent the percentage of the entire class that correctly identified the sub-problem according to the criteria above. Those that were still unscorable in this round were considered incorrect.
Here, there is clearly discrimination among the groups. The classes without special instruction in collisions, 131 and H131, performed at about the same low level. The class which had extra instruction in springs-pushing-objects performed better on 3S, with a spring pushing an object. Therefore, while overall Decomposition scores were much the same in all three groups, the reasons for their errors differed. Different "blind spots" had been addressed in the two honors classes, with the apparent result that their PDD responses reflected those treatments.

5.6. Suggestions for further development

To summarize the comments made in the previous section, study of the results of the administration of the PDD v1.1 suggested the following changes.

1) Reduce the cognitive load presented by the instructions. This was seen as probably the single most important change to be made, as a test where only half the students are able to follow the instructions will not do its job very well, if at all. Two primary strategies were developed for helping reduce the cognitive load. First, the
classification tasks would be separated from the decomposition tasks. By setting up more focused classification tasks, the options presented to the students could be reduced to a more manageable number. Second, write the items so that a standard five-option scannable test response sheet could be used. In addition to the obvious time-savings benefit in scoring tests, the limitations of the scannable response forms would help reduce the cognitive load on students, as the author would be forced to make items more concise.

2) Find a way to score decomposition tasks so that discrimination was improved. Given the results of examining springs and collisions, a scoring system that looked primarily at those “blind spots” (as well as any others like them, if such could be found) would improve the test’s discrimination.

In addition to these two areas suggested by research on the PDD v1.1, another topic was suggested in personal conversation with Dr. Edward F. Redish. Since the separation of decomposition and classification tasks meant that a decomposition task could have more than one right answer without making scoring problematic, Dr. Redish suggested that an additional item be introduced in which application of energy conservation would lead to a different solution than would application of kinematics. This could help determine if students had realized the power of energy conservation as a problem-solving tool, as such students would be more likely to choose the solution with less sub-problems. Because the classification section of the test would also include energy conservation options, the two sections could be compared. Would students who consistently classified problems as involving conservation of energy also use it to determine that a “complex” problem could be solved in a single step with conservation of energy?

The next chapter will deal with the early efforts to create a new version of the Problem Decomposition Diagnostic that attempts to address these three issues: cognitive load, test discrimination and energy as a problem-solving tool.
CHAPTER 6

DEVELOPMENT OF THE PDD VERSION 2

In the spring and summer of 1999, the author took the lessons learned from the PDD v1.1 and set out to create a new version of the Problem Decomposition Diagnostic. This work can be divided into three segments:

6.1. The PDD v2.0, a first attempt at a multiple choice PDD, which was administered only in an interview format. Serious changes resulted from this work.

6.2. Administration of the PDD v2.1 at the beginning of the summer term and further development of the test based on those results.

6.3. Administration of the PDD v2.2 at the end of the summer and the determination whether version 2.2 was stable enough to administer to a large group.

6.1. The Problem Decomposition Diagnostic v2.0

This version (see Appendix H) was the first attempt at creating a PDD that could be administered using a standard five-option scannable response sheet. The decomposition and classification tasks were separated into two distinct sections, and a multiple choice of five options was devised for the classification task, reducing the cognitive load that the old 26 options had placed on the students.
However, this version of the test still sought to have a more or less open response format for the decomposition task. This would be accomplished by changing from 7 intervals to 5 intervals, then instructing the students to mark the sheet based on how many sub-problems had been completed by that interval. The switch to 5 intervals was not necessary for this method, and was an artifact of a previous design which proved impractical. In any case, this format was never actually presented to students.

For the classification task, a single-step problem would be presented to the student, followed by five possible solution strategies. Each strategy could potentially generate a correct answer, and this was emphasized in the instructions. The five options were based on the three types of classification found in earlier research: surface feature matching, use of deep structure and goal-oriented or “equation hunt” strategies.

A small number of students (N=3) volunteered to be interviewed as they attempted to complete the classification section of the PDD v2.0. The goal of these interviews was to determine if any items could be worded in a way that would make them more attractive as distractors. In addition to such clarifications to both text and diagrams, the interviews revealed that five options was a bit of a stretch, leaving most problems with an option that either could not be made attractive or which was too similar to another option. As a result, the author reduced the options to four in the subsequent versions of the PDD. This problem arose in the very first interview, so all subjects were asked explicitly which option seemed weakest or “most wrong” to them, and these opinions were factored into the decision of which options to eliminate during revision.

After the interviews for the second section, the problem still remained of what to do with the first section, the decomposition section. During discussions with Dr. Van Heuvelen, the author realized that the PDD v1.1 responses could be used to construct
distractors for a multiple choice version of the decomposition task. Rather than attempt to leave the choices fully open, the PDD would present distractors based on popular responses on the large number of PDD v1.1 test response sheets.

The PDD v1.1 test sheets were reexamined and the decomposition choices were tallied. As with the examinations of springs and collisions, most of the test sheets were usable for this task, as it was primarily in the classification section where student difficulties with the instructions surfaced. The four most popular incorrect answers were determined for each item, although sometimes a more interesting fifth-place contender was chosen over the fourth-place answer.

However, the new item 3, an item where a kinematics approach would yield a different answer than would an energy conservation approach, had not been part of version 1.1, and so the author had no data on popular distractors. In addition, there would have to be at least two correct answers for that item (in fact, there were three). The author created distracters based on the sort of answers encountered on the other four items, but this made it vital to perform trial runs before administering this new item to a large class.

It was decided to use the summer term to administer the new PDD twice. Once at the beginning of the quarter to students taking second term calculus-based physics (Physics 132), and once at the end of the quarter to students taking Physics 131. This would provide the opportunity for two rounds of revision, if necessary, before administering the test to a larger class in the fall.

6.2. The Problem Decomposition Diagnostic v2.1

A total of 35 students in Physics 132 completed the PDD v2.1 (see Appendix I) at the beginning of the summer term. In addition to the ten items of the test, two more items
were appended to the end, asking the students to report when they had taken Physics 131 (or its equivalent calculus-based introductory mechanics course at another institution) and what grade had been attained in that class. This step was necessary because not all students had come into summer classes directly from their last physics course, and not all could be counted on to have been students at Ohio State for their first term of college physics. While experience suggests that there might be two distinct populations involved, the ambitious year-round students and students struggling to get off academic probation, the class was not large enough to see any real effect from this potential split population.

All students had taken Physics 131 or its equivalent within the last 9 months, with the majority having taken the course in the spring quarter of that year. 6 reported attaining the grade of A, 13 reported B’s, 14 reported C’s, one a grade less than a C and one chose not to report a grade. Additionally, one student had taken the course pass/fail. Three of the students took the equivalent of Physics 131 at an institution other than Ohio State. Tables 6.1 through 6.4 show how student responses sort out by their self-reported grades in Physics 131 (the pass/fail student’s responses are not included).

It is interesting to note that the students reporting A’s tended to cluster, with the majority of them picking the same response on items 1, 2, 3, 6, 7, 8 and 10, while the B and C students were more spread out. Oddly, C students outperformed other groups on item 4, but this may simply be a chance result. While the number of students was too small to draw any solid statistical conclusions from, the clustering of the students with high grades would suggest that the PDD is measuring something that contributes to or at least correlates with student success in physics courses. It’s also interesting to note that on item 2, the A students clustered on an incorrect response, which might indicate that the “pinball paradigm” indicated by that response is something that better students are more likely to buy into. Of course, these results are all only suggestive.

74
Table 6.1 - Results of the PDD v2.1 on items 1 through 3. The letters across the top are responses on the PDD (correct responses are underlined), compared against course grades.
Table 6.2 - Results of the PDD v2.1 on items 4 and 5. The letters across the top are responses on the PDD (correct responses are underlined), compared against course grades.
Table 6.3 - Results of the PDD v2.1 on items 6 through 8. Only four options are listed across the top, as this section had only four options per item, and no “wrong” answers.
Table 6.4 - Results of the PDD v2.1 on items 9 and 10. Only four options are listed across the top, as this section had only four options per item, and no "wrong" answers.
Examining the results in tables 6.3 and 6.4 demonstrates a rather encouraging result. Namely, students would admit to strategies in the classification section which might be considered “novice,” such as scanning the textbook for help or hunting for appropriate equations. There had been some concern that students would not pick answers that contradicted what the instructor might have told them was correct problem-solving technique.

Two major revisions came out of this administration of the PDD. The first involved, as expected, the new item in the decomposition section. It turned out that three of the five options were correct, making the item almost useless. One of the correct responses was replaced with a distractor chosen to test springs more carefully, given that the A-grade students seemed to have trouble with springs on item 2. The other revision was to significantly reword several of the less popular distractors in the classification section, such as item #9 option A.

Additionally, during analysis of the data, a number of small wording changes and other improvements suggested themselves, and were incorporated into version 2.2. Distractors beyond just the unpopular ones were also slightly reworded.

6.3. The Problem Decomposition Diagnostic v2.2 trial run

After making the changes suggested by the results of the PDD v2.1, the PDD version 2.2 (see Appendix J) was administered to 37 students in the summer term Physics 131. Overall course grades were also collected and the student responses were divided by course grade as was the data for the PDD v2.1. However, since one generally has to pass Physics 131 before taking Physics 132, the reported grades for the PDD v2.1 were generally higher, with all but two students falling into the A, B or C categories. In the case
of this data, 8 of the 37 students got grades less than a C-, spreading the students more across the grade spectrum. Additionally, one student had an Incomplete, and was not included in the data for Tables 6.5 through 6.8. As with the previous data set, the course grade was the only other measure of ability that was collected. The Force Concept Inventory was not given in the summer quarter to Physics 131 students.

As can be seen by examining Tables 6.5 through 6.8, the students with course grades of A did not cluster quite as much as was the case on the PDD v2.1, nor did they always pick the same answers when they did cluster. Clearly, a larger sample would be needed to help determine if this behavior was more than an artifact of random data.

Astute readers may notice that the numbers in Table 6.8 for item 10 do not total to 36 as they do for the other nine items. For some reason, one student (grade of D-) chose the non-existent option e for item 10, and another student (grade of A-) left that item blank. It is possible that the student picking e was not taking the test seriously, or was answering randomly without reading the test. However, the student may have also been opting for “none of the above” despite the admonition in the instructions to pick the best option from those presented. In any case, one student out of 36 does not represent a large source of noise, although such responses will be discussed more fully in the next chapter, as they occurred in other administrations of the PDD v2.2.

The results of this administration convinced the author that no changes were required before moving on to a larger sample (although a few typos were caught and corrected). The next chapter will present the PDD v2.2, as adjusted based on this chapter’s studies, in greater detail.
Table 6.5 - Results of the PDD v2.2 on items 1 through 3. The letters across the top are responses on the PDD (correct responses are underlined), compared against course grades.
Table 6.6 - Results of the PDD v2.2 on items 4 and 5. The letters across the top are responses on the PDD (correct responses are underlined), compared against course grades.
Table 6.7 - Results of the PDD v2.2 on items 6 through 8. Only four options are listed across the top, as this section had only four options per item, with no "wrong" answers.
Table 6.8 - Results of the PDD v2.2 on items 9 and 10. Only four options are listed across the top, as this section had only four options per item, with no “wrong” answers.
CHAPTER 7

THE PROBLEM DECOMPOSITION DIAGNOSTIC VERSION 2.2

This chapter presents the PDD version 2.2, the "final" administered version at the time of this writing, but by no means the final version overall, as all testing instruments undergo evolution. In fact, version 2.3 will be discussed in Chapter 10.

The PDD that emerged from the work of the summer of 1999 was a ten item test. The first section contained five decomposition tasks, in which the right answer or answers were included along with distractors based mostly on the open response PDD v1.1 (the exception being item 3, as mentioned in Chapter 6). The second section contained five classification tasks, in which the potential strategies were based on the three types of classification discussed in Chapters 4 and 5: deep structure, surface features and goal-oriented behavior. In all of these items, two deep structure options were presented. In all items except item 9, one of these options focused on energy conservation and the other on kinematics/dynamics and calculus-intensive solutions. All options could lead to successful completion of the problem, so there were no "wrong" answers, although some responses certainly demonstrate less expert-like strategies.

See Appendix J for a copy of this version of the test. In the next two sections, each item will be discussed in detail, including potential sources of error. Note that this test isn't really designed to be reduced to a numerical score, rather it is a diagnostic tool where each
response should be looked at. But if the user insists on having a numerical score, suggestions will be provided, using a model where 5 points is full credit for the answer. More general sources of error will follow the item-by-item analyses for each section.

7.1. Section I: decomposition tasks

Each item in this section presents a multiple-part physics problem from the domain of mechanics, with five multiple-choice options concerning how best to decompose the problem. Items 1 through 5 of the PDD v2.2 are in this section.

1) Pendulum-Box-Bash.

In this item, a pendulum swings down to hit a box. The pendulum stops after hitting the box, and the box slides a distance before friction slows it to a stop.

The correct decomposition of this item is option b. The pendulum swinging takes place in intervals 1 through 3, the collision takes place in interval 3, and the box slides in intervals 3 through 7. A little more than a quarter of students used this decomposition on the PDD v1.1.

Option a is the only one to lack any indication of a collision, but it was the single most popular result in the open response test, with approximately a third of all students choosing it.

Options c, d and e all represent “boundary problems.” in which the students either do not recognize that the center of mass of each object is fully within interval 3 at the time of the collision, or have some difficulty in accepting that the sub-problems can overlap. No other student error garnered a significant number of responses in the v1.1 administration.
Of the three, d is most interesting, as it suggests that students view the collision as taking place over an extended period, despite the usual simplification in class of treating it as a very short (almost instantaneous) event.

This item does not seem to have any unique sources of error, but it does share several sources with other items, and they will be discussed later.

In scoring this item, it is suggested to grant the full five points to b, no points to a and partial credits (3 or 4) to the other three options.

2) Spring Launcher.

A spring pushes a ball up a frictionless ramp, reaching total relaxation just as the ball leaves the ramp. The ball then flies through the air and lands in a box placed some distance away.

The correct answer to this item is b, with the spring pushing the ball up the ramp in intervals 1 and 2 as one sub-problem (albeit one with multiple simultaneous events) and the ballistic motion of intervals 2 through 7 being the other sub-problem. Approximately half of the students used this decomposition on the PDD v1.1.

Three basic student difficulties were found in the most common incorrect responses to this item on the PDD v1.1. The first was the "pinball paradigm" mentioned earlier in this paper, where students see the spring as acting only in interval 1, despite multiple representations showing that it extends into interval 2. Second is the boundary problem similar to that seen in item 1, where students do not consider the ballistic arc to begin until interval 3. Finally, some students seem to think that the ramp plays no real part in the problem.

Option a seems to suggest all of these difficulties. By assigning the first sub-problem to only interval 1, the student appears to be saying that the spring "hits" the ball at
an instant and launches it into a ballistic arc that doesn't even travel on the ramp. Yet the
arc doesn't start in interval 1, indicating also a boundary issue, or perhaps a belief that the
ramp isn't important, only guiding the ball without performing any "physics" on it.

Option c was perhaps the most popular incorrect answer, and included only the
"pinball paradigm" issue. The spring acts in interval 1, the ball slides up the ramp in
intervals 1 and 2, then it flies through the air in intervals 2 through 7.

Option d is similar to c, with the addition of a total commitment to border
separation. There is no overlap at all: the spring acts in interval 1, the ball slides up the
ramp in interval 2, and the ball flies through the air in intervals 3 through 7.

Option e is also similar to c, but with a lesser expression of the boundary problem.
In interval 1 the spring acts, in interval 2 the ball slides up the ramp, and in intervals 2
through 7 the ball flies through the air.

The decomposition of intervals 1-2 and 3-7 was also a fairly popular incorrect
answer, but not as popular as the ones used.

There do not seem to be any unique sources of error for this item.

For scoring: b gets full points, the rest get zero. All four incorrect answers involve
the "pinball paradigm," which is considered a worse error than simple boundary
separation.

3) Block Catcher I.

In this item, a block slides down a frictionless incline, then across a flat surface
with friction where it eventually meets a massless spring and compresses it before
stopping. This item was intended to have two correct answers, one using conservation of
energy and the other using kinematics/dynamics. The options were derived from
expectations of student errors rather than from student responses on the PDD v1.1, as this
was not an item on that test. Some refinement was done during the work described in
Chapter 6.

Options a and e are both correct. Option a is the conservation of energy solution.
The gravitational potential energy of the block at the beginning is converted into spring
potential energy and internal energy of friction. Therefore, it's a single-step problem.
However, depending on how energy is taught, students may not believe that energy can be
conserved on a surface with friction. This will be discussed further in a moment. Option e
uses kinematics/dynamics, in which the change of forces in intervals 3 and 5 form the
boundaries for the three sub-problems.

Option b was intended to mix the energy solution with the "pinball paradigm,"
although it seemed unlikely that students sophisticated enough to use energy would fall
prey to the spring-related error. It was included, however, to confirm whether this
assumption was true. Also, it is worth noting here for future options that the "pinball
paradigm" didn't seem to apply as often when the spring was catching an object on the
PDD v1.1 as when it was pushing.

Option c was intended to see if students would try to include a collision where there
was none, due to the massless nature of the spring.

Finally, option d mixed some surface features (frictionless versus with friction) and
the "pinball paradigm."

Obviously, this item had a number of unique sources of error. For one, as it
wasn't based on student responses on the PDD v1.1, it was more difficult to judge what
sort of items would make good distractors. In fact, the version of this item on the PDD
v2.1 was something of a disaster in that regard. Secondly, although this was not an issue
in the administration of the PDD v2.2 in fall 1999, the way energy is taught can influence
whether option a is seen as viable. Often, energy is taught so that non-conservative forces like friction are implied to invalidate energy conservation completely. One cannot use conservation of energy if energy is not conserved. Q.E.D. The Physics 131E course, however, included thermal energy (one form of internal energy) in the discussion of energy conservation (Van Heuvelen(2000)), so such a problem was not expected to appear in the results for that group. However, it should be kept in mind that students may have been essentially taught that option a would be wrong because there’s friction.

For scoring, options b, c and d should get zero points, as they contain serious errors. Option a should get 5 points. Depending on how much the instructor is stressing energy as a problem-solving tool, option e should get either 4 or 5 points.

4) Block Catcher II.

In this item, a block slides across a frictionless surface for some distance before hitting and sticking to another block that is attached to a spring. The combined mass compresses the spring.

The correct answer for this item is option e. with the initial slide in intervals 1 through 6, the collision in interval 6 and the spring compression in intervals 6 and 7 being the three sub-problems. Only about a quarter of students taking the PDD v1.1 used this decomposition.

Option b is the primary distractor, representing slightly less than a third of all student responses on the PDD v1.1 for this item. The collision is omitted for whatever reason, despite the fact that blocks sticking together should trigger a fairly strong “inelastic collision” response.
Option a is similar to b, but with the added trouble of a boundary problem. While there was enough space in interval 6 for the entirety of the first block to enter before hitting the second block, some students still insisted on a separation of sub-problems.

Option c was not actually one of the top four incorrect responses on this item. However, it was close (within a few students) and seemed to be more interesting than the choices ahead of it, which were primarily boundary problem mistakes. It removes the collision and adds a "pinball" effect, or perhaps stretches the collision into two intervals. As it could represent two different student difficulties, it should be compared to a student's choices on other items.

Finally, option d was the second most popular incorrect answer, and it either represents a boundary problem or a "pinball" effect.

Having the collision take place in interval 6 seems to introduce a number of uncertainties in the results for this item. It is impossible to distinguish a minor boundary separation difficulty from a case where the student believes that the spring only acts in interval 7, for instance. Or "lack of collision plus pinball effect" from "extended collision" beliefs. Unfortunately, this flaw was not caught until after administration of the PDD v2.2 to the Physics 131E classes.

Scoring this item by partial credit is not recommended, due to the difficulty in distinguishing major errors from minor errors. Alternately, if looking specifically for student use of collisions, give 5 points to e, 3 points to c or d, and 0 to the others.

5) Stunt Ramp.

A ball rises up a ramp and launches into the air, landing on a ramp of the same height some distance away. This was the only item in which the problem did not extend into interval 7. It was decided not to list interval 7 in any of the distractors.
The correct answer for this item is **c**, with the ball moving up the ramp in intervals 1 and 2, then flying through the air in intervals 3 through 6. This was by far the easiest item for students on the PDD v1.1, with three quarters of the combined classes solving the decomposition task with this set of two sub-problems.

Option **a** represents the idea that once the ball hits the ramp, it immediately goes into a ballistic path for intervals 1 through 6. While not very popular on the PDD v1.1, it turned out to be surprisingly popular on the v2.2.

Option **b** is for students who do not consider the ramp to be a necessary part of the problem at all, ignoring the change in velocity that the ball will have as it rises up the ramp. This option is a single-sub-problem solution, covering only intervals 2 through 6.

Option **d** appeals to the boundary problem issue, starting the ballistic path in interval 3 instead of 2. Otherwise, it is correct.

Option **e** looks at the boundary problem from the other end, placing the beginning of the ballistic arc in interval 2, but having the ramp only affect the ball in interval 1.

It should be noted that of the 170 students whose test sheets were used to construct the distractors, no option other than the correct answer garnered more than 9 students’ support. As a result, it’s difficult to say with much confidence beforehand whether these are good distractors or not. Fortunately, as will be discussed later, they were.

For scoring: **a** - 0, **b** - 0, **c** - 5, **d** - 3 or 4, **e** - 3 or 4.

All items in section I share a source of error that the PDD v1.1 did not have. Namely, with the correct answers displayed, some students who would not have generated a correct solution in the open response version of the test will be able to recognize the proper decomposition when they see it. This is generally the risk of a multiple choice test, where both random guessing and "I don't know the answer now, but I'll recognize it when
I see it” behaviors inflate the scores. Only very good distractors can prevent this from happening, and not all items in this section had very good distractors. Most had at least one good distractor, however. In any case, student recognition of collisions (if you include all responses including collisions, not just the completely correct responses) was expected to rise on this test, and as will be discussed later, it seemed to do just that.

Additionally, the format of the diagrams seems to have contributed to a number of possible errors. Objects are often large enough that they are barely inside an interval before a sub-problem ends, leading to confusion over the boundaries of sub-problems. In two problems, springs only extend or compress over two intervals, making it more difficult to judge whether the students really are exhibiting the “springs only act at a point” belief.

Finally, what the student has learned and how he or she has learned it can affect responses on a number of the items. This should be more of a problem on the pretest responses examined in Chapter 9, as students came from a wide variety of high school physics backgrounds.

7.2. Section II: classification tasks

One thing that should be noted is that while the “goal oriented” strategy emphasized the variable to be solved for in the PDD v1.1, the emphasis has now shifted to the “equation hunt” variant seen in earlier interviews. In other words, students are given the option of looking for an equation that contains the correct variables.

For scoring items in this section, a “deep structure” option is marked fully correct (5 points), a goal-oriented or “equation hunt” response is marked partly correct (2-3 points)
and a surface feature response is totally incorrect. However, this section really should not
be scored with points, at least if students are ever to see the results, because they are told in
the instructions that all responses are correct.

6) A block is sliding down a ramp that has nonzero friction. The goal is to
determine how fast the block is moving after it has gone a particular distance.

Option a is the conservation of energy solution, one of two “deep structure”
choices. It is a fairly straightforward application of energy conservation, although it does
include friction, which as previously mentioned, not all students will consider a possibility
with the conservation of energy.

Option b is the “surface features” choice for this item.

Option c is the “goal oriented” strategy, set up as an “equation hunt.”

Option d is the other “deep structure” option, using Newton’s Second Law and
kinematics to solve the problem.

7) A pendulum swings from a horizontal position down to an arbitrary angle. The
goal is to determine the speed of the pendulum at that angle.

Option a involves use of deep physics understanding at the kinematics/dynamics
level, finding force as a function of angle and then integrating acceleration to get velocity.
It is mathematically perhaps the most difficult solution option. The need for integration is
not mentioned in the problem statement.

Option b is the “equation hunt.” The expectation is that the problem has been
solved already in the book or in class, and the student needs only find the right equation
and enter the numbers or symbols.
Option c uses student understanding of conservation of energy to solve the problem. This is mathematically the easiest option, with a very straightforward application of the conversion of gravitational potential energy to kinetic energy.

Option d is the surface feature “pattern match” option.

8) A cannon fires a projectile into the air, and information about the cannonball has specific restrictions on it. The goal is to find the speed of the ball at the top of its arc.

Option a is one of the “equation hunt” options for this item, trying to find a pre-derived ballistics equation to fit this problem. Given the way the problem is stated, it’s unlikely that such an equation will be present, but it’s not impossible for such a problem to have been worked in lecture.

Option b uses conservation of energy as one of the steps, although there is not enough information to rely completely on it, as another mathematical step (there are several possibilities here) must be performed as well.

Option c is the other “equation hunt” choice, sending the student on a quest through all two-dimensional kinematics equations.

Option d uses kinematics/dynamics to solve the problem, and may not be any more mathematically difficult than conservation of energy in this specific case.

Note that there was no surface feature option, partly because the surface feature of “ballistic motion” is so intimately tied to specific equations in most texts. Therefore, the “equation hunt” choices could also be seen as surface feature strategies.

The way in which the givens are presented in this item may cause some confusion for students and introduce error in the results.
9) First a sport utility vehicle (SUV) pushes a car that has stalled. Then, with the driver applying the same accelerator pedal pressure, it accelerates without encumbrance. The goal is to determine how fast the SUV would be traveling without pushing the smaller car. This item was an attempt to use conservation of momentum equations without having a collision or an explosion, two surface features closely tied to conservation of momentum. However, as a result it was fairly ambiguous and could not be made solvable with conservation of energy, and the author has never felt terribly confident about it.

Option \textbf{a} is the surface feature solution.

Option \textbf{b} uses kinematics/dynamics to determine accelerations and forces to solve the problem. Probably the most straightforward way of solving it.

Option \textbf{c} tries to evoke an “equation hunt” for collision equations. This is the only place conservation of momentum is invoked, and seeks to see if the students are trying to use it inappropriately. While the collision equations could give the correct answer, this isn’t a case where momentum is being conserved.

Option \textbf{d} uses Newton’s Second Law to try and solve the problem. This solution is equivalent to option \textbf{b}, but with the explicit addition of diagrams.

Despite reservations, the author left this item on the PDD v2.2 to see if perhaps he was incorrect in disliking it. No suitable substitute involving conservation of momentum had been devised at the time of administration. Because there is no really “good” solution, it is difficult to judge this item, and it lacks the energy versus kinematics/force decision of the other four items in this section.

10) A block slides across a frictionless surface and hits a massless spring, which compresses. The goal is to find the mass of the block. This could be considered a vastly simplified version of Block Catcher II with an altered goal.
Option \textbf{a} was the "equation hunt" for this item.

Option \textbf{b} used Newton's Second Law to determine the force exerted on the block while in contact with the spring, then would require integration to solve. The integration is not mentioned in the problem statement.

Option \textbf{c} invokes conservation of energy to solve the problem in a very clean manner. Kinetic energy is converted to spring potential energy, with no friction present to "muddy the waters," so to speak.

Option \textbf{d} is the "equation hunt" of this item, although more detail than normal is given regarding the equation that is sought. By giving more detail on a solution that otherwise seemed marginal (the equation used is for a related but not identical situation), the author wished to see if this could make this a more effective distractor. The results were quite interesting, as will be seen later.

As will be seen in the next two chapters, some of the difficulties discussed here expressed themselves in either the interviews or the test, but others were not present in the groups studied. It is possible, however, that these "missing" concerns would still present problems in deriving information from this test when administered to other groups. And, if nothing else, such concerns have a negative impact on face validity.

Chapter 8 will present the results of student interviews regarding the PDD v2.2. Chapter 9 will discuss the results of administering the PDD v2.2 to a large class of students, and Chapter 10 will consider improvements that could be made to the PDD, as well as other areas of future research regarding this instrument.
CHAPTER 8

PDD v2.2 INTERVIEWS

In the fall of 1999, six students from the honors physics course Physics H131 volunteered to take the PDD v2.2 and be interviewed. The reason for these interviews was to help establish some validity for the test as currently constructed, to determine if students were making the choices they did for the reasons the author expected. Obviously, with only six students, the results could not be broadly generalized, a problem compounded by the fact the students were from an honors course intended for physics majors. However, the interviews were seen as a way to determine if there were any gross invalidities in the test.

The six subjects were interviewed singly over the course of several weeks, as their schedules permitted it, for up to an hour each. The first two interviews took place before energy was discussed in H131 lecture, although all subjects had been exposed to the concept in high school physics. The progression of the course could account for some of the trends in the responses, but the sample is too small to say much with confidence. Additionally, energy was taught according to the Work-Energy Theory in H131 rather than using a "system"-based approach (Van Heuvelen (2000)) as used in Physics 131E. As a result, it is possible that students in this study would not consider energy conservation to be valid in a situation with nonconservative forces such as friction.
This chapter will be split into three sections.

8.1. Interview protocol
8.2. Summary of student responses
8.3. Discussion of conclusions that can be drawn from this study

8.1. Interview protocol

The following protocol was used. Each interview was conducted in four parts:

1) The subject was given a general description of the project and promised that more detailed information would be available later in the session. At this point the subject signed consent forms.

2) The subject took the PDD v2.2 under circumstances close to those normally seen in test administration. The main difference was that the subjects were allowed to write on the test sheet directly, rather than on Scantron™-style answer forms. Because it could have significantly altered the subject's performance, a "think aloud" protocol was not used.

3) With a recording device activated, the interviewer and subject discussed each item on the PDD. The subject was asked why particular answers were given, as well as probing questions to help determine the robustness of that answer. In most interviews, the subjects were also asked which option on each item was the "most obviously wrong" for items in Section One of the test. This question was added during the first interview when the interviewer realized it would be a useful test of the power of various distractors. The interviewer took written notes in addition to the audio recording, which was only used later to help flesh out the notes.
4) The recording device was turned off and any questions the subject had about the test were answered, in accordance with informed consent guidelines. Any additional information gained during this discussion is also added to the interviewer’s written notes.

For his or her time, each subject was then paid $6. All subjects signed both consent forms and receipts for the monies received.

8.2. Interview Summaries

These summaries are the result of the notes taken during the interview process. The taped discussions were used only to clarify points in the notes, and no transcripts were made. The summaries use a number of abbreviations and shorthands, as described below:

A number followed by a left parenthesis represents an item on the PDD. 1) refers to item 1.

A letter in boldface represents an option on the item under discussion.

A number or range of numbers enclosed in square brackets represents an interval or range of intervals. [1-2] indicates intervals one through two are being discussed.

These codes are primarily intended to reduce the repetition of common words and phrases and generally condense the information.

In addition to the summaries of item-by-item responses, a small amount of demographic information about each subject will be presented. Gender, major area of study, performance on the Force Concept Inventory, high school physics background and time spent on the PDD v2.2 will be given for each subject.
S6-1

Date: October 26, 1999. 7:30 PM

Gender: Male

Major: Mathematics

FCI: Did not take either pretest or posttest

Physics Background: a year of non-AP physics two years ago, remembers little of it.

Time on test: 15 minutes

1) Guessed b, which is the right answer. Wrote friction equation on his test sheet. Intuition told him something happened in [3], and he explained it in terms of interacting forces. Force was “transferred” from pendulum to box. Not sure of his physics.

2) Chose b, the right answer. Still not sure of his physics, but he had good enough physical intuition that attempts to steer him towards c in the interview did not sway him. He saw that the spring had to do something all the way through [1-2].

3) Chose e, the kinematics/dynamics answer. He didn’t really remember energy from high school, and hadn’t gotten to it yet in H131. He decomposed the problem on the basis of forces, could not be convinced to combine [1-3] and [3-5], but was less sure on mixing [3-5] with [5-7]. When asked which looked most obviously wrong, he chose a. It was at this point that the interviewer decided that the question “Which response appears to be most clearly wrong?” should be included more often, and always on this item.

4) Initially chose c on the basis of a vague third force he couldn’t articulate, but switched to b when he realized he couldn’t justify the [7] part. He had no boundary problems as in a or d. During discussion, he realized that the collision in [6] was a necessary part, but he did not come to this conclusion until reminded of his response for 1).
5) Chose c, the correct answer, and stated that the problem was pretty obvious. When pressed on the issue of whether a could be correct as well, he said that c was definitely easier, but couldn’t say for sure that a was wrong.

6) Quickly picked d. He followed a memorized procedure from homework, and said he didn’t trust himself to try and use book examples or equation hunting. In general, he stated a lack of confidence in his own ability to generalize from one example to another, and was more comfortable sticking with a rote procedure.

7) Chose a, again preferring the rote force/acceleration procedure. When told of the troublesome calculus this might involve, he was still comfortable with this procedure. His opinion was that it was better to go with tough math than a new strategy he didn’t yet trust.

8) Picked d because it resembled the way he had solved a homework problem that week. Despite the possibility that a pre-derived ballistics equation would work here, he stated a mistrust of the equation hunt, for reasons given above.

9) Waffled a bit, but eventually picked b over d on the basis of not wanting to do Free Body Diagrams (FBDs). In this case he said he’d try the surface feature hunt in the textbook if his ability to do the problem without FBDs fell short.

10) Picked b, again preferring rote net force procedure. As before, he was not daunted by calculus concerns. Found the unfamiliar d to be least favored.

Preliminary conclusions: Subject 6-1 has good physical intuition, but doesn’t trust in his own ability to apply new concepts and strategies. preferring to remain with the tried and true, even when it means more work.
S 6-2

Date: November 1, 1999. 1:30 PM

Gender: Male

Major: Physics

FCI: pretest 16/30, posttest 21/30, normalized gain of 0.36

Physics Background: took AP Physics, found it not very useful and did not take the advanced placement test.

Time on test: 10 minutes initially, realized he'd missed one, took 1 minute more.

1) Picked a. The collision was recognized physically, but treated as an unnecessary or implicit step. Seemed to conflate force and velocity and treated both as conserved quantities. Saw d as obviously wrong. When pressed, he admitted the collision step in [3] could be performed separately, but didn’t feel it to be necessary.

2) Picked c. [1] was the action of the spring, [1-2] the motion of the ball up the ramp, as intended by the author. He continued to conflate force with velocity. When pressed, he wasn’t sure if b was actually wrong. Did not have difficulties with the boundaries between sub-problems (i.e. a response of d).

3) Picked e, the kinematics/dynamics solution. He wasn’t sure how to deal with the spring, however, compared to his performance on 2). Claimed a (energy approach) was obviously wrong, and that you needed to break everywhere the force changed, although he wasn’t totally sure the break in [5] was absolutely necessary.

4) Picked e, the correct answer. Admitted he was being inconsistent with previous problems, but could not (would not?) articulate why. In this case, he felt the collision couldn’t be ignored, says he thought it was because the problem was all on a straight line, like most collision homework exercises.
5) Initially picked e. When asked why, he looked at the problem again and couldn’t remember why he made that choice, quickly changing his answer to c, the correct answer. Could not be shaken on boundary issues otherwise, but wasn’t sure why a or b were wrong.

6) Picked d. Subject had spent the previous weekend in intensive drills on FBDs and felt very comfortable with them. Energy was only starting to be covered in class, and he was not comfortable with it yet. He considered the equation hunt tactic to be a waste of time, and would only use it if all else failed.

7) Chose b. He knew there was an equation for that item, but could not recall what it was exactly or what methods were used to find it, so he would look for it in the textbook. If he couldn’t find the equation, his next option would be to look at similar surface feature exercises.

8) Picked d. He knew the method in that choice already, so didn’t see a point in looking for another strategy.

9) Again picked d. Another case of comfort with recently-drilled FBDs.

10) While none of the options seemed unreasonable to him, b used FBDs. so he selected that option.

Preliminary Conclusions: A case of familiarity breeding comfort. He trusts what he knows well, even if it’s not the easiest way to solve the problem.
S6-3 (Nov 8, 1:30 PM)
Date: November 8, 1999. 1:30 PM
Gender: Male
Major: Physics
FCI: pretest 21/30, posttest 26/30, normalized gain of 0.56
Physics Background: took non-AP Physics. no particular complaints about it.
Time on test: 12 minutes

1) Picked e, which includes the collision correctly. However, his view of the intervals was that a sub-problem had to take place either entirely within an interval or over the majority of that interval in order to be counted as including that interval, supporting the view that the boundary problems were relevant to at least some students. He did not consider the collision as “leaking over” into [4].

2) Chose the correct answer, b. He did not think that there was a separate sub-problem involving the spring in [1]. When asked how c and e were different, he couldn’t really say.

3) Picked e, the kinematics/dynamics solution. Pointed to a (energy) as obviously wrong. Uncomfortable with energy conservation as a solution strategy. Thought d might work, but he was reluctant to leave regions of different forces together. He had studied energy for about a week at this point. S6-3 said that using energy conservation in this case made him feel like he was skipping important steps in the problem.

4) Picked b. He did not have any strong opinions on rightness or wrongness of the various options, and claimed his only real difficulty with this item was determining where
the calculations had to start, rather than expressing any difficulty with physics concepts. When asked about the collision, he said that because the blocks stuck together, the collisions took place over all of [6-7].

5) Chose c, the correct answer. He felt that a was obviously wrong, because the projectile motion didn’t start until [2]. When it was pointed out to him that this seemed to contradict his boundary treatment in 1), he became uncertain and couldn’t explain why the two were different. Possibly a matter of surface features.

6) Chose d based on his familiarity and comfort with force diagrams.

7) Picked a. S6-3 would look in the book if he got stuck, but preferred to try things on his own first. When the difficult math of his choice was pointed out, he said he’d still try the forces approach first anyway, preferring an older method despite the fact that it involved more work on his part.

8) Picked d. The large amount of time required to adequately search his resources discouraged him from attempting the “equation hunt.”

9) Despite previous preferences for FBDs. S6-3 picked b. on the grounds that without the FBD, it seemed to be an easier solution.

10) Picked c. He quickly recognized this as a “stock” energy conservation one-step problem.

Preliminary conclusions: If given a choice between a computationally difficult, familiar strategy and a computationally simpler, but unfamiliar one, he seemed to prefer the familiar strategy. However, given a choice between two familiar methods, he went with the procedurally easier one, even though it might give him a higher chance of error on the problem.
S6-4 (Nov 8, 3:30 PM)

Date: November 8, 1999. 3:30 PM

Gender: Female

Major: Physics

FCI: pretest 24/30, posttest 27/30, normalized gain of 0.50

Physics Background: took non-AP Physics, no particular complaints about it.

Time on test: 18 minutes

1) Chose e, demonstrating somewhat inconsistent views on what made a sub-problem exist in an interval or not. Made extensive notes on the sheet, identifying what would be needed for each sub-problem. Preferred a very reductionist strategy, breaking things into as many parts as is feasible, feeling this would decrease error rate in calculations. S6-4 did not see the collision as a necessary sub-problem, rather she thought that excluding it might only make it harder to get the calculations right. Wrote about "mv=p" on the test sheet, but talked about transfer of energy from the ball to the box. She generally seemed to want to decompose the problem based on what equations were available (for example, she felt d was obviously wrong, because there was "no equation for" what happened in [3-4]).

2) Chose c. S6-4 wrote energy conservation equations on the test sheet over the ramp in [1-2] and separated the velocity at the top of the ramp into components. She said that the order of sub-problems would be as follows: [1-2], [2-7], [1]. The final part was to go back and calculate the desired quantity, a mathematical step rather than a physical one (and in fact, not a very good mathematical step, as the distance would be a variable in [1-2], not in [1] alone). Option d looked wrong to S6-4, but she couldn’t say why. She said b was wrong because the problem required three equations, so it had to have three parts.
3) Chose the kinematics/dynamics answer, e. Again, the energy conservation answer a was picked as most obviously wrong. Expressed the opinion that a change in forces requires a new sub-problem. She wrote “need $v_1$” in two places on the diagram, at the breaks in the kinematics/dynamics solution. Underlined “massless” in the problem description, but the interviewer did not ask if that information was used to eliminate c as a viable answer.

4) Picked the correct answer, e. S6-4 wrote “mv” on the test sheet with a line to the collision point, indicating again that she understood that there was a collision problem there. However, when asked if she thought an expert (the interviewer) could solve the problem by breaking it up as in b, she said it might be possible. She considered inclusion of the collision to be more a “security blanket” (interviewer’s term, not subject’s) than a necessary step.

5) Picked the correct answer, c. She added extra “givens” notation to the diagram. including an apparent “alpha” or possibly a stylized x or d that was the same as the distance between ramps. S6-4 continued to show an inclusive boundary view. rather than S6-3’s more exclusive view. She found b to be most obviously wrong, as it ignored part of the ramp. However, when asked about e, she wasn’t totally sure it was wrong.

6) Picked d. It was a block on an incline, and to her that meant FBDs.

7) Picked c. The inclusion of a change in height (albeit implicitly) triggered energy conservation ideas. Reluctant to try the problem with kinematics/dynamics, as she was concerned over her own computational skills.

8) Answered d. She claimed it was similar to homework problems she had solved. Additionally, in the past she had lost points on a similar problem because she forgot that the x-component of velocity was invariant, so this particular solution had been reinforced in her mind.
9) Picked b. She claimed that in this case, a FBD would be superfluous, since the problem took place on a horizontal surface. To her, FBDs are for "harder problems" that involve surfaces at an angle. Her link between FBDs and inclined planes seems to be a bit too strong.

10) Picked b. When originally working this item, she incorrectly assumed that the spring's force would be constant. However, when informed of the necessary integral due to changing force, she said she'd still try to do it with forces. The surface features did not seem to trigger energy ideas in her.

Preliminary conclusions: S6-4 seems to have developed strong ties between surface features and physics content areas. She said that she really only used energy when things were changing height on odd paths, and that FBDs should be used in problems with inclined planes. Additionally, she was not convinced that one had to work through conservation of momentum in order to solve a problem, merely considering it an easier way to do things.

S6-5 (Nov 15, 3:30 PM)
Date: November 15, 1999, 3:30 PM
Gender: Male
Major: Physics
FCI: pretest 25/30, posttest 27/30, normalized gain of 0.4
Physics Background: took non-AP Physics, considered it a very bad program.
Time on test: 11 minutes
While taking the test, S6-5 was uncertain what the intervals meant, and what qualified a sub-problem to be considered as part of the interval. The interviewer did not answer the subject's questions during the test, directing the subject to make those decisions on his own.

1) Picked a. The pendulum "has force" which is transferred to the box. He conflated force with either energy or momentum. S6-5 said that including [3] as a separate part wouldn't help any, since "all the energy is shifted, there's no real calculations to do." This seems to be an assumption that velocity is conserved, as the masses of the two objects were stated in the problem to be different.

2) Picked b. Hung up on the issue of partial intervals for a while. S6-5 would have preferred to draw in his own intervals. Once he got past this, however, he had no difficulty seeing that the spring motion had to take place in both [1] and [2], and that there was no difference in its behavior in the two intervals. so splitting them made no sense.

3) Followed the same pattern as everyone else. Kinematics answer, e, was chosen, and the energy conservation solution, a, was seen as most obviously wrong. He was of the opinion that the change in forces heralded the beginning of a new sub-problem, and believed that one could not put all of the kinematics/dynamics pieces together as a single sub-problem. S6-5 was not distracted by the "collision" in [5].

4) Picked b. "The force continues onwards" because the blocks stick together, so the collision seems to be extended over [6-7] in his opinion. The mass of the system changed, but the force (which was treated as a transferable quantity by S6-5) remained the same. As the second student to do this, he raises the point that perhaps the "missing" collisions on this item were actually extended collisions. This means two different errors were combining in a way that could not be easily teased apart. He saw c as the most obviously wrong, because the spring didn't do anything special in [7].
5) Initially selected b, since the frictionless surface of the ramp wouldn't slow the ball. During discussion, he realized that the change in height would slow the ball, and changed his answer to c and claimed b was the most obviously wrong answer. When asked about a, he was uncertain. He thought it might work, but wouldn't want to try it himself.

6) Went with d. He said that, had he taken the test earlier in the quarter, he definitely would have gone with the surface feature strategy b. He was comfortable with FBDs at the time of the interview, but not with energy, despite two weeks of instruction on the topic.

7) Picked a, but he wasn't very comfortable with it, and stated that he probably would complete a surface feature check for confirmation. When confronted with the integral involved in a. S6-5 didn't see it as a serious obstacle, still preferring to use forces rather than energy.

8) After wondering about whether air friction was an issue, he picked d.

9) Initially picked b and then didn't read the subsequent options. After being told to read them, he decided he'd need a FBD and picked d.

10) Picked b, again preferring this method even after being told about the integral. All options looked okay to him, however.

Preliminary conclusions: Similar to previous subjects, S6-5 preferred familiar concepts and difficult computations over unfamiliar concepts and simpler computations. Showed somewhat consistent incorrect views regarding collisions, treating quantities other than momentum as being conserved. He had difficulty accepting the way the intervals worked, but once he reached a conclusion regarding this, he seemed to be okay with them.
S6-6 (Nov 22, 3:30 PM)

Date: November 22, 1999. 3:30 PM

Gender: Male

Major: Physics

FCI: pretest 13/30, posttest 16/30, normalized gain of 0.18

Physics Background: took non-AP Physics, considered it a “so-so” program.

Time on test: 15 minutes

1) Picked b, the correct answer, but S6-6 was unsure of his ability to solve the third sub-problem. He thought a might be a practical solution if the problem was worked by an expert, so the collision was not seen by S6-6 as a vital part of the problem. He considered d most wrong, as he felt that collisions are not extended events.

2) Picked c. He read the problem as having the spring only “launch” at [1]. When confronted with the actual problem statement, he amended slightly, but considered [1] to be a mathematical sub-problem, and stuck to his answer of c. Didn’t think any option was really obviously wrong, but considered b (the correct answer) to be wrong because it didn’t include a sub-problem for the calculation of the initial velocity.

3) Picked e, the kinematics/dynamics solution. S6-6 called a (the energy solution) most obviously wrong, claiming the problem simply could not be done in one step. He was unsure whether a division of [1-3],[3-7] would work, but thought that a value of velocity at [5] was needed. He treated sub-problems as conservation of energy (except [3-5]), but didn’t seem to want to put them together into a single state problem using conservation of energy.

4) Initially chose b, but on his own realized that a momentum transfer was involved, and he crossed out that answer and circled e. Felt a was most obviously wrong,
because of boundary issues. Did not think [7] was a subproblem on its own...however, when presented with the idea of a time-reversed version of the problem, he then considered [7] to be a vital subproblem, as it represented some sort of “initial conditions” physics problem.

5) Picked c. Change of potential energy told him that velocity would change on the ramp, so single-step options wouldn’t work. As with other problems, he was clear on boundary issues.

6) Chose a, and he was very comfortable with conservation of energy in this case. However, when the interviewer suggested adding friction to the problem, the subject said he would then choose a kinematics/dynamics solution. This pattern persisted through other items in this section...the addition of friction caused a switch from conservation of energy principles to kinematics/dynamics, suggesting that he considered energy to not be conserved when friction was involved.

7) Picked c, and his responses were similar to those on 6). The existence of an integral in the kinematics/dynamics solution did not stop him from preferring it in the event of a similar problem using nonzero friction.

8) Initially chose d, but he didn’t feel very confident in the area of ballistics overall, and he switched his answer to a, the strategy of checking the textbook for surface features. S6-6 was uncertain how energy could even be applied in this situation. although when the angle was changed so that the cannon fired straight up, he immediately claimed he would use energy principles.

9) Considered c, but picked d because he claimed he could already see a solution strategy. As this item did not have an energy option, his views on that were not probed here.
10) Quickly picked c. When nonzero friction was introduced, he faltered and said he’d prefer to use forces and kinematics/dynamics here, despite the existence of an integral in the solution.

Preliminary Conclusions: This student had learned energy principles and was very comfortable with them...but only within very limited boundaries. Only certain surface features triggered his use of energy conservation, and friction rendered it unusable regardless of other surface features. In addition to the H131 curriculum’s effect mentioned earlier in this chapter, student may not have seen the value of the “state process” concept in being able to ignore everything but initial and final states.

8.3. Discussion of interviews

There were enough subject difficulties with the interval boundaries to suggest that it might be preferable to alter the diagrammatic representations of the items or the explanation in the instructions of what the intervals were. However, most of the time the subjects were able to cope well enough that other factors dominated their responses.

Despite being bright, well-motivated students who largely had a firm grasp on the material covered by the Force Concept Inventory, the subjects generally did not view collisions as being necessary sub-problems when they were present. In fact, the idea that working conservation of momentum separately was useful but not necessary appeared multiple times. This could be seen as validating the emphasis on collisions in the PDD, as it is a piece of physics content with which the students seem to have many difficulties.

The spring-related difficulties, however, did not seem to appear as often with this group of subjects. Granted, they were drawn from the same class (albeit a different year
and different professor) as the one that did so well with springs on the PDD v1.1, but examination of these results seem to indicate that the “pinball”-like responses arise more from a mathematical decomposition of the problem. citing the necessity of determining behavior at the very beginning for the solutions. Clearly, this issue warrants further investigation.

Finally, the subject of energy conservation yielded very interesting results. Every one of the six subjects felt that the energy conservation solution of item 3 was the most obviously incorrect, and most subjects did not even pause to think about their answer, so great was their certainty. Even subjects who were comfortable with energy solutions in Section II of the test would not use them for item 3. As previously mentioned, one possible reason for this is that these students were taught that friction was a non-conservative force, and may have felt that summing energies would not work in this case. However, judging from the way some subjects applied energy, it is also possible that they did not think of energy conservation as being a problem-solving technique. Rather, it was simply another tool to use in working a single-step exercise.

Because the 131E course had a heavy emphasis on a system approach to energy that allowed energy conservation to be applied even when friction was involved, the author was quite interested in seeing how this group’s responses changed between administrations of the PDD v2.2. If student responses shifted to energy conservation solutions on both item 3 and the relevant Section II items, then it would suggest that the “friction means energy is not conserved” idea was likely the culprit. However, should item 3 remain largely unaffected, it could suggest that even a good conceptual understanding of energy does not necessarily mean students would see it as a useful problem-solving strategy.

Chapter 9 will present the results of this pretest/posttest administration of the PDD v2.2.
CHAPTER 9

LARGE-SCALE ADMINISTRATION OF THE PDD v2.2

In the fall quarter of 1999 at The Ohio State University, the author administered the PDD v2.2 as both a pretest and a posttest to students enrolled in the Physics 131E Freshmen Engineering Honors class. This chapter concerns the results of this testing cycle, and will be split into the following sections:

9.1. Discussion of the 131E class, the procedure for test administration, and other measures taken for comparison.
9.2. Discussion of the goals of this administration.
9.3. Examination of pretest results as correlated to other measures.
9.4. Examination of posttest results as correlated to other measures.
9.5. Comparison of pretest and posttest results.

9.1. The experimental group

As with previous years, the Physics 131E course was part of the Freshman Engineering Honors program, so all students had to meet the requirements of that program in order to enroll in 131E. All students in 131E had previous physics experience in high
school, although naturally the details of their backgrounds varied significantly. Therefore, the students under consideration here should not be considered to be representative of the "normal" introductory calculus-based physics student. However, gross invalidities of the PDD should show up in a group of this size (N=157 for pretest, N=153 for posttest) even considering their higher demonstrated abilities.

131E students were subjected to a large number of treatments that are believed to improve student learning, such as cooperative learning (Heller (1992b, 1992c)), interactive engagement (Hake) and innovative instruction in energy concepts (Zou, Van Heuvelen (2000)). They also used the Toys In Motion labs (Appendix A). With such intensive treatment, the author claims that there should be measurable changes in student understanding of physics over the course of the term, and therefore there should be noticeable differences in pretest versus posttest performances on the PDD v2.2.

In addition to the PDD, the 131E students also completed the Force Concept Inventory as a pretest and a posttest. The class mean on the pretest was 66%, while the posttest mean score was 80%, resulting in a normalized gain of 0.41. These exceptionally high scores help underscore the earlier point made regarding the atypical nature of the experimental group.

In the next section, PDD pretest results will be compared against FCI pretest scores. In the section following that, PDD posttest responses will be compared against both FCI posttest scores and final course grades.

9.2. Goals of this administration

The primary goal of the work done administering the PDD v2.2 to the Physics 131E students in the fall quarter of 1999 was to establish some measures of validity. As
discussed in Chapter 1, any test or measure can be considered on two major criteria, validity and reliability. Because of the nature of the PDD (small number of items, lack of strong divisions between right and wrong answers, etc), statistical reliability was not measured for this iteration of the test. Rather, the goal was to concentrate on validity: whether the instrument tested what the author intended for it to test, or at least measured something worth knowing.

As previously mentioned, without establishing reliability for the PDD, it is difficult if not impossible to establish validity for the PDD as a whole. However, it is possible to examine individual items to determine whether they have validity. Since statistical reliability isn’t applicable to single tasks (and, of course, single-test reliability measures are impossible to apply to a single item), validity is not as tied to reliability in this case.

Section II of the PDD has some construct validity to help it, as it was based on Chi’s card sorting task work, as well as additional sorting task experiments described previously in this paper. While the surface features of the tasks are different (multiple choice rather than open-response sorting), the underlying construct that non-experts tend to classify problems more by surface features or goal features informs Section II.

For the remainder of this chapter, attempts will be made to demonstrate predictive validity for the items on the Problem Decomposition Diagnostic v2.2. Some discussion of face validity will also take place.

9.3. Pretest comparisons with other measures

While the PDD was not intended to be administered as a pretest due to its use of physics content knowledge. in the case of the 131E course such a use was possible. All students had been exposed to the content knowledge during secondary education, so the
items would have been comprehensible to them, if not necessarily easy. Due to the fact that
treatment of energy in high school physics varied from school to school and was likely
weaker for some students than others, Section II presented the largest potential barrier to
administration of the PDD as a pretest for this group. Section I, the decomposition task,
was not considered to be an excessive challenge to these students.

To look for correlations between FCI scores and responses on the PDD, the data
was dichotomized. FCI scores were broken into “low” and “high” on the basis of the class
average. PDD responses were broken into “right” and “wrong” by criteria that will be
discussed below. A dichotomous correlation was then calculated, along with the t-score
for statistical significance. N=157 for this group, so a t-score of 1.97 indicates
significance at the P<0.05 level, and a t-score of 2.60 indicates significance at the P<0.01
level. The pretest mean score on the FCI was 66% with a standard deviation of 18.6%.

The FCI’s domain is comprehension of introductory Newtonian mechanics
concepts, rather than problem-solving skills. However, one premise of the PDD’s
development has been that content knowledge is required for meaningful problem-solving
in the physics domain, so there should be some correlation if items on the PDD are valid.
Students with poor content knowledge would be expected to have poor problem-solving
skills, although students with strong content knowledge might still have poor problem-
solving skills. It is possible to do well on Section I with less content knowledge, so lower
correlations were expected for that section. However, as seen in Chapter 8, a student’s
comfort with the material can affect responses in Section II, so was hope for a stronger
correlation in Section II.

1) For this item, a “right” response was any that included a collision in some form,
meaning that only option a was “wrong.” The dichotomous correlation was nearly zero,
and was not significant.
2) For this item, only option b was considered to be a "right" answer. It is possible that a response of a was due only to boundary issues and could have been counted as correct for purposes of this dichotomization, but as it is impossible to separate the reasons for choosing a in this version of the test, that option was counted as wrong. As with item 1, the dichotomous correlation was nearly zero, and was not significant.

3) Options a and e were correct for this item. A dichotomous correlation of 0.12 was found with a t-score of 1.50, not significant either statistically or colloquially. As no students picked option a, the possibility of comfort levels with energy affecting the results could not be quantitatively determined. Qualitatively, it can be said that the students did not feel energy conservation was a valid tool to use in the manner suggested by option a.

4) Because of the confusion over the nature of errors involving the spring on this item, only the collision was considered. Options a and b were incorrect, and the remaining options were considered correct, as they included some form of collision. A dichotomous correlation of 0.20 was found, with a t-score of 2.60. This is statistically significant at P<0.01, although only 4% of variance is accounted for. It is possible that this correlation represents the following: students whose previous physics experience covered inelastic collisions adequately also had a firmer grounding overall, thus scoring higher on the FCI.

5) As the boundary problems are not considered serious student difficulties, the correct answers were all options that broke the problem into two sub-problems. With that dichotomization, there was no significant correlation.

6) For this item and all others in Section II, the "right" answers for dichotomizing the PDD responses were those evidencing deep structure. Surface feature-based and goal-oriented strategies were marked as "wrong." Options a and d were marked correct for this item. The dichotomous correlation for this item was 0.24, with a t-score of 3.11, indicating statistical significance at the P<0.01 level. As mentioned earlier, this correlation
probably results from student comfort with the concepts: those who score higher on the FCI are less likely to admit to preferring a non-expert solution. Of course, this does not mean that they necessarily wouldn’t use a non-expert solution when left to their own devices, but the fact that 60 students out of the 157 admitted to non-expert solutions suggests that there was at least not an overwhelming tendency to give the “expected” answer on this item.

7) Options a and c were the “right” answers for this item. The dichotomous correlation was 0.25, with a t-score of 3.19 (P<0.01). Another moderate effect, possibly due to the comfort factor.

8) For this item, options b and d were marked as correct. The incorrect answers represented only goal-oriented strategies. The effect size was a moderate correlation of 0.31, with a t-score of 4.06 (P<0.01).

9) The options marked correct on this item were b and d. The effect size was another moderate 0.31 correlation, with a t-score of 3.97 (P<0.01).

10) Options b and c were “right” for this item. The dichotomous correlation was 0.27 with a t-score of 3.46 (P<0.01).

As expected, Section I did not correlate highly with FCI scores, although item 4 did have a statistically significant correlation. Every item in Section II had a moderate correlation (between 6% and 10% of variance accounted for) with strong statistical significance. It is quite possible that this correlation merely measures how well students were prepared by their high school physics courses. Students who were more comfortable with non-expert strategies also tended to perform below average on the FCI, and this could be explained by the below-average students being less well-prepared overall. As Section I is more dependent on basic logic and problem-solving skills, the merits of previous physics courses would have less effect on the results.
In any case, the presence of statistically significant correlations on all items suggests that Section II of the PDD may be measuring something useful in the domain of physics, and face validity suggests that it is measuring what the author intends it to.

9.4. Posttest comparisons with other measures

For the posttest, 153 students took the PDD v2.2 and the Force Concept Inventory. In addition, student course grades were compared to PDD responses. Significance levels remain the same for this group’s t-scores.

Of the 153 students who took the PDD, 70 had overall course grades of A or A-. 67 had course grades of B, B+ or B-, and 16 had course grades of C+ or worse. When dichotomizing this variable, an A or A- was labeled “High” and all other grades “Low.”

The FCI mean score was 80%, with a standard deviation of 15.2%. This higher mean and narrower spread meant that it would be harder to find any correlations, as there simply was less variance to be explained by correlations. “High” FCI scores for purposes of dichotomization were those over 80%. As with the grades, performance was so high overall that the variance was quite limited.

PDD responses were categorized as “right” or “wrong” in the same way as in the analysis of pretest responses.

Because students in 131E performed so well on other measures of ability at the end of the term, very little correlation was expected, but some was hoped for.

None of the items had a significant correlation with High versus Low course grades. In fact, none had a correlation with magnitude greater than 0.15. Qualitatively examining the patterns as was done in Chapter 6, we find the following results:
1) No real difference in patterns. The students with grades of C+ or worse clustered a little more tightly than other groups, but their lower numbers (16 total) make it hard to tell if this clustering is real or merely a statistical artifact.

2) 56% of the A/A- students picked the correct response, 48% of the B+/B/B- students picked the correct response, and 38% of C+ or worse students, with corresponding spread of percentages to the various distractors. While the actual correlation is not significant, this does suggest that the better students follow the pattern established in Chapter 6 of clustering more on a single answer while other groups are more spread out.

3) Only 11 of the 153 students picked the energy solution for this item, but it's worth noting that 9 of those 11 earned an A or A- in the class, with the other two being in the B range. 95% of students scoring B+ or below picked the correct kinematics decomposition. This is somewhat promising, as it indicates that the better a student absorbs the lessons of 131E (which include energy as a problem-solving tool), the more likely he or she is to use energy to solve a state system problem.

4) Response patterns almost identical for all three course grade groups. Either this item does not test anything the students were learning, or it tests something that was independent of course grading.

5) All three grade levels had the same percentage of correct respondents, but the higher the overall course grade, the more likely a student would be distracted by option b or e instead of option a. Treating the problem as a single step running from interval 1 through interval 6 was the only distractor chosen by the C+ or lower grade group, and as course grades improved, the other three distractors picked up adherents.

6) While all grade levels were equally likely to pick deep structure strategies, the higher the student's course grade, the more likely he or she was to pick a, the strategy using energy conservation. 70% of A students picked a, while only 38% of C or lower
students did. Conversely, only 19% of A-level students chose the forces solution, compared to 44% of the lowest group of students. B-level students were in between the two extremes in both cases. In fact, looking at a as “right” and d as “wrong,” the dichotomous correlation is 0.17 with a t-score of 1.99, which is barely statistically significant.

7) Similarly to item 6, overall use of deep structure is about the same for all grade levels. But once again there is a shift from kinematics/forces to energy strategies that occurred with increasing course grade. With energy solutions “right” and non-energy deep structure strategies “wrong,” the dichotomous correlation between strategy and course grade is 0.20, with a t-score of 2.36 (P<0.05).

8) This item lacked any real trending of a statistically significant nature. Even the choice of energy versus non-energy deep structure strategies was not particularly influenced by course grade. This may be due to the nature of the item, which can be solved just as easily with kinematics as with energy.

9) No patterns at all that could be discerned, nor was there a choice of an energy conservation strategy for this item.

10) The vast majority (130 out of 153) of the class picked the energy strategy for this item, but there was still a correlation to course grade of 0.18, and a t-score of 2.17 (P<0.05). There was no correlation between overall deep structure use and grades, however.

Even with the restricted range of grades, there were a number of patterns that could be seen based on differences in overall course grade. Most cannot be considered robust results, but a number of statistically significant correlations were found when examining which sort of deep structure strategy a student would apply. Use of energy conservation
was heavily emphasized in Physics 131E, and it seems that good performance in the class went hand in hand with choice of energy conservation strategies in Section II of the PDD.

Now to consider the FCI posttest scores. A “High” FCI score was 25 out of 30 or better. All correlations are once again dichotomous.

1) Correlation 0.19, t-score 2.05 (P<0.05). The students who recognize the collision also seem to have performed better on the FCI, although this only accounts for 4% of the variance.

2) Correlation of 0.19, t-score of 2.37 (P<0.05). It should be noted at this point that the covariances for these items may overlap. Items 1 and 2 do not necessarily explain 8% of the variance between them.

3) Correlation of 0.22, t-score of 3.15 (P<0.01). Again, of the 11 students who opted for the single sub-problem energy solution on this item, 9 of them were in the “high” group. It should also be noted that overall “correctness” on this item was very high, so the variation in the pattern of distractors chosen represents only a small number of total students.

4) No significant correlation. The A-level and B-level student groups had nearly identical response patterns, but the students scoring C+ or less on the final course grade had a markedly different response pattern, being twice as likely to pick option b (about 50% versus around 25% likelihood), one of the distractors without a collision. Unfortunately, this group is very small compared to the entire class (16 students), and it’s hard to make too firm of a statement based on their responses. A dichotomous correlation calculated using this group as “Low” and everyone else as “High” did not yield a significant result.
5) Essentially the same results as in the comparison based on grades. No difference between groups in terms of picking the correct option, but students with higher FCI scores tended to pick a different distractor (b) than the students with lower scores did (a).

6) The likelihood of picking the energy solution increased slightly as FCI scores increased, but not to a statistically significant level. There was no other pattern that could be discerned.

7) Correlation between “high/low” FCI and “deep/not-deep” strategies was 0.30, with a t-score of 3.89 (P<0.01). Additionally, when looking at a choice of the energy strategy versus the kinematics strategy, the higher-scoring students were more likely to pick the energy strategy, with a correlation of 0.26 and a t-score of 3.07 (P<0.01). This item correlates very well with FCI scores, at least for this group of students.

8) As with the grades, there is no real pattern or correlation here, likely due to the kinematics strategy being as simple to implement as the energy strategy, if not moreso.

9) There is a slight improvement in student performance as FCI scores increase, but it is not statistically significant.

10) There is a fairly strong pattern when looking at the percentages of each response by groups (0-18, 19-24 and 25-30 out of 30), but the increase in both deep structure responses and energy strategy responses with FCI scores is not statistically significant. Again, given that 130 out of 153 students picked the energy strategy in the posttest, it’s not surprising to see very little correlation, as there’s not much variation.

Again, there’s a number of interesting patterns and correlations in the posttest data, but the generally high performance of the students on all measures makes it more difficult to see correlations. Some items seem to have very strong ties to other measures of ability, and those the author would prefer to retain in later versions of the test, perhaps slightly
modified. However, some items consistently showed no correlations, suggesting very
careful consideration of whether they should be retained should be exercised. The validity
of these items is not utterly destroyed, but neither can it be seen as high.

9.5. Comparison of pretest PDD to posttest PDD results

Finally, since this run of the PDD did have the advantage of a pretest and a posttest,
the patterns of student responses between the two administrations should be compared.
The students learned a great deal over the course of the quarter, so it is expected that their
performance on the PDD should change noticeably as well.

Because these inter-test differences are considered by the author to be more
important than the small correlations (anything under 0.3 is fairly small) between the PDD
and other measures of ability, this section will go into more detail and present full data in
tabular form for ease of examination. Each item will get its own table. Tables 9.1 through
9.10, with five columns and five rows.

The columns will be:

Option - a, b, c, d or e. There is no e in Section II, but this did not stop some
students from picking it a few times.

Description - A short description of what that option represents. Correct answer,
type of student difficulty, or type of strategy.

Pretest - The number of students out of 157 who picked an option on the pretest.

Posttest - The number of students out of 153 who picked an option on the posttest.

Change - This will be a change in absolute numbers from pretest to posttest. Keep
in mind that the total of this column should be -4, since four less students took the posttest
than took the pretest.
The rows will cover the five options possible for each item. Following each table will be a short discussion of its meaning.

The validity the author is concerned with here is a sort of enhanced face validity. It is known that the students learned a great deal, and any test which does not reflect such a change in student ability levels should be seen as somewhat suspect. At the very least, care should be taken to determine whether the course actually taught the skill or material tested by the item.
### Item 1

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Missing collision</td>
<td>55</td>
<td>31</td>
<td>-24</td>
</tr>
<tr>
<td>b</td>
<td>Correct</td>
<td>74</td>
<td>90</td>
<td>+16</td>
</tr>
<tr>
<td>c</td>
<td>Boundary problem</td>
<td>16</td>
<td>18</td>
<td>+2</td>
</tr>
<tr>
<td>d</td>
<td>Extended collision</td>
<td>4</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>e</td>
<td>Boundary problem</td>
<td>8</td>
<td>14</td>
<td>+6</td>
</tr>
</tbody>
</table>

Table 9.1 - Pretest/Posttest responses for Item 1 on the PDD v2.2

As expected, the percentage of correct responses on this item is higher in the multiple choice format than it was in open response on the PDD v1.1. There is a small shift from the "no collision" answer to more correct answers, and the small number who felt that the collision took place over an extended period vanished on the posttest.

The results were somewhat disappointingly muted for this item, but that could be in part due to the generally higher ability of these students and the fact that multiple choice items present the correct answer to a student quick enough to look for it. As this item is based on a laboratory that prompted much of the work on this thesis, however, the face validity is considered to still be high enough to retain the item. However, changes will be seriously considered.
**Item 2**

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Boundary problem?</td>
<td>20</td>
<td>22</td>
<td>+2</td>
</tr>
<tr>
<td>b</td>
<td>Correct</td>
<td>71</td>
<td>77</td>
<td>+6</td>
</tr>
<tr>
<td>c</td>
<td>Pinball paradigm</td>
<td>41</td>
<td>38</td>
<td>-3</td>
</tr>
<tr>
<td>d</td>
<td>Multiple errors</td>
<td>11</td>
<td>4</td>
<td>-7</td>
</tr>
<tr>
<td>e</td>
<td>Multiple errors</td>
<td>14</td>
<td>12</td>
<td>-2</td>
</tr>
</tbody>
</table>

Table 9.2 - Pretest/Posttest responses for Item 2 on the PDD v2.2

Very little changed on this item. Two possibilities come to mind. One, the item simply isn’t testing what the author thinks it is testing. Two, Physics 131E did not spend enough time on springs to really affect student ideas of how they worked. In any case, as previously mentioned, this item should be altered to help clarify the sort of errors involved in each option.
### Item 3

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Energy correct</td>
<td>0</td>
<td>11</td>
<td>+11</td>
</tr>
<tr>
<td>b</td>
<td>Pinball, energy</td>
<td>2</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>c</td>
<td>Mistaken Collision</td>
<td>8</td>
<td>1</td>
<td>-7</td>
</tr>
<tr>
<td>d</td>
<td>Pinball, dynamics</td>
<td>14</td>
<td>3</td>
<td>-11</td>
</tr>
<tr>
<td>e</td>
<td>Dynamics correct</td>
<td>133</td>
<td>137</td>
<td>+4</td>
</tr>
</tbody>
</table>

Table 9.3 - Pretest/Posttest responses for Item 3 on the PDD v2.2

While it's a small number, 11 is certainly better than zero, and some students did shift to picking the single-sub-problem solution energy conservation would allow. Most of the change came out of the distractors, which didn't seem to really distract very well in the first place. The author is generally pleased with the results on this item, but feels that the distractors need further modification. It's also possible that the "friction doesn't conserve energy" concept is still hindering these students, and an item not using friction should be developed (and was for version 2.3).
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>No coll. boundary</td>
<td>14</td>
<td>10</td>
<td>-4</td>
</tr>
<tr>
<td>b</td>
<td>No collision</td>
<td>65</td>
<td>41</td>
<td>-24</td>
</tr>
<tr>
<td>c</td>
<td>No coll. pinball</td>
<td>6</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>d</td>
<td>Collision, pinball?</td>
<td>11</td>
<td>7</td>
<td>-4</td>
</tr>
<tr>
<td>e</td>
<td>Correct</td>
<td>61</td>
<td>89</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 9.4 - Pretest/Posttest responses for Item 4 on the PDD v2.2

For some reason, this collision continues to prove more difficult for students to identify than the one in item 1, although it’s possible that the “extended collision” mentioned in interviews (Chapter 8) came into play here. However, this item did show some marked improvement between pretest and posttest, possibly in part due to the emphasis on collisions and internal energy transfers in the 131E course. Validity for this item seems to be within the realm of possibility.
**Item 5**

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>One step</td>
<td>31</td>
<td>27</td>
<td>-4</td>
</tr>
<tr>
<td>b</td>
<td>One step, no ramp</td>
<td>12</td>
<td>11</td>
<td>-1</td>
</tr>
<tr>
<td>c</td>
<td>Correct</td>
<td>94</td>
<td>104</td>
<td>+10</td>
</tr>
<tr>
<td>d</td>
<td>Boundary problem</td>
<td>11</td>
<td>3</td>
<td>-8</td>
</tr>
<tr>
<td>e</td>
<td>Boundary problem</td>
<td>9</td>
<td>8</td>
<td>-1</td>
</tr>
</tbody>
</table>

Table 9.5 - Pretest/Posttest responses for Item 5 on the PDD v2.2

This item was one of the easier ones for the 131E students, even in the pretest, and it's not too surprising to see no change in performance. At least it was a small improvement rather than a small backslide. Despite the lack of evidence to support the validity of this item, it will be retained in some form. It is possible that such evidence will arise when the PDD is administered to non-honors students.

The next five tables concern items from Section II of the test. These items did not have an option e given, but that did not stop a few students from picking that option.
### Item 6

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Energy</td>
<td>11</td>
<td>94</td>
<td>+83</td>
</tr>
<tr>
<td>b</td>
<td>Surface Features</td>
<td>8</td>
<td>6</td>
<td>-2</td>
</tr>
<tr>
<td>c</td>
<td>Equation Hunt</td>
<td>52</td>
<td>14</td>
<td>-38</td>
</tr>
<tr>
<td>d</td>
<td>Force/Kinematics</td>
<td>83</td>
<td>39</td>
<td>-44</td>
</tr>
<tr>
<td>e</td>
<td>Not an option</td>
<td>1</td>
<td>0</td>
<td>-1</td>
</tr>
</tbody>
</table>

Table 9.6 - Pretest/Posttest responses for Item 6 on the PDD v2.2

Well, here’s a large pretest/posttest difference, which makes a strong case for the validity of this item. After instruction, students were much more comfortable with using energy conservations to solve this problem. This could be accounted for by the fact that many students came in with only a sketchy idea of how energy conservation could be used. However, the strong reduction in the number of students using the “equation hunt” strategy suggests an overall greater comfort with the material.
### Item 7

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Kinematics</td>
<td>48</td>
<td>26</td>
<td>-22</td>
</tr>
<tr>
<td>b</td>
<td>Equation Hunt</td>
<td>64</td>
<td>12</td>
<td>-44</td>
</tr>
<tr>
<td>c</td>
<td>Energy</td>
<td>38</td>
<td>111</td>
<td>+73</td>
</tr>
<tr>
<td>d</td>
<td>Surface Features</td>
<td>5</td>
<td>4</td>
<td>-1</td>
</tr>
<tr>
<td>e</td>
<td>Not available</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 9.7 - Pretest/Posttest responses for Item 7 on the PDD v2.2

Another strong shift into an energy conservation strategy and away from both kinematics and equation searching. This item also seems to be worth retaining for later iterations of the PDD.
### Item 8

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Equation Hunt</td>
<td>30</td>
<td>19</td>
<td>-11</td>
</tr>
<tr>
<td>b</td>
<td>Energy</td>
<td>13</td>
<td>31</td>
<td>+18</td>
</tr>
<tr>
<td>c</td>
<td>Equation Hunt</td>
<td>27</td>
<td>30</td>
<td>+3</td>
</tr>
<tr>
<td>d</td>
<td>Kinematics</td>
<td>84</td>
<td>72</td>
<td>-12</td>
</tr>
<tr>
<td>e</td>
<td>Not available</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 9.8 - Pretest/Posttest responses for Item 8 on the PDD v2.2

As discussed earlier, the kinematics solution to this item is not particularly difficult, which could explain why the energy solution did not gain much in popularity. The rise that was observed may be due to the "use the most recently-learned concept" strategy that some students develop, rather than from seeing energy as an inherently better problem-solving tool. The item should be altered to increase the difference in the mathematical difficulty of the energy and kinematics solutions, as well as to add a genuine surface features option if possible.
### Item 9

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Surface Features</td>
<td>15</td>
<td>10</td>
<td>-5</td>
</tr>
<tr>
<td>b</td>
<td>Kinematics</td>
<td>60</td>
<td>56</td>
<td>-4</td>
</tr>
<tr>
<td>c</td>
<td>Equation Hunt</td>
<td>22</td>
<td>21</td>
<td>-1</td>
</tr>
<tr>
<td>d</td>
<td>Newton’s 2nd Law</td>
<td>58</td>
<td>64</td>
<td>+6</td>
</tr>
<tr>
<td>e</td>
<td>Not available</td>
<td>0</td>
<td>2</td>
<td>+2</td>
</tr>
</tbody>
</table>

Table 9.9 - Pretest/Posttest responses for Item 9 on the PDD v2.2

Essentially no changes in performance to speak of on this item. Combined with the somewhat odd nature of the problem and the lack of a viable energy conservation strategy, the author doubts this item should remain on the PDD, as it has poor face validity and no comparison validity. Attempts will be made to fashion a different item that probes conservation of momentum, or perhaps that concept should not be included in the Section II of later versions.
### Item 10

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Surface Features</td>
<td>16</td>
<td>3</td>
<td>-13</td>
</tr>
<tr>
<td>b</td>
<td>Newton's 2nd Law</td>
<td>49</td>
<td>15</td>
<td>-34</td>
</tr>
<tr>
<td>c</td>
<td>Energy</td>
<td>37</td>
<td>130</td>
<td>+93</td>
</tr>
<tr>
<td>d</td>
<td>Equation Hunt</td>
<td>52</td>
<td>3</td>
<td>-49</td>
</tr>
<tr>
<td>e</td>
<td>Not available</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 9.10 - Pretest/Posttest responses for Item 10 on the PDD v2.2

This item had perhaps the most striking shifts in student responses. While the specific nature of option d is probably what attracted so many students initially, the students were also probably able to tell it was a somewhat odd use of the relevant equation by the time they completed the quarter, hence the precipitous drop in support for d. Future versions of the PDD might do well to be more specific in the distractors for the goal-oriented or “equation hunt” options.

Overall, looking back at all the examinations of this chapter, it appears that while the items on the PDD v2.2 do a good job of probing student problem-solving strategies, the test as a whole has many holes that could be plugged. Some items need serious revision or outright replacement, but others are in good shape, with reasonable predictive validity and face validity. Possible alterations will be discussed in the next chapter.
CHAPTER 10

FUTURE DIRECTIONS, CONCLUSION

Now that the PDD v2.2 has been shown to have a reasonable amount of validity, the author is confident that it is worth the effort to further develop the PDD in the multiple choice format. There are three basic ways in which the PDD can be further enhanced.

First, the deficiencies in the existing test can be corrected. Several of the items had formatting choices that seemed to obscure the desired results, and one simply didn’t seem to work at all. A revision to version 2.3 could focus on repairing these flaws and replacing the one troublesome item in Section II.

Secondly, additional items should be created. If a large enough bank of items can be created, then reliability can be measured reasonably well. Additionally, item banking would allow for retesting throughout an academic year without the students simply answering the same items over and over.

Third, the domain of the PDD could be expanded. All of the items currently come from introductory mechanics concepts, but it is certainly possible to create items with context drawn from electrodynamics or even modern physics.

In all three cases, the PDD should be administered at other institutions in order to get a broader base of responses.
This chapter will be divided into five sections:

10.1. The PDD v2.3 revisions.
10.2. New Section I items.
10.3. Additional content areas.
10.4. Administration at other institutions.
10.5. Conclusions.

10.1. The Problem Decomposition Diagnostic v2.3

As a result of the analyses performed in Chapter 9, the PDD has been revised in an attempt to correct what is seen as its largest flaws. Appendix K contains the full text of this revision, which will now be discussed item by item.

1) The only real change made to this item was to make the pendulum and the block smaller so that they more clearly fit inside interval 3. If this change results in options c, d, and e losing all support, then the author may consider making item 1 a 2-option question. Breaking the pattern of five options per item might have some negative consequences, but if the other distractors do not distract anymore, then it would make sense to eliminate them. This item tests a single student difficulty, leaving little room for other errors.

2) For this item, the main problem on v2.2 was that it was difficult to distinguish between the "pinball paradigm" error and simple boundary problems. So the ramp and the spring extension were stretched to cover intervals 1 through 3. Additionally, the landing point was raised to be equal to the top of the ramp. The difference in height was originally picked to see if the asymmetric flight path would draw students into splitting the arc into
two sub-problems, but that did not happen in the v1.1 administration. Therefore, the decision was made at this point to simplify the problem by raising the landing point. As the changes to this item affect the nature of the responses, it may be necessary to administer it in an open response format.

3) The fact that friction is often taught as being "non-conservative" meant that many students would think it impossible to use energy conservation to solve the problem in this item. By replacing the frictioned surface with a ballistic arc, the revised item 3 retains the same basic physical structure, but now avoids the difficulty with friction. Instead of ramp-friction-spring, it is ramp-ballistic-spring. As none of the distractors were particularly distracting in the v2.2 administrations, the author altered them slightly for this revision.

4) This item shared item 2's problem with the spring compression length. The spring now compresses from interval 5 through interval 7. Additionally, the blocks are smaller so as to clearly touch while fully inside interval 5. As with item 2, the changes in this item could invalidate some of the distractors, so it may be necessary to implement an open response version. However, the issue of whether the collision was omitted or extended to cover the entire compression of the spring could not be addressed here. Since both errors are serious, though, this is not seen as a major problem.

5) In order to clarify some of the student errors, the launch ramp was extended to cover intervals 1 through 3, and the landing ramp was replaced by a landing block in interval 7. As this is a fairly straightforward item that most students were able to get correct, the author does not think it will be necessary to use an open response format to check the distractors again.

6) No changes.

7) No changes.
8) The rather baroque wording of this item was simplified and the requirements altered slightly. The conditions have changed so that the energy strategy is computationally simpler than the kinematics strategy, unlike the original version where both strategies were roughly equal in computational difficulty. Finally, one of the “equation hunt” options was replaced with a surface features option to bring this item more in line with the rest of the section.

9) Totally replaced. The author has, for now, abandoned the idea of involving a conservation of momentum item in Section II. The new item involves a new surface feature, the pulley. It has the advantage of fitting the “surface/goal/deep kinematic/deep energy” structure of the other four items. Additionally, it is a surface feature strongly associated with a forces approach, even though energy could be used to more easily solve the problem. Students not totally confident with energy would therefore be more likely to pick the forces option.

10) No changes.

This revision reflects the generally stronger results found in Section II, in that most of the revision was done on Section I items.

10.2. New mechanics-content items

While the current length of the PDD seems to be fairly good for ease of administration, it is desirable to create additional items, if only to allow for an item bank that would give the tester greater flexibility. And, of course, a longer exam would lend itself more to analysis of reliability. The goal is to create items with different surface features but which cover essentially the same student difficulties as the original five items.
One quintet of new items was created in the fall of 1999 and administered in open form to the Physics H131 class. 57 students completed the test in such a way that their responses could be used in an attempt to generate multiple choice distractors. Three response sheets were unscorable, and three more were marginally scorable and were included in the 57 above.

Appendix L shows the open response format items, along with the instructions the students were given. Note that these items were created before the data from the PDD v2.2 was analyzed, so some of the “sticking points” found in that test (such as too-short ramps) can be found in these items.

Each item is meant to be analogous to the identically-numbered item in the PDD v2.2, but with different surface features or order of actions. Appendix M shows the tentative multiple choice versions of the items.

1) Item 1 essentially time-reverses the Pendulum-Box-Bash item, although it changes the collision to one which is completely inelastic. 27 of 57 students identified the collision in this item in some fashion, performance somewhat better than the 35% scored by their predecessors in 1998. This suggests that, at least for the block and pendulum surface features, a totally inelastic collision is more readily identified as important than a partially inelastic one. The curved path followed by the pendulum may have defused the “extended collision” student difficulty. For this item, the distractors were chosen based on popularity of responses.

2) The new Spring Launcher actually simplifies the situation by removing the change of height from the first sub-problem, but it is still essentially the same problem as its counterpart in the PDD v2.2. The spring is longer, however, as even before analyzing the v2.2 data the author realized the problems inherent in the shorter springs of the v2.2. Additionally, a “phantom” image of the fully extended spring was added in hopes of
visually clarifying what the spring did. The responses were more scattered on this item, so
the distractors were not simply chosen from the most popular responses. Rather, three
common errors were identified: the “pinball paradigm” where interval 1 was singled out as
a sub-problem; some sort of launching effect where interval 3 was singled out; and a
“something happens here” effect where interval 7 was labeled a sub-problem, perhaps
because of the “hit” on the water. 12 students had interval 1 separate, 19 students had
interval 3 separate, and 12 had interval 7 separate. No other error came close to these
three. To generate a fourth distractor, the author looked at pairs of these errors, and found
that 8 students had both interval 3 and interval 7 as sub-problems on their own. Further
open response work is probably necessary for this item.

3) This item was an attempt to follow Block Catcher I’s multiple solution pattern.
However, the curved paths in intervals 1-3 and 5-7 were intended to convince students that
energy conservation was necessary to solve those sub-problems, leaving this more of a test
of whether students believed that the non-frictionless section in intervals 3-5 prevents them
from using energy conservation for the entire item. As with Block Catcher I, a majority of
students (32 out of 57) broke the item up into three sub-problems. Only one identified it as
a single part problem. Due to the small number of wrong answers, the same method as
seen for item 2 above was used to construct distractors. However, most student errors fell
into the not-very-interesting category of boundary problems. Three students labeled
interval 7 as a sub-problem, perhaps due to the “collision” with the bell. It may be
worthwhile to restate the problem as having no friction anywhere and see if the students
still prefer breaking it into three sub-problems.

4) This item is similar to Block Catcher II, but it adds some complications. First,
the orange travels on a ballistic path, but all vertical velocity would be lost on collision with
the cart. Therefore, while it is essentially moving at a constant velocity (as does the block
in Block Catcher II), it has extra complications. Second, the spring is stretched rather than compressed. So far, all springs have either been compressed by something or started off compressed and extended until they reached relaxed position. The fact that this spring starts relaxed and gets extended could trigger different student ideas. Finally, the spring was made longer to avoid some boundary problems. Only one student exhibited any sort of "pinball paradigm" that could be conclusively identified, although it is possible that some of the "collisions" were actually students thinking the spring acted in interval 4 alone. However, as the surface features are different from the compression of a pinball plunger, it is possible that this idea was not activated. 18 out of 57 students could be considered to have identified the collision, although some of these may have been describing the spring action. This does seem to suggest that student recognition of collisions is very context dependent, with the pendulum being a surface feature that helps trigger student recognition. Of course, testing such a hypothesis rigorously is outside the bounds of this paper, but it seems a worthwhile subject to pursue. Distractors for this item were chosen based on popularity of responses.

5) This item replaces the simple ramp in the Stunt Ramp item with a cylindrical half-pipe ramp. The fact that the skateboarder goes down and then up while still on the ramp changes the situation enough to result in the ramp being seen (incorrectly) as involving two sub-problems. However, even when using kinematics and integrals to solve for the launch velocity, it is a single integral (from negative 90 degrees to theta). A few students (8 out of 57) did break the ramp into two or more sub-problems, while an almost equal number (7 out of 57) identified the landing in interval 7 as a separate sub-problem. Four students considered interval 3 to contain its own sub-problem, but that may have been simply the mathematical sub-problem of figuring out the direction of the launch velocity (it is theta, but that's not necessarily immediately obvious to all students).
All five of these items look promising, but most of them need revision and all should get further open-response testing to solidify the distractors.

Additional Section II items were not created for this administration. This was for two reasons. One was time. Open response items take longer for students to complete than multiple choice (in general), and the author did not wish to place undue time strain on the subjects by also including Section II items. The other reason was that, at the time of administration, Section II seemed to need no additional items. Since then, of course, item 9 has been replaced as a result of research analysis. And, naturally, Section II items really aren't meant to be tested with an open response format, although new multiple choice items could certainly have been added to the end of the test.

10.3. Additional content areas

Mechanics is not the only area in physics with complex problems. It would certainly be useful to be able to test a student's complex problem-solving ability in an electricity and magnetism class, or a waves and modern physics course. If nothing else, being able to track problem decomposition and classification skills over the course of a full academic year would help an instructor determine if a new form of problem-solving instruction was effective. Additionally, each area of physics comes with its own "blind spots," just as mechanics has with collisions.

Section II is probably more easily expanded into other subject areas than Section I, simply because it does not depend as much on the specific student difficulties. It would be worthwhile to check student comfort with conservation of energy through the entire first year of introductory physics. However, later courses also have their own new material which should be examined for useful deep structure strategies. For instance, appeals to
symmetry in electrostatics, when a student might be more comfortable just summing
everything and having it cancel in the end.

Section I could be expanded in two ways. The simpler way would be to simply
create mechanics-like multi-part problems in which the forces are electric instead of
gravitic, magnets instead of springs. However, many of the more difficult E&M problems
aren’t simply strung-together sub-problems, they involve simultaneous solutions.
Including such problems would require altering the format of the PDD to allow for
simultaneous sub-problems, which could make it much more difficult to interpret.

As the author’s primary expertise has been in the domain of introductory
mechanics, little has been done to date towards expanding the PDD into other content
domains. The input of those with greater experience in non-mechanics domains would be
greatly welcomed as work goes forward on the Problem Decomposition Diagnostic.

10.4. Administration at other institutions

Even when examining multiple courses at The Ohio State University, there will be
certain common factors that could bias results taken only at OSU. Admissions
requirements, course prerequisites, degree program course requirements and other aspects
of The Ohio State University’s programs would be similar for all students involved in work
at that institution.

Therefore, it is necessary to involve other institutions in "proving" the PDD. There
are two ways to go about this.

The first is the simpler from a procedural point of view. Copies of the PDD v2.3
would be sent to volunteer instructors at as many universities, colleges and perhaps even
high schools as the author can manage. Data would be collected on Scantron™ sheets, and where possible, scanned onto computer disk at the participating institution. Then all data would be entered (along with other information where available) and analyzed.

The second way would certainly be more time-intensive, but would likely be more valid, especially if the distractors used on the PDD v2.3 result from any local bias. Version 2.3 items would have the multiple choice options removed from Section I, and changed to an open response format. In this case, it is likely that only Section I would be tested for reasons of time, but Section II cannot be ruled out entirely. Data from such a trial would all be entered by the researcher, increasing the time requirement. However, the large and diverse sample would yield superior distractors for a potential v2.4.

If the open response format were to be used, it would be followed the next year by a multiple choice version incorporating the results of the open response administration.

Additionally, the researcher would like to involve other institutions in the examination of any new items or new subject areas mentioned in this chapter.

10.5 Conclusions

As implied by the previous four sections, this is not a completed work, so these conclusions should not be taken as the final word on the Problem Decomposition Diagnostic. However, there are a number of things that can be said about the PDD as a result of the research that went into this paper, both about the test itself and about related issues raised by the development process.

In the course of developing the PDD, a number of very interesting questions were raised. How do people really view collisions? How do they think about the action of springs? What good is energy conservation, anyway? These and other questions raised in
the course of development certainly merit attention, but they were outside the scope of this particular project. Some have been addressed in other works already (Zou), but others remain fertile ground for further research.

As for the test itself, while it shows a great deal of promise, it has yet to be transformed into an instrument that can be used in the classroom to measure the success or failure of a curriculum. However, most of the individual items have been shown to have at least some level of validity, and the general approach seems to be workable. Expansion to ten items in each section should allow for a reasonable test reliability without making the PDD too long to administer. Once the further measures discussed above have been implemented, the PDD should make a valuable addition to Physics Education Research.

Essentially, the Problem Decomposition Diagnostic is still in its childhood. But it looks like it will grow up into a healthy adult.

In closing, the author believes that a test of physics problem-solving skills is both necessary and possible. The Problem Decomposition Diagnostic has demonstrated itself as being reasonably valid and worth pursuing further. Such further development will hopefully form the core of this author's research for the next few years, and should involve researchers from other institutions as well.
INTRODUCTION

The labs in this appendix were designed originally for an honors course of Engineering students, with the intent of helping the students learn not just mechanics but also skills related to group work and design, as well as the ability to make reasonable approximations. Since then, versions of this course have also proven usable in non-honors calculus-based physics classes. Each lab presents the students with a goal (or goals) to obtain, such as "Devise an experiment to measure acceleration," and a number of items they can use in reaching this goal. The students then work as a group to design the experiment, and once the design has been approved, they perform the experiment and get results.

This version of the Toys In Motion lab manual includes all laboratory exercises used in the previous versions of the course. Several labs rely on equipment that is no longer available, and this will be noted in each lab so affected. In some cases, multiple variants of the same lab will be presented.
The equipment used in most of these labs consists of a digital timer, a number of rulers and meter sticks, and various Hot Wheels™ products, mostly cars and tracks. Other labs, however, do use equipment beyond this, such as action figures or battery-operated trucks and airplanes. This use of toys in all the labs seems to have a positive motivating effect on the students, engaging their sense of play and removing the negative connotations of standard lab apparatus (such as “boredom” and “unreality.”). Not all students will react positively, but most seem to.

It is suggested that students work in groups of 3-4 for these labs. preferably 4 so that they can split into two groups of two for some of the earlier labs. Grading should emphasize process, and if possible be performed during class...the goal of these labs is to get students to work together and think on their feet, not to give them reports to write alone and outside of class.
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>150</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>152</td>
</tr>
<tr>
<td>Format of Labs</td>
<td>153</td>
</tr>
<tr>
<td>Lab 1 - Average and Instantaneous Velocity</td>
<td>155</td>
</tr>
<tr>
<td>Lab 2 - Basics of Graphing</td>
<td>159</td>
</tr>
<tr>
<td>Lab 3 - Acceleration and Force, Gravity and Friction</td>
<td>161</td>
</tr>
<tr>
<td>Lab 4 - Ballistic Motion</td>
<td>167</td>
</tr>
<tr>
<td>Lab 5 - Forces in Balance, Tension and Friction</td>
<td>172</td>
</tr>
<tr>
<td>Lab 6 - Collisions (Graphing)</td>
<td>177</td>
</tr>
<tr>
<td>Lab 7 - Centripetal Force</td>
<td>181</td>
</tr>
<tr>
<td>Lab 8 - Energy</td>
<td>185</td>
</tr>
<tr>
<td>Lab 9 - Analysis (E-Yo)</td>
<td>192</td>
</tr>
<tr>
<td>Grading Philosophy</td>
<td>196</td>
</tr>
</tbody>
</table>
Format of Lab Writeups:

The lab writeups in this manual will follow a standard format, which will be explained here.

Task(s): A listing of the goals for the students in this lab. Labs will either have a number of short tasks or one large task. These tasks should be written on the blackboard or otherwise given to students at the start of the class. In some cases, it may be desirable to only give the students the first task of the day initially, then write down the second one once students start to finish the first, to keep attention on one design at a time.

Equipment: A list of the equipment provided to the students, as well as potential problems to consider. You may wish to label all cars and special track pieces like loops and launchers so that students can find the same ones week after week.

Main Point(s): What this lab specifically is trying to accomplish with respect to the students. Some labs focus more on the physics, while others are more concerned with the students learning to include the effects of approximations or to make estimations.

Sample Design(s): One or more possible designs to meet the goal of the lab. Students need not reproduce these specific designs, and may surprise you with new ways of accomplishing things. However, there are going to be designs students come up
with that won’t work with the equipment provided (such as any which require timing more accurate than 0.1 seconds). Ideally the instructor should be able to guide the students to a workable design through Socratic questioning. The sample designs are also helpful as reminders to keep in hand while helping students.

**Pitfalls:** Many labs will have attractive design possibilities that the students will wish to use, but which are for whatever reason untenable. In that case, this section will describe some of the more common design flaws seen in the previous administrations of the particular lab exercise.

Suggestions for running the lab sessions will be included at the end of Appendix 1. As a minimum, however, all students should have a lab notebook in which they keep all their data.
Lab 1: Average and Instantaneous Velocity, Acceleration

Tasks: A) Devise and perform two experiments to measure average velocity given the materials provided. B) Devise and perform two experiments to find an instantaneous velocity (other than zero) given the materials provided. C) Devise and perform two experiments to measure constant acceleration.

Equipment: For each group, enough track and connectors to make two sections between one and two meters long, two labeled toy cars (label with marker or paint), four meter sticks, two digital timers and something that a meter stick can be propped up on to form an incline. Masking tape should also be available in the room.

Main Points: Learn the use of the basic definitions of velocity, average velocity and acceleration. Get used to working in groups and designing experiments (as opposed to working from a “cookbook” style manual). Learn that sometimes simpler is better, and that one doesn’t need expensive equipment or elaborate mathematics to get data.

Sample Designs: See next page for details. Part A is simply measuring a distance and a time the car takes to move that distance, and dividing distance by time. The two designs should find two different ways to get the car moving. Part B hinges on the idea that when velocity at one time is zero, then the velocity at another time is twice the average velocity between those times. So designs should find a way to get a zero velocity at some point. Since this only works when acceleration is constant.
students should be sure to have flat tracks (tape them to rulers for slopes). Part C can use the same setup as Part B, but uses the data to find acceleration instead.

**Pitfalls:** No major pitfalls for part A, students should fairly quickly realize that they can just measure a distance and a time, and divide them. However, in part B, many will want to take limits as time intervals get smaller and otherwise use powerful mathematical methods to find instantaneous velocity in the manner they are used to doing in homework from graphs. Some may want to use sonic motion detectors if they are set up in the lab room. Students may also misunderstand and think they're required to find instantaneous velocity at all points, rather than at some arbitrary point. In part C, decide if you want to let students use the same data as in part B.

**Important timing note:** If you decide to let students use data from this lab in finding their results in Lab 3, you may wish to move part C of this lab into Lab 3, since the established Lab 3 exercises go very quickly when students are allowed to use old data.

A) **Measure Average Velocity two ways:**

**Method 1:** (Refer to Figure A1.1 a) Given a set distance \( \Delta x \) (use a ruler next to the track), time the car over that distance. Try to push the car the same amount each trial. The average velocity will be \( \Delta x \) divided by the interval \( \Delta t \).

**Method 2:** (Refer to Figure A1.1 b) Tape a section of track to a ruler to keep it straight. Then, prop one end of the ruler on a book or some other object. Release the car from a measured point on the track. The average velocity will be the \( \Delta x \), the distance traveled, divided by the time interval \( \Delta t \) needed to travel that distance.
Figure A.1: Diagrams for possible setups for Lab 1. a) Flat surface. b) Sloped surface.
B) Measure instantaneous velocity at a point two ways:

Method 1: (Refer to Figure A1.1 a) Push the car just hard enough that it stops before the end of the track, the distance it rolls being \( Ax \). Start the timer at a given point and stop it when the car stops. Instantaneous velocity at point 2 is twice the average velocity, or \( 2 \cdot \frac{Ax}{At} \).

Method 2: (Refer to Figure A1.1 b) Start the timer when you let go of the car and stop the timer when the car reaches the bottom. The instantaneous velocity at the bottom will be twice the average velocity, or \( 2 \cdot \frac{Ax}{At} \).

C: Measure Constant Acceleration Two Ways:

Method 1c: (Refer to Figure A1.1 a) Push the car gently enough that it will stop near the end of the track, and time it from a set point to where it stops. Find starting velocity as above, then use it to find the acceleration. \( \Delta x = v_{initial} \cdot At + a \cdot \frac{At^2}{2} \)

Method 2c: (Refer to Figure A1.1 b) Release the car from a measured point on the track, measure time \( At \) that the car takes to move the distance \( \Delta x \). Since initial velocity is zero. \( \Delta x = a \cdot \frac{At^2}{2} \).
Lab 2: Graphing Motion Under Constant Acceleration

Tasks: Analyze the motion of cars rolling on a flat track and on a sloped track. Graph position, velocity and acceleration versus time for both situations. If time allows, show that the area under the velocity curve equals the distance traveled.

Equipment: For each group, 2 cars, track sections, metersticks, 1 or 2 timers, small rulers, graph paper (students should bring their own, but have some on hand anyway). Work in groups of 2 students to reduce chances of one student doing the work and the other three copying that student's graph.

Main Points: Reinforcement of the basic kinematics, but mainly this lab is about making sure all students have a certain minimum level of graphing skill. If time allows, demonstration of integration's validity (although not all students will be at the stage in Calculus that they know how to integrate).

Sample Designs: No real design work. Students should repeat the setups they used for the velocity/acceleration lab (Lab 1). Acceleration graphs should have horizontal lines, velocity graphs should have straight lines of the appropriate slope, and position graphs will have parabolas.
**Pitfalls:** The main problem with this lab will be that there’s bound to be a group where some students have plenty of experience graphing while other students will have almost no experience. Splitting the work groups into subgroups of two students each will help, although it may be necessary to guide the groups in splitting so the split doesn’t result in one bored subgroup and one confused subgroup.

Many students have poor graphing skills but think they have good graphing skills, and will cheerfully make a large number of bad graphs, only to be angry when told they need to redo the graphs. Be sure to emphasize good graphing skills such as axis choice and labeling, filling the entire space available (when told to do six graphs, many students will try to fit all six in the top half of one sheet of graph paper).

On a technical level, the position graphs are likely to be the hardest for students unfamiliar with graphing, since they’re not straight lines, but rather parabolas.
Lab 3: Acceleration and Force: Gravity and Friction

Tasks: A) Devise and perform two experiments to find the acceleration of gravity, \( g \). B) Devise and perform two experiments to find the coefficient of friction, \( \mu \), between a car and a piece of track. Students may take \( g = 9.8 \text{ m/s}^2 \) for this second task. (Do Lab 1. Task C first if you've moved it to this lab.)

Equipment: All equipment from Lab 1 is required for Lab 2 as well. In addition, students may want to mass their cars, so a balance should be provided. Finally, students should have something reasonably unbreakable to drop should they decide to measure \( g \) by timing a falling object. Rubber balls can be provided for this purpose.

Main Points: Learn the use of the equations for acceleration. Newton's Second Law and the force of friction. As this is still only the second lab, the Main Points of Lab 1 also apply, since the students will require repeated exposure to and practice with them. Finally, students should discover in this lab that the shorter the time interval measured, the less accurate the results are, due to the limitations of human reaction speed.

Sample Designs: See next two pages. For gravity, the simplest experiment is to drop an object a measured distance, taking the time of descent and calculating \( g \) from that. A second likely method is to roll a car down a ramp, determine the acceleration, then use trigonometry to find \( g \) from that.
For friction, the simplest method is to roll a car along a flat surface, find the acceleration from distance traveled and time taken, and multiply by the mass to get the force of friction. From there, divide by the Normal force (weight in this case) to get $\mu$. The more complicated method uses a slope and finds the difference in acceleration between expected (from gravity) and experimental to get the acceleration of friction.

**Pitfalls:** In the gravity experiments, drops of less than 3 meters are likely to result in times close to the students’ reaction time, so expect almost meaningless numbers. And for shallow slopes, friction will be enough of a factor that the obtained $g$ will differ greatly from known values. Also, because reaction time is still a factor even in the slope experiments, students may get $g > 10 \text{ m/s}^2$ here, which will imply negative friction.

In the friction experiments, students may have enough data collected from the past two weeks to get numbers without actually taking new data. If you don’t mind them doing this, say nothing and many will mine old data for this experiment. However, if you want them to practice data taking, ask them to use different angles or different cars in this experiment. In this situation, it helps to shuffle the cars between weeks.
Figure A.2: Diagrams for Lab 3, Part A. a) A car rolling down a slope. b) A falling object.
A: Finding the Acceleration of Gravity:

Method 1a: (Refer to Figure A1.2 a) Measure acceleration as in Lab 1C. This acceleration will be proportional to the acceleration of gravity, \( a = g \cdot \sin \theta \).

Repeat for several heights \( y \) and two different cars. Do the results seem to be better for greater or lesser angles?

Method 2a: (Refer to Figure A1.2 b) Drop an object, preferably one which won’t break (like an eraser or a track connector) from a set height \( y \). Start the timer when it is dropped, stop the timer when it hits the ground. Since initial velocity is zero, \( y = gT^2 \). Do this several times and average the results. Are you closer to the accepted value than with using method 1a? Why or why not? Do you get better results for larger or smaller \( y \)?

Method 3a: Some students will hit upon the method of using a pendulum to measure gravity’s acceleration. Be sure that everyone in the group understands the pendulum equation...if not, encourage use of another method.
Figure A.3: Diagrams for Lab 3, Part B. a) A car rolling on a horizontal surface.  b) A car rolling down a slope.
B: Measuring the Coefficient of Friction:

Method 1b: (Refer to Figure A1.3 a) Push the car just hard enough that it stops before the end of the track. Start the timer at a given point and stop it when the car stops. Acceleration between point 2 and point 3 can be found from $\Delta x = v_{\text{initial}} \cdot \Delta t - a \cdot \Delta t^2/2$. (minus sign since it’s slowing down and you only need the magnitude) where initial velocity is twice the average velocity of $\Delta x/\Delta t$.

The force of friction is then $F_k = m \cdot a = \mu \cdot m \cdot g$ (since normal force $N$ is just $mg$ in this case). Note how mass cancels out of the equation.

Method 2b: (Refer to Figure A1.3 b) Find the acceleration down the slope as $\Delta x = a \cdot \Delta t^2/2$ (initial velocity is zero), this acceleration will be proportional to the net force on the car, $F = m \cdot a$.

The net force on the car is a component of weight downhill, minus friction trying to keep the car from moving.

$$F_{\text{net}} = W \cdot \sin \theta - F_k = m \cdot g \cdot \sin \theta - \mu \cdot m \cdot g \cdot \cos \theta.$$ 

Given mass, the accepted value of $g$, the angle and the measured acceleration, solve for the coefficient of friction.

Students can check this against data for the acceleration of gravity as determined by Method 1a. The difference between your observed acceleration and the accepted value of $g \cdot \sin \theta$ will be the acceleration due to friction. Place this into the equation for Method 1b above to solve for the coefficient of friction. Are the two values close? If not, are your results better for larger angles or smaller?
Lab 4: Ballistic Motion

Task: Find two ways to determine the angle of the provided missile launcher. One of these methods will be very easy (i.e. use a ruler), the other must involve measurements of the missile in flight. ALTERNATE VERSION: Using a Hot Wheels™ car, track, launcher and ramp, determine the angle of the ramp in two different ways.

Equipment: For each group of four, one action figure with missile launcher, plus missiles. Also available in the room should be masking tape and boxes or books (including the student textbooks) to use as a landing platform, if required by the design. ALTERNATE VERSION: Each group will require a car, a short length of track, a "stunt ramp" and a launcher. Mattel changes the style of launcher periodically, but for this lab, any launcher that can be hooked up to a track is usable. This version also requires the non-action-figure equipment of the main version.

Main Points: Students should learn that symmetric ballistic paths are much easier to work with than non-symmetric ones. Additionally, while many textbooks provide numerous equations for ballistic paths, this lab exercise presents a problem not easily solved by the complicated equations found in most books, requiring the students return to more basic kinematic equations and find their own solution. Finally, skills at taking hard-to-obtain data (such as that of a missile hurtling through the air) will be practiced.
**Sample Design:** Successful designs will have the missile or car land at the same height as it launched from (the height of the end of the launcher or ramp), so that the midpoint of the path will be on the axis of symmetry. This can be accomplished by either raising the landing point by piling up books or other objects, or by lowering the launch point to the edge of the table. The former method is generally easier. Students measure how high the missile is at the midpoint in flight, then use kinematics to find time of flight and initial vertical component of velocity. Given those and the measured horizontal distance, the horizontal component of velocity and hence the angle of launch can be found.

**Pitfalls:** Students will want to directly time the missile or car in flight. While they should have been exposed to the idea that very short times are almost impossible to measure by hand, they may need reminders. Students who do not hit upon the idea of landing at the same height as launching may bog down for some time in unsolvable equations, so keep an eye out for students falling into this trap. Some students may insist that the height of the missile in flight cannot be measured and will need to be convinced that a rough measurement is good enough. Finally, some of the launchers are strong enough to fire the missile farther than the length of the table at some angles, so they may need to adjust the angle so that the missile lands on the table. If using the Hot Wheels™ variant, determine in advance a good setting for the launcher that will not send the car off the end of the table and provide that information to the students.

Most successful designs for this exercise will require at least three people to run the experiment.
Figure A.4: Diagrams for Lab 4. a) Missile launcher version. b) Hot Wheels™ version.
Determine the angle of the ramp:

**Method 1:** (Refer to Figure A1.4 a) Set up the missile launcher, then set up a "landing strip" as tall as the end of the end of the launcher is high. One person will launch the missile so that it lands on the books. A second measures how far from the launcher the missile lands, while a third measures how high in the air the missile flies. These measurements will be very hard to get with any accuracy, so repeat the trials several times with different observers. be sure to set the launcher at the same angle each time (if it's adjustable).

From how high the missile flies, \( h \), you can determine both the time spent in the air and the initial vertical component of velocity:

1) \( h = g \cdot t_{\text{down}}^2/2 \) (second half of arc)

2) \( h = v_{y,\text{initial}} \cdot t_{\text{up}} - gt_{\text{up}}^2/2 \) (first half of arc)

Since the path is roughly symmetric, \( t_{\text{down}} = t_{\text{up}} \). In other words, the time the missile takes to travel from the top of the arc to the landing pad is the same as the time it takes to go from the end of the ramp to the top of its arc. Total time in the air is twice \( t_{\text{down}} \).

Knowing the total time of flight and the total distance traveled, you can assume horizontal velocity remains constant throughout the flight (air friction is small enough you can ignore it for now), and find the initial horizontal component of velocity:

3) \( v_{x,\text{initial}} = \Delta x/t_{\text{total}} \) (\( t_{\text{total}} = 2t_{\text{down}} \))

170
With the horizontal and vertical components of the initial velocity known, it's just trigonometry to find the angle:

4) \( \theta = \arctan \left( \frac{v_{y,\text{initial}}}{v_{x,\text{initial}}} \right) \)

The above method also works for the Hot Wheels\textsuperscript{TM} version of the lab, although the arc will of course be shallower.

Method 2: The second method is the easier one, and it is simply to measure the angle geometrically, either using a protractor or by measuring rise and run. This method provides a check on the first one.
Lab 5: Forces in Balance - Tension and Friction

Tasks: A) Measure the maximum force which can be exerted by the battery operated truck in the given experimental setup (see next page for diagram). B) Measure sliding coefficient of friction $\mu$ between the ramp and a cart dragged behind the truck.

Equipment: Each group of four needs the following items: 1 wooden ramp (either painted or varnished, but smooth), 1 battery-powered “monster truck”, 1 box with paper bottom (either a cardboard box or a plastic box with paper taped to the bottom), 1 pulley, 1 post-and-bar setup for setting the ramp at an angle, 1 weight hanger, an assortment of small masses (about 100-150g worth, in increments of 1g, 5g, 10g, 50g recommended), meter sticks, string.

Note, depending on the nature of the trucks used, it may be necessary to alter them. Inexpensive toy trucks often have motors that lock up at rather low torques. Wrapping the wheels of these trucks in masking tape reduces the coefficient of friction to the point that the motors will not be overwhelmed, as the wheels will slip before the maximum output of the motor is reached.

Main Points: The main purpose of this lab is simply to gain practice with force diagrams and using tension, as well as give the groups one last week to get used to working as a team before the labs start to get more difficult. If the instructor wishes to focus more closely on design in the labs, this lab may be eliminated.
Sample Design: There is little design work to do in this lab, although students will need to determine for themselves that they can find the maximum force the truck can exert by hanging weight off the pulley until the truck no longer moves. They will also need to be careful in setting up force diagrams to determine the exact value of that force. See the next page for the equations involved.

Pitfalls: Students may have trouble with the idea that the force of friction can be in the same direction as motion, as is the case with a wheeled vehicle under power. They may need to be led through their force diagrams carefully to determine that the friction force needs to be uphill to make the forces add up. Additionally, it will probably take some effort in Socratic dialogue to lead them to the idea that friction in this case is opposing the forces downhill, not the motion uphill. Or, more specifically, that it is opposing the free slippage of the wheel against the ramp, and that is being turned into uphill force.

As far as equipment is concerned, the battery-powered trucks aren’t exactly precision instruments, and there will be a fairly wide range of weight which will stop the truck (different trials will yield different numbers). The gears tend to seize up when too much force is applied, which is why the tires should be replaced with a layer of masking tape, to reduce friction and hence the force needed.
Figure A.5: Lab 5 diagrams. a) Truck pulling mass hung from a pulley. b) Truck pulling mass in cart tied behind it on slope.
A: Determine Maximum Force Exerted By Truck:

**Method:** (Refer to Figure A1.5 a) Draw a force diagram for the truck and the mass. Assuming the string is relatively ideal, the tension force will be the same for both. The force, $F_{\text{truck}}$, exerted by the truck will be uphill, while the tension and part of gravity will pull the truck downhill. Add mass until the truck cannot move uphill anymore, this mass will be $m_{\text{max}}$.

$$F_{\text{truck}} = \text{Tension} + m_{\text{truck}}g \cdot \sin\theta \quad \text{(from truck force diagram)}$$

$$\text{Tension} = m_{\text{max}}g \quad \text{(from mass force diagram)}$$

$$F_{\text{truck}} = m_{\text{max}}g + m_{\text{truck}}g \cdot \sin\theta \quad \text{(final equation)}$$

B: Measure Sliding $m$ Between Cart and Ramp:

**Method:** (Refer to Figure A1.5 b) Draw force diagrams for both the truck and the cart. The tension force can again be assumed to be the same magnitude for both objects. Three forces will act on the truck in the direction of the ramp ($F_{\text{truck}}$, $m_{\text{truck}}g \cdot \sin\theta$, Tension) and three will act on the cart (Tension, $\mu m_{\text{cart}}g \cdot \cos\theta$, $m_{\text{cart}}g \cdot \sin\theta$). Add mass to the cart until the truck is no longer able to pull it uphill (be sure to tap the cart to get it moving, so you know the sliding friction is more than the motor can move). At this point, presuming the setup
hasn't changed, the tension should be the value found in Part A. The force of friction is the difference between the tension and the component of gravity, solve for $\mu$.

$$\mu \cdot m_{\text{cart}} \cdot g \cdot \cos \theta = \text{Tension} - m_{\text{cart}} \cdot g \cdot \sin \theta \quad \text{(Tension known)}$$

Note: While the cart is not actually moving in the final part, this is an estimate of sliding friction. By tapping the cart, you move into sliding friction, and if the truck still can't keep the cart moving, you know the force of friction is too high.
Lab 6: Collisions, Graphing

Task: Plot the position versus time and velocity versus time graphs for two cars involved in a collision. Put both cars on each graph together so they can be compared.

Equipment: A length of track (~1.5m), two cars, rulers and two digital timers for each group. Each student should also have graph paper to plot the data, and there should be a scale or balance in the room for massing the cars. Note: you will want to make sure the pair of cars each group gets can collide reasonably well, without one of the cars being forced up and over the other.

Main Points: This is a cumulative lab, to test the ability of the students to use the kinematic principles they’ve been using. Additionally, the students will practice the conservation of momentum in this lab, as well as skills in determining how to get needed data. Finally, this lab will help the students with their graphing skills.

Sample Design: Place one car at the middle of the track and push the other car towards it. Start the timers when the cars collide and stop each timer when the car that student is watching comes to a stop. Using the distance traveled and the force of friction (find by method of Lab 2B), determine the velocities of the cars right after the collision, then use conservation of momentum to determine the velocity of the pushed car right before collision. This velocity and the force of friction can be used to determine the velocity of the car right after being pushed. Finally, given all these velocities and distances, plot versus time for distance and velocity.
**Pitfalls:** Sometimes the combination of cars will be just right that the moving car comes to a complete stop at the point of impact. Since this is a somewhat trivial case, try to make sure the pair of cars each group gets won’t do this. Also, unless you have plenty of timers to go around, encourage students to calculate the time it takes from pushing the first car until the two cars impact, and only directly time the post-impact motion. Finally, despite its apparent simplicity, this lab can take more than two hours, so you may want to let students work on their graphs as homework, rather than insist they finish the graphs in class.
Figure A.6: Diagrams for Lab 6. a) Setup for Lab 6. Initial positions are shown by dashed lines, final positions by solid lines. All velocities not explicitly indicated are zero. b) Sample velocity versus time graph for Lab 6. c) Sample position versus time graph for Lab 6.
Graph the position and velocity of both cars versus time:

**Method:** (Refer to Figure A1.6 a) Measure positions from the front of Car 1 and the back of Car 2, so that at the collision point they have the same position.

Start Car 1 with a push. When it hits Car 2, two observers start their timers. One is watching Car 1, the other is watching Car 2. The distances and times from collision point to where the cars stop are measured.

Since the final velocities are known, as are the distances and times, the acceleration (due to friction) can be found, as can the instantaneous velocities at the time just after impact. Acceleration will be found from \( \Delta x = 0.5 at^2 \), while instantaneous velocity will be twice the average velocity, or \( 2\Delta x/t \).

Using conservation of momentum, \( m_1v_{1,\text{initial}} = m_1v_{1,\text{final}} + m_2v_{2,\text{final}} \), find the velocity of Car 1 just before the collision. The known acceleration due to friction for Car 1 can then be used to determine the velocity of Car 1 at its initial position, by the kinematic equation \( 2a\Delta x = v_f^2 - v_i^2 \). Additionally, \( \Delta v = at \) will yield knowledge of the time from start to impact. **Note:** If students have only two timers and need to calculate one of the time intervals, this interval (from the initial push until the collision) is recommended as the one they calculate.

Knowledge of constant acceleration and velocities at the various times (initial, final, right before and right after impact) will allow position and time graphs to be constructed. The graphs in Figure A1.6 (b and c) show one possible set of data.
Lab 7: Centripetal Force

Task: Design two different experiments to determine the tangential velocity (speed) of a toy airplane on a string that’s moving in a conical pendulum path.

Equipment: For each group - 1 toy airplane, stands and string to set up the conical pendulum, rulers and timer. For each student - Safety goggles. For the room - spare batteries, scale. If unable to find the toy airplane commercially, one can be constructed using battery-powered electric motors, toy propellers and a suitable base block.

Main Points: Use of centripetal force principles and force diagrams. Connection of linear and centripetal ideas of speed.

Sample Design: See next page for details. The simplest method is to simply measure the diameter of the plane’s path to find the circumference, then time the plane through several cycles to get an average speed. A more advanced method is to use the angle of the string to determine the centripetal acceleration, and hence the speed. Students may devise other methods, but one method should use centripetal principles.

Pitfalls: The plane will only fly well in one direction. students may be frustrated if they fail to get circular motion in their preferred direction. Students may also have trouble deciding where the center of mass is for purposes of working out the radius of motion and the effective length of the string: let them use any reasonable approximation, but be sure they realize the center of mass should be somewhere
under the point where the string connects. Since the plane balances well at rest.
Finally, this problem can be found in many physics textbooks, so it may be
inappropriate as a design lab if your textbook solves this problem.
Figure A.7: Diagrams for Lab 7. a) Physical representation. b) Force diagram.
Determine the Tangential Velocity:

**Method A:** (Refer to Figure A1.7 a) Determine the period $T$ of revolution using a timer (preferred method is to time the airplane through ten revolutions, then divide the total time by ten). Then measure the radius $r$ as well as possible. The speed will be the total distance $2\pi r$ divided by the period $T$.

$$v = \frac{2\pi r}{T}$$

**Method B:** Determine the angle $\theta$ by measuring any two of $l$, $r$ and $y$ and then applying trigonometry. Alternately, just measure $r$ and $y$ and keep them available.

Next, by use of force diagram or other methods (refer to Figure A1.7 b), determine that the centripetal acceleration and acceleration due to gravity are related by the equation:

$$\tan \theta = \frac{a_c}{g}$$

Centripetal acceleration is also equal to the square of the speed over the radius, so the equation can be solved for speed, thusly:

$$v = \sqrt{g \cdot r \cdot \tan \theta}$$

Finally, if $r$ and $y$ were measured, the ratio of the two can be substituted for the trigonometric expression, resulting in:

$$v = r \cdot \sqrt{\frac{g}{y}}$$

Note regarding the Force Diagram (Figure A1.7 b): Tension and gravity are the only forces acting on the toy plane (friction is being ignored), and the unlabeled net force is centripetal force.

Tension's vertical component equals gravity and its horizontal component is the centripetal force, so the ratio of the centripetal force to the magnitude of gravity is $\tan \theta$. 
Lab 8: Energy, the Vertical Loop

Tasks: A) Find energy of rubber band car-launcher at all four settings. B) Determine the minimum setting which will allow a specially-weighted car to make it through the loop without falling. ALTERNATE VERSION: If time allows, a ramp may be placed at the end of the track following the loop, and the students required to determine the setting required to get the car off the end of the ramp.

Equipment: For each group, 1 launcher, 1 regular car, 1 weighted car (at least 50g and loop-capable, see below), track sections, Hot Wheels™ loop, rulers, masking tape. There should also be a scale or balance for massing the cars, and if possible, corrugated cardboard sheets for use as landing pads in the event a group decides on a horizontal launch. ALTERNATE: Also include a ramp.

A “loop-capable” car is one which has its wheels set close enough to its bumpers that the wheels remain in contact with the track in a loop. A quick way to test whether a car is loop-capable is to hold a loop upside down and place the car inside it. If the car rolls freely, it’s loop-capable. If it skids on its bumpers, it should not be used in this lab as the weighted car. Cars which are not loop-capable can still be used for experimenting on the launcher energy.

Equipment Note: The four-setting launchers used in this lab are no longer manufactured by Mattel. This lab is included for completeness, and in case the four-setting launchers ever become available again.
Main Points: The value of approximation, specifically ignoring a factor in order to approximate that same factor (in this case, ignoring friction to get an estimate of normal force, and hence of friction). Students should also become more familiar with the use and usefulness of energy conservation.

Note: because this lab has a definite yes/no result, it can help with student motivation, since they will instantly get feedback.

Sample Designs: Only Part A really has a design, strictly speaking, as Part B is calculation based on a given design. At least four workable designs exist for determining the energy of the launcher, two of which are detailed on the next page. Those not detailed include launching the car horizontally and using the distance traveled and the height fallen to determine initial velocity, or launching the car at an angle and finding initial velocity as in Lab 4.

The main approximation of Part B is to assume no friction in order to find an average normal force. Given a needed velocity at the top of the loop (centripetal force = weight), normal force at the top is zero. Then, assuming no friction, energy conservation is used to find what velocity at the bottom of the loop would have yielded the minimum needed velocity at the top, and the normal force can be found from this (centripetal force + weight). Averaging the two and multiplying by the half-circumference and μ (which students should be able to easily find by now) will give an approximate energy loss due to friction in the loop.

Once students determine energy losses due to friction on the straightaway and on the loop as well as energy loss due to change in height, they can add those on to the minimum kinetic energy (or velocity) which will keep the car on the loop.
This will yield the minimum starting energy needed to get the car through the loop. Referring to their chart of launcher energies from Part A, students can determine what the minimum setting is to get the car through the loop.

**ALTERNATE:** Students will also have to determine if the energy left in the car after passing through the loop is sufficient to get the car up the ramp. More friction will need to be taken into account, as well as the gravitational potential energy of going up the ramp.

**Pitfalls:** On a purely equipment-related side, rubber bands break. Especially right after students have finished making all their calculations based on the old rubber band. So be sure to have spare rubber bands handy and be ready to help students quickly get new data if the old band broke at a bad time. Also, be sure the cars used in the loop are heavy enough that it requires at least the second-lowest energy setting to get them through the loop, or else it will be hard to check the results.

On a theory side, plan to spend a fair amount of time helping students understand why they can approximate the friction by ignoring friction. If they have had series calculus, you can use the concept of an expansion, but it’s unlikely they will have gotten this far in math by the time they take this course. One way to get them to accept the approximation is to point out the small μ values they have, and note that even a bad approximation won’t be all that far off overall. Note that it is very unlikely that the students will hit upon this approximation themselves. you will need to lead them to it by Socratic dialogue, or simply give them the approximation if their math skills aren’t up to it yet.

Once the students see how well this lab works, however, they should be more willing to accept approximations later on.
A: Find Energy of Launcher at all four settings:

**Method 1a:** Placing a car in the launcher, fire it nearly straight up. Best-guess measure the maximum height of the arc, use conservation of energy principle to determined that the energy of the launcher equals the potential energy, \( mgh \). Be sure to measure from where the car started prior to the launch, as this point will change with the setting. **Fastest method, but be careful.**

**Method 2a:** Launch car up a sloped track. From known coefficient of kinetic friction, determine force of friction, energy dissipated from friction will be \( Fd \), where \( d \) is the distance traveled along the track. \( Fd + mgh \) will equal the entire energy.

**Method 3a:** Aim the launcher at a measured angle and launch the car so that it lands at the same level as the end of the launcher. Using the ideas developed in Lab 4, determine the velocity from the angle and the distance traveled. The energy of each setting will equal \( 0.5mv^2 \).

**Method 4a:** Launch the car off the edge of the table so that it lands on something reasonably soft (like a notebook or jacket). Given the height of the table, \( h \), and the distance from the table the car lands, \( d \), the velocity at the top can be determined.

\[
t = \sqrt{\frac{2h}{g}}, \quad v = \frac{d}{t}
\]

From \( v \), kinetic energy can be found as in Method 3a.
Figure A.8: Setup for Lab 8, the Vertical Loop

h (or 2r)
B: Find minimum setting to get weighted car through loop:

Method: The first approximation is to ignore the friction inside the loop. Thus, the total energy needed to get through the loop is $\mu Nd$ (energy lost to friction) plus $mgh$ (energy to rise to top of loop).

However, there is significant friction inside the loop, and if the students aren't already rushed for time, they should find a reasonable way to approximate this. One way is to assume that the energy loss due to friction (which constantly changes through the loop) would be about the same as if the friction were a constant value through the loop. A useful value for this constant is the average of $\mu$ times the normal force at the top of the loop (zero, for minimum energy case) and the normal force at the bottom of the loop ($m(g + a_c)$, where $a_c = v^2/r$). Students will not know the velocity just before the car enters the loop, but by setting the centripetal force at the top of the loop equal to the car's weight, they can find the minimum velocity the car needs at the top of the loop. Then, approximating no friction, students can use energy conservation to determine how fast the car should be going at the bottom of the loop after adding in the extra $mgh$ energy.

Once the students have all the sources of energy change ($\mu Nd$, $mgh$ and $\mu N_{loop} \pi r$), they can add these to the energy needed at the top ($0.5mv_{minimum}^2$) to get the energy needed from the launcher, and hence the minimum setting needed.

ALTERNATE: In addition to making sure that the kinetic energy at the top of the loop is sufficient, students will also have to add up the friction along the entire track (two straight sections and the entire loop) and the $mgH$ of the ramp. Unless the coefficient of friction is very high, a setting that will get the car through the loop will also get it to the top of the ramp.
Checking results is easy. If the car makes it through the loop at the students’
predicted setting for the launcher, but not at one setting lower, they got it right.
Lab 9: Analysis - The E-Yo™

Tasks: Given an E-Yo™ toy, determine what’s wrong with the measurements given by the toy’s digital displays. The displays that can be tested are time, distance and (tangential) speed.

Equipment: For each group, 1 E-Yo™, plus the usual assortment of timers and rulers and so forth. Note: It is uncertain if Tiger Toys will continue to produce the E-Yo™. At the time of this writing, it seems the toy has been discontinued.

Main Points: Developing both physical intuition (i.e. no way is that yo-yo going 150 miles per hour) and analysis skills.

Sample Designs: This is a more open-ended lab and will not require designs in advance. At least, not complete designs. Students should present possible ways to test the claims of the digital readout, however. These include pulling the E-Yo™ along the floor, using a ruler to trigger the photogate, etc.

For distance, a ruler or card could be passed in front of the photogate a large number (100 or more) of times, and the distance in kilometers divided by the repetitions to give the effective diameter of the E-Yo™.

For speed, a ruler or card could be passed through the photogate at a regular beat (might help to have someone trained in music in the group) for a length of time, then find the period by dividing the time by the repetitions. Using the speed given by the E-Yo™ and the period the students have found can give an effective diameter that will be different from that found in distance.
Note that the sample designs on the next page are simply some methods the author discovered, and do not represent all possible ways to analyze the system.

**Pitfalls:** Being essentially open-ended, it’s possible students will spend the entire time messing around and find out nothing of value. Also, there’s no guarantee every E-Yo™ is set to the same numbers.

Also, be sure students clear the results before starting a new trial, especially with distance. Velocity gives the highest “point to point” velocity attained since the last time the velocity button was pushed, but distance just gives the cumulative distance “traveled” since last time it was cleared.
A: Determine what the time reading corresponds to:

Method: Start by clearing the time display. Throw the E-Yo™ so that it "sleeps" at the bottom of the string for some period of time, which you measure independently with a timer or a watch. Compare the two time readings. The time reading should be reasonably accurate.

B: Determine what the distance reading corresponds to:

Method 1: Pull the E-Yo™ along the floor by the string a known distance. Divide this distance by the real circumference of the E-Yo™ to get the number of cycles completed in that distance. Now, divide the distance displayed on the E-Yo™ by this number of cycles to get the effective circumference. This method has problems with slipping.

Method 2: Pass a card or ruler through the photogate $N$ times (a large number). Divide the displayed distance by $N$ to get the effective circumference. This should be the circumference you'd find if the position of the photogate was used for the radius.

C: Determine what the speed reading corresponds to:

Method 1: Pull the E-Yo™ along the floor a known distance while timing it. Compare this average speed to the displayed speed. However, since the display gives the highest instantaneous speed, not the average, this is a very inaccurate method.
Method 2: Let the E-Yo™ fall from your hand and unwind as it goes. The length of the string gives a change in height and thus a change in gravitational potential energy. Ignoring friction, all of this energy has to go into rotational kinetic energy. Calculate the speed from this, and it should be the highest instantaneous speed the E-Yo™ reaches. However, friction is not insignificant.

Method 3: Pass a card or ruler through the photogate at intervals as regular as you can manage. Count the number of "cycles" (the number of passes minus one) and measure the time this takes. Divide the time by the cycles to get a period $T$.

There’s two ways to go from here:

A) Assume the E-Yo™ is measuring the speed at the same radius as it was found to measure the distance. Divide the circumference corresponding to that radius by $T$ to get a speed. Compare to the displayed speed (which may be about 10 times larger).

B) Multiply the displayed speed by $T$ to get an effective circumference, compare this circumference to both the outer circumference of the E-Yo™ and to the circumference corresponding to the radius at which the E-Yo™ measures distance.
Grading Philosophy:

These labs were designed with a specific grading philosophy in mind, and some of the procedures assume this grading policy will be followed. If using a significantly different policy, the instructor is advised to carefully read the labs with an eye towards how this policy will affect them.

1) All grading is done in class: Rather than collect lab reports to grade and hand back, students are evaluated as they work. Groups are to check their designs with the instructor before proceeding to performing the experiment, and must check their experimental results with the instructor before either leaving or moving on to the next task. If the design or results are acceptable, the group is given credit and allowed to move on. If either is unacceptable and there is ample time remaining, the instructor should point out (or help the students discover) any flaws, then let them attempt another design or run another experiment. Provided the group eventually produces acceptable work, no credit is lost. However, if there is not enough time to try again, the students should be guided quickly to an acceptable design or result, and points (half of the points for that portion is suggested) taken off. If, for some reason, the students cannot generate workable solutions even with help, all points should be lost for that portion.

Note that just because nothing is collected doesn’t mean students don’t have to write down what they’re doing. Students should have a lab notebook in which they keep all their data and write down all their designs (especially since several designs come up more than once). If a group is not clearly keeping track of their work, the instructor can refuse to give credit for designs until they are written out properly.
2) **Group work means group grades:** A significant portion of a student’s grade in this lab should be based on the overall group performance. The group grade should be split evenly between design and experiment components, as evaluated in class.

3) **Some individual assessment is necessary:** If only to motivate those who would otherwise “ride the coattails” of their group members, some percentage of the class grade should be based on individual merits. Generally, a combination of lab quizzes and notebook checks will be given at the end of class to act as individual assessment tools, worth 20% of the lab’s grade.
APPENDIX B
THE ZONE TESTS

The "Zone" tests were administered in three different versions, each with its own set of items to be decomposed. This appendix will present each version in order, including instructions. Diagrams have been reformatted slightly to fit the page where necessary.

Tests 1 and 2 were given simultaneously, randomly distributed to students. Test 3 was given a quarter later, to all students in the class being studied.
Zone Test 1 - Instructions

This exercise and others you will do over the course of this quarter are part of research on how students work problems, especially more complex problems than those normally seen in the textbook. While it is asked that you put your name and Student ID Number on each sheet, this is only for purposes of tracking your progress, and all information about your personal “score” will remain private.

In this first exercise, you are presented with a number of problems drawn out in diagrams with a few notes attached. On each problem, please identify the different parts, or Zones.

To find the Zones, please first label (preferably with capital letters) every point in the diagram where something changes. Examples include: when something starts moving, where an impact happens, when something reaches the highest point on a hill before rolling back, etc. In the worked example following these instructions, the labels are:

A - Pendulum starts to drop
B - Pendulum hits cart
C - Pendulum reaches top of its swing
D - Cart rolls to a stop
Once the points are labeled, Zones can be described. A Zone is either a point or a region between points where something happens. The boundary between Zones is any place where what’s happening changes. Using the points from the sample problem, we have the following Zones:

A-B - Pendulum swings down (motion without friction)
B - Pendulum hits cart (impact)
B-C - Pendulum swings up (motion without friction)
B-D - Cart rolls along and finally stops (motion with friction)

In later exercises, you will use what you have learned about Physics to solve these problems for some unknown, such as how far the cart rolls or how slippery the surface it rolls on is. For now, simply label the points of interest and define your Zones.

If you don’t recognize a symbol on the diagram (such as m, the coefficient of friction), don’t worry about it. Those are only important in later exercises, and by the time you do those exercises you will have been introduced to those symbols.

Figure B.1 is the worked sample item. Figures B.2 through B.6 are the five test items.
1-0: Pendulum-Cart Bash

Solve for: $\mu$

A: PENDULUM RELEASED, STARTS FALLING
A-B: PENDULUM FALLS UNDER INFLUENCE OF GRAVITY
B: IMPACT BETWEEN PENDULUM AND CART
B-C: PENDULUM RISES TO NEW MAXIMUM HEIGHT, WILL KEEP SWINGING BACK AND FORTH TO THIS HEIGHT
B-D: CART, STARTED MOVING AT IMPACT, ROLLS ALONG UNTIL IT SLOWS TO A STOP DUE TO FRICTION

Figure B.1 - Worked sample problem for Zone Test 1.
1-1: Torsion Spring Launcher

Notes: All surfaces are frictionless. Spring is torsional (twisting). Both angles \( q \) are the same. Stopper block is at the equilibrium position of the torsion spring, and simply prevents "follow through" of the launcher arm. Launcher arm is massless.

Solve for: \( h \)

Figure B.2 - Zone Test 1, item 1.
1-2: Spring-Ramp-Passer

Notes: All surfaces are frictionless. Both springs have same spring constant $k$. Black triangles are fixed supports. The block never leaves the surface.

Solve for: $x_2$

Figure B.3 - Zone Test 1. item 2.
1-3: Pendulum Whacker

Notes: All surfaces are frictionless. Launched ball hits pendulum at the top of its arc (all velocity horizontal).

Solve for: $d$

Figure B.4 - Zone Test 1. item 3.
1-4: Clay Ball Drop

released just as other ball starts to fall

Notes: All surfaces are frictionless. The black triangle is a fixed support. The balls are made of clay and stick together on impact. Both springs have the same spring constant $k$.

Solve for: $x$ (initial compression of launcher), $y$ (final compression of catcher)

Figure B.5 - Zone Test 1. item 4.
1-5: Stunt APC

Notes: An Armored Personnel Carrier (APC) starts with mass $M+m$. Projectile fired at top of arc has mass $m$. The APC falls straight down after firing projectile. All surfaces are frictionless and the APC is coasting at start of problem. Ignore path of projectile after it has separated from the APC (it falls past the edge of the page).

Solve for: $v_2, h_2$

Figure B.6 - Zone Test 1, item 5.
Zone Test 2 - Instructions

This exercise and others you will do over the course of this quarter are part of research on how students work problems, especially more complex problems than those normally seen in the textbook. While it is asked that you put your name and Student ID Number on each sheet, this is only for purposes of tracking your progress, and all information about your personal “score” will remain private.

In this first exercise, you are presented with a number of problems drawn out in diagrams with a few notes attached. On each problem, please identify the different parts, or Zones.

To find the Zones, please first label (preferably with capital letters) every point in the diagram where something changes. Examples include: when something starts moving, where an impact happens, when something reaches the highest point on a hill before rolling back, etc. In the worked example following these instructions, the labels are:

A - Block starts to slide
B - Block reaches top and stops sliding for a moment

Once the points are labeled, Zones can be described. A Zone is either a point or a region between points where something happens. The boundary between Zones is any place where what’s happening changes. Using the points from the sample problem, we have the following Zones:

A-B - Block slides (motion without friction)
B - Block stops momentarily
In later exercises, you will use what you have learned about Physics to solve these problems for some unknown, such as how far the cart rolls or how slippery the surface it rolls on is. For now, simply label the points of interest and define your Zones.

If you don’t recognize a symbol on the diagram (such as m, the coefficient of friction), don’t worry about it. Those are only important in later exercises, and by the time you do those exercises you will have been introduced to those symbols.

Figure B.7 is the worked sample item. Figures B.8 through B.12 are the five test items.
2-0: Motion Up Slope

Note: Block is moving at start of problem, problem ends when the block momentarily stops at the top of the slope.

Solve for: \( h \)

**A-B: BLOCK MOVES UP SLOPE**  
**B: BLOCK STOPS MOMENTARILY**

Figure B.7 - Worked sample problem for Zone Test 2.
2-1: Slope-Slide-Skid

Notes: Surface is frictionless until the shaded region. Ramp is a circular arc.

Solve for: $x_2$
2-2: Pendulum-Ramp Splat

Notes: Block and clay glob each have mass $m$. Ramp is frictionless. Block and glob remain firmly stuck together after impact. Glob is initially suspended straight down.

Solve for: $h_2$

Figure B.9 - Zone Test 2. item 2.
2-3: Torsion Spring Catcher

Notes: \( \kappa \) is the torsion spring constant. All of the surfaces are frictionless. The ramp is fixed to the tabletop. You need only label Zones on Top View (upper picture).

Solve for: \( \theta \)

Figure B.10 - Zone Test 2, item 3.
Notes: All surfaces are frictionless. The black triangle is a fixed support. The collision between \( M \) and \( m \) is totally inelastic, but the blocks separate without sticking when the spring extends. \( M > m \).

Solve for: \( m \)
2-5: Clay Block Bounce

Notes: Spring constant $k$ and spring compression $x_i$ are the same for both blocks. The surface is frictionless. $x_f$ is the maximum compression of the spring when the combined mass hits it. Everything but the blocks is massless. Black triangles are fixed supports.

Solve for: $x_f$, and which spring will be hit first by the combined blocks.

Figure B.12 - Zone Test 2, item 5.
Zone Test 3 - Instructions

This exercise and others you will do over the course of this quarter are part of research on how students work problems, especially more complex problems than those normally seen in the textbook. While it is asked that you put your name and Student ID Number on each sheet, this is only for purposes of tracking your progress, and all information about your personal “score” will remain private.

In this first exercise, you are presented with a number of problems drawn out in diagrams with a few notes attached. On each problem, please identify the different parts, or Zones.

To find the Zones, please first label (preferably with capital letters) every point in the diagram where something changes. Examples include: when something starts moving, where an impact happens, when something reaches the highest point on a hill before rolling back, etc. In the worked example in Figure B.13, the labels are:

A - Pendulum starts to drop
B - Pendulum hits box
C - Pendulum reaches top of its swing
D - Box slides to a stop
Once the points are labeled, Zones can be described. A Zone is either a point or a region between points where something happens. The boundary between Zones is any place where what’s happening changes. Using the points from the sample problem, we have the following Zones:

A-B - Pendulum swings down (motion without friction)
B - Pendulum hits box (impact)
B-C - Pendulum swings up (motion without friction)
B-D - Box slides along and finally stops (motion with friction)

In later exercises, you will use what you have learned about Physics to solve these problems for some unknown, such as how far the cart rolls or how slippery the surface it rolls on is. For now, simply label the points of interest and define your Zones.

**Remember, do not attempt to solve the problems. Simply label the points of interest and describe the Zones.**

Figures B.13 and B.14 show the worked example. Figures B.15 through B.18 show the four items of this test.
**3-0: Pendulum Box Bash**

![Diagram of pendulum box bash](image)

**Description:** Lead pendulum is released and falls to hit the box, which slides for a while until stopped by friction. Pendulum keeps going after it has hit the cardboard box.

**A-B:** PENDULUM FALLS UNDER INFLUENCE OF GRAVITY  
**B:** IMPACT BETWEEN PENDULUM AND BOX  
**B-C:** PENDULUM RISES TO NEW MAXIMUM HEIGHT, WILL KEEP SWINGING BACK AND FORTH TO THIS HEIGHT  
**B-D:** BOX, STARTED MOVING AT IMPACT, SLIDES ALONG UNTIL IT SLOWS TO A STOP DUE TO FRICTION

As you can see, there are four important points in this problem, places where things happen or change significantly. Starting, stopping, hitting, reaching new heights, etc. Likewise, there are four significant Zones in which things happen. Moving in arcs, moving in lines, collisions and so forth.

Figure B.13 - Zone Test 3 Worked Example. Continued on next page.
KEY TO TEST DIAGRAMS: Some common pieces of problems.

Spring: Problem will say if the spring is currently stretched or compressed, and by how much. Springs are not attached to moving blocks.

Fixed Block: Usually attached to a spring, this block is firmly attached to the surface and will not move.

Moving Block: Black at initial position, grey at later positions. Usually a square, but not always.

Path Arrow: These dotted or dashed arrows will appear in the problems and give a basic idea of the path things take and the direction of motion.
**3-1: Spring Ramp Pusher**

**Description:** Spring pushes on block, which slides up ramp and hits a sticky surface, where it slows to a stop.

---

Figure B.15 - Zone Test 3, item 1.
3-2: Slide Skid Stopper

Description: Block slides down ramp, eventually stops on the rough part of the surface.

Figure B.16 - Zone Test 3, item 2.
3-3: Half-Pipe Slammer

Description: Block 1 slides down and hits block 2, both stick together and slide up the other side.

Figure B.17 - Zone Test 3, item 3.
3-4: Bed Of Nails

**Description:** Spring pushes on block, which slides up ramp, flies off through the air, and lands on a bed of nails which stops it where it lands.
APPENDIX C

THE PROBLEM DECOMPOSITION DIAGNOSTIC VERSION 0.5

Instructions:

In this test you will read descriptions of several situations which could be analyzed with basic physical principles. These situations will be multi-part problems which do not lend themselves to solution in a single step. **You will not solve these problems:** instead [insert whichever version is settled on]

VERSION 1)

...after each description you will be presented with four possible pieces of the problem. If you think a piece could reasonably be worked in a single step, mark A on the answer sheet. If you think the piece is too big to manage in one step, mark B on the answer sheet. Finally, if you think a piece is trivial and requires no work at all, mark C on the answer sheet.

[perhaps reproduce the following bit on each page:

A) Single step
B) More than a single step
C) Trivial]
VERSION 2)

...after each description you will be presented with four possible pieces of the
problem. If you think a piece could reasonably be worked in a single step, mark A on the
answer sheet. If you think a piece is either too complex to work in a single step, or is
trivial and need not be worked out at all, mark B on the answer sheet.

1 - 4)

You’re watching a movie in which an unpleasant main character’s life is changed
after an automobile accident, and the scene with the accident is coming up. The “hero” is
driving down a straight stretch of country road while talking on his cellphone, and doesn’t
see the stop sign ahead. A driver on the cross road sees him and hits the brakes, but the
two cars collide anyway. The twisted mass skids off the road and up an embankment
where it comes to rest a short ways up the slope. Assuming this stunt was done all in one
take, it’s possible to analyze the situation using physics, once it’s broken down into
manageable parts.

Possible pieces:

1) Motion of the main character’s car from the beginning of the stunt until just before the
impact.

2) Motion of the other vehicle from the beginning of the stunt until just before the impact.

3) The crash as both cars hit and lock together.

4) Motion of the combined mass of cars from just after the impact until they come to a
stop.
A kid is practicing with his slingshot in the back yard of his parents’ house, shooting rocks at some empty soup cans lined up near the back fence. He puts a rock in the pocket of the slingshot, pulls back and lets go, lobbing the rock into the air in a lazy arc. On a lucky shot, one rock goes into a can and the two slide along the grass until they hit the fence, where they stop dead. This can be broken into several parts and analyzed physically.

Possible parts include:

5) From just after letting go of the slingshot to when the rock is at the high point in its flight.
6) Motion of the rock from the high point until just before it hits the can on the ground.
7) Motion of the rock and can from just after the rock lands in the can, until just before they run into the fence.
8) The collision of the can with the fence.
A new rollercoaster called "The Steam Demon" has been proposed. Instead of the standard method of cranking the coaster up a tall hill to start it going, it uses a steam catapult like the kind used on aircraft carriers to launch the coaster out of the gate at a high speed. The coaster then goes straight for a bit and flips sideways for a horizontal half-loop in which the riders' heads are on the outside, what they call a "negative gee" experience. From there it's into a traditional vertical loop and up a hill to a plateau where the riders can catch their breath for a second before a dramatic plunge down into a tunnel below the ground and finally back up to the surface, where the track turns around 180° on a banked curve before returning to the station.

Here's some pieces the roller coaster might be broken up into for study:

9) Motion while the steam catapult is pushing the coaster.
10) Motion from the point where the catapult cuts out until the start of the half-loop.
11) Motion from the beginning of the horizontal half-loop until the end of the vertical loop.
12) Motion on the banked curve near the end.
You’re enjoying a day at the park with friends, and are currently sitting in an old tire swing (a tire hung from a rope or chain) at rest. Someone tosses a softball to you, but catching it causes you to start swinging. You try to throw the softball back while at the bottom of the swing’s arc, but your aim is off and the softball hits a soccer ball that rolls to a stop nearby.

Here’s some pieces the situation could be broken into if you wanted to figure out where the softball finally landed:

13) Catching the softball while in the swing.
14) Swinging back after catching the softball.
15) Throwing the softball.
16) Motion of the soccer ball before it gets hit.
You and some friends are out playing miniature golf, and have reached the inevitable “windmill hole,” where the blades of a small motor-driven windmill are one of the obstacles you need to get your ball past. Just then, someone yells, “FORE!” and a ball flies through the air, bouncing off one of the windmill blades and making the windmill turn in the opposite direction for a few seconds before the friction of the drive belt stops it. The ball bounces almost straight up into the air before falling onto the green and rolling into the cup. While watching the person who hit the ball be escorted out of the course, you wonder just how hard he hit the ball.

If you were to actually try and calculate this, there’s several parts you might break the problem into. Here are some pieces which might be single step parts:

17) Motion of the golf club (a putter being swung like a driver).
18) Rotation of the windmill before it gets hit.
19) Rotation of the windmill after it gets hit.
20) Motion of the ball from after the impact until it reaches the top of its arc.
Splatman, the Ductile Detective, is getting a tan at the entrance to his above-ground Splatcave when he sees a helicopter approaching in the distance. A quick look through his telescope reveals it to be a robot drone carrying a bomb at top speed towards his headquarters! Rushing to his spring-loaded Splatapault, Splatman fires a gooey glob at the drone-copter. The glob flies through the air and, at the top of its arc, strikes the drone. The glob wraps around the helicopter blades and the gooey drone-copter falls into the lake outside the Splatcave. Well, hopefully. Assuming you knew all the appropriate information (masses, speeds, etc.) it would be possible to calculate where the helicopter lands.

Here's some parts the problem could be broken into:

21) From when Splatman triggers the Splatapault until just before the glob hits the drone-copter.

22) The motion of the drone-copter during the time that the glob is in the air.

23) The glob hitting the drone-copter.

24) The motion of the globbed-up copter as it falls out of the sky.
Alternate Section: Could replace one of the six large problems.

This section contains several short situations. Each question will ask for a comparison between two of them. Answer:

A) If the first situation breaks into more parts than the second.

B) If both situations break into the same number of parts.

C) If the first situation breaks into less parts than the second.

Note: assume there is friction.

**Situation 1:** A box is pushed across the floor by a student, speeding up, then released. It eventually skids to a stop.

**Situation 2:** A box is pushed across the floor by one student, speeding up, then released. A second student catches the box and slows it to a stop.

**Situation 3:** A spring pushes a box up a slope. The box is released and slides upwards until it hits another spring, which stops it.

**Situation 4:** A spring pushes a box up a slope. The box is released and slides upwards until it stops.

21) Situation 1 versus Situation 2
22) Situation 1 versus Situation 3
23) Situation 1 versus Situation 4
24) Situation 2 versus Situation 3
25) Situation 2 versus Situation 4
26) Situation 3 versus Situation 4
APPENDIX D

ACTIVPHYSICS INTERVIEW RESULTS

This section includes the analyses of the individual student interviews performed on April 22 and 23 of 1998. The names have been removed to protect the subjects' privacy. Under "131 Lab Type," TIM means the student was in a Toys In Motion laboratory section, while CACP means the student was in a Constructing and Applying the Concepts of Physics laboratory section.
Subject: S1-1 (male)

Date: April 22, 1998
Time: 1:30-2:30 PM
Interviewer: Dave Van Domelen
131 Lab Type: TIM

Task 1: Skier Into Cart

S1-1’s initial impression was that the problem could be divided into three parts, but when he examined the simulation more carefully, he decided on two parts. Unknown if the abandoned third part was the collision, but he didn’t include that part in his final analysis. Described everything in terms of “force,” even when he was talking about properties which sounded more like momentum or energy.

Task 2: Pendulum Box Bash

S1-1 continued to call many different things “forces.” This time he realized that the collision was a separate part of the problem, possibly because the appearance of the velocity arrow on the box is a visual cue that something happens in that frame. S1-1 was labeling the parts not in terms of the surface features or the deep concepts, but rather in terms of the quantity he would need to calculate.
Task 3: Pendulum Person Bowling

S1-1 decided that the pendulum motion was a centripetal force problem, continuing his "force is all" idea that objects "have" a force rather than exerting it. He saw that the collision between pendulum and man was important right away, possibly because of similarity to the previous task, but S1-1 almost missed that there was a collision between the man and car (he caught it at the end). By this point he had settled in to very consistently labeling the "type" of problem by the quantity he thought he needed to find, especially the "velocity problems" of ballistic motion and frictionless motion.

Overall:

Observing this subject strongly suggested that students might not actually categorize problems by surface features when solving them, but instead categorize by goals.

Also, S1-1 consistently seemed to associate the term "force" with any physical concept other than velocity.
Subject: S1-2 (male)

Date: April 22, 1998

Time: 2:30-3:30 PM

Interviewer: Dave Van Domelen

131 Lab Type: Control

Task 1: Skier Into Cart

S1-2 considered finding the givens to be the first part of the problem and assigned it the frame number of zero. He thought the initial velocity of the skier might be zero, since the figure doesn’t move before one starts the simulation. He initially missed the collision entirely, then later went back and added it as a change in acceleration at the same time as another part. S1-2 had a lot of trouble with the idea that a part took up more than a single instant in time, and once told that he had to assign all the frames to a part, generally stopped considering any part to have only a single frame. He saw force as proportional only to change in velocity, not change in time...seemed to confound force and momentum.

Task 2: Pendulum Box Bash

Continued problems with “givens are a part” and the continuum/point nature of the parts in time were exhibited. S1-2 talks about a “hit” occurring, but by this time was firmly convinced that no single frame was a valid part, so he listed no collision. S1-2 labeled parts by the quantity to be found. S1-2 resisted either talking into the tape recorder or writing any sort of explanation in the “Explain” column for this task.

234
**Task 3: Pendulum Person Bowling**

S1-2 tried to find "an equation" for the problem before he started to break it up. He continued his position of "no single frames" from above. Confused by the simulation’s use of a short ledge and thought there was friction (despite the "Friction is zero" note on the worksheet) adding a part for motion on the ledge. Labeled momenta as forces. S1-2 did not consider the frames after the collision between the person and the cart to be a part, possibly because they occur after the point where it's possible to solve the problem, despite having been told repeatedly that all frames must be assigned to some part.

**Overall:**

Similar patterns to S1-1.

It was difficult to get to S1-2 to think-aloud, he had his own way of doing things and reverted to them very quickly after being reminded of the instructions (stayed silent while moving his lips, kept trying to label parts as either only points or only ranges, etc).

His difficulties with frames suggested a paper-and-pencil test where the time labels go not on the hash marks, but between them.
Subject: S1-3 (male)

Date: April 23, 1998

Time: 1:30-2:30 PM

Interviewer: Dave Van Domelen

Lab Type: Experimental

Task 1: Skier Into Cart

S1-3 was another case of a student defining problem type by the quantity to be solved for. He immediately applied work-energy concepts to this problem, but tacitly assumed that kinetic energy is always conserved during collisions, and didn’t consider the collision to be a separate part. S1-3 ignored the mass of the cart.

Task 2: Pendulum Box Bash

S1-3 explicitly included the collision this time, again probably due to the visual cue of the velocity arrow appearing. However, for some reason he decided the collision happened a frame later than shown in the animation. S1-3 demonstrated fairly strong energy conservation skills, but continued to omit momentum conservation or even consider that kinetic energy might not be conserved in a collision.
Task 3: Pendulum Person Bowling

S1-3 went back to omitting the collision for the pendulum/person system, reinforcing the idea that it was the arrow which cued him in Task 2. Initially tried to use energy conservation in very inappropriate situations (such as the ballistic motion phase). This seemed to be a demonstration of the "If your only tool is a hammer, every problem looks like a nail" principle.

When encouraged to write more in the "Explain" column, he started to take more time and notice holes in his method. He held several contradictory convictions, including that all collisions conserve kinetic energy and that he needed to break the velocity of the falling man into $x$ and $y$ components and use only the $x$ component. After several minutes of trying to determine where the extra energy went, he convinced himself that the kinetic energy from the $y$ direction velocity was absorbed into the car, while the $x$ component kinetic energy was still conserved. He finally decided that this meant the problem was unsolvable, since he didn’t think he could know for sure how much energy actually went into the car.

S1-3 implicitly included both collisions as elements of adjacent parts, "stretching" the first and last parts to cover the two collisions. He considered all parts to be energy conservation problems.

Overall:

It's possible that the emphasis on energy conservation in the later TIM labs rubbed off on S1-3 too well.
**Subject:** S1-4 (female)

Date: April 23, 1998

Time: 2:30-3:30 PM

Interviewer: Dave Van Domelen

131 Lab Type: Control

**Task 1: Skier Into Cart**

S1-4 initially broke the problem into three parts (after an abortive four-part solution caused by thinking the slow motion in the last few frames was no motion): before, during and after the collision. She later changed her mind and decided that all frames where the skier and cart were in contact were “during” the collision. Confounded force and velocity, shared S1-1’s view that everything was a force. S1-4 seemed confused at times by the fact that the time slider moved to the right but the action of the animation moved to the left. She couldn’t remember “the spring constant equation,” but felt it was very important, and seemed to imply there was only one such equation.

**Task 2: Pendulum Box Bash**

S1-4 fell into the “Finding initial conditions is a part” viewpoint here, although later admitted that this first part might actually be part of what she called the second part. She labeled all parts based on what quantity was to be found. S1-4 recognized that the collision was a separate part in this task and held strongly to this position (possibly because the items do not remain in contact this time), but didn’t seem to know how to approach it.
labeling that part as “All conditions at time of hit.” She initially considered the pendulum’s downward swing to be a “velocity problem,” since she needed to find the velocity, but later amended that to being an “acceleration problem.”

**Task 3: Pendulum Person Bowling**

S1-4 seemed to be mixing up acceleration and momentum. She thought in terms of equations that she needed to solve. S1-4 worked backwards from the end, and started to approach conservation of momentum, calling everything from the collision to the end of the animation a “center of mass problem.” Like S1-3, she included the two collisions in the parts adjacent to them, no longer recognizing them as separate problems.

**Overall:**

S1-4 followed the pattern of defining parts based on the goals, and of calling many different things forces. She admitted to some difficulty due to having taken 131 two quarters previously, rather than in the immediately preceding quarter. S1-4 had a definite preference for working backwards from the desired answer to the initial conditions.
APPENDIX E

CARD-SORTING TASK

Two card sorting interviews were carried out using this set of problems. In each exercise, the students were given the following instructions:

“In front of you are 12 pieces of paper with a physics problem on each piece of paper. You are not to actually solve these problems. Instead, what we would like you to do is to sort the problems into groups, based on what kind of problems you think they are. In other words, consider the way you would solve the problems and group together the problems that you think would be solved in similar ways.

“Once you’ve sorted the cards into groups, put your name on each card and give each group a number, putting that number on each card in the group. This is mainly insurance for us, in case the cards get accidentally shuffled.

“Finally, fill out the table below. Write the group number, the problems you placed in that group, and most importantly, why those problems form a group. In other words, what is similar about how you’d solve those problems? If you run out of space on the front of this page, please continue on the back of the sheet.”

The cards themselves follow. The cards are ordered here by surface features, but the numbers were randomly assigned before the cards were printed out. The cards were presented to the students in numbered order.
An object of mass $m$ starts at an unknown point and slides without friction (it doesn't roll) down a circular ramp of radius $R$ which is fixed to the table. When the object gets to the point on the ramp corresponding to angle $\theta$, its velocity ($v$) is known.

Find the normal force, $N$, on the object when it is at the point corresponding to angle $\theta$. 

Figure E.1 - Card Sorting Task: Card 11
An object of mass \( m \) starts at an unknown point and slides without friction (it doesn’t roll) down a circular ramp of radius \( R \) which is fixed to the table. When the object gets to the point on the ramp corresponding to angle \( \theta \), the normal force \( N \) is known.

Find the object’s velocity, \( v \), at the point corresponding to angle \( \theta \).
A block (drawn as a circle above, but it could be any shape) initially at rest slides without friction (it doesn't roll) down a circular ramp of radius $R$, which is free to move without friction. By the time the block reaches the bottom, both it and the ramp are moving. The masses of both movable objects ($m, M$), the starting position of the block (at $\theta$) and its final velocity ($v_1$) are known.

Find the final velocity of the ramp, $v_2$. 

Figure E.3 - Card Sorting Task: Card 10
A block (drawn as a circle above, but it could be any shape) initially at rest slides without friction (it doesn’t roll) down a circular ramp of radius $R$, which is free to move without friction. By the time the block reaches the bottom, both it and the ramp are moving. The masses of both objects ($m$, $M$), the starting position of the block (at $\theta$), its final velocity ($v_1$) and the time ($t$) that the block is on the ramp are known.

Find the average horizontal force ($F_{x, \text{avg}}$) exerted on the ramp by the block.
An object initially at rest slides without friction (it doesn’t roll) down a fixed circular ramp of radius $R$. The mass ($m$) of the object and its initial position (corresponding to angle $\theta$) on the ramp are known.

Find the velocity ($v$) of the object once it has left the ramp and is sliding without friction on the table.
An object initially at rest slides without friction (it doesn’t roll) down a fixed circular ramp of radius $R$. The mass ($m$) of the object, its initial position (corresponding to angle $\theta$) on the ramp and its final velocity ($v$) are known. The acceleration of gravity ($g$) is not known. (The problem takes place in a computer simulation where gravity can be changed.)

Find the Weight ($W$) of the object.
A golf club (which can be treated as a mass \( M \) on the end of a massless rod) hits a golf ball of mass \( m \). The ball travels a known distance \( d \) and reaches a known height \( h \). Assume that there is no air resistance (air "friction").

Find the velocity \( v \) of the golf ball immediately after it is hit by the golf club.
A golf club (which can be treated as a mass $M$ on the end of a massless rod) hits a golf ball of mass $m$. You know the velocity ($v$) of the golf ball immediately after it is hit by the club. If you were to assume that there is no air resistance (or air "friction") you could calculate that the ball would travel a distance $D$ and rise to a height $H$. But the ball actually travels a distance $d$ (less than $D$) and rises to a height $h$ (less than $H$), so there must be air resistance.

Find the average force ($F_{x,\text{avg}}$) of air resistance in the horizontal direction.
A golf club (which can be treated as a mass \( M \) on the end of a massless rod) hits an initially stationary golf ball of mass \( m \). You know the velocity of the ball immediately after it is hit by the club (\( v_{b,f} \)) and you also know the velocity of the club head right before it hits the ball (\( v_{c,i} \)).

Find the velocity of the club immediately after it hits the ball (\( v_{c,f} \)).
A golf club (which can be treated as a mass $M$ on the end of a massless rod) hits an initially stationary golf ball of mass $m$. You know the velocity of the ball right after it stops touching the club head ($v_{b,f}$). You also know the amount of time ($t$) the ball and club were in contact, thanks to video analysis. The time of contact is small enough that you can make the approximation that the ball and club head travel in a straight line during the time that they’re touching.

Find the average force ($F_{avg}$) exerted on the ball by the club during the time of contact.
Problem Number: 5
Group Number:
Name:

\[ M , m , v_{b,f} \]

A golf club (which can be treated as a mass \( M \) on the end of a massless rod) hits a golf ball of mass \( m \). You know the velocity of the ball right after the club hits it (\( v_{b,f} \)) and the distance (\( x \)) over which the club and ball are in contact. The distance (\( x \)) over which the club head and ball are in contact is small enough that you can make the approximation that they travel in a straight line while touching. (The distance (\( x \)) is too small to show on the diagram.)

Find the average force (\( F_{\text{avg}} \)) exerted on the ball by the club.
A golf club (which can be treated as a mass \( M \) on the end of a massless rod) hits a golf ball of weight \( W \). The ball travels an unknown distance \( d \) and reaches a known height \( h \). You know the speed \( v_0 \) of the ball immediately after it is struck by the club. **You do not know the acceleration of gravity (g),** but you know there’s no air resistance (air “friction”). (The problem takes place in a computer simulation where gravity and air resistance can be changed.)

Find the velocity of the ball at the highest point in its arc \( v_{\text{top}} \).
This version of the Problem Decomposition Diagnostic was created to explore the possibilities of the format, and also to help expand the list of potential classifications for subproblems. In addition to the written instructions, students were encouraged strongly to add their own multiple choice options on the test sheet.

Instructions

This test consists of several multiple-part physics problems, which you will not be solving. Rather, you will be determining how to break them up into pieces which can be solved more easily.

Each problem has a timeline below it, breaking the events into an arbitrary number of "slices." Each slice represents some period in time during the events of the problem. It is possible for several distinct events to take place in the same slice.

Your first task will be to break the problem into pieces, and determine which slices each piece covers. For example, if you decide that part 1 happens in slices 1, 2 and 3, and part 2 happens in slices 3 and 4, you would write "1" on the lines for slices 1-3, and "2" on the lines for slices 3-4. A part might only happen during a single slice, or it might be stretched out over several slices. Problem 1 will provide an example.
You only need consider events which occur during the time slices shown. Aspects of solving the problem which happen outside of the time slices, such as reading the problem or checking your answer, don’t count for purposes of this exercise.

Once you have decided what slices each part covers, look to the next page and decide what sort of problem each part is, putting the letter for that type on the worksheet. If you think that more than one answer fits, put all of your choices in the space provided, but circle the answer you think best describes the part.

For example, when working a hypothetical problem, you may think that Part 1 is both a “Constant Acceleration Kinematics” problem (choice c) and an “Object on an Inclined Plane” problem (choice p), so put both c and p on the answer sheet. Then circle one of the two as best describing the way you’d solve the sub-problem.

Multiple Choices

These are the types of problems you have to choose from in the second part of each task. You can pick as many as you think describe the subproblem, but be sure to circle the one choice you think best describes how you’d solve the part.

a) Conservation of Momentum
b) Conservation of Energy
c) Constant Acceleration Kinematics \((x-x_0 = v_0t + (1/2)at^2, \text{ etc})\)
d) Find a Position or Distance horizontally \((x)\)
e) Find a Velocity \((v)\)
f) Find an Acceleration \((a)\)
g) Find a Force of Friction ($F_k$)

h) Find a Normal Force (N)

i) Find a Weight (W)

j) Find a Force (other than Friction, Normal or Weight)

k) Newton’s First Law (velocity constant unless acted on by unbalanced force)

l) Newton’s Second Law (net force proportional to acceleration)

m) Newton’s Third Law (action/reaction)

n) Object is a Pendulum

o) Object on a Spring

p) Object on an Inclined Plane (Ramp)

q) Object Moves on a Curved Path

r) Object Moves on a Straight Path

s) Object Moves with Projectile Motion

t) Object Falls

u) Object Moves with Friction

v) Object Moves without Friction

The following reminder was in boldface at the bottom of the first page (figure F.1, omitted from that figure for reasons of space).

**Remember: you can pick more than one type of problem from the list. Be sure to circle the choice you think best represents how you’d solve the part.**

Figures F.1 through F.4 are the four items of the test.
Problem 1: Sliding Box

This problem is broken into parts for you, to help show what you’re supposed to do in the later problems. It has two parts, one covering intervals 1-4, the other covering intervals 4-6. The parts overlap, which is allowed. Also, nothing happens in interval 7, which is also okay.

For the rest of this page, determine what kinds of problems parts 1 and 2 are.

A box of mass $m$ starts at the top of a frictionless ramp, a height $h+y$ above the table. It slides down the ramp until it reaches a height $h$ above the table, at which point it drops off the end of the ramp and lands on the table a distance $d$ from the end of the ramp, stopping immediately.

The goal of this problem would be to find $d$, the position of the block when it lands.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
<th>Part</th>
<th>Type of Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1, 2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure F.1 - Problem Decomposition Diagnostic v1.0 item 1
Problem 2: Pendulum-Box-Bash

The pendulum (mass $M$ at the end of a string of length $L$) is released from rest and swings down to hit the box (mass $m$). The pendulum stops after it hits the box, and the box slides a distance $d$ along a surface with friction until it stops.

The goal of this problem would be to determine the coefficient of sliding friction $\mu$ between the box and the surface.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part</th>
<th>Type of Problem</th>
</tr>
</thead>
</table>

Figure F.2 - Problem Decomposition Diagnostic v1.0, item 2
Problem 3: Spring Launcher

A ball rests on a compressed spring attached to a frictionless ramp. When released, the ball should fly through the air and land in the tray at the right side of the page. When the spring is relaxed, it reaches the top of the ramp.

The goal of this problem would be to find the amount $x$ that the spring must be compressed in order for the ball to land in the tray. You would know or be able to measure $k$, $m$, $h$, $\theta$, and the distance $d$. The ball is dense enough that air friction can be ignored.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part</th>
<th>Type of Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure F.3 - Problem Decomposition Diagnostic v1.0, item 3
Problem 4: Block Catcher

A block of mass \( M \) has an initial velocity \( v_0 \). It slides a distance \( d \) along a surface which has friction. The coefficient of friction \( \mu \) is known. The first block then sticks to a smaller block \( m \) which is attached to a spring with spring constant \( k \). The second block has the same coefficient of friction with the surface, and the spring is initially at relaxed length. The black triangle at the far right is attached to the surface.

The goal of this problem would be to determine the speed \( v_0 \) necessary for the spring to be compressed by \( x \) before the blocks start moving back to the left.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure F.4 - Problem Decomposition Diagnostic v1.0, item 4
APPENDIX G

THE PROBLEM DECOMPOSITION DIAGNOSTIC VERSION 1.1

This version of the Problem Decomposition Diagnostic resulted from the work done examining version 1.0. It has an additional item, plus an extended list of multiple choice options.

Instructions

This test consists of several multiple-part physics problems, which you will not be solving. Rather, you will be determining how to break them up into smaller problems that can be solved more easily.

Each problem has a timeline below it, breaking the problem into a number of intervals of arbitrary length. Each interval represents some period in time during the course of the problem; the intervals are not necessarily the same amount of time. It is possible for several subproblems to occupy part of the same interval, such as when one subproblem ends and another begins.
Your first task will be to break the problem into pieces, and determine which intervals each subproblem covers. For example, if you decide that part A happens in intervals 1, 2 and 3, and part B happens in intervals 3 and 4, you would write "A" on the lines for intervals 1-3, and "B" on the lines for intervals 3-4. A part might only happen during a single interval, or it might be stretched out over several intervals. Problem 1 will provide an example.

You only need consider events that occur during the time intervals shown. Aspects of solving the problem that happen outside of the time intervals, such as reading the problem or checking your answer, don’t count for purposes of this exercise.

The second task you will perform on each problem involves classifying the subproblems by choosing from a menu of possibilities. The next page has a list of choices, each of which is a possible way to solve a subproblem. If you think that more than one of these choices is appropriate, list all your choices on the answer sheet, but circle the answer you think best describes the way you’d solve the subproblem.

For example, when working a hypothetical problem, you may think that in Part A you would both “Apply constant acceleration kinematics” (choice c) and “Solve as an inclined plane problem” (choice t), so you would put both c and t on the answer sheet. Then you’d circle one of the two as best describing the way you’d solve the sub-problem.
Problem Types

You will choose from the list below (Table G.1) for the second part of each problem. You may pick as many items as you think describe each subproblem, but be sure to circle the one choice you think best describes how you'd solve the subproblem.

PLEASE ANSWER THE FOLLOWING QUESTION:

Did the length of the multiple choice list cause you any difficulty in working the problems on this test? Please circle one of the numbers below, where 1 means “The list didn’t cause me any trouble” and 5 means “The list caused me a great deal of trouble.”

1 2 3 4 5

In addition to these instructions, the warning below also appeared at the bottom of item 1, as it did on version 1.0.

Remember: you can pick more than one type of problem from the list. Be sure to circle the choice you think best represents how you’d solve the part.

Figures G.1 through G.5 are the five items on this version of the PDD.
a) Apply conservation of momentum
b) Apply conservation of energy
c) Apply constant acceleration kinematics \(x-x_0 = v_{0x}t + \frac{1}{2}a_xt^2,\)
\(y-y_0=v_{0y}t + \frac{1}{2}a_yt^2,\) etc

d) Find a horizontal position or distance \((x\ or\ d)\)
e) Find a vertical position or distance \((y\ or\ h)\)
f) Find a velocity \((v)\)
g) Find an acceleration \((a)\)
h) Find the force of friction \((F_k)\)
i) Find the normal force \((N)\)
j) Find the weight \((W)\)
k) Find a force other than friction, normal or weight
l) Find the kinetic energy
m) Find the gravitational potential energy
n) Find the spring potential energy

o) Apply Newton’s First Law (velocity is constant unless the object is acted on by a net force)
p) Apply Newton’s Second Law (net force is proportional to acceleration)
q) Apply Newton’s Third Law (the force of A on B is equal and opposite to the force of B on A)
r) Solve as a pendulum problem
s) Solve as a spring problem
t) Solve as an inclined plane (ramp) problem
u) Solve as a curved path problem
v) Solve as a straight path problem
w) Solve as a projectile problem
x) Solve as a falling object problem
y) Solve as a zero-friction problem
z) Solve as a problem with non-zero friction

Table G.1 - Multiple choice options for the Problem Decomposition Diagnostic v1.1

263
Problem 1: Sliding Box

This problem is broken into parts for you, to help show what you’re supposed to do in the later problems. It has two parts, one covering intervals 1-4, the other covering intervals 4-6. The parts overlap, which is allowed. Also, nothing happens in interval 7, which is also okay.

For the rest of this page, determine what kinds of problems parts 1 and 2 are.

A box of mass $m$ starts at the top of a frictionless ramp, a height $h+y$ above the table. It slides down the ramp until it reaches a height $h$ above the table, at which point it drops off the end of the ramp and lands on the table a distance $d$ from the end of the ramp, stopping immediately.

The goal of this problem would be to find $d$, the position of the block when it lands.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
</tr>
<tr>
<td>4</td>
<td>A, B</td>
</tr>
<tr>
<td>5</td>
<td>B</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Part</th>
<th>Type of Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
</tr>
</tbody>
</table>

Figure G.1 - Problem Decomposition Diagnostic v1.1 item 1

264
Problem 2: Pendulum-Box-Bash

The pendulum (mass $M$ at the end of a string of length $L$) is released from rest and swings down to hit the box (mass $m$). The pendulum stops after it hits the box, and the box slides a distance $d$ along a rough surface until it stops.

The goal of this problem would be to determine the coefficient of sliding friction $\mu$ between the box and the surface. The values of $M$, $m$, $L$ and $d$ are known.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure G.2 - Problem Decomposition Diagnostic v1.1, item 2
Problem 3: Spring Launcher

A ball rests on a frictionless ramp, pushed back against a spring which is compressed by an amount $d$. When released, the ball should move up the ramp and then fly through the air to land in the tray at the right side of the page. The spring stops pushing the ball when it reaches the end of the ramp. The top of the ramp is a height $h$ above the landing pad, and the ball starts out a height $y$ below the top of the ramp.

The goal of this problem would be to find the amount $d$ that the spring must be compressed in order for the ball to land in the tray. You would know or be able to measure $k$, $m$, $h$, $\theta$ (relating $y$ and $d$) and the distance $L$. Air friction can be ignored.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure G.3 - Problem Decomposition Diagnostic v1.1, item 3
A block of mass $M$ has an initial velocity $v_0$. It slides a distance $d$ along a rough surface with coefficient of friction $\mu$. The first block then sticks to a smaller block $m$, which is attached to a spring with spring constant $k$. The second block has the same coefficient of friction with the surface, and the spring is initially relaxed. The black triangle at the far right is attached to the surface.

The goal of this problem would be to determine the speed $v_0$ necessary for the spring to be compressed by $x$ before the blocks start moving back to the left. You know or can find the quantities $\mu$, $m$, $M$, $k$, $d$ and $x$.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
<th>Part</th>
<th>Type of Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure G.4 - Problem Decomposition Diagnostic v1.1, item 4
Problem 5: Stunt Ramp

A ball starts out at the bottom of a frictionless ramp with an initial velocity of $v_0$ in the direction shown. You want it to land at the top of another frictionless ramp on the right side of the page, a distance $d$ away from the end of the first ramp.

The goal of this problem would be to determine $d$ so that the ball lands at the top of the right-hand ramp. You know or can find $v_0$, $m$, $\theta$ and $h$.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure G.5 - Problem Decomposition Diagnostic v1.1, item 5
APPENDIX H

THE PROBLEM DECOMPOSITION DIAGNOSTIC VERSION 2.0

This version of the PDD was never completed or administered to students, but is included for completeness. It represents an attempt to retain a relatively open response format in decomposition like that of version 1.1 while still being usable with a scannable test sheet. The classification section was presented to students in an interview situation.

Figures H-1 and H-2 are the instructions for the decomposition component of the test, while figures H-3 through H-7 are the items of the classification component. The following are the instructions for the classification component:

INSTRUCTIONS: There are 5 items in this section. Each item will present a less complex problem and then several options for solution strategies. All of these strategies will work. In this section, pick the strategy that you think would give you the best chance of solving the problem. In other words, which are you most comfortable with? There is no “wrong” answer in this section, as the items are intended to build a picture of the way you prefer to solve problems.

Do not worry about trying to figure out how the instructor might solve the problem, or how anyone else in your study group would solve the problem. Simply pick the answer which best fits how you would solve the problem.
Part 1: Breaking Problems Up

In this section, you will simply be breaking a physics problem into subproblems that you could solve more easily. **You will not solve these problems.**

The following sample problem demonstrates how to complete the test sheet.

A box of mass $m$ starts at the top of a frictionless ramp, a height $h+y$ above the table. It slides down the ramp until it reaches a height $h$ above the table, at which point it flies off the end of the ramp and lands on the floor a distance $d$ from the end of the ramp. The goal of this problem would be to find $d$, the position of the block when it lands.

The big problem can be broken into subproblems. The first subproblem is called A, the second part B, and so forth. Each segment of the "timeline" above corresponds to a number on the scoring sheet. The sample problem uses Roman numerals i through v for its five segments. The first problem of the test will have segments 1-5, the second problem will have segments 6-10, etc.

Figure H.1 - PDD v2.0 Decomposition instructions page one
In this example problem, subproblem A occurs in timeline segments i, ii, and iii, as the block slides down the ramp. Subproblem B occurs in timeline segments iii, iv, and v, as the block flies through the air.

The two subproblems overlap in timeline segment iii. That’s okay. When you get an overlap, only fill in the last subproblem in that timeline segment.

Additionally, in some problems, a subproblem may only occur in a single timeline segment. If an extra subproblem occurred in timeline segment iii, then the last three lines of the sheet would have C darkened, instead of B.
6) A block slides down an incline with gravity \( g \) and a coefficient of friction \( \mu \) known. How fast is the block moving after it has slid a distance \( x \)?

I would:

A) Apply energy principles at the beginning and the end of the block's slide and figure out where the energy goes.
B) Look through the book for an inclined plane problem with friction and follow its example.
C) Find an equation with velocity, change in height and friction and see if I could solve for velocity.
D) Draw a force diagram, sum the forces and find the acceleration of the block and then use the acceleration to find the speed.
E) Check the chapter on friction for examples that involve pushing something against friction.

Figure H.3 - PDD v2.0, first item in the Classification section. Would have been numbered either 6 or 26 depending on the method used in the Decomposition section.
7) A pendulum swings down from a horizontal position to an angle \( \theta \). You know the mass of the pendulum bob and the length of the string, and the acceleration is \( g \). How fast is the pendulum moving at \( \theta \)?

I would:

a) Find the tangential component of acceleration due to gravity as a function of angle and use that to find the velocity as a function of angle.

b) Look through the textbook for an equation relating pendulum speed to angle.

c) Consider the energy at the beginning and at the end and figure out how much kinetic energy the pendulum must have at \( \theta \).

d) Check the textbook and my lecture notes for anything with pendulum problems and see if there's anything helpful in them.

e) Determine the torque due to gravity and use that to find the angular velocity as a function of angle.

Figure H.4 - PDD v2.0, second item in the Classification section.
8) A cannon fires its shot with an initial velocity $v_0$ which is known (both direction and magnitude). It reaches an unknown maximum height $h$ above the mouth of the cannon after traveling a horizontal distance $d$. What is the speed of the shot at the top of the arc?

I would:

a) Use the equation relating maximum height to velocity for a projectile.

b) Find the maximum height $h$ and use conservation of energy to find the change in velocity.

c) Check all the two-dimensional kinematic equations with initial and final velocities and find the one that fits best.

d) Reread the section of the book covering projectile motion and see if I can find a worked example of a similar cannon problem.

e) Break $v_0$ into horizontal and vertical components and then use kinematics to see how they change.

Figure H.5 - PDD v2.0, third item in the Classification section.
9) A Sport Utility Vehicle (SUV) of mass $M$ is helping push a stalled compact car of mass $m$ off the road. After a short period of constant acceleration $t$, the two vehicles reach a speed $v$. If the SUV driver pushed the gas pedal down by the same amount while not pushing a car, how fast would it be going after the same amount of time, $t$?

\[ \text{Figure H.6 - PDD v2.0, fourth item in the Classification section.} \]
10) A block slides along a frictionless surface with a velocity $v$ until it hits a massless spring with spring constant $k$. It compresses the spring by a distance $d$ (the two pictures are before and after shots). What is the mass, $m$, of the block?

I would:

a) Find an equation that relates initial velocity to stopping distance and solve it for mass.
b) Look through my book and notes for a "spring catching block" problem.
c) Find the force of the spring and use Newton's Second Law to relate the acceleration and mass.
d) Apply conservation of energy principles to the "before" and "after" states.
e) Use the equation for a mass on the end of a spring to find the maximum speed as a function of maximum compression.

Figure H.7 - PDD v2.0, fifth item in the Classification section.
APPENDIX I

THE PROBLEM DECOMPOSITION DIAGNOSTIC VERSION 2.1

This was the first fully multiple-choice version of the PDD to be administered to students. Figures I.1 through I.10 will show the ten items of this test, and will be followed by two additional information-gathering items used when this test was administered to students in Physics 132 at the beginning of the summer term in 1999. This version reflects both the results of the open responses on version 1.1 and a limited number of student interviews regarding the new classification section. Section II.

A note on presentation: the instructions for Section II were placed at the beginning of that section in the test as given to students. All instructions have been condensed into one block for the purposes of this appendix.
The Problem Decomposition Diagnostic

The Problem Decomposition Diagnostic, or PDD, is a test designed to measure two aspects of problem-solving skill. The two sections of the PDD each have a number of multiple-choice items that will probe these aspects. The PDD is meant as a diagnostic, to help determine each student's strengths and weaknesses, and is not a part of your final grade. However, while it is not a part of your grade, please take it seriously, as it is intended to help you.

INSTRUCTIONS

ANSWER SHEET: Fill out your name and student ID number on the answer sheet, and be sure to fill in the circles below your name and ID number. If your name is too long to fit on the sheet, fill out your family name and then as much of your first name as will fit. You do not need to fill out the other information, such as age, birthdate, and so forth.

SECTION ONE:

There are five problems in this section. Each presents a complex problem which may need to be broken into smaller sections to be solved. The diagram for the problem is broken into 7 arbitrary parts, or intervals.

After the problem is presented, there will be five options for breaking the problem into manageable sub-problems. Pick the one you think would best describe how you would go about breaking the problem up. Your preferred method may not be an available choice (there are many ways to solve any complex problem) so choose the best one of the alternatives given.
Sample choice:

a) Three sub-problems: 1-2, 2-5, 5-7

This means that the first subproblem takes place in intervals 1-2, the second subproblem takes place in intervals 2-5, and the third subproblem takes place in intervals 5-7. Note that it is possible for a subproblem to happen entirely inside an interval.

SECTION TWO:

There are 5 items in this section. Each item will present a less complex problem and then several options for solution strategies. All of these strategies will work. In this section, pick the strategy that you think would give you the best chance of solving the problem. In other words, which are you most comfortable with? There is no “wrong” answer in this section, as the items are intended to build a picture of the way you prefer to solve problems.

Do not worry about trying to figure out how the instructor might solve the problem, or how anyone else in your study group would solve the problem. Simply pick the answer which best fits how you would solve the problem.

PLEASE DO NOT MAKE ANY MARKS ON THE TEST SHEET, AND ERASE ANY YOU FIND.
SECTION ONE:
1) Pendulum-Box-Bash

The pendulum (mass \( M \) at the end of a string of length \( L \)) is released from rest and swings down to hit the box (mass \( m \)). The pendulum stops after it hits the box, and the box slides a distance \( d \) along a rough surface until it stops. There are only one pendulum and one box, the leftmost drawings of each are the starting positions, and the rightmost drawings are the final positions.

The goal of this problem would be to determine the coefficient of sliding friction \( \mu \) between the box and the surface. The values of \( M, m, L, \) and \( d \) are known.

(1) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-3, 3-7
b) Three subproblems: 1-3, 3, 3-7
c) Three subproblems: 1-2, 3, 4-7
d) Three subproblems: 1-3, 3-4, 4-7
e) Three subproblems: 1-3, 3, 4-7

Figure I.1 - Problem Decomposition Diagnostic v2.1 item 1
A ball rests on a frictionless ramp, pushed back against a spring which is compressed by an amount \( d \). When released, the ball should move up the ramp and then fly through the air to land in the tray at the right side of the page. The spring stops pushing the ball when it reaches the end of the ramp. The top of the ramp is a height \( h \) above the landing pad, and the ball starts out a height \( y \) below the top of the ramp. Both the tray and the ramp are attached to the floor and do not move.

The goal of this problem would be to find the amount \( d \) that the spring must be compressed in order for the ball to land in the tray. You would know or be able to measure \( k, m, h, \theta \) (relating \( y \) and \( d \)), and the distance \( L \). Air friction can be ignored.

(2) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1, 2-7
b) Two subproblems: 1-2, 2-7
c) Three subproblems: 1, 1-2, 2-7
d) Three subproblems: 1, 2, 3-7
e) Three subproblems: 1, 2, 2-7
A block of mass \( m \) has an initial velocity of zero, starting a height \( h \) up a frictionless ramp inclined at angle \( \theta \). The block slides down the frictionless ramp and reaches a rough surface with coefficient of friction \( \mu \). After sliding a distance \( d \) on this surface, the block hits a massless spring and compresses it a distance \( x \). The spring is initially relaxed and the black triangle at the far right is attached to the surface.

The goal of this problem would be to determine the spring constant \( k \) necessary for the spring to be compressed by \( x \) before the block starts moving back to the left. You know or can find the quantities \( \mu, m, \theta, h, d \), and \( x \).

(3) How many subproblems are there in this problem, and what intervals do they cover?

a) One subproblem: 1-5
b) One subproblem: 1-7
c) Two subproblems: 1-3, 3-7
d) Three subproblems: 1-3, 3-5, 7
e) Three subproblems: 1-3, 3-5, 5-7
A block of mass $M$ has an initial velocity $v_0$. It slides a distance $d$ along a rough surface with coefficient of friction $\mu$. The first block then sticks to a smaller block of mass $m$, which is attached to a spring with spring constant $k$. The coefficient of friction between the second block and the surface is also $\mu$, and the spring is initially relaxed. The black triangle at the far right is attached to the surface.

The goal of this problem would be to determine the speed $v_0$ necessary for the spring to be compressed by $x$ before the blocks start moving back to the left. You know or can find the quantities $\mu$, $m$, $M$, $k$, $d$, and $x$.

(4) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-5, 6-7
b) Two subproblems: 1-6, 6-7
c) Three subproblems: 1-6, 6-7, 7
d) Three subproblems: 1-5, 6, 7
e) Three subproblems: 1-6, 6, 6-7
A ball starts out at the bottom of a frictionless ramp with an initial velocity of $v_0$ in the direction shown. You want it to land at the top of another frictionless ramp on the right side of the page, a distance $d$ away from the end of the first ramp. Both ramps are attached to the table and do not move.

The goal of this problem would be to determine $d$ so that the ball lands at the top of the right-hand ramp. You know or can find $v_0$, $m$, $\theta$, and $h$.

(5) How many subproblems are there in this problem, and what intervals do they cover?

a) One subproblem: 1-6  
b) One subproblem: 2-6  
c) Two subproblems: 1-2, 2-6  
d) Two subproblems: 1-2, 3-6  
e) Two subproblems: 1, 2-6
6) A block slides down an incline with gravity $g$ and a coefficient of friction $\mu$ known. How fast is the block moving after it has slid a distance $x$?

I would:

a) Apply energy principles at the beginning and the end of the block's slide and figure out where all the energy goes.
b) Look through my textbook for an inclined plane problem with friction and follow its example.
c) Find an equation with velocity, change in height and friction and see if I could solve for velocity.
d) Draw a force diagram, sum the forces and find the acceleration of the block, and then use the acceleration to find the speed.
7) A pendulum swings down from a horizontal position to an angle \( \theta \). You know the mass of the pendulum bob and the length of the string, and the acceleration is \( g \). How fast is the pendulum moving at \( \theta \)?

I would:

a) Find the tangential component of acceleration due to gravity as a function of angle and use that to find the velocity as a function of angle.

b) Look through my textbook and notes for an equation relating pendulum speed to angle.

c) Consider the energy at the beginning and at the end and figure out how much kinetic energy the pendulum must have at \( \theta \).

d) Check the textbook and my lecture notes for anything with pendulum problems and see if there’s anything helpful in them.

Figure I.7 - Problem Decomposition Diagnostic v2.1 item 7
8) A cannon fires its shot with an initial velocity \( \mathbf{v}_0 \) which is known (both direction and magnitude). It reaches an unknown maximum height \( h \) above the mouth of the cannon after traveling a horizontal distance \( d \). What is the speed of the shot at the top of the arc?

I would:

a) Use the equation from my textbook relating maximum height to velocity for a projectile.

b) Find the maximum height \( h \) and use conservation of energy to find the change in velocity.

c) Check the two-dimensional kinematic equations with initial and final velocities and find the one that fits best.

d) Break \( \mathbf{v}_0 \) into horizontal and vertical components and then use kinematics to see how they change.
9) A Sport Utility Vehicle (SUV) of mass $M$ is helping push a stalled compact car of mass $m$ off the road. After a short period of constant acceleration $t$, the two vehicles reach a speed $v$. If the SUV driver pushed the gas pedal down by the same amount while not pushing a car (exerting the same force), how fast would it be going after the same amount of time, $t$?

I would:

a) Look at problems involving cars and trucks pushing each other and see if any of them ask the same question.
b) Use kinematics to find the force exerted by the SUV, then use that to find an acceleration on just the SUV.
c) Look at equations that have two different masses and two different velocities, pick one that seems to best fit the situation.
d) Draw force diagrams for both vehicles and find all the forces, then remove the car's forces and apply Newton's Second Law.

Figure 1.9 - Problem Decomposition Diagnostic v2.1 item 9
10) A block slides along a frictionless surface with a velocity $v$ until it hits a massless spring with spring constant $k$. It compresses the spring by a distance $d$. What is the mass, $m$, of the block?

I would:

a) Look through my textbook and notes for a problem in which a spring catches a block and apply that example to this situation.

b) Find the force of the spring and use Newton's Second Law to relate the acceleration and mass.

c) Apply conservation of energy principles to the "before" and "after" states.

d) Use the equation for a mass oscillating on the end of a spring to find $v$ as a function of $d$. 

Figure I.10 - Problem Decomposition Diagnostic v2.1 item 10
Finally, here are the two information-gathering items appended to the PDD v2.1 when it was administered.

The following two questions are to help us gather information for research purposes. Your answers will be kept confidential.

11) When did you take Physics 131 (or the equivalent calculus-based mechanics course)?

   a) Spring 1999
   b) Winter 1999
   c) Autumn 1998 or earlier
   d) In the 1998-9 Academic year, but not at the Ohio State University
   e) I did not take Physics 131 or its equivalent (skip #12)

12) What was your letter grade in Physics 131 or its equivalent?

   a) A- or better
   b) B- to B+
   c) C- to C+
   d) Below C-
   e) Took the course pass/fail or other non-letter-grade
APPENDIX J

THE PROBLEM DECOMPOSITION DIAGNOSTIC VERSION 2.2

This version of the test represented alterations to the PDD based on the administration of version 2.1. It proved stable enough in summer of 1999 that it was not altered for administration to 131E students in autumn 1999. Figures J.1 through J.10 will show the ten items of this test. No additional items for information-gathering were included, as full grade information was available for all students involved.

A note on presentation: the instructions for Section II were placed at the beginning of that section in the test as given to students. All instructions have been condensed into one block for the purposes of this appendix.
The Problem Decomposition Diagnostic

The Problem Decomposition Diagnostic, or PDD, is a test designed to measure two aspects of problem-solving skill. The two sections of the PDD each have a number of multiple-choice items that will probe these aspects. Your score on the PDD will not affect your final grade, but please give it your best effort.

INSTRUCTIONS

ANSWER SHEET: Fill out your name and student ID number on the answer sheet, and be sure to fill in the circles below your name and ID number. If your name is too long to fit on the sheet, fill out your family name and then as much of your first name as will fit. You do not need to fill out the other information, such as age, birthdate, and so forth.

SECTION ONE:

There are five problems in this section. Each presents a complex problem which may need to be broken into smaller sections to be solved. The diagram for the problem is broken into 7 arbitrary parts, or intervals.

After the problem is presented, there will be five options for breaking the problem into manageable sub-problems. Pick the one you think would best describe how you would go about breaking the problem up. Your preferred method may not be listed, as there are many ways to solve any complex problem. Please choose whichever one of the alternatives given comes closest to your preferred solution.
Sample choice:

a) Three sub-problems: 1-3, 3-7, 7

This means that the first subproblem takes place in intervals 1-3, the second subproblem takes place in intervals 3-7, and the third subproblem takes place only in interval 7.

SECTION TWO:

There are 5 items in this section. Each item will present a less complex problem and then several options for solution strategies. All of these strategies will work. In this section, pick the strategy that you think would give you the best chance of solving the problem. In other words, with which are you most comfortable? There is no "wrong" answer in this section, as the items are intended to build a picture of the way you prefer to solve problems.

Do not worry about trying to figure out how the instructor might solve the problem, or how anyone else in your study group would solve the problem. Simply pick the answer which best fits how you would solve the problem.

PLEASE DO NOT MAKE ANY MARKS ON THE TEST SHEET, AND ERASE ANY YOU FIND.
The pendulum (mass $M$ at the end of a string of length $L$) is released from rest and swings down to hit the box (mass $m$ less than $M$). The pendulum stops after it hits the box, and the box slides a distance $d$ along a rough surface until it stops. There are only one pendulum and one box, the leftmost drawings of each are the starting positions, and the rightmost drawings are the final positions.

The goal of this problem would be to determine the coefficient of sliding friction $\mu$ between the box and the surface. The values of $M$, $m$, $L$, and $d$ are known.

(1) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-3, 3-7
b) Three subproblems: 1-3, 3-7

c) Three subproblems: 1-2, 3-4-7
d) Three subproblems: 1-3, 3-4, 4-7
e) Three subproblems: 1-3, 3, 4-7

Figure J.1 - Problem Decomposition Diagnostic v2.2 item 1

294
A ball sits on a frictionless ramp, resting against a spring which is compressed by an amount \( d \). When released, the ball should move up the ramp and then fly through the air to land in the tray at the right side of the page. The spring stops pushing the ball when it reaches the end of the ramp. The top of the ramp is a height \( h \) above the landing pad, and the ball starts a height \( y \) below the top of the ramp. Both the tray and the ramp are attached to the floor and do not move.

The goal of this problem would be to find the amount \( d \) that the spring must be compressed in order for the ball to land in the tray. You would know or be able to measure \( k, m, h, \theta \) (relating \( y \) and \( d \)), and the distance \( L \). Air friction can be ignored.

(2) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1, 2-7
b) Two subproblems: 1-2, 2-7
c) Three subproblems: 1, 1-2, 2-7
d) Three subproblems: 1, 2, 3-7
e) Three subproblems: 1, 2, 2-7

Figure J.2 - Problem Decomposition Diagnostic v2.2 item 2
A block of mass \( m \) has an initial velocity of zero, starting a height \( h \) up a frictionless ramp inclined at angle \( \theta \). The block slides down the frictionless ramp and reaches a rough surface with coefficient of friction \( \mu \). After sliding a distance \( d \) on this surface, the block hits a massless spring and compresses it a distance \( x \). The spring is initially relaxed and the black triangle at the far right is attached to the surface.

The goal of this problem would be to determine the spring constant \( k \) necessary for the spring to be compressed by \( x \) before the block starts moving back to the left. You know or can find the quantities \( \mu, m, \theta, h, d, \) and \( x \).

(3) How many subproblems are there in this problem, and what intervals do they cover?

a) One subproblem: 1-7
b) Two subproblems: 1-7, 7
c) Three subproblems: 1-5, 5, 5-7
d) Three subproblems: 1-3, 3-7, 7
e) Three subproblems: 1-3, 3-5, 5-7
A block of mass $M$ has an initial velocity $v_0$. It slides a distance $d$ along a rough surface with coefficient of friction $\mu$. The first block then sticks to a smaller block of mass $m$, which is attached to a spring with spring constant $k$. The coefficient of friction between the second block and the surface is also $\mu$, and the spring is initially relaxed. The black triangle at the far right is attached to the surface.

The goal of this problem would be to determine the speed $v_0$ that the larger block has to have in order for the spring's maximum compression to be $x$. You know or can find the quantities $\mu$, $m$, $M$, $k$, $d$, and $x$.

(4) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-5, 6-7
b) Two subproblems: 1-6, 6-7
c) Three subproblems: 1-6, 6-7, 7
d) Three subproblems: 1-5, 6, 7
e) Three subproblems: 1-6, 6, 6-7
5) Stunt Ramp

A ball starts out at the bottom of a frictionless ramp with an initial velocity of $v_0$ in the direction shown. You want it to land at the top of an identical frictionless ramp on the right side of the page, a distance $d$ away from the end of the first ramp. Both ramps sit on a level tabletop in normal Earth gravity.

The goal of this problem would be to determine $d$ so that the ball lands at the top of the right-hand ramp. Once the second ramp is placed on the table, both ramps are clamped down and do not move. You know or can find $v_0$, $m$, $\theta$, and $h$.

**Figure J.5 - Problem Decomposition Diagnostic v2.2 item 5**

(5) **How many subproblems are there in this problem, and what intervals do they cover?**

a) One subproblem: 1-6
b) One subproblem: 2-6
c) Two subproblems: 1-2. 2-6
d) Two subproblems: 1-2. 3-6
e) Two subproblems: 1. 2-6
6) A block slides down an incline. Values for gravity (g) and a coefficient of friction μ are known. How fast is the block moving after it has slid a distance x?

I would:

a) Apply energy principles at the beginning and the end of the block's slide and figure out where all the energy goes.

b) Compare this situation to other inclined plane problems in my textbook and notes and see if the problem has been solved there.

c) Find an equation with velocity, change in height and friction and see if I could solve for velocity.

d) Draw a force diagram, sum the forces and find the acceleration of the block, and then use the acceleration to find the speed.
7) A pendulum swings down from a horizontal position to an angle $\theta$. You know the mass of the pendulum bob and the length of the string, and the acceleration of gravity, $g$, is known. How fast is the pendulum moving at $\theta$?

I would:

a) Find the tangential component of acceleration due to gravity as a function of angle and use that to find the velocity as a function of angle.
b) Look through my textbook and notes for an equation relating pendulum speed to angle.
c) Consider the energy at the beginning and at the end and figure out how much kinetic energy the pendulum must have at $\theta$.
d) Look at the solutions of other pendulum problems in my textbook and notes first, to see if this problem has already been solved.
8) A cannon fires its shot with an initial velocity $\mathbf{v}_0$ which is known (both direction and magnitude). It reaches an unknown maximum height $h$ above the mouth of the cannon after traveling a horizontal distance $d$. What is the speed of the shot at the top of the arc?

I would:

a) Use the equation from my textbook relating maximum height to velocity for a projectile.

b) Find the maximum height $h$ and use conservation of energy to find the change in velocity.

c) Check the two-dimensional kinematic equations with initial and final velocities and find one that involves the givens of this problem ($d$, etc).

d) Break $\mathbf{v}_0$ into horizontal and vertical components and then use kinematics to see how they change.
9) A Sport Utility Vehicle (SUV) of mass $M$ is helping push a stalled compact car of mass $m$ off the road. The SUV driver holds the gas pedal down halfway for a time $t$ (measured in seconds). At the end of this time, the two vehicles have reached a speed $v$. If the SUV driver pushed the gas pedal down halfway while not pushing a car, how fast would it be going after $t$ seconds?

I would:

a) Go over examples in the textbook and my notes that involve things pushing other things and see if the worked problems are any help.
b) Use kinematics to find the force exerted by the SUV, then use that to find an acceleration on just the SUV.
c) Look at equations that have two different masses and two different velocities, pick one that seems to best fit the situation.
d) Draw force diagrams for the vehicles and find all the forces, then remove the compact car's forces and apply Newton's Second Law.

Figure J.9 - Problem Decomposition Diagnostic v2.2 item 9
10) A block slides along a frictionless surface with a velocity \( v \) until it hits a massless spring with spring constant \( k \). The maximum compression of the spring is the distance \( d \). You know the values for \( v \), \( k \) and \( d \). What is the mass, \( m \), of the block?

I would:

a) Look through my textbook and notes for a problem in which a spring catches a block and apply that example to this situation.
b) Find the force of the spring (as a function of position) and use Newton's Second Law to relate the acceleration and mass.
c) Apply conservation of energy principles to the "before" and "after" states.
d) Use the equation for a mass oscillating on the end of a spring, which relates \( v \), \( d \) and \( m \), then solve for mass.
APPENDIX K

THE PROBLEM DECOMPOSITION DIAGNOSTIC VERSION 2.3

This version of the test represents alterations to the PDD based on the administration of version 2.2 in the fall of 1999. It has not been administered to any students at the time of this writing. Figures K.1 through K.10 will show the ten items of this test.

A note on presentation: the instructions for Section II would be placed at the beginning of that section in the test as given to students. All instructions have been condensed into one block for the purposes of this appendix.
The Problem Decomposition Diagnostic

The Problem Decomposition Diagnostic, or PDD, is a test designed to measure two aspects of problem-solving skill. The two sections of the PDD each have a number of multiple-choice items that will probe these aspects. Your score on the PDD will not affect your final grade, but please give it your best effort.

INSTRUCTIONS

ANSWER SHEET: Fill out your name and student ID number on the answer sheet, and be sure to fill in the circles below your name and ID number. If your name is too long to fit on the sheet, fill out your family name and then as much of your first name as will fit. You do not need to fill out the other information, such as age, birthdate, and so forth.

SECTION ONE:

There are five problems in this section. Each presents a complex problem which may need to be broken into smaller sections to be solved. The diagram for the problem is broken into 7 arbitrary parts, or intervals.

After the problem is presented, there will be five options for breaking the problem into manageable sub-problems. Pick the one you think would best describe how you would go about breaking the problem up. Your preferred method may not be listed, as there are many ways to solve any complex problem. Please choose whichever one of the alternatives given comes closest to your preferred solution.
Sample choice:

a) Three sub-problems: 1-3, 3-7, 7

This means that the first subproblem takes place in intervals 1-3, the second subproblem takes place in intervals 3-7, and the third subproblem takes place only in interval 7.

SECTION TWO:

There are 5 items in this section. Each item will present a less complex problem and then several options for solution strategies. All of these strategies will work. In this section, pick the strategy that you think would give you the best chance of solving the problem. In other words, with which are you most comfortable? There is no “wrong” answer in this section, as the items are intended to build a picture of the way you prefer to solve problems.

Do not worry about trying to figure out how the instructor might solve the problem, or how anyone else in your study group would solve the problem. Simply pick the answer which best fits how you would solve the problem.

PLEASE DO NOT MAKE ANY MARKS ON THE TEST SHEET, AND ERASE ANY YOU FIND.
SECTION ONE:
1) Pendulum-Box-Bash

The pendulum (mass $M$ at the end of a string of length $L$) is released from rest and swings down to hit the box (mass $m$ less than $M$). The pendulum stops after it hits the box, and the box slides a distance $d$ along a rough surface until it stops. There are only one pendulum and one box, the leftmost drawings of each are the starting positions, and the rightmost drawings are the final positions.

The goal of this problem would be to determine the coefficient of sliding friction $\mu$ between the box and the surface. The values of $M$, $m$, $L$, and $d$ are known.

(1) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-3, 3-7
b) Three subproblems: 1-3, 3, 3-7
c) Three subproblems: 1-2, 3, 4-7
d) Three subproblems: 1-3, 3-4, 4-7
e) Three subproblems: 1-3, 3, 4-7

Figure K.1 - Problem Decomposition Diagnostic v2.3 item 1
A ball sits on a frictionless ramp, resting against a spring which is compressed by an amount $d$. When released, the ball should move up the ramp and then fly through the air to land in the tray at the right side of the page. The spring stops pushing the ball when it reaches the end of the ramp. The ball starts a height $h$ below the top of the ramp. Both the tray and the ramp are attached to the floor and do not move.

The goal of this problem would be to find the amount $d$ that the spring must be compressed in order for the ball to land in the tray. You would know or be able to measure $k, m, h, \theta$ (relating $h$ and $d$), and the distance $L$. Air friction can be ignored.

(2) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-3, 3-7
b) Two subproblems: 1-3, 4-7
c) Three subproblems: 1, 1-3, 3-7
d) Three subproblems: 1, 2-3, 3-7
e) Three subproblems: 1, 2-3, 4-7
A block of mass $m$ has an initial velocity of $v$, moving a distance $d$ up a frictionless ramp that ends a height $h_1$ above the starting position of the block. After leaving the ramp, the block flies through the air across a chasm $L$ wide and hits a spring that's $h_2$ higher than the end of the ramp. When the block hits the spring, it has no vertical component of velocity. The massless spring compresses by $x$.

The goal of this problem would be to determine the spring constant $k$ necessary for the spring to be compressed by $x$ before the block starts moving back to the left. You know or can find the quantities $m$, $h_1$, $h_2$, $d$, $L$ and $x$.

(3) How many subproblems are there in this problem, and what intervals do they cover?

a) One subproblem: 1-7  
b) Two subproblems: 1-7, 7  
c) Three subproblems: 1-5, 5, 5-7  
d) Three subproblems: 1-5, 5-7, 7  
e) Three subproblems: 1-3, 3-5, 5-7

Figure K.3 - Problem Decomposition Diagnostic v2.3 item 3
4)

Block Catcher II

A block of mass $M$ has an initial velocity $v$. It slides a distance $d$ along a rough surface with coefficient of friction $\mu$. The first block then sticks to a lighter block of mass $m$, which is attached to a spring with spring constant $k$. The coefficient of friction between the second block and the surface is also $\mu$, and the spring is initially relaxed. The black triangle at the far right is attached to the surface, and the dotted square shows the position of the first block when it sticks to the second.

The goal of this problem would be to determine the speed $v$ that the heavier block has to have in order for the spring’s maximum compression to be $x$. You know or can find the quantities $\mu, m, M, k, d,$ and $x$.

(4) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-4, 5-7
b) Two subproblems: 1-5, 5-7
c) Three subproblems: 1-5, 5-7, 7
d) Three subproblems: 1-4, 5, 6-7
e) Three subproblems: 1-5, 5, 5-7

Figure K.4 - Problem Decomposition Diagnostic v2.3 item 4
A ball starts out at the bottom of a frictionless ramp with an initial velocity of $v_0$ in the direction shown. You want it to land at the top of a frictionless block on the right side of the page, a distance $d$ away from the end of the ramp. Both the ramp and the block have a height of $h$ and sit on a level tabletop in normal Earth gravity.

The goal of this problem would be to determine $d$ so that the ball lands at the top of the right-hand block. Once the block is placed on the table, both the ramp and the block are clamped down and do not move. You know or can find $v_0$, $m$, $\theta$, and $h$.

(5) How many subproblems are there in this problem, and what intervals do they cover?

a) One subproblem: 1-7
b) One subproblem: 3-7
c) Two subproblems: 1-3, 3-7
d) Two subproblems: 1-3, 4-7
e) Two subproblems: 1, 3-7

Figure K.5 - Problem Decomposition Diagnostic v2.3 item 5
6) A block slides down an incline. Values for gravity (g) and a coefficient of friction \( \mu \) are known. How fast is the block moving after it has slid a distance \( x \)?

I would:

a) Apply energy principles at the beginning and the end of the block’s slide and figure out where all the energy goes.

b) Compare this situation to other inclined plane problems in my textbook and notes and see if the problem has been solved there.

c) Find an equation with velocity, change in height and friction and see if I could solve for velocity.

d) Draw a force diagram, sum the forces and find the acceleration of the block, and then use the acceleration to find the speed.
7) A pendulum swings down from a horizontal position to an angle $\theta$. You know the mass of the pendulum bob and the length of the string, and the acceleration of gravity, $g$, is known. How fast is the pendulum moving at $\theta$?

I would:

a) Find the tangential component of acceleration due to gravity as a function of angle and use that to find the velocity as a function of angle.

b) Look through my textbook and notes for an equation relating pendulum speed to angle.

c) Consider the energy at the beginning and at the end and figure out how much kinetic energy the pendulum must have at $\theta$.

d) Look at the solutions of other pendulum problems in my textbook and notes first, to see if this problem has already been solved.

Figure K.7 - Problem Decomposition Diagnostic v2.3 item 7
8) A cannon fires its shot with an initial speed \( \mathbf{v}_0 \) which is known, but at an unknown angle. It reaches an known maximum height \( h \) above the mouth of the cannon after traveling an unknown horizontal distance \( d \). What is the speed of the shot at the top of the arc?

I would:

a) Look over the problems in my text and notes that involve projectile motion and see if any seem to fit this situation.

b) Apply conservation of energy to find the change in kinetic energy.

c) Check the two-dimensional kinematic equations with initial and final velocities and find one that involves the givens of this problem (\( d \), etc).

d) Use kinematics to figure out the initial vertical component of velocity, then use that to find the horizontal component of velocity.

Figure K.8 - Problem Decomposition Diagnostic v2.3 item 8
9) Two blocks of masses \( m \) and \( M \) (\( M > m \)) hang from a massless, frictionless pulley (radius \( R \)), connected by an unstretchable, massless string. The blocks are released and start to move. After one second, the more massive block has dropped by \( H \) and the less massive block has risen by \( H \). How fast are the two blocks moving after one second?

I would:

a) Go over examples in the textbook and my notes that involve pulleys and see if the worked problems are any help.
b) Draw Free Body Diagrams for both blocks, find the acceleration of each block using Newton's Second Law, then determine the speed.
c) Look at equations that involve moving pulleys and pick one that seems to best fit the situation.
d) Apply conservation of energy to find the change in kinetic energies as the potential energies of the blocks change.

Figure K.9 - Problem Decomposition Diagnostic v2.3 item 9
10) A block slides along a frictionless surface with a velocity $v$ until it hits a massless spring with spring constant $k$. The maximum compression of the spring is the distance $d$. You know the values for $v$, $k$, and $d$. What is the mass, $m$, of the block?

I would:

a) Look through my textbook and notes for a problem in which a spring catches a block and apply that example to this situation.
b) Find the force of the spring (as a function of position) and use Newton’s Second Law to relate the acceleration and mass.
c) Apply conservation of energy principles to the “before” and “after” states.
d) Use the equation for a mass oscillating on the end of a spring, which relates $v$, $d$ and $m$, then solve for mass.
APPENDIX L

PROBLEM DECOMPOSITION DIAGNOSTIC
EXPANSION ITEMS, OPEN RESPONSE

The items presented here were devised for possible expansion of the PDD or to replace items that were not salvageable. The instructions below were for the open response version of the test administered to the H131 students in fall of 1999.

Instructions

This test consists of several complex physics problems, which you will not be solving. Rather, you will be determining how to break them up into smaller sub-problems that can be solved more easily.

Each problem has a timeline below it, breaking the problem into a number of intervals of arbitrary length. Each interval represents some period in time during the course of the problem; the intervals are not necessarily the same amount of time. It is possible for several sub-problems to occupy part of the same interval, such as when one sub-problem ends and another begins.

Your task will be to break the problem into pieces, and determine which intervals each sub-problem covers. For example, if you decide that sub-problem A happens in
intervals 1, 2 and 3, and sub-problem B happens in intervals 3 and 4. You would write “A” on the lines for intervals 1-3, and “B” on the lines for intervals 3-4. A sub-problem might only happen during a single interval, or it might be stretched out over several intervals.

You only need consider events that occur during the time intervals shown. Aspects of solving the problem that happen outside of the time intervals, such as reading the problem or checking your answer, don’t count for purposes of this exercise.

Figure L.1 is the sample problem presented to the students. Figures L.2 through L.6 are the five items on this version of the PDD.
Sample Problem: Sliding Box

This problem is broken into parts for you, to help show what you're supposed to do in the later problems. It has two parts, one covering intervals 1-4, the other covering intervals 4-6. The parts overlap, which is allowed. Also, nothing happens in interval 7, which is also okay.

A box of mass $m$ starts at the top of a frictionless ramp, a height $h+y$ above the table. It slides down the ramp until it reaches a height $h$ above the table, at which point it drops off the end of the ramp and lands on the table a distance $d$ from the end of the ramp, stopping immediately.

The goal of this problem would be to find $d$, the position of the block when it lands.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$A$</td>
</tr>
<tr>
<td>2</td>
<td>$A$</td>
</tr>
<tr>
<td>3</td>
<td>$A$</td>
</tr>
<tr>
<td>4</td>
<td>$A, B$</td>
</tr>
<tr>
<td>5</td>
<td>$B$</td>
</tr>
<tr>
<td>6</td>
<td>$B$</td>
</tr>
</tbody>
</table>

Figure L.1 - Problem Decomposition Diagnostic expansion sample item
Problem 1: Ballistic Pendulum

A block of mass \( m \) slides along a frictionless surface (over a distance \( d \) and for a time \( t \)) until it hits a hanging clay ball of mass \( M \) (the position of the block at the moment of impact is shown by the dotted square). The block sticks to the ball and both swing up until the massless string forms an angle \( \theta \) with the vertical.

The goal of this problem would be to find the angle \( \theta \). It is assumed you’d know or could measure \( m, d, t, M \) and \( L \).

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure L.2 - Problem Decomposition Diagnostic expansion item 1
Problem 2: Spring Launcher

A block of mass $M$ is pushed by a spring along a frictionless surface. The spring is compressed by $x$ and has a spring constant of $k$. The block flies off the end of a cliff of height $h$, and the intention is for the block to land in a river which is a distance $d$ from the base of the cliff.

The goal of this problem would be to find out the smallest spring constant $k$ that will get the block into the river. Assume you know or can measure $M$, $x$, $h$ and $d$.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure L.3 - Problem Decomposition Diagnostic expansion item 2
Problem 3: Bell Ringer

A bead of mass \( M \) slides along a metal rod, with an initial velocity of \( v \). At first it slides on a frictionless curved rod, which rises by \( h \). Then it reaches a horizontal stretch of rod of length \( x \) and coefficient of friction \( \mu \). After that, it slides up another curved frictionless rod until its center of mass has risen by \( h \), assuming it’s going fast enough.

The problem would be to find the minimum \( v \) needed in order for the bead to make it all the way up to the bell.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure L.4 - Problem Decomposition Diagnostic expansion item 3
Problem 4: Orange Catcher

An orange of mass $M$ and initial horizontal velocity $v$ starts out $h$ above the level of a cart on the tabletop. The cart has mass $m$ and is attached to a spring with spring constant $k$. The other end of the initially relaxed spring is attached to the table. When the orange lands on the bed of nails atop the cart, it rolls along a frictionless surface until stopped by the spring.

The problem would be to find out how far ($d$) the cart and orange would travel before the spring caused it to start moving back to the left. You would know or be able to measure $M$, $m$, $v$, $k$ and $h$.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure L.5 - Problem Decomposition Diagnostic expansion item 4
Problem 5: Half Pipe Jump

A skateboarder of mass $M$ starts at the position shown on a frictionless half-pipe ramp of radius $r$. He slides down and flies off the end of the ramp, which is $h$ above the ground. He then flies through the air and lands on a stack of mattresses as tall as the end of the ramp.

The goal of this problem would be to figure out in advance how far away the mattresses have to be, since trial and error would be...painful. You know or can measure $M$, $r$, $\theta$ and $h$.

<table>
<thead>
<tr>
<th>Interval</th>
<th>Parts in Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

Figure L.6 - Problem Decomposition Diagnostic expansion item 5
APPENDIX M

PROBLEM DECOMPOSITION DIAGNOSTIC
EXPANSION ITEMS, MULTIPLE CHOICE

The items presented here were devised for possible expansion of the PDD or to replace items that were not salvageable. The multiple choice options are based on the result of administration of the open response version (Appendix 12). The instructions would be identical to those for the PDD v2.3 (Appendix 11).

Figures M.1 through M.5 present the items with tentative multiple choice options.
Problem 1: Ballistic Pendulum

A block of mass \( m \) slides along a frictionless surface (over a distance \( d \) and for a time \( t \)) until it hits a hanging clay ball of mass \( M \) (the position of the block at the moment of impact is shown by the dotted square). The block sticks to the ball and both swing up until the massless string forms an angle \( \theta \) with the vertical.

The goal of this problem would be to find the angle \( \theta \). It is assumed you’d know or could measure \( m, d, t, M \) and \( L \).

(1) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-5, 5-7
b) Three subproblems: 1-5, 5-6, 7
c) Three subproblems: 1-5, 5, 6-7
d) Three subproblems: 1-5, 5, 5-7
e) Three subproblems: 1-4, 5, 6-7
Problem 2: Spring Launcher

A block of mass $M$ is pushed by a spring along a frictionless surface. The spring is compressed by $x$ and has a spring constant of $k$. The block flies off the end of a cliff of height $h$, and the intention is for the block to land in a river which is a distance $d$ from the base of the cliff.

The goal of this problem would be to find out the smallest spring constant $k$ that will get the block into the river. Assume you know or can measure $M$, $x$, $h$ and $d$.

(2) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-3, 3-7
b) Three subproblems: 1, 1-3, 3-7
c) Three subproblems: 1-3, 3, 3-7
d) Three subproblems: 1-3, 3-7, 7
e) Four subproblems: 1-3, 3, 3-7, 7

Figure M.2 - Problem Decomposition Diagnostic expansion item 2: multiple choice

327
Problem 3: Bell Ringer

A bead of mass $M$ slides along a metal rod, with an initial velocity of $v$. At first it slides on a frictionless curved rod, which rises by $h$. Then it reaches a horizontal stretch of rod of length $x$ and coefficient of friction $\mu$. After that, it slides up another curved frictionless rod until its center of mass has risen by $h$, assuming it’s going fast enough.

The problem would be to find the minimum $v$ needed in order for the bead to make it all the way up to the bell.

(3) How many subproblems are there in this problem, and what intervals do they cover?

a) One subproblem: 1-7
b) Three subproblems: 1-3, 3-5, 5-7
c) Three subproblems: 1-2, 3-5, 5-7
d) Three subproblems: 1-3, 3-5, 6-7
e) Four subproblems: 1-3, 3-5, 5-7, 7

Figure M.3 - Problem Decomposition Diagnostic expansion item 3: multiple choice
Problem 4: Orange Catcher

An orange of mass $M$ and initial horizontal velocity $v$ starts out $h$ above the level of a cart on the tabletop. The cart has mass $m$ and is attached to a spring with spring constant $k$. The other end of the initially relaxed spring is attached to the table. When the orange lands on the bed of nails atop the cart, it rolls along a frictionless surface until stopped by the spring.

The problem would be to find out how far (d) the cart and orange would travel before the spring caused it to start moving back to the left. You would know or be able to measure $M$, $m$, $v$, $k$ and $h$.

(4) How many subproblems are there in this problem, and what intervals do they cover?

a) One subproblems: 1-7
b) Two subproblems: 1-3, 3-7
c) Two subproblems: 1-3, 4-7
d) Two subproblems: 1-4, 4-7
e) Three subproblems: 1-4, 4, 4-7

Figure M.4 - Problem Decomposition Diagnostic expansion item 4: multiple choice
Problem 5: Half Pipe Jump

A skateboarder of mass $M$ starts at the position shown on a frictionless half-pipe ramp of radius $r$. He slides down and flies off the end of the ramp, which is $h$ above the ground. He then flies through the air and lands on a stack of mattresses as tall as the end of the ramp.

The goal of this problem would be to figure out in advance how far away the mattresses have to be, since trial and error would be...painful. You know or can measure $M$, $r$, $\theta$ and $h$.

(5) How many subproblems are there in this problem, and what intervals do they cover?

a) Two subproblems: 1-3, 3-7
b) Two subproblems: 1-3, 4-7
c) Three subproblems: 1-3, 3-6, 7
d) Three subproblems: 1-2, 2-3, 3-7
e) Three subproblems: 1-3, 3, 3-7

Figure M.5 - Problem Decomposition Diagnostic expansion item 5: multiple choice
BIBLIOGRAPHY


331


