

INVESTIGATING THE DEVELOPMENT OF PROBLEM SOLVING  
SKILLS DURING A FRESHMAN PHYSICS SEQUENCE

DISSERTATION

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

By

Kathleen Andre Harper, B.S.E., M.S.

\*\*\*\*\*

The Ohio State University  
2001

Dissertation Committee:

Professor Alan Van Heuvelen, Co-Advisor

Professor Richard Furnstahl, Co-Advisor

Professor Thomas Humanic

Professor Lei Bao

Approved by

---

Co-Advisor

---

Co-Advisor  
Department of Physics



## ABSTRACT

Problem solving research to this point has primarily focused on differences between expert and novice problem solvers. Only a few studies have made any effort to describe how the transition from novice to expert occurs. Also, most problem solving studies have been set in laboratories, so few have taken place in actual classrooms. In an effort to learn more about these unexplored areas, a series of studies was designed.

One portion of this work tracked the development of a set of physics problem solving skills with documented differences between experts and novices. These included the drawing and use of physical representations, writing general or specific initial equations, writing algebraic or numerical initial equations, insertion of numbers, word usage, and fractionation. The evolution of these skills was followed by examining written exam data from two quarters of physics. A small probe was made for indications of potential transfer by examining a final exam problem from an engineering statics course in which a small portion of the initial research sample was enrolled.

To gain additional information about students in the transitional stage, a series of interview tasks were designed. Two of the tasks, which were also administered to faculty for comparison purposes, involved evaluation of previously written solutions. The first of these provided the subjects with a solution written by a student in an introductory course and asked whether the solution was a good one or not. The second

asked the subjects to classify various problem solutions as written by either students or instructors. A third task for students was to solve two complex electricity and magnetism problems in a think-aloud interview. This task was designed to check the validity of the methods developed to analyze the written exam solutions.

This research was carried out in the context of a two-quarter introductory physics sequence for honors engineering freshmen. It was not the purpose of this study to evaluate the instruction of the course, but rather to describe the development of the students in this particular environment. The interviews and probe into the transfer issue took place the following quarter. Both qualitative and quantitative research methods were used to analyze the data.

Dedicated to my husband



## ACKNOWLEDGMENTS

As anyone who has undertaken a work of this magnitude knows, it cannot be completed without the support and assistance of many other people. While it is impossible to publicly acknowledge the help of everyone, there are many individuals without whom this effort would have turned out much differently.

I would like to first thank my parents, Eileen and Bill Andre, for their constant encouragement in anything I've undertaken, no matter how out of reach the goal may have seemed (or continued to be). My husband, Jeff, has been there throughout this process, listening to me ramble about developments in my work, asking appropriately skeptical questions, being supportive during momentary bad times, and helping with many daily tasks to enable me to complete this dissertation in a timely manner. His contributions to this effort have been invaluable. I could not have done this without his love and support.

I would like to thank the many people in the Physics Education Research group at The Ohio State University for their input along the way: Alan Van Heuvelen, Lei Bao, Gordon Aubrecht, Len Jossem, Bill Ploughe, Ken Wilson, Jim Stith, Constance Barsky, Seth Rosenberg, Leith Allen, Xueli Zou, David May, Keith Oliver, Dave Van Domelen, Tom Kassebaum, Yuh-Fen Lin and Dave Torick. I particularly need to acknowledge my advisor, Alan, who provided gentle guidance but gave me the freedom

to make and learn from my own mistakes. It has truly been a privilege to learn from such a wise and modest man. I am very grateful to Xueli for sharing some of her vast statistical knowledge, for her assistance in data collection, for her aid in interview design, and for being the trail blazer for the rest of us at Ohio State; thank you for setting such a high standard for the rest of us to attempt to measure up to. I thank Leith for the long and productive discussions which shaped my research, Keith for his brilliant flashes of insight, and David May for his assistance in interpreting and presenting my data. I also want to thank two visitors to the group during this last quarter of my graduate work: David Mills for his kind support and Eugenia Etkina, both for her assistance in keeping me focused on my research questions and for her help with validity issues.

I would be remiss in not thanking the many students and faculty at OSU who gave of their time and participated in the interview portion of this work. Additionally, I want to thank the people in the OSU physics department who agreed to be part of my expert problem solving sample. Also, I need to thank professors Frank DeLucia, Douglass Schumacher, George Staab, Alan Van Heuvelen, and Phil Wigen for graciously allowing me access to data from their classes. I would also like to thank the OSU Freshman Engineering Honors staff, particularly John Demel and Rick Freuler, for their support, friendship, and encouragement along the way.

Thanks to professors Lei Bao, Dick Furnstahl and Tom Humanic for graciously agreeing to serve on my dissertation committee.

Special thanks go to Tom Foster for his friendship, support, and advice throughout my years of graduate work. I am particularly grateful that the focus of his



dissertation ended up being different than mine. I also want to thank Michael Wittmann for suggesting I use the Diana technique in my interviews.

I need to thank many friends who helped in a variety of ways: Rebecca and Scott Butler, Matt Fulkerson, Jared Gump, Derek Seeber, and Julie and Jarvis Smallfield for providing breaks from physics with their delight in card playing; the 20something group at the St. Thomas More Newman Center for helping me grow spiritually while stretching my mind; Rich Andre, Kathy Perevosnik, Lisa Radke, and Sharon Sawl for the e-mails that brightened many days otherwise filled with data analysis; and Jared and Carolyn Sawl, Steven and Anthony Boyd, Heather, Emily, and Wes Utsler, Carson Richter, and Sara Stephens for helping me keep a perspective of what is really important.



## VITA

- October 4, 1971..... Born – Pittsburgh, Pennsylvania
- 1993.....B. S. Electrical Engineering and Applied Physics,  
Case Western Reserve University
- 1996.....M. S. Physics, Case Western Reserve University
- 1996 – present..... Graduate Research Trainee Fellow,  
The Ohio State University

## PUBLICATIONS

1. Andre, Kathleen M., “Cooperative Learning: An Inside Story,” *The Physics Teacher* 37 (6), 1999.
2. Van Heuvelen, A., and Andre, K. M., “Calculus-Based Physics and the Engineering ABET 2000 Criteria” *American Physical Society Forum on Education newsletter*, Spring/Summer 2000.

## FIELDS OF STUDY

Major Field: Physics

Studies in physics education research: Professor Alan Van Heuvelen



## TABLE OF CONTENTS

	<u>Page</u>
Abstract.....	ii
Dedication.....	iv
Acknowledgments.....	v
Vita.....	viii
List of Tables.....	xii
List of Figures.....	xiii
Chapters:	
1. Introduction.....	1
Context of the Research.....	1
Research Methods.....	3
Analysis Techniques.....	5
Overview of the Dissertation.....	7
Endnotes of Chapter 1.....	9
2. Review of the Literature.....	11
The Role of Problem Solving in Life and in Physics Instruction.....	11
The Role of Knowledge Structure in Problem Solving.....	15
Transfer.....	16
Assessing Problem Solving.....	18
Expert Problem Solving Strategies.....	22
Representational Tools in Problem Solving.....	26
Expert-Novice Comparisons.....	27
Novice-Novice Comparisons.....	33
Problem Solving Methods in the Physics Classroom.....	34
Gaps in the Literature.....	38
Endnotes of Chapter 2.....	41

3. The Diana Problem.....	46
Description of the Task.....	46
Analysis and Results.....	48
Summary.....	56
Endnotes for Chapter 3.....	58
4. The Card Sorting Task.....	59
Design and Description of the Task.....	59
Broad Analysis.....	62
Solution-by-Solution Analysis.....	76
Summary.....	82
Endnotes for Chapter 4.....	84
5. Trends From Exam Data.....	85
Context and Overview of the Trends Study.....	85
Diagram Usage.....	87
General Versus Specific Initial Equations.....	96
Algebraic Versus Numerical Initial Equations.....	99
Number Insertion.....	102
Use of Words.....	106
Fractionation.....	109
Summary.....	112
Endnotes for Chapter 5.....	114
6. Validity and Applicability of Experimental Techniques.....	115
Written Exam Analysis.....	115
Think-Aloud Interviews.....	122
Solution Evaluation Tasks.....	125
Summary.....	126
Endnotes for Chapter 6.....	128
7. Transfer.....	129
Context of This Probe.....	129
Analysis.....	130
Summary.....	134
Endnote for Chapter 7.....	135

8. Conclusions and Future Research.....	136
Conclusions.....	136
Future Research.....	139
Endnotes for Chapter 8.....	142
Appendices:	
A. Protocol and Representative Transcripts From the Diana Task.....	143
B. Protocol, Problem Solutions, and Representative Transcripts From the Card Sorting Task.....	149
C. Problems From Physics 131E, 132E, and Engineering Mechanics H210 Exams Used in The Study.....	176
D. Comparison of Researcher and External Evaluations.....	183
E. Think-Aloud Interview Protocol and Representative Transcripts.....	185
Bibliography.....	194





## LIST OF TABLES

<u>Table</u>		<u>Page</u>
4.1	Characteristics of the Solutions Selected for the Card Sorting Task.....	61
4.2	Criteria Groupings for Card Sorting Task.....	63
4.3	Broad Categories for Card Sorting Criteria.....	64
4.4	Top Five Decision Criteria Used by Each Subject (Including Ties) .....	71
4.5	Dominant Decision Making Categories for Novices and Experts.....	73
5.1	Rating Scale for General/Specific Initial Equations.....	96
5.2	Rating Scale for Algebraic/Numerical Initial Equations.....	100
5.3	Rating Scale for Number Insertion Analysis.....	103
6.1	Comparison of Expert/Novice Ratings on Spring Problem.....	118
6.2	Comparison of Expert Ratings on Two Different Problems.....	119
6.3	Intra-Exam Comparison Summary.....	121
D.1	Comparison of Researcher and External Evaluator Ratings.....	184



## LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
3.1	“Diana’s” Solutions to the Spring and Sliding Block Problems.....	47
3.2	Comparison of Expert-Novice Responses to Diana Task.....	51
4.1	Analysis of Student-Instructor Decision Making on Card Sorting Task.....	67
4.2	Focused Analysis of Student-Instructor Card Sorting Decisions.....	68
4.3	Card Sorting Summary for the Spring Problem, Separated by Novices and Experts.....	77
4.4	Card Sorting Summary for the Sliding Block Problem, Separated by Novices and Experts.....	80
5.1	Problem Solution Using a Work-Energy Bar Chart.....	88
5.2	Tracking the Number of Free-Body Diagrams During the Two Quarters.....	90
5.3	Stratified Development of Free-Body Diagram Usage.....	92
5.4	Development of Bar Chart Usage During the Two Quarters.....	93
5.5	Development in Quality of Free Body Diagrams.....	95
5.6	General/Specific Initial Equation Development Over Two Quarters ...	97
5.7	Stratified Analysis of General/Specific Equation Behavior.....	98
5.8	Algebraic/Numerical Equation Development for the Two Quarters....	100
5.9	Stratified Analysis of Algebraic/Numerical Initial Equations.....	101

5.10	Number Insertion Development Over Two Quarters.....	104
5.11	Stratified Analysis of Number Insertion.....	105
5.12	Word Usage Development Over Two Quarters.....	107
5.13	Stratified Analysis of Word Usage.....	108
5.14	Fractionation Development Over the Two Quarters.....	110
6.1	Evaluation Sheet for Written Exam Analysis .....	116
7.1	Algebraic/Numerical Comparison of Statics Sample to Large Sample.....	130
7.2	Number Insertion Comparison of Statics Sample to Large Sample.....	131
7.3	Algebraic/Numerical Equation Summary for Statics Sample.....	132
7.4	Number Insertion Summary for Statics Summary.....	133
B.1	Spring Solution O.....	150
B.2	Spring Solution P.....	150
B.3	Spring Solution Q.....	151
B.4	Spring Solution R.....	151
B.5	Spring Solution S.....	152
B.6	Spring Solution T.....	152
B.7	Spring Solution U.....	153
B.8	Spring Solution V.....	153
B.9	Spring Solution W.....	154
B.10	Spring Solution X.....	154
B.11	Spring Solution Y.....	155
B.12	Spring Solution Z.....	155

B.13	Sliding Block Solution O.....	156
B.14	Sliding Block Solution P.....	156
B.15	Sliding Block Solution Q.....	157
B.16	Sliding Block Solution R.....	157
B.17	Sliding Block Solution S.....	158
B.18	Sliding Block Solution T.....	158
B.19	Sliding Block Solution U.....	159
B.20	Sliding Block Solution V.....	160
B.21	Sliding Block Solution W.....	161
B.22	Sliding Block Solution X.....	161
B.23	Sliding Block Solution Y.....	162
B.24	Sliding Block Solution Z.....	162



# CHAPTER 1

## INTRODUCTION

### Context of the Research

Physics education research began attracting national attention about thirty years ago and has blossomed in the past decade. Originally focusing on identifying common student difficulties and developing techniques for overcoming them, [1, 2, 3, 4, 5, 6, 7] more recent studies now investigate such diverse areas as epistemology, [8, 9] cooperative learning, [10, 11] and workplace skills which can be developed (in addition to content knowledge) in the physics classroom. [10, 11, 12]

This last area is receiving more attention recently, in light of studies from various organizations, including one by the American Institute of Physics [13]. The AIP study surveyed former physics majors now in academia, industry, governmental jobs, and private industry, asking what types of skills they used most frequently in their jobs. Except for the roughly twelve percent in academia, physics content knowledge was ranked at or near the bottom of the list. The skills that were most useful to the majority of those surveyed were those such as interpersonal skills, communications skills, and real-world problem solving skills. Similar studies of broader cross-sections of the work force, such as those by the National Science Foundation [14] and the United States Department of Labor [15], reinforce these findings. More motivation to help develop these skills through physics comes from the recent revisions to the criteria used

by the Accreditation Board of Engineering and Technology when reviewing college engineering degree programs. [16] The list of core competencies barely mentions specific mathematical or scientific knowledge, instead concentrating on workplace skills. These changes in the educational goals for future engineers were sparked in part by feedback from their future employers in the workplace.

Perhaps the one workplace skill that fits the most naturally with physics instruction is problem solving. Most research in problem solving done prior to this time has focused on expert-novice comparisons, [17, 18] which will be described in detail in Chapter 2. There has been little exploration of how one transforms from novice problem solving approaches to more expert-like behavior. In an attempt to fill this gap, a study was designed to determine how this transition might occur. The primary research questions under investigation were:

1) How do the skills associated with effective problem solving develop in students enrolled in a particular physics sequence?

2) What differences are there in the ways experts and novices perceive problem solving? What characteristics do they feel make for a good written problem solution?

The first research question encompasses many more specific issues, including:

a) How does the use of physical representations in problem solving on exams evolve as students progress through the physics sequence?

b) Can the analysis of written exam data give researchers useful information about student problem solving methods?

To answer these questions, the problem solving skills of a group of students enrolled in a special two-quarter introductory physics sequence for freshmen honors



students in engineering were tracked over the course of the sequence. The primary data source was the written solutions to exams given in these physics courses. The solutions were analyzed by looking at how well certain key skills of physics problem solving were executed while working on that particular problem.

A portion of the class also came in for interviews in the quarter following the sequence; in these interviews, they solved two problems and completed two tasks involving their evaluation of other people's problem solutions. For comparison purposes, a group of university physics instructors participated in interviews focused on similar problem solution evaluation tasks.

A small transfer study was attempted via a third-quarter exam from an engineering statics course in which a small portion of the sample population was enrolled. One problem from the exam was reviewed for evidence of some skills developed during the first two quarters possibly transferring to the third-quarter class.

## **Research Methods**

Data was acquired using the following methods, common in physics education research:

1) Informal observations: Watching the behavior of students in classrooms, talking with students outside of the classroom, and evaluating written assignments are several methods of gaining information about how and what students learn. Many physics education research studies, including this one, have their origins here. Informal observations spark ideas about potential student difficulties, possible improvements in

instructional methods, and new techniques for researching issues related to both of these.

2) Individual student and faculty interviews: The interviews done as part of this study utilized two different techniques. The first portion of the student interviews was conducted in a think-aloud format. In a think-aloud interview, the interviewer gives the participant a task to accomplish and asks the participant to voice his or her thoughts out loud while working on it. The interviewer does not interrupt the thought process with any sort of follow-up questions, but may make notes of some questions to ask once the task is done. If the subject is silent for too long, the interviewer simply reminds them to keep talking. It is important to note that the participant is not asked to describe or explain anything. Think-aloud interviews are accepted as the way to get the most accurate portrait of what is actually going on in the participant's mind. This makes think-aloud interviews a natural tool for investigating problem solving processes. [19]

Other portions of the interviews were not specifically think-aloud. That is, the participants were asked questions which they answered more directly, and occasionally the interviewer asked for some clarification. By the same token, the interviewer is more likely to be asked for and give clarifying information in a more conventional interview setting. Often written materials are produced in these interviews, which also contain useful information.

Individual interviews are time-intensive to administer, transcribe, and analyze, but they provide the richest information of the experimental techniques discussed here. It is not practical to administer interviews to large numbers of subjects, but the depth of information gained via a small number of interviews is considerable.

3) Written exams: Much of the data used in this study comes from analysis of written exams. This technique consists of crafting a question or problem and including it on an exam or quiz in a class. The advantage of this technique is that it can be administered to entire classes, and so the sample sizes here are generally much larger than those of interviews, and so more appropriate for statistical analysis. However, the data gained this way is usually less rich than the interview data; there is a trade-off.

A common combination of techniques is to administer interviews to a relatively small number of students to probe particular aspects of physics conceptual understanding, design test questions and problems based upon the interviews which are then administered to a larger number of students, and then analyze the responses to the questions based upon information gained in the interviews. Often a second set of interviews based upon the exam questions will follow, to insure that responses on the exam are being properly interpreted.

Given the complexity of problem solving, using just one of these techniques would result in only seeing certain aspects of the process. To more fully understand student problem solving, all three of these techniques were employed in this study.

### **Analysis Techniques**

To analyze the data, a combination of qualitative and quantitative methods were used. Each set of techniques is more applicable in a particular domain, and since this research crosses domains, both groups are necessary.

Qualitative methods are best used to analyze descriptive, observational data which are not readily handled by quantitative statistical techniques. Since much of the

research done for this dissertation is based on behaviors and characteristics, most of the initial analysis utilized such techniques. Many qualitative studies are inductive, rather than deductive. That is, they do not always begin with a specific hypothesis to test, but build theories based upon patterns seen in the data which is collected. It is not the purpose of a purely qualitative study to generalize the results to a larger population, but to describe in depth the research sample. Therefore, it is not necessarily important to choose a completely random sample. [20]

In this work, once the raw data was processed qualitatively, much of it was converted to numerical data by counting or applying a numerical rubric. These numbers were analyzed using fairly common statistical methods, such as computing means and variances. In several cases, comparisons were done utilizing standard statistical tests. Which test was most appropriate depended upon the type of data being tested. In some parts of the study, two different samples were compared; for example, some interview questions were given to both instructors and students. In these situations, a Chi-squared independence test was applicable. This test applies when frequency data in distinct categories has been generated. Essentially, it compares the proportions of the populations exhibiting a certain behavior to see if there is a statistically significant difference between the two.

Most of the data in this work is ordinal, that is, generated from a ranking scale. An example of this is assigning a number to indicate the generality of an equation, with one being the most specific and five being the most general. To appropriately identify changes in behavior which are coded by ranked data, one must use the Wilcoxon matched-pairs signed-ranks test. To apply this test, the differences between the two

measurements for each individual in the sample are computed and ranked, regardless of the direction of the change. After the ranking is done, signs indicating the change direction are affixed to the rankings. These rankings are then used to generate a standard z score from which the statistical significance of any change can be determined. [21] In most research, a p-value of .05 or less is considered highly significant. In educational research, p-values up to .1 are often reported as fairly significant, so that is what this work will do.

### **Overview of the Dissertation**

This dissertation is laid out in two sections. The main body describes the research process; the appendices contain exam questions, interview protocols, and selected transcripts.

Chapter 2 further sets the context for this series of studies by giving a brief review of the relevant literature. Following that, Chapters 3 and 4 discuss two problem solving evaluation tasks which were part of student and faculty interviews. Chapter 5 describes a whole-class analysis focusing on the development of several problem solving skills during a two-quarter physics sequence. Chapter 6 contains information on the usefulness of the experimental techniques employed in this work for acquiring specific kinds of data and also addresses the reliability of the methods unique to this study. Chapter 7 discusses a brief probe into the transfer issue in the quarter following the physics sequence. The main body of the dissertation ends with Chapter 8, Conclusions and Future Research.

Information pertinent to the main body of the dissertation which the reader may want to refer to can be found in the appendices. Appendices A and B contain the protocols and some representative transcripts from the interviews described in Chapters 3 and 4. The exam problems used in the whole-class study and transfer probe are in Appendix C. A comparison of two evaluators' ratings of a sample of problems using the methods developed for this dissertation is in Appendix D. Finally, Appendix E consists of the protocol and representative transcripts from the think-aloud problem solving interviews.

## ENDNOTES OF CHAPTER 1

- [1] Ibrahim Abou Halloun and David Hestenes, "Common Sense Concepts About Motion," *American Journal of Physics* **53**, 1056-1065 (1985).
- [2] John K. Gilbert, Michael D. Watts, and Roger J. Osborne, "Students' Conceptions of Ideas in Mechanics," *Physics Education* **17**, 62-66 (1982).
- [3] Christina Karrquist and Bjorn Anderson, "How Swedish Pupils, Age 12-15, Understand Light and its Properties," in *Proceedings of the International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*, edited by Hugh Helm and Joseph D. Novak (Cornell University Press, Ithica, NY, 1983), 380-392.
- [4] David P. Maloney, "Charged Poles?" *Physics Education* **20**, 310-316 (1985).
- [5] David E. Trowbridge and Lillian C. McDermott, "Investigation of Student Understanding of the Concept of Velocity in One Dimension," *American Journal of Physics* **48**, 1020-1028 (1980).
- [6] Peter W. Hewson, "Diagnosis and Remediation of an Alternative Conception of Velocity Using a Microcomputer Program," *American Journal of Physics* **53**, 648-690 (1985).
- [7] Jim Minstrell, "Explaining the 'At Rest' Condition of an Object," *the physics Teacher* **20**, 10-14 (1982).
- [8] Andrew Elby, "Another Reason that Physics Students Learn by Rote," *Physics Education Research Supplement to American Journal of Physics* **67**, 7, S52-S57 (1999).
- [9] David Hammer, "Two Approaches to Learning Physics," *The Physics Teacher* **27**, 12, 664-670 (1989).
- [10] Patricia Heller, Ronald Keith, and Scott Anderson, "Teaching Problem Solving Through Cooperative Grouping. Part 1: Group Versus Individual Problem Solving," *American Journal of Physics* **60**, 7, 627-636 (1992).

- [11] Patricia Heller and Mark Hollabaugh, "Teaching Problem Solving Through Cooperative Grouping. Part 2: Designing Problems and Structuring Groups," *American Journal of Physics* **60**, 7, 637-644 (1992).
- [12] Thomas M. Foster, *The Development of Students' Problem-Solving Skills From Instruction Emphasizing Qualitative Problem-Solving*, unpublished doctoral dissertation (University of Minnesota, 2000).
- [13] D. Rosil, "What are Masters Doing?" (AIP Statistics Division, College Park, MD, 1996).
- [14] "Shaping the Future: New Expectations for Undergraduate Education in Science, Mathematics, Engineering, and Technology," (NSF Directorate for Education and Human Resources Review of Undergraduate Education, 1996).
- [15] A. P. Carneval, L. J. Gainer, and A. S. Meltzer, *Workplace Basics: The Essential Skills Employers Want* (Jossey-Bass Publishers, San Francisco, CA, 1990).
- [16] ABET criteria can be found at <http://www.abet.org/eac/2000.htm>.
- [17] Michelene T. H. Chi, Paul J. Feltovich, and Robert Glaser, "Categorization and Representation of Physics Problems by Experts and Novices" *Cognitive Science* **5**, 2, 121-152 (1981).
- [18] J. H. Larkin and H. A. Simon, "Why a Diagram is (Sometimes) Worth 10,000 Words," in *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, edited by Janice Glasgow, N. Hari Narayanan, and B. Chandrasekaran (AAAI Press/MIT Press, Menlo Park, CA, 1995).
- [19] K. Anders Ericsson and Herbert A. Simon, *Protocol Analysis* (Massachusetts Institute of Technology Press, Cambridge, MA, 1996).
- [20] Robert C. Bogdan and Sari Knopp Biklen, *Qualitative Research for Education* (Allyn and Bacon, Needham Heights, MA, 1998).
- [21] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill Book Company, New York, NY, 1956).



## CHAPTER 2

### REVIEW OF THE LITERATURE

#### **The Role of Problem Solving in Life and in Physics Instruction**

Problem solving is a fundamental part of every human's existence. Nearly each day from childhood on, everyone uses the complex cognitive skills involved in this process. [1] To a degree, most people solve problems as naturally as breathing. [2] There is an element of problem solving in almost every human activity, whether it is vocational or recreational. Rubenstein even asserts that problem solving is the area where all disciplines find the most common ground. [3]

Since problem solving pervades human existence, it is somewhat surprising that it did not become a topic of intense research until the mid 1950's. Prior to this time, several individuals investigated problem solving issues, but there were no unifying theories or foci. In the late 1950's, two major theories of problem solving emerged. The stimulus-response, or behaviorist, theory essentially viewed the solver as a "black box" and did not attempt to describe the processes involved in problem solving. The alternative, developed primarily by Newell, Simon, and Shaw, was an information processing theory, focusing more on what happens inside the box. Research guided by this theory did more than simply present subjects with problems and look at the final answer. Rather, it looked more at how individuals interpret problem situations and proceed to a solution. By the mid-1960's, the information processing approach had become the dominant one, and continues to be so today.

Much of this initial push in problem solving research occurred in psychology departments or in newly evolving artificial intelligence groups. [4] Most of the work centered on logic puzzles, rather than content-specific problems. It has been found since then that there are definite differences in the ways these two types of problems are approached, even by the same person. [1] It was not until the 1970's that physics began to be a field of much interest in content-specific problem solving. For people in many disciplines, physics became recognized as an excellent area in which to study problem solving, since its well-defined organization around a small number of basic principles hopefully makes it easier to observe more about problem solving than would be possible in other scientific domains. [5] Around this same time, as physics education research began to grow, physicists also became interested in problem solving issues. As Fuller tells it, "Before the early 1970's it was thought that physicists knew how to solve physics problems and they knew how to teach other people to solve physics problems." [6] Unfortunately, most of the research done since that time has shown that every discipline, including physics, has done a much poorer job of teaching problem solving to their students than was previously thought. Some researchers even feel that most students, after working some 3,000 homework problems and seeing countless others worked as examples for them in their four years of college, graduate and exhibit no noticeable change in their problem solving ability. [7, 8]

Significant progress has been made in the past thirty or so years in understanding problem solving, both in general and in physics. Extensive studies have investigated areas such as what differences exist between expert and novice physics problem solvers, how explicit instruction in problem solving techniques affects students' behavior, and what seem to be effective strategies for solving physics problems. The application of these results should have a significant impact on the quality of students graduating from our universities and entering the work force. In addition to this, research continues in

understanding the psychological factors involved in problem solving. This work is also important,

“because it provides us with a degree of flexibility which we might not otherwise have. If we can examine our own problem solving processes with some degree of understanding, then we have a better chance of improving them. Further, if we have some understanding of how people think, we can be more effective in helping others. Anyone who is to teach, or to tutor, or even help a child with homework can benefit....” [9]

What exactly is problem solving? Posing this question to twelve different people will almost certainly result in twelve different answers. One possible definition, based on work of Mayer, is “cognitive processing directed at transforming a given situation into a desired situation when no obvious method of solution is available.” [10] In language perhaps more familiar to physicists, a problem can be thought of as “a situation where a person desires to resolve a gap between a goal state and an initial state. Some blockage in the gap prevents the person from immediately seeing a course of action.” [11] In this way, problem solving can be thought of as applying a series of operators to a given state to convert it to a desired final state. [12] Selecting the appropriate operators is part of the solution process. [3]

It is important to distinguish between problem solving and the working of exercises. A problem must be a genuine challenge that cannot be executed automatically. [7] Referring to the above definition of a problem, if the blockage is not present, and the transformation from initial to final state is clear, the situation is not a problem, but merely an exercise. [11] Whether a situation is an exercise or a problem depends not only on the situation, but on the person to whom the situation is presented. One man’s exercise may be another man’s problem, and the degree of complexity in the problem will also vary with the background and perceptions of the solver. [1] An illustrative example of this is getting dressed before leaving home in the morning. As a young child, this is a very complex and difficult problem; not only does a toddler not have the motor skills necessary

to dress him or herself, but he or she also has not learned which clothing items match and what sorts of clothes are most appropriate for the outdoor climate and the trip's destination. For the most part, by the time adulthood is reached, one can get dressed without giving it much conscious thought. At some point it is no longer necessary to think about all the manipulations needed to tie shoelaces or fasten buttons.

Most instructors agree that it is important for students to learn problem solving skills as part of their college education. To achieve this in a science course, instruction needs to address both the content knowledge that students must acquire and the methods by which they can apply this knowledge to new situations. [13] The teaching of problem solving skills may be even more important than the teaching of factual knowledge when one considers that instructors have a very small chance of correctly predicting all of the content knowledge a student is going to need during his or her professional career. In many cases, excellent problem solving skills ought to be able to compensate for not knowing some specific content knowledge. [14] At this time, anticipated changes in the work environment indicate that problem solving skills, along with teamwork and communications skills, will be extremely important. Companies need employees that can acquire new technological skills as systems change, can learn needed information on their own, can set and monitor their own goals, and can make decisions on their own. [15, 16, 17, 18, 19] These skills are all essential in problem solving. Unfortunately, "even when the teaching of problem solving skills is deemed a major goal of instruction, the means for achieving this goal are [often] far from obvious." [12] As mentioned earlier, most instruction does not have an impact in this area. There are cases where students acquire large amounts of formal knowledge but are unable to apply it correctly to a slightly novel situation. Most efforts to improve this result have not succeeded. [20, 21] Strategies that are not effective (at least in isolation) include many homework problems, many worked examples in class, open-ended problems, and peer instruction in problem solving. [8]

## **The Role of Knowledge Structure in Problem Solving**

One area of research in problem solving has been how knowledge and knowledge structures impact problem solving. While most problem solving researchers believe there are general problem solving skills that are applicable in any domain, they may not do much good if one does not possess the relevant content knowledge. [22] This works both ways; problem solving skills without content knowledge will not be sufficient to solve the problem, but neither will content knowledge alone suffice to resolve the situation. [23] Focusing on content knowledge, the key issue is not only what is known, but how it will be communicated to the problem solving process. [22]

There are different types of knowledge stored in the brain. One classification scheme, used by Ferguson-Hessler and de Jong, breaks knowledge into four different types: situational, declarative, procedural, and strategic. Situational knowledge is information regarding how situations typically appear or are presented in a particular domain. Declarative knowledge is a collection of facts. Procedural knowledge is how to execute procedures. Strategic knowledge, which is more general than the other types, involves basic plans of action. [24] How these knowledge types inform each other is a large factor in a person's problem solving success. As most instructors can attest, many students possess a large amount of declarative knowledge which is easily regurgitated upon cue, but these same students are unable to access this knowledge when it is needed to solve a problem. [24] This is true even of students who have completed physics courses with high grades - they often make incomplete or incorrect descriptions of basic mechanics problems, which lead to wrong solutions. They can recite the facts, but lack the strategies for applying those facts in problem situations. [25]

Most knowledge is probably initially encoded as declarative. The student at this point concentrates on the facts and gives little thought as to how to use them. As the

declarative knowledge is applied repeatedly, it begins to not be recalled as explicitly, and over time, becomes part of procedural knowledge. [26] To return to the example of getting dressed, few adults when looking for a shirt to wear with some plaid shorts would explicitly think, “Stripes don’t go well with plaid.” They would simply not consider their striped shirts.

Most experts organize their knowledge hierarchically. That is, their knowledge is organized about general principles, with progressing levels of detail. A hierarchical knowledge structure leads to better knowledge recall, modification of existing knowledge, and correcting of errors. [12] When people with hierarchical knowledge structures were compared to those with single-level structures, it was found that the hierarchical structure resulted in better performance on complex tasks involving knowledge from several different areas. [27] It is believed by most researchers that knowledge hierarchies make for better problem solvers. [28]

Another theory of knowledge organization involves problem schemas. A problem schema contains all the knowledge necessary to solve a particular class of problem: procedural, declarative, and situational. Some work done by Ferguson-Hessler and de Jong found that students who organized their knowledge around problem types did better on problem solving exams. [29] The important commonality between these two different views of knowledge is that the *organization* of the knowledge is the important factor. The structure is more important than the amount of knowledge contained within the structure.

## **Transfer**

On the surface, different types of learning can look similar. One place where differences can be observed is a transfer situation. [30] Asking a student to solve a (genuine) problem is an opportunity to test transfer of skills, as well as the depth of his or her understanding. [10] In some sense, transfer is always occurring in learning, since

previous knowledge must be transferred to the current learning situation to make it meaningful. [30]

There are actually two types of transfer possible: positive and negative. When a strong positive transfer of a solution method occurs, it is called *Einstellung*, or the *Einstellung* effect. This is highly desirable. Related inappropriate, or negative, transfers cause many problems in the educational process. [4] It is thought that *Einstellung* results when the learner has a knowledge structure containing both general concepts and specific examples, so a sufficient content base is required before meaningful transfer is possible. [30]

Whether transfer to a more general situation is possible or not depends largely on how a person has identified the cues in the initial problem. Sometimes the strategies students develop for one problem are too specialized or too context dependent to lend themselves well to generalization. [2] If a procedure is learned by rote, it is more context dependent, but if it is learned with more attention to the meaning of what is being done, it will be more closely related to concepts, facilitating appropriate transfer. [32] In cases where procedures are more context than concept dependent, external prompts or cues can aid transfer. This is because the prompts help the solver access the relevant prior knowledge for transfer. [33]

Some studies have investigated transfer specifically in physics. Bassok compared transfer of skills from solving banking problems and from physics problem solving. She found that transfer to new situations occurred much more easily from the banking context than from physics. [34] In another study, she investigated the transfer of solution methods from algebra to physics problem solving, and vice versa. In the limited domain of her study, significantly more transfer occurred from algebra to physics than from physics to algebra. [33] Transfer from physics to algebra was facilitated when the variables involved in each domain more clearly “matched up” with each other. [34] Also,

as the amount of specific instruction in a physics topic increased, the spontaneous transfer to algebraic situations decreased. [33]

### **Assessing Problem Solving**

When studying problem solving, a key issue is assessment. One difficulty is that there is no generally accepted definition of problem solving that naturally leads to a reasonable measure of problem solving. [15] In assessing problem solving, there are several aspects to consider. Should the product or the process be tested? Should the assessment focus only on nonroutine problems, or should these be combined with some routine ones? Should problem solving skills be looked at in isolation or in more complex contexts? As stated previously, transfer is a critical component of any problem solving assessment, so when designing an assessment of problem solving, the researcher must have a specific transfer domain in mind. [10]

Standardized tests generally do not adequately measure problem solving ability. Typically the items on these tests are artificial. Even if the questions on the exams are genuine problems, simply looking at multiple-choice responses to them does not convey to the evaluator the thinking processes of the subject. [16] As a result, one-shot tests such as these may underestimate the amount of possible transfer. [30]

On a slightly different tact, problem solving-based evaluations tell teachers or researchers far more about learning situations for students than any standardized test. Not only do these evaluations aid the teacher in overall instructional development, but they suggest ways to teach specific individuals in a class. A standardized test can assess “where” a student is in the learning process, but a problem solving assessment yields information on “how” that student is learning and how he or she might best be taught. [2]

What is necessary to make a good problem solving assessment? First the material must be cognitively complex - it must involve more than simple declarative or procedural



knowledge. The material must be beyond the recognition or recall level, and ideally the solution will require multiple cognitive steps. The test must be constructed so that success cannot be attributed to unrelated factors. [10] To be a diagnostic assessment, the instrument must contain problems involving multiple skills in multiple contexts. [2] If the purpose of the assessment is to find out whether learning occurred while solving a problem, a new sort of performance should be observable which was not exhibited before. [31]

It is not recommended to use a puzzle-type test as the sole measure of problem solving skill, since these puzzles do not have a clear relation to a subject's previous learning. [22] Ideally, a test applies to the higher order, more general, problem solving skills, as well as specific content knowledge in the discipline of interest. [2] Tasks that have several possible solution strategies give more information than problems with only one acceptable solution method. [2] Additionally, the assessment must be structured so that cross-checking between items is possible. [2] Reasoning is more important than the specific final answer.

As with any test, validity and reliability are issues that must be addressed. If an instrument is valid, it is consistent with the skills being tested, consistent with any previous tests of those skills, and fits with theories about the material. There are many types of reliability to strive for, including internal - whether the test is consistent with itself, test-retest - whether the same test given to the same subject at two different times will yield the same results, and evaluator - whether the score of the test is independent of the evaluator. [35]

When assessing problem solving, it is also important to assess the subject's understanding. Chi suggests a number of ways to do this. Students can be asked to solve problems which are isomorphic, that is, they require the same basic manipulations to solve, but the contexts of the problems are different. Another method of investigation

is to look for transfers to novel situations, or “far field” transfers. Student understanding can also be probed by examining the explanations that students generate for themselves while studying or solving problems. These explanations give insight into what conditions cause students to take various actions, how students see the consequences of their actions, the relationships between the actions and the goals aimed for, and the relationship of the goals and actions to the governing principles of the discipline. [36]

Several people have done psychometric analyses of student problem solving. Essentially, a psychometric method determines whether the subject’s final answers to a set of problems is right or wrong, then compares the student’s performance to various norms and covariances calculated from a larger population. The results of this are used to assess the subject’s underlying problem solving traits. One major weakness with this approach is that no information is gained about the processes used in attempting the problem. This method of analysis is almost a relic from the days of the behaviorist theories of problem solving. The second major flaw is that it is based on the assumption that the underlying traits supposedly being assessed are related to whether a given final answer is correct or not. [37]

That said, here is an example of how a study based primarily on whether final answers were right or wrong gleaned much useful information. Wankat designed a study of an advanced mastery-based course in chemical engineering laboratory techniques. He chose the series of tests in one subject area for analysis. In a mastery-based course, the students are periodically tested in a given area until they master the material. He grouped students by the number of failures before passing the test, analyzed and classified their errors, and used these results to draw conclusions about both the exam problems and the students.

Not only did the error analysis help classify the exam problems as either difficult or easy, but he was also able to identify the characteristics that made a problem either

simple or difficult for the students. This resulted in a series of exams in the subsequent year that was of more uniform difficulty. Additionally, the student analysis produced some interesting results. He found that the initial failures looked similar for all students; it was not possible to predict based on the first exam failure how many more attempts a student would need before he or she mastered the material. Differences, however, were noticed on repeat exams. Students who failed one or two times before passing were similar to each other. Students who failed three or more times before passing (if they ever passed), were slower to learn from their errors, repeating the same mistakes from previous exams. The bulk of these students seemed unable to evaluate their answer. Other trends among the students in this classification included sloppier work, less logical solutions, more basic-level errors, repeated errors, and difficulties with transfer. Additionally, these students were not able to see the solution path while working the problem - they would often just try something to see if it would work. [38]

A more process-based assessment is the method used by Forehand. Subjects were presented with problems in which needed information was missing. They could ask the examiners for any additional information. Periodically throughout the solution process, subjects would be asked to give their current analysis of the problem, the list of additional information they needed, and their current solution strategy. In studying the subjects' analyses, the researchers looked to see whether one or many hypotheses were generated, as well as how strongly attached to any particular hypothesis the subject was. The desired information was analyzed by looking at the amount of information requested, as well as the appropriateness of that information. Strategies were classified as either open (leaving open alternative courses of action based upon any new information) or unidirectional (plans which would not be affected by the results of the information search). In general, as a solution progressed, less hypotheses were developed and strategies became more unidirectional. [37]

## **Expert Problem Solving Strategies**

What sorts of processes and strategies seem to be effective in problem solving?

Much of this research is based on observing or getting advice from problem solving experts. Most of the methods described are similar, including many of the same ideas in varying combinations. "There is no rigid separation of strategies into different stages of the problem-solving process." [39] The way in which one decides to describe the processes of problem solving depends on his or her basic view. Problem solving contains aspects which are cognitive (What the problem solver is doing can often only be inferred from observations.), process-based (It involves the manipulation and transformation of knowledge.), directed (The solver is working towards a goal.), and personal (The skills and knowledge the solver brings to the situation are a large factor.). [10]

One psychologically based version of problem solving, described by Rubenstein, consists of preparation - processing and organizing the given information, incubation - thinking about the ideas and possible solutions, inspiration - where the "correct" solution path is chosen, and verification. [3] This is a procedure that will apply only to genuine problems and not to exercises, since the solution paths to exercises are obvious and do not require the incubation and inspiration steps. [4]

The description of Mayer is more based on information processing. His procedure for expert problem solving is translation - interpreting the statement, integration - forming mental models and representations of the situation, solution planning, and solution execution. This description may be preferable, since it seems to better capture the mental struggles involved in solving problems and does not portray the process as a single flash of insight. [10] The method of Hayes is similar to this, adding a step of

consolidating gains, or learning from the experience. [9] Larkin's process of assembling information, planning, solving, and checking, is also quite similar. [40]

All of these approaches agree that the beginning of the process is understanding the problem. In the words of C. F. Kettinger, "A problem well stated is a problem half solved." The initial description generated by the solver is often the crucial factor determining the ease of the solution (or indeed whether the problem will be solved at all!) [25] An effective problem solver thoroughly explores the problem and related information, extracts what is relevant, makes that information personally relevant, and then organizes it into a coherent representation for use in solving the problem. [16] To do this, one must understand exactly what is given, what assumptions can be made, and what can be neglected. This is one of the areas where novices encounter difficulties, because they want to dive right in to the problem instead of thinking about these issues. [6] An expert problem solver typically does not construct an internal representation of the problem right away, but waits until he or she has read most or all of the problem statement. [41] Then a detailed internal representation leads to an external representation. For instance, in mechanics, a good problem description should identify the system of interest, describe the system, describe interactions, explain how the laws governing the interactions connect the system's objects, and explain how correctness can be checked. [25]

There are a variety of ways people may choose to search for information or solutions. One is trial and error, which can be executed either blindly or systematically. In a blind search, one does not keep track of which paths have been investigated. In a systematic search, one keeps records and explores only the unexplored. Searches by proximity can also be used. Proximity searches include means-end methods. In means-end, an initial state and goal state are identified. A difference between the two states is identified and reduced, then the two currently existing states are compared and the process

is repeated. A problem with this sort of search is that unnecessary detours down dead-end paths can occur, so it is not the most efficient. [9]

Search patterns that are more knowledge-based are advised. Pattern matching is one example. Of course, it is possible to match patterns which are not useful, so informed pattern matching is a more successful technique. Fractionation, or breaking the problem into smaller parts that are solvable, is a very successful solution method for many experts. [9] In general, when expert problem solvers are searching for possible solutions, they do so using a depth-first search in a restricted area of the problem space. [1] They apply their knowledge to identify which solutions are most likely to be appropriate and then investigate those in depth, ignoring (at least temporarily) methods they believe will not be as productive. Novices either do not possess that knowledge or do not apply it and search more randomly.

A further refinement in approach is a schema-based approach. A useful problem schema contains information based upon the application of a simple relation, the conditions of its applicability, how relevant quantities are related and can be found from each other, various representational tools, and how to interpret all the symbols in the relations. [12] Problem solving processes which are schema driven are more direct, because they limit the initial search area. As a person gains knowledge in an area, related problem solving will become more schema-driven. [39] This is especially a big advantage in physics, where selecting the proper principle(s) to apply seems to be much more difficult than actually executing the application of it. [42]

Another strategy often observed in experts is subgoaling - setting an intermediate goal and working to meet it along the way to the ultimate solution. One reason it is effective is because it reduces the search space. [1] This is related closely to decomposing a problem into smaller conceptual pieces to work through and then put back together. In order for one to be successful with this method, one needs to have a good decomposition

strategy, a good knowledge base of soluble problems, and a good organization of knowledge for retrieval. [12]

A final device that is useful for experts but not often observed in novices is the use of analogies. An analogy actually constitutes a change of representations, a creative means for changing some of the perceived constraints of a problem. In this sense, it is almost like taking a “side trip” along the journey to the desired final solution. [43] It may be for this reason that novices do not tend to use analogies much. Clement has developed the following classification scheme for analogies: transformation (modifying the original situation to obtain the analogous one, like stretching a spring into a rod), association (one is reminded of an analogous case, which is not a conscious transformation of the given situation), and from principles (one notices that a potentially pertinent concept also applies to another situation). [43] In order to successfully and confidently use an analogy, one must be able to generate it, establish confidence in the relation, understand the analogous case, and apply the results of reasoning in the analogous case to the given situation. In Clement’s study of ten expert physics problem solvers, he found that eight of them spontaneously generated analogies while solving a difficult problem. A total of 38 analogies were generated, 31 of which he judged to be part of a serious solution attempt. Over half of these were analogies by transformation, eight were by association, one was from general principles, and a few were not able to be classified. [43]

In addition, one process which distinguishes experts from novices is the effort to learn from the problem solving process. Some typical questions experts may ask themselves after working a problem include, “Should I have used a different representation?” “Why did I miss critical clues?” “What errors did I make?” “What important discoveries did I make?” and “What problems would be similar to this?” [9]

## Representational Tools in Problem Solving

As mentioned previously, experts rely heavily on representational tools in problem solving. There are two basic classifications of representations: internal and external. Internal representations are within the solver's mind, where external ones are either on paper or physical objects. Internal representations are not merely copies of external ones - each representational type codes the information in different ways. Some relevant pieces of information are more obvious in one representation than another. However, an external representation will give the solver no help at all if he or she does not have an internal representation of the situation. [9] As an example, consider a common external representation in physics, free-body diagrams. Unless the person using the diagram understands what the arrows and symbols represent, it will be meaningless to him or her. Indeed, it would be impossible to draw a free-body diagram for a situation without making an internal representation first. One must isolate the system, identify the forces acting on the system, and determine the directions in which those forces act.

External representations can be further classified as sentential, where each component of the diagram has a one-to-one correspondence with a sentence in the complete problem description, or diagrammatic, where the diagram preserves characteristics of the situation such as geometry, topology, or hierarchy. Diagrammatics are especially useful because they can reduce the time spent in a solution search. If diagrams are presented along with the problem statement, most of the solution work has already been done. [40] Diagrams are especially useful because they provide a summary of the main points of a problem situation while eliminating distracting details. Practiced problem solvers can string various sorts of diagrams together to reason qualitatively towards a solution, and may use these diagrams to later construct a mathematical representation of the situation. [28]



The choice of representation not only impacts the ease of a solution, but also what knowledge is gained during the solution process. However, it is not natural for most people to look for new representational tools or solution methods. In some work at the University of Massachusetts-Amherst, it was found that students make more progress in using some of the often less-emphasized representational tools, such as motion maps and graphs, than they do in using algebraic representations. This is in spite of the fact that students seem to initially be much more comfortable with equations! An implication for instruction is that teaching students to use multiple representations in problem solving should be a goal of physics classes. [21]

### **Expert-Novice Comparisons**

Much of the problem solving research done to date involves comparisons of expert and novice problem solvers. A few of the results from these studies have already been alluded to earlier in this chapter. Experts are usually faculty members or graduate students. Novices in these studies are typically students who are either taking or have just completed an introductory course. There are differences between novices and experts at every point in the solution process, although some similarities have also been observed.

Starting at the beginning of the process, experts and novices perceive problem solving differently. Where experts see problems as situations for analysis and reason, novices often perceive problems as pure recall tasks. [44] Better problem solvers know problems can be solved with care and persistence. Less experienced students do not realize the problem may be confusing at first and believe they will either know the answer right away or they will never know it. [45] Additionally, some novices think they must have all of the solution worked out internally before writing anything down. [46] When presented with a problem, there are perceptual differences - the two groups see the same stimuli differently, including text, slides, and videos. [30] Often novices will see what

they want to see in a problem situation, rather than what is actually there. [47] This may partially explain a difficulty observed over varied studies and topic areas, that the most typical trouble people have when solving problems is failing to use all the information available. [48] If information is not correctly interpreted by the solver, it will be difficult for him or her to use it.

Once the initial information has been processed, an expert will recognize a basic class into which the problem falls. [49] Novices also categorize problems, but the categories described by experts and novices differ greatly. An expert bases his or her classification on guiding scientific principles and conditions of applicability. This sort of classification is a “deep structure” categorization. Novices distinguish problem types based upon surface features, often describing the categories in terms of potential variables that need to be determined. In the major study revealing this difference, designed by Chi et al., experts and novices who had just completed a physics course were asked to sort cards with physics problems on them according to how they would solve them. The number of piles produced by each group was roughly the same, but the categorizations differed, reflecting the behavior described above. Experts also tended to sort more slowly than the novices. When each group was asked to repeat the sorting procedure, all sorted much more quickly than the initial sort, probably because they were looking for familiar cues in the problem statements. Novices seemed to cue on specific words in the problem, often describing surface features, like springs or inclined planes. Experts tended to cue more on second-order features like states or conditions described in the problem. [50]

Although some words are cues for both experts and novices, experts cue more on underlying knowledge than do novices. [50] In Simon’s think-aloud studies, he found that novices verbalize more separate steps than do experts. For example, an expert almost never explicitly states the formula he or she is going to use. Typically the processes of choosing a formula, putting it in the correct form, and substituting in numbers for the

variables is encompassed by one statement from an expert. A novice verbalizes each bit of this separately. [51] Larkin found additional evidence that expert knowledge is more “chunked” than novice knowledge. She analyzed the time between the verbalizing of ideas in think-aloud protocols. With experts, after a delay in speaking, the initial statement would be followed quickly by several others, indicating that these ideas were accessed together. By contrast, novices seemed to generate an idea, then almost randomly search for the next one. [52] Chunking of knowledge is essential to becoming a proficient problem solver in any area because the human brain is limited to recalling about seven unrelated elements. More information than this is usually required to solve a problem. [3] By chunking the information, it is possible to recall more. Even if each chunk only contains two elements, the amount of knowledge available for retrieval doubles.

Knowledge organization is related to the methods in which experts and novices attack problems. Experts allow the underlying concepts to guide their solution, where novices often rely on memorized algorithms. [53] Furthermore, where experts often spend a great deal of time reasoning qualitatively about the situation before even looking for an equation [50], most novices generate an equation immediately (which may not even be appropriate) and begin trying to calculate something (which may not even be useful in the solution). [52, 54] This particular novice behavior is not only observed in physics, but in small children solving simple word problems. [55] In some cases, novices may state several physics principles initially in their solution attempt, but often will not think about actually applying those principles until much later in the process. [5]

The type of overall approach also differs between the two groups. Experts tend to work forward and build their knowledge about the problem as they work through it. In contrast, novices more often work backwards, using a means-end approach. [5] Also, if one asks an expert about what he or she is planning to do next, the answer will usually

contain very explicit solution techniques, for example, “I’m going to apply conservation of momentum to the system to find the change in position of the system’s center of mass.” A novice’s response to the same question might be, “I’m going to find  $x$ ,” an answer which is not very specific and does not include any details about the process needed to accomplish the goal stated. [50]

In the quantitative portion of physics problem solving there are similarities and differences between experts and novices. Simon found that the two groups use the same basic procedure when performing calculations: find the equation, modify the equation, solve the equation. However, the equations chosen are different between the two groups. Where novices use many equations that are only applicable to specific cases (e.g. the range equation for projectiles), experts solve problems starting with more general equations. In this particular study, a book was available for reference. Experts chose not to use the book, but the novices often looked up or verified equations in it. [51] There is also experimental evidence that students need to “bind” values to variables when solving problems. In Larkin’s think-aloud transcripts, statements such as, “ $y$  equals two meters” occur much more frequently with novices than experts. [5]

When there are multiple solution methods that a subject considers, it can be either a hindrance or a help for either group. Generally, experts and novices consider many of the same options, but experts find the optimum path sooner. They then proceed until reaching a solution, and only consider alternative approaches afterward, as a method for checking answers. [1, 20] Novice approaches vary in whether they are too persistent or not persistent enough. Sometimes a novice runs into difficulties while following a valid path and quits, due to lack of confidence. [51] In other cases, novices fail to solve a problem because they persist with one unsuccessful approach and ignore other possibilities. [48]

Another difference between the two groups centers about the use of representational tools. In solving true physics problems, experts need qualitative representations to reason with and to aid in building correct mathematical representations. [28] Good examples of this are free-body diagrams, work-energy bar charts, and Feynman diagrams. Novices in science have a very difficult time constructing or employing valid scientific representations. [56] In fact, studies show that only between ten and twenty percent of introductory physics students use diagrams to solve problems on final exams. [28]

There are other differences that have been observed between experts and novices. Experts will take large problems and break them into smaller problems that are soluble. Poorer problem solvers tend not to make sub-problems. [45] Experts are more likely to use analogies in solutions than novices, and when novices generate analogies, they are often superficial and therefore not appropriate. [39] As an expert works through a problem, there is care to properly understand the facts, relationships, and problem statement. An expert typically re-reads the statement several times. Novices do not exhibit this same care for accuracy and are not particularly bothered if they do not entirely understand the problem statement. [45] One reason novices make more errors than experts is because experts are more likely to test their hypotheses in ways that expose invalid assumptions. Novices tend to look for additional information to verify a rule, rather than disprove it. [32] Psychological studies show it takes more evidence for a person to change an incorrect hypothesis than to verify a correct one, [48] so if one does not actively test hypotheses, it may be difficult to find errors. In addition, although good problem solvers may use intuition for guidance, they avoid guessing, working in careful steps when in an unfamiliar area. Poor problem solvers are more careless, jumping to conclusions and not checking their intuitive moves for correctness. [45] In the area of self-monitoring, experts are much more active, asking themselves questions, creating

mental pictures, and checking for consistency. Poorer students sit back and hope the solution will present itself to them instead of actively looking for it, [45] and they do not show signs of what Redish calls, “internalizing the other person,” as experts do. [57] Many times a student will reach an answer, but will not execute any of the checking procedures that are routine to most scientists, such as checking units, magnitude, or consistency with fundamental principles. [58]

The amount of time needed to reach a solution also varies between experts and novices. In general, experts need less time to complete problems. Although think-aloud protocols for experts from a study by Simon show that they speak more quickly than novices, their protocols contained significantly less words. This effect was heightened when the groups were given a relatively more complex problem. Part of this was because the novices did not recognize the need to involve inexplicit variables until ten minutes into the protocol. [51] However, it is not always true that an expert will need less time. Although experts recall relevant information quickly, they may take their time and be more methodical in other portions of the solution process. [30] Larkin even based one of her studies on a comparison between an expert and novice who took about the same amount of time to solve the given problems. [52]

In considering these expert-novice comparisons it is “unwise to *assume* [emphasis added] that the performance of experts is necessarily optimal.” [25] All that is known is that the experts are more successful than the novices. Also, being an expert problem solver does not necessarily mean one can teach problem solving to others. Additional evidence shows that experts in particular fields vary in their flexibility in situations outside their area of study. [30]

## Novice-Novice Comparisons

Studies have also compared students who performed at different levels in problem solving in the ways they studied and learned from example problems. The basic format of these studies was to observe students while studying the text (with embedded examples), assess their declarative knowledge, then to conduct think-aloud problem solving interviews. Based on the results of the interviews, students were categorized as good or poor problem solvers. One major result is that the differences in good and poor novice physics problem solving are not due to differing familiarity with the basic factual knowledge. [59]

When studying, the two groups of students were equally active, but the approaches were quite different. In studying and problem solving, the explanations of good performers tended to be more complete, more correct, and more related to the principles discussed in the text. [36] When solving problems, good students mentioned procedural knowledge roughly twice as much as the poor students. Poor students did nearly twice as much superficial processing as the better students, where the good students spent more of their time integrating, or giving structure to the new knowledge they were acquiring. As a result, good students studied in more depth, and the poor students took more things for granted. When exercises were embedded in the text, the good students worked significantly more of them than the poor students. [24]

In another study concentrating only on the studying of examples and how it affected problem solving, good students spent significantly more time and thought (as measured by protocol lines) when studying. This was because good students generated more of their own ideas when studying and spent more time thinking about any given idea, presumably because they wanted to understand it better. Protocols of example studying for good students contained nearly twice as many physics explanations than those of the poor students. Additionally, good students stated failure to comprehend

much more frequently than their poorer counterparts. Looking at where students had trouble understanding, the poor students almost always voiced poor comprehension at mathematical points, but when better students felt they were not understanding something, it took place about half the time at conceptual points. Also, the questions formed by the good students at these points were more specific than those of the poor students. [36]

When solving problems after studying and learning the material, both groups of students spent roughly equal amounts of time and thought when solving problems, but the way that time was spent differed greatly. Both good and poor students referred back to the examples, but there were major differences in the methods of reference. Poorer students referred back to the example more times overall. When they did, they did much more rereading than the better students, lots of copying and direct mapping of the solutions, and much less comparing or checking. The good students referred to the examples for specific bits of information and help, where the poor students were usually looking for a solution to copy. In general, the good students learned from studying the examples. The poor students did not get much benefit from studying the examples, and barely accessed the factual knowledge they had previously demonstrated they knew. [36]

In another study, Zajchowski analyzed think-aloud protocols of better and worse problem solvers in a physics class for what he calls “instances of intentionality,” or moving the solution process forward. The protocols of better novices contained more of these than the poorer novices, so it seems that good students have a clearer picture of where the solution process is taking them. [59]

### **Problem Solving Methods in the Physics Classroom**

Problem solving researchers to date have made some interesting recommendations to instructors. Piaget and some Russian psychologists found that thinking skills develop



in children because they have to give reasoning or to justify their actions to adults and peers. [48] There are a variety of learning theories, many of which can complement the others. The theory of practice simply states that once students have been presented with declarative knowledge, they must actively practice using it. A similar technique is elaboration, where students systematically and progressively practice using more complex versions of a skill. In categorization theory, one explicitly teaches the students an organization scheme. A mastery approach creates an order for teaching the skills and insures that the student has mastered the basic skills before moving on to more advanced ones. Teaching by induction is showing students examples and asking them to work on isomorphic problems. [60] Certainly all these techniques sound like they ought to be effective; if there is a particular combination of these approaches which especially promotes learning problem solving, it is still unidentified.

A few researchers have looked at the role of the textbook in problem solving. For some students (classified as the “good” ones in the work of Chi, Ferguson-Hessler, etc. described previously), the textbook helps them, but for the vast majority, “the textbook seems to have been more successful in teaching equations than in inducing a high level of physical intuition.” [51] This is probably largely due to the fact that examples in the texts rarely explain WHY actions are taken in solving problems, so the solution steps are not explicit enough for the student to learn from. [36]

Ward and Sweller have been researching whether examples worked for students can be a beneficial component of instruction in physics. They assert that properly structured worked examples can help focus the students’ attention to relevant points and help reduce cognitive load. They found that to be helpful, an example must be focused; a “split attention” example - one which causes the student to divide their attention between multiple sources of information (e. g. statements and equations) - may actually be detrimental. Also they found that properly structured worked examples help students do

better on similar exercises and assert that the students also have better success on transfer problems. [61]

The usefulness of repetition has also been investigated, with varying results. Sweller conducted a study where subjects solved six problems, then solved thirteen problems five times in a row, then solved another six problems one time each. The final six were matched to the first six. As the subjects worked through the repeated problems, the tendency to work backwards decreased and more forward working was observed. Additionally, there was more tendency to work forward on the final six problems than on the initial six. [62] Simon, however, in a study where he had novice subjects solve a series of twenty-five problems, noticed no significant evidence of any behavior changes. Two minor trends were noticed: one, subjects referred to the textbook slightly less as the session progressed, and two, they focussed on the equations slightly less toward the end of the session. [51]

There are also mixed opinions on whether teaching content via problem solving is a good instructional design or not. In some cases, certain types of problem solving seem to interfere with learning. For instance, trying to teach students to organize their physics knowledge in schemas using a problem solving method does not seem to be effective. [49] Sweller believes this is because students will largely use means-end analysis in a new situation, and that concentrating on this means-end approach prevents them from acquiring the key features of the schema which will be needed later. Since a means-end approach concentrates on differences, this prevents the students from seeing the similarities that are key in schema development. Also, the cognitive load during a means-end process is quite high, leaving less mental processing space available for the schema construction. [49, 62] However, there is contradictory evidence, from several sources, that problem based learning can lead to better problem solving in new situations. At medical schools it has been seen that students who are taught in a problem-based

environment are better at diagnosing patients and understanding their abilities than their counterparts who were taught in a lecture setting. [30]

Goal-free problems have been designed to steer students away from means-end problem solving. These problems present the student with a situation, but rather than giving a particular task, such as, “Find the velocity of the ball when it reaches the top of the ramp,” the students are told to find as many quantities as they can. When students solve these sorts of problems, they cannot use means-end methods, and it does appear that they notice the key characteristics of a problem more easily. [49] After exposure to goal-free problems, students tend to work forward more than backward, [62] make less mathematical errors, [49] and need less solution steps. [62]

Many problem solving researchers believe not only that good problem solving behavior should be modeled for students, but that the skills should be specifically taught. As Reif says, “General cognitive skills necessary for effective performance in a science can be taught and should be considered a proper subject for explicit instruction, just as much as the facts and principles ordinarily taught.” [54] In thinking about which skills to teach and how to teach them, the behavior of experts is a good starting point, but simply telling novices what experts do and expecting them to automatically act the same way is ineffective. [30] Reif did find that just mentioning certain cognitive skills as course goals could have a positive effect, but the skills must also be modeled and taught. [54] In some of his early work, Maier found that simply telling students to keep an open mind, as well as some other similar cognitive advice, resulted in more correct solutions. [48] How much time needs to be spent working on these skills in order to have a substantial effect is still not agreed upon, but in at least one circumstance, just spending an hour acquainting students with qualitative analysis and chunking theory led to a substantial improvement in their problem solving. [52] In a class where an explicit strategy was taught, it was observed that students used more diagrams, performed more intelligent algebra before

inserting numbers, exhibited more extensive and more successful planning, accessed more relevant principles, and ventured down less blind alleys. [52, 54] Similar results have been observed in teaching other problem solving strategies and knowledge organization techniques. [27, 39, 25]

Another aspect of problem solving instruction that has seen a small amount of research is evaluation and feedback. In a study of engineers conducted by Hyman, subjects were split into two groups to review solutions of a design problem. The first group was instructed to constructively evaluate the solutions of the others, where the second group evaluated critically. Then each subject produced a solution to the problem. The resulting final designs were more creative in the group that had evaluated constructively than the group that had been critical of each other's work. [48] This leads to some interesting implications for instruction, and further research is needed to determine what about the process was key in producing this creativity.

### **Gaps in the Literature**

A review of the literature reveals that much has been discovered about problem solving processes and skills utilized by expert problem solvers in physics. However, there are still gaps which must be explored.

One area of problem solving that has had very little research reported to date is the area of solving complex problems. A good working definition of a complex, or ill-defined, problem is one where any of the initial state, final state, or permissible operators are not well defined. [1] Some exploratory work by Simon, Chi, and others indicates that experts use the same basic skills to attack poorly defined problems as they do with simpler ones. [1, 63] To solve a complex problem, one needs to be able to interpret, organize, apply multiple criteria, deal with uncertainties, self-regulate, consider the worth of multiple solutions, and apply intense mental effort. [64] Although little is known at

this point about how to develop the skills needed for solving complex problems, [41] some research by Kohler with apes in the 1920's indicates that it is necessary to know how to solve the related subproblems before they can be put together in a more complex situation. [4]

Another area which is especially intriguing is the absence of much discussion of the process of becoming an expert problem solver. Virtually nothing is known at this point about any transitional stages between expert and novice. Although Chi's and Zajchowski's studies describe some differences between good and poor problem solvers, almost nothing is published regarding how individual problem solving skills develop in the transition from novice to expert. There is one brief mention of some "intermediate" students in Chi's card-sorting study, stating that there are students who use a hybrid categorization scheme. These students will identify the physics principle(s) involved, but also mention the surface features the novices use as the basis for their sorts. [50] It is known that children start with rather ineffective strategies in problem solving. As they mature, the strategies change and problem solving skills improve. Also, the types of mistakes they make change. [55] People's uncertainty about their ability to problem solve decreases as young adulthood is approached, but gradually increases again. [65]

It is agreed that expertise in any area can only develop if a person invests major amounts of time. This is even true for people who may possess natural talent in an area. The time investment alone is not the important factor, though; deliberate practice, including some sort of monitoring and feedback mechanism, is essential. [30] Still, "extensive practice and study is a necessary but not sufficient condition for becoming an expert." [36] It is true that experts will recall extensive practice, but the actual mechanism that helped them learn through this practice does not seem to be clear. No one seems able to describe it. [66]

All of the studies described here took place in artificial settings; none of them actually took place in classrooms. There are a few studies which have evaluated attempts to teach problem solving skills in introductory physics courses. [67, 68, 69] These have not been discussed in any detail here, because the focus of this dissertation is not to evaluate problem solving instruction.

Only one study prior to now has endeavored to track students throughout introductory physics to determine the impact of these courses on problem solving development, and this is the recent work of Foster. [69] As this is also the major focus of this dissertation, some fundamental differences between these two studies are described here. Foster's work centered about a comparison between two groups of students to evaluate a particular instructional method in problem solving. The work which will be described here focuses on the evolution of particular problem solving skills in a group of freshman and does not attempt to assess the methods by which they were taught. Another difference between the two studies is in the manner of assessment; while Foster evaluated problem solving development in several broad categories (general approach, specific application of physics, logical progression, and appropriate mathematics), this work focuses on more specific problem solving skills.

With the previous work related to this study described and the specific gaps which this dissertation attempts to fill identified, the context for this research has been established.

## ENDNOTES OF CHAPTER 2

- [1] Michelene T. H. Chi and Robert Glaser, "Problem Solving Ability" in *Human Abilities: An Information Processing Approach*, edited by Robert J. Sternberg (W. H. Freeman and Company, New York, NY, 1985).
- [2] Avish and Nancy Dworkin, *Problem Solving Assessment* (Academic Therapy Publications, Novato, CA, 1988).
- [3] Moshe F. Rubenstein, *Patterns of Problems Solving* (Prentice Hall, Incorporated, Englewood Cliffs, NJ, 1975).
- [4] Bert F. Green, Jr., "Current Trends in Problem Solving" in *Problem Solving: Research, Method, and Theory*, edited by Benjamin Kleinmuntz (Wiley and Sons, Incorporated, New York, NY, 1966).
- [5] Jill H. Larkin, John McDermott, Dortehea P. Simon, and Herbert A. Simon, "Models of Competence in Solving Problems," *Cognitive Science* 4, 4, 317-345 (1980).
- [6] Robert G. Fuller, "Solving Physics Problems - How Do We Do It?" *Physics Today* 35, 9, 43-47 (1982).
- [7] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* 12, 6, 446-448 (1983).
- [8] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* 23, 3, 157-158 (1994).
- [9] John R. Hayes, *The Complete Problem Solver* (Franklin Institute Press, Philadelphia, PA, 1981).
- [10] E. L. Baker and R. E. Mayer, "Computer-based Assessment of Problem Solving," *Computers in Human Behavior* 15, 3, 269-282 (1999).
- [11] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* 15, 2, 148-150 (1986).
- [12] Frederick Reif, "Teaching Problem Solving - A Scientific Approach," *The Physics Teacher* 19, 5, 310-316 (1981).
- [13] Frederick Reif, J. H. Larkin, and G. C. Brackett, "Teaching General Learning and Problem Solving Skills," *American Journal of Physics* 44, 3, 212-217 (1976).
- [14] Herbert A. Simon, "Problem Solving and Education" in *Problem Solving and Education: Issues in Teaching and Research*, edited by D. T. Tuma and F. Reif (Lawrence Erlbaum Associates, Hillsdale, NJ, 1980).

- [15] H. F. O'Neil, Jr., "Perspectives on Computer-Based Performance Assessment of Problem Solving," *Computers in Human Behavior* **15**, 3, 255-268 (1999).
- [16] J. Schacter, H. E. Herl, G. K. W. K. Chung, R. A. Dennis, and H. F. 15, Jr., "Computer-based Performance Assessments: A Solution to the Narrow Measurement and Reporting of Problem Solving," *Computers in Human Behavior* **15**, 3, 403-418 (1999).
- [17] The criteria can be found at <http://www.abet.org/eac/2000.htm>.
- [18] A. P. Carneval, L. J. Gainer, and A. S. Meltzer, "Workplace Basics: The Essential Skills Employers Want" (Jossey-Bass Publishers, San Francisco, CA, 1990).
- [19] D. Rosil, "What are Masters Doing?" (19 Statistics Division, College Park, MD, 1996).
- [20] Yuichiro Anzai and Tohru Yokoyama, "Internal Models in Physics Problem Solving," *Cognition and Instruction* **1**, 4, 397-450 (1984).
- [21] Robert J. Dufresne, William J. Gerace, and William J. Leonard, "Solving Physics Problems With Multiple Representations," *The Physics Teacher* **35**, 5, 270-275 (1997).
- [22] R. Bhaskar and H. A. Simon, "Problem Solving in Semantically Rich Domains: An Example From Engineering Thermodynamics," *Cognitive Science* **1**, 2, 193-215 (1977).
- [23] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* **26**, 1, 74-76 (1997).
- [24] Monica G. M. Ferguson-Hessler and Ton de Jong, "Studying Physics Texts: Differences in Study Processes Between Good and Poor Performers," *Cognition and Instruction* **7**, 1, 41-54 (1990).
- [25] Joan I. Heller and F. Reif, "Prescribing Effective Human Problem-Solving Processes: Problem Description in Physics," *Cognition and Instruction* **1**, 2, 177-216 (1984).
- [26] John R. Anderson, James G. Greeno, Paul J. Kline, and David M. Neves, "Acquisition of Problem Solving Skill" in *Cognitive Skills and Their Acquisition*, edited by John R. Anderson (Lawrence Erlbaum Associates, Incorporated, Hillsdale, NJ, 1981).
- [27] Bat-Sheva Eylon and F. Reif, "Effects of Knowledge Organization on Task Performance," *Cognition and Instruction* **1**, 1, 5-44 (1984).
- [28] Alan Van Heuvelen, "Learning to Think Like a Physicist: A Review of Research-Based Strategies," *American Journal of Physics* **59**, 10, 891-897 (1991).
- [29] Monica G. M. Ferguson-Hessler and Ton de Jong, "On the Quality of Knowledge in the Field of Electricity and Magnetism," *American Journal of Physics* **5**, 6, 492-497 (1987).
- [30] John D. Bransford, Ann L. Brown, and Rodney R. Cocking, editors, *How People Learn: Brain, Mind, Experience, and School* (National Academy of Science, 1999).



- [31] Robert M. Gagne, "Human Problem Solving: Internal and External Events" in *Problem Solving: Research, Method, and Theory*, edited by Benjamin Kleinmuntz (John Wiley and Sons, Inc, New York, NY, 1966).
- [32] Stephen K. Reed, *Word Problems: Research and Curriculum Reform* (Lawrence Erlbaum Associates, Mahwah, NJ, 1999).
- [33] Miriam Bassok and Keith J. Holyoak, "Interdomain Transfer Between Isomorphic Topics in Algebra and Physics," *Journal of Experimental Psychology: Learning, Memory, and Cognition* **15**, 1, 153-166 (1989).
- [34] Miriam Bassok, "Transfer of Domain-specific Problem-solving Protocols," *Journal of Experimental Psychology: Learning, Memory, and Cognition* **16**, 3, 522-533 (1990).
- [35] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* **24**, 3, 212-213 (1995).
- [36] Michelene T. H. Chi, Miriam Bassok, Matthew W. Lewis, Peter Reimann, and Robert Glaser, "Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems," *Cognitive Science* **13**, 2, 145-182 (1989).
- [37] Garlie A. Forehand, *Individual Differences in Problem Solving Processes of College Students* (Carnegie-Mellon University, Pittsburgh, PA, 1967).
- [38] Phillip C. Wankat, "Analysis of Student Mistakes, and Improvement of Problem Solving on McCabe-Thiele Binary Distillation Tests" in *Problem Solving*, edited by J. T. Sears, D. R. Woods, and R. D. Noble (American Institute of Chemical Engineers, New York, NY, 1983).
- [39] Mary L. Gick, "Problem-Solving Strategies," *Educational Psychologist* **21**, 1 & 2, 99-120 (1986).
- [40] J. H. Larkin and H. A. Simon, "Why a Diagram is (Sometimes) Worth 10,000 Words," in *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, edited by Janice Glasgow, N. Hari Narayanan, and B. Chandrasekaran (AAAI Press/MIT Press, Menlo Park, CA, 1995).
- [41] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* **17**, 4, 317-321 (1988).
- [42] Jill H. Larkin, "Skilled Problem Solving in Physics: A Hierarchical Planning Model," *Journal of Structural Learning* **6**, 3, 271-297 (1980).
- [43] John Clement, "Observed Methods for Generating Analogies in Scientific Problem Solving," *Cognitive Science* **12**, 4, 563-586 (1988).
- [44] Ron Good, "Scientific Problem Solving by Expert Systems," *Journal of Research in Science Teaching* **21**, 3, 331-340 (1984).
- [45] A. Whimbey and J. Lochhead, *Problem Solving and Comprehension - A Short Course in Analytical Reasoning* (Franklin Institute Press, Philadelphia, PA, 1980).

- [46] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* **14**, 3, 204-206 (1985).
- [47] James L. Adams, *Conceptual Blockbusting* (W. W. Norton and Company, New York, NY, 1979).
- [48] Ray Hyman and Barry Anderson, "Solving Problems" in *The R&D Game: Technical Men, Technical Managers, and Research Productivity*, edited by David Allison (MIT Press, Cambridge, MA, 1969).
- [49] John Sweller, "Cognitive Load During Problem Solving: Effects on Learning," *Cognitive Science* **12**, 2, 257-285 (1988).
- [50] Michelene T. H. Chi, Paul J. Feltovich, and Robert Glaser, "Categorization and Representation of Physics Problems by Experts and Novices" *Cognitive Science* **5**, 2, 121-152 (1981).
- [51] Dortha P. Simon and Herbert A. Simon, "Individual Differences in Solving Physics Problems" in *Children's Thinking: What Develops?* edited by Robert S. Siegler (Lawrence Erlbaum Associates, Hillsdale, NJ, 1978).
- [52] Jill H. Larkin, "Processing Information for Effective Problem Solving," *Engineering Education* **70**, 3, 285-288 (1979).
- [53] William C. Robertson, "Detection of Cognitive Structure With Protocol Data: Predicting Performance on Physics Transfer Problems," *Cognitive Science* **14**, 2, 253-280 (1990).
- [54] F. Reif, J. H. Larkin, and G. C. Brackett, "Teaching General Learning and Problem Solving Skills," *American Journal of Physics* **44**, 3, 212-217 (1976).
- [55] Diane J. Briars and Jill H. Larkin, "An Integrated Model of Skill in Solving Elementary Word Problems" *Cognition and Instruction* **1**, 3, 245-296 (1984).
- [56] Jill H. Larkin, "Understanding Problem Representations and Skill in Physics," in *Thinking and Learning Skills, Vol. 2*, edited by Susan F. Chipman, Judith W. Segal, and Robert Glaser (Lawrence Erlbaum Associates, Incorporated, Hillsdale, NJ, 1985).
- [57] E. F. Redish, personal communication.
- [58] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* **14**, 2, 134-136 (1985).
- [59] Richard Zajchowski and Jack Martin, "Differences in the Problem Solving of Stronger and Weaker Novices in Physics: Knowledge, Strategies, or Knowledge Structure?" *Journal of Research in Science Teaching* **30**, 5, 459-470 (1993).
- [60] J. L. Wheeler and J. W. Regian, "The Use of a Cognitive Tutoring System in the Improvement of the Abstract Reasoning Component of Word Problem Solving," *Computers in Human Behavior* **15**, 3, 243-254 (1999).

- [61] Mark Ward and John Sweller, "Structuring Effective Worked Examples," *Cognition and Instruction* 7, 1, 1-39 (1990).
- [62] John Sweller, Robert F. Mawer, and Mark R. Ward, "Development of Expertise in Mathematical Problem Solving," *Journal of Experimental Psychology: General* 112, 4, 639-661 (1983).
- [63] Norman Fredericksen, "Implications of Cognitive Theory for Instruction in Problem Solving," *Review of Educational Research* 54, 3, 363-407 (1984).
- [64] Donald R. Woods, "PS Corner," *Journal of College Science Teaching* 18, 5, 338-340 (1989).
- [65] H. J. A. Rimoldi, Hermelinda M. Fogliatto, James B. Erdmann, and Michael B. Donnelly, *Problem Solving in High School and College Students* (Loyola University Press, Chicago, IL 1964).
- [66] Jill H. Larkin, "Enriching Formal Knowledge: A Model for Learning to Solve Textbook Physics Problems" in *Cognitive Skills and Their Acquisition*, edited by John R. Anderson (Lawrence Erlbaum Associates, Incorporated, Hillsdale, NJ, 1981).
- [67] J. H. P. van Weeren, F. F. M. de Mul, M. J. Peters, H. Kramers-Pals, H. J. Roosink, "Teaching Problem-solving in Physics: A Course in Electromagnetism," *American Journal of Physics* 50, 8, 725-732 (1982).
- [68] Patricia Heller, Ronald Keith, and Scott Anderson, "Teaching Problem Solving Through Cooperative Grouping. Part 1: Group Versus Individual Problem Solving," *American Journal of Physics* 60, 7, 627-636 (1992).
- [69] Thomas M. Foster, *The Development of Students' Problem-Solving Skills From Instruction Emphasizing Qualitative Problem-Solving*, unpublished doctoral dissertation (University of Minnesota, 2000).



## CHAPTER 3

### THE DIANA PROBLEM

#### **Description of the Task**

As described in the literature review, expert-novice comparisons have been a valuable tool in uncovering methodological differences between groups with differing experience in problem solving. [1, 2, 3] Two expert-novice comparison activities were designed as part of this research; this chapter describes the first of these tasks, called the Diana problem. The related portion of the interview protocol, as well as some representative transcripts, are in Appendix A.

The concept for this task is based upon Elby's work. [4] In studying student epistemology, he created a fictitious student, Diana. She was described to the interview subject as being "just like you," but was taking physics with the main objective of learning the subject more deeply. In the epistemology interviews, subjects were then asked questions about how Diana should divide her study time among various options. In the problem solving interviews described here, students were told to imagine that they were in class with Diana, who has just worked a problem solution, which the interviewer hands to them. They are told that Diana shows them the solution and asks whether they think it is a good solution or not.

In the instructor interviews, the subjects were told that Diana was a student in their introductory physics course who was taking the course pass/fail with the primary objective of learning physics more deeply. Diana has worked out the solution and brings it to the instructor and asks if the instructor thinks it is a good solution or not. The interview subject (whether instructor or student) was asked to tell the interviewer how he or she would answer Diana.

Diana's solution was a copy of an actual solution written on a final exam by a 131E student. Since students from both sections of 131E were being interviewed, two Diana solutions were chosen, one from each class. These are shown below in Figure 3.1.

$\frac{1}{2}kx^2 = \frac{1}{2}mv^2$ $v = 8.95 \text{ m/s}$ $mv = (M+m)v$ $v = \frac{2 \cdot 8.95}{.1}$ $v = 179 \text{ m/s}$ <p style="text-align: center;">(A)</p>	$P_a + P_w = P_{sw}$ $(.05kg)(150m/s) + 0 = (1.50)v$ $v = 5 \text{ m/s}$ $\frac{1}{2}mv^2 = \Delta U_{\text{int}} \quad \Delta U_{\text{int}} = F_R d$ $\frac{1}{2}(1.5)(5)^2 = (15)(25)d$ $\frac{18.75}{5m} = 3.75d$ $\underline{\underline{5m = d}}$ <p style="text-align: center;">(B)</p>
---	---

Figure 3.1: "Diana's" solutions to the spring problem (A) and the sliding block problem (B)

Subjects were given the solution to the problem from the section in which they had not been enrolled. There were two main reasons for this. First, the subjects would then see a problem which they had not worked previously. Second, since the card sorting portion of the interview to follow would also involve student-generated solutions to the same problem, there would be no chance that a subject would see his or her own solution to the problem, which would confound the results.

These particular solutions were selected for this exercise for several reasons. First, these solutions are somewhat sparse, making some of the solution steps unclear. Second, they obtain the correct numerical answer; it was thought that having Diana's solution actually be wrong might cause the subjects to focus more on that aspect of her solution than they would otherwise.

### **Analysis and Results**

A total of 38 student interviews were conducted. However, the thirteen instructor subjects were only asked to evaluate solution A, which involved the spring. Therefore, for the purposes of the expert-novice comparisons on the Diana task, only the student subjects who also evaluated this same solution were compared to the instructors. The original students viewing solution A numbered twenty, but the tape from one of these interviews was inaudible in places, so that subject had to be discarded from the student sample for this part of the study, bringing the student sample for the Diana exercise to nineteen.

To analyze the results, the interviews were recorded and transcribed. The transcripts were first analyzed qualitatively to find trends in the responses. The observations, comments, and behaviors of the subjects fell into the categories listed and described below. Note that some categories are fairly broad, while others are more specific. This is because both general and specific comments were made.

- **Ambiguous notation** – The subject commented on the fact that Diana’s notation was not accurate enough, usually that her v’s and m’s were not distinguishable from each other.
- **Ask Diana** – The subject phrased a question for Diana about her solution.
- **Checking** – The subject made a comment on how Diana did not show any evidence of checking her work.
- **Fractionation** – The subject commented on breaking the problem into parts to solve.
- **Hard to follow** – The subject said something indicating Diana’s solution was difficult to understand, unclear, or confusing. Any questions the subject asked about what Diana was doing, for example, “Was she trying to conserve energy?” fit into this category.
- **Method** – Subject talked about the methods Diana used to solve the problem.
- **No words** – Subject commented on the lack of words in Diana’s solution.
- **Missing work** – Subject stated that Diana had not shown very much work on the paper or that she needed to show more work.



- **Number insertion** – Subject spoke about when Diana inserted numbers in her solution.
- **Numbers right** – Subject evaluated whether Diana’s numerical work was correct.
- **Right/wrong** – Subject made a statement about whether or not Diana’s final answer solution was correct.
- **Sloppy** – Subject said Diana’s work was sloppy.
- **Two answers** - Subject commented that it appeared as if Diana had calculated two answers.
- **Units** – Subject talked about Diana’s use of units in her solution.
- **What principles** – Subject wanted Diana to indicate what principles she was using in her solution.
- **Looks good** – Subject said after some analysis that Diana’s solution looked good to him or her.

Additionally, there were observed behaviors not directly related to Diana’s solution which were analyzed. These are listed below.

- **Am I doing right?** – The subject asked a question at some point in the interview asking the interviewer if he or she was doing what the interviewer wanted him or her to do.
- **Worked problem** – Subject worked out a solution of his or her own as part of evaluating Diana’s.

The data analysis began by tallying for each interview the number of times an event occurred which fit into one of the categories described above. Then the percentage of subjects in each sample who exhibited that particular behavior was computed for each category and compared. The results of this analysis are summarized in Figure 3.2, shown below.

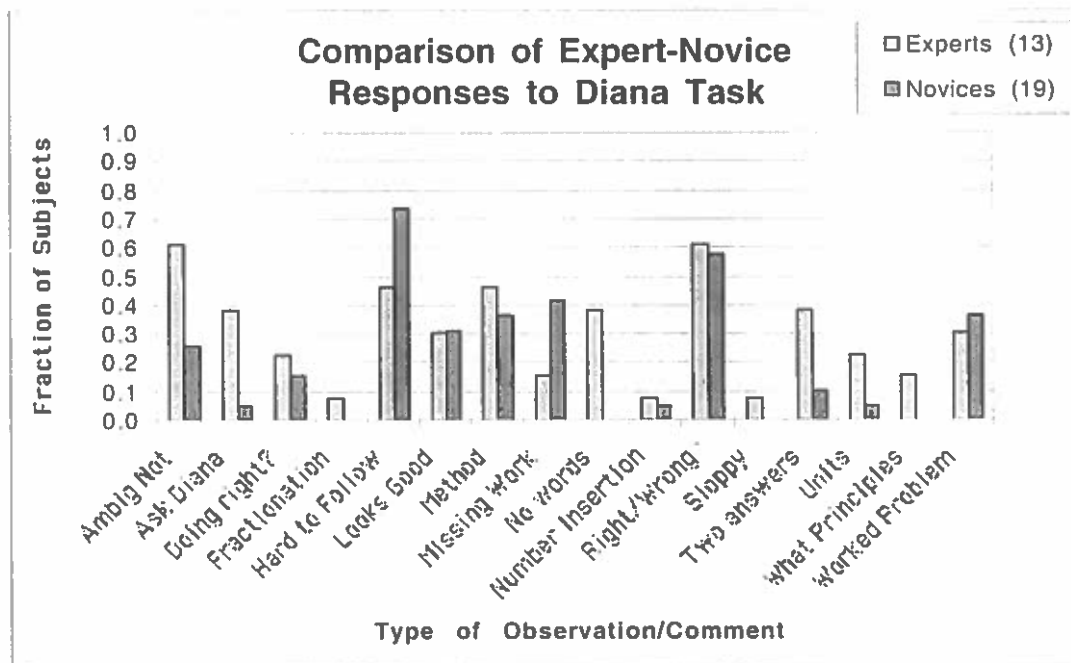


Figure 3.2: Comparison of Expert-Novice Responses to the Diana Task

The most significant difference was in the “Ask Diana” category ( $p < .025$ ). Only one student talked about asking Diana any questions, but 38 percent of the instructors had specific questions they wanted to ask Diana about her solution. One of the expert

subjects even asked the interviewer to play the part of Diana, so that he could have a conversation like one he would have with a student. Most of the questions were about what Diana was thinking, indicating that these expert subjects were aware that problem solving is a process which is not completely captured by what is written on paper. None of the students explicitly displayed this awareness.

The category of “Ambiguous Notation” also contained a significant difference, at the  $p < .05$  level. Sixty-two percent of the expert subjects, compared to 26 percent of the novice subjects, commented that Diana did not distinguish well between the various masses and velocities in her solution, leading to some confusing equations.

Instructors were also more likely to comment on how Diana’s solution appeared to have two answers ( $p < .1$ ) and how she did not identify the physical principles she was applying ( $p < .1$ ). Although not statistically significant, students stated that Diana’s solution was hard to follow and missing work more often than instructors did. Additionally, nearly 40 percent of the instructor sample commented on the lack of words in Diana’s solution, but no students did.

The differences outlined above show that the students tended to speak in generalities about Diana’s solution. On the other hand, the instructors, although sometimes making general comments, would elaborate on the generalities and point out specific areas of Diana’s solution that were missing information or aspects that made it difficult to follow. This should not be surprising, given the different levels of problem solving experience each group brought to the interview.

Before the study, it was hypothesized that novices would focus more on whether the solution was correct or not and base their answer to Diana’s question primarily on

the correctness. Referring back to Figure 3.2 shows this was not the case; about 60 percent of each population commented on whether Diana's solution was right or wrong. Also, an equal portion of each group, roughly thirty percent, said at the end of their analysis that the solution "looked good" to them. The two populations also came out fairly equal in the percentage of subjects who worked out a solution of his or her own to the problem, with just six percent more of the student sample exhibiting this behavior.

Although the raw numbers make it appear that the populations behaved similarly in these two respects, there actually were fundamental differences in the behaviors related to these measurements. Most of the instructors who said Diana's solution looked fine to them were those who had little difficulty figuring out how her solution worked. They made the statement that the solution was a good one after just a few other comments on it. It should also be noted that three instructors stated directly that they did NOT think Diana's solution was good, one even saying that he would not call it a solution. By contrast, no students directly stated that the solution was poor. Some of the students who said the solution was good or right did so after a long struggle of trying to make sense of the solution, as the following excerpt from Bob's interview shows:

Bob: OK. So she set the bullet's  $v$  equal to the – the bullet's kinetic energy equal to the spring energy. There's no friction or anything. I don't see what the kinetic energy would be. OK, um,  $x$  is .4,  $m_1$  is 1.9. (calculator, pause) Why does she have  $v = 4.9$  m/s? Then she just solved this for  $v$ . (pause, calculator) I'd say the answer's, it's not right. I think, uh... (pause) Oh, OK. Momentum conserved. Oh, they have the wrong mass up there. We have 1000 times  $.4^2$ , divide by a mass of 1.9 and square root. (pause) The energy of the bullet has to be the same as the spring energy... which you have at the end. (pause) I don't know what the 8.95 is. Oh, OK, I get it. She did the momentum for the, the block. (calculator) Entering the mass of the block. There's her 8.95. It's just the momentum of the block after. (pause, calculator) That's, that works, but it didn't work my way. (calculator,

pause) I guess her way is right, but it doesn't work my way. It should, because the energy of the bullet should be equal to the energy when the spring is compressed. I got a different answer that way. (pause, staring at page) Oh, well. Her answer's right.

This behavior is indicative of the low abilities of students when it comes to reading mathematical "language" with understanding. Instructors have learned this language and do not have the same difficulties.

Although a similar percentage of each sample wrote a solution of his or her own, the way in which this solution was written and utilized differed greatly between the two groups. Most instructors who wrote one decided to do so after looking at Diana's solution for a short while and worked the solution quickly. Two of the subjects then used the self-generated solution to illustrate points about what was lacking in Diana's. Most of the students working solutions worked theirs alongside Diana's, comparing their work to hers frequently along the way. Three of the students arrived at an incorrect answer (by forgetting to conserve momentum) and decided that *Diana's* solution was incorrect. By contrast, the two instructors who initially made errors in their solutions corrected them.

It was possible to do some further comparisons of the novices to themselves at a prior time in their physics development. All the student interview subjects had solved the similar problem shown in Figure 3.1B approximately fifteen weeks earlier. The observations each subject made about Diana's solution to problem A were compared to his or her solution of problem B. One of the results of this comparison is especially striking: any student who singled out particular aspects of Diana's solution to comment on exhibited favorable versions of those behaviors in his or her own solution from the

first-quarter final. For example, four students expressed that it was difficult to follow Diana's solution because she did not show on paper where she was inserting numbers into her equations; all four of those students did an excellent job in their solutions of showing exactly which numbers were being substituted for which variables. This same pattern was observed for the three students who commented on lack of subscripts. One student said that Diana did not do a good job of displaying the equations she was working with; his solution clearly showed the equations he used to solve the problem. Also, the six students who made a statement to the effect that Diana had left work out of the solution had all written solutions containing few, if any, omissions. Ten of the nineteen novice subjects who viewed Diana's solution to the block and spring problem focused entirely on the method Diana used. However, many of these students had written solutions on their final exams which were quite easy to follow. The predictive power of the Diana exercise is that it can reveal some aspects of good problem solving which the subjects have been utilizing for a while. The areas of Diana's problem solving which were described as poor by the students were the ones which those students had been using well for over a quarter. However, failing to mention deficiencies in Diana's problem solving process told the researcher nothing about whether the subject had those same deficiencies or not. Also, it was not possible to predict a subject's responses to the Diana question by looking at his or her solution to the similar problem from the mechanics final exam.

## Summary

To summarize the results of this particular expert-novice comparison, when students and instructors were asked to evaluate a student solution which was correct but had steps missing, differences and similarities were observed between the two groups. The students were more likely to make general comments that the solution needed to have more work shown or was difficult to follow. Instructors' observations were more specific, making significantly more comments about the solution having ambiguous notation, an apparent two answers, and no identification of the physics principles involved. This difference indicates that although the students had developed some useful problem solving skills, they had not yet acquired the skills necessary to read mathematics with proper understanding. Keeping in mind that these students were honors students in a class with some emphasis on interpreting the mathematical language of physics, it can only be assumed that this difference would be even more pronounced for a more general class. Instructors were also far more likely to tell the interviewer specific questions they would ask Diana.

Similar percentages of the student and instructor samples commented on whether the solution arrived at the correct answer or not, and similar percentages felt that the solution was good. However, instructors who stated that the solution was a good one did so after a very short time; many of the students who liked the solution stated so after struggling a long time to comprehend it. Once again, this indicates a fundamental student difficulty in reading the symbolic representations of physics. Also, similar percentages of each sample wrote solutions of their own during the evaluation process, but the manner in which they were used differed between the populations.

Instructors wrote the solution to either familiarize themselves with the correct method for solving the problem or to illustrate particular characteristics of a solution, and, for the most part, did not refer to Diana's solution while writing theirs; students often referred to Diana's work and used their solutions to make sense of Diana's.

Additionally, the student responses were compared to their solution of a similar problem they had solved over a quarter earlier. Almost without exception, when a student criticized a specific aspect of Diana's solution, it was an area where they themselves had performed superbly in their solution of the similar problem. The Diana exercise not only allows for some new comparisons between expert and novice physics problem solvers in the area of evaluating a solution, but, in the case of novices, tells the researcher something about an individual's problem solving skills which are firmly in place.



### ENDNOTES FOR CHAPTER 3

[1] Michelene T. H. Chi, Paul J. Feltovich, and Robert Glaser, "Categorization and Representation of Physics Problems by Experts and Novices" *Cognitive Science* 5, 2, 121-152 (1981).

[2] Jill H. Larkin, "Processing Information for Effective Problem Solving," *Engineering Education* 70, 3, 285-288 (1979).

[3] Herbert A. Simon, "Problem Solving and Education" in *Problem Solving and Education: Issues in Teaching and Research*, edited by D. T. Tuma and F. Reif (Lawrence Erlbaum Associates, Hillsdale, NJ, 1980).

[4] Andrew Elby, "Another Reason that Physics Students Learn by Rote," *Physics Education Research Supplement to American Journal of Physics* 67, 7, S52-S57 (1999).



## CHAPTER 4

### THE CARD-SORTING TASK

#### **Design and Description of the Task**

A second problem solution evaluation task, inspired by Chi's card sorting study, was designed and included in the third quarter interviews. [1] It was administered to both the student and instructor samples after completion of the Diana task. The subject was told that Diana's solution was for a 131E final exam problem. Then the interviewer presented a stack of twelve solutions to the exact same problem, stating that some of the solutions were written by students in the class, while others were written by instructors from the physics department. The instructors had been given a small booklet of problems and asked to solve them as if they were taking a test. The subject's job was to go through the stack and determine which solutions were by students and which were by instructors, explaining his or her reasoning. No information was given about the distribution of instructor and student solutions in the stack. Appendix B contains the protocol for this portion of the interview, the problem solutions, and some representative transcripts.

This task, as well as the Diana task, was designed to give some insight into how novices and experts perceive problem solving. The instructor responses would provide information on what physics instructors thought made for good problem solving, while the student responses would reveal their perceptions of what a good problem solution

looked like. Differences between the two groups would indicate instructional aspects of problem solving in need of development. Also, keeping in mind that evaluation is a high-level skill in Bloom's taxonomy of educational development [2], similarities between certain students and the instructors might indicate that these students were becoming more expert.

The solutions to sort were carefully selected to contain many different potential cues, based primarily on the research discussed in Chapter 2. Hypothesized triggers included diagram usage, checking of the answer, correctness of the answer, listing of given information, number insertion, neatness of the solution, clarity of fractionation, logical progression of the solution, written explanations in the solution, method of solution, and omissions. A summary of these characteristics for each solution is given in Table 4.1, on the next page. A check mark indicates that the solution possesses the characteristic listed at the top of that column

Solution	Energy Method	Diagram/Chart	Checking	Wrong Answer	Givens	Numbers Late	Messy	Clearly Fractionated	Words	Omissions
O (instructor)	√		√			√		√	√	
P (student)	√									√
Q (student)	√				√	√		√		
R (student)	√			√						
S (instructor)	√	√				√		√	√	
T (student)	√	√			√	√				√
U (student)	√			√						√
V (instructor)	√	√	√			√		√	√	
W (instructor)	√				√	√				
X (student)	√				√			√	√	
Y (student)	√	√				√	√	√		√
Z (student)	√							√		

Solution	Energy Method	Diagram/Chart	Checking	Wrong Answer	Givens	Numbers Late	Messy	Clearly Fractionated	Words	Omissions
O (instructor)	√	√	√			√		√	√	
P (student)							√			
Q (student)		√		√						√
R (student)		√			√			√		
S (instructor)				√		√		√		
T (student)	√	√					√			
U (student)		√								
V (instructor)	√	√	√			√	√		√	
W (instructor)		√			√	√				
X (student)	√					√				√
Y (student)		√			√					
Z (student)				√				√		√

Table 4.1: Characteristics of the Solutions Selected for the Card Sorting Task

The instructor solutions in each set were written by the same four people. Also notice that although instructor S's solution to the sliding block problem contains explicit fractionation and late number insertion, his final answer is wrong; the analysis of this particular solution by the subjects promised to be interesting.

Student subjects sorted solutions to the problem which they had not solved on their mechanics final exam. Twenty students evaluated the problem containing the spring, while a different eighteen students worked with the sliding block problem. Instructor subjects were given both sets of solutions to sort. There were thirteen instructor subjects, all of whom sorted the spring problem solutions. However, one

instructor was very uncomfortable in the interview and so was not asked to sort the sliding block problem, bringing the instructor sample size for that problem to twelve.

### **Broad Analysis**

When the spring and sliding block problems were written for their respective sections' final exams, they were written to be equivalent problems, and this was part of the reason these particular problems were selected for this study. However, some analysis showed that these problems are quite different, resulting in different types of observations from the subjects, depending on which set of solutions they viewed. Therefore, the sets were separated for analysis.

To process the data, the interviews were first transcribed. Then the reasons each subject used in arriving at a decision were grouped into 32 different criteria, listed in Table 4.2, on the next page, with a brief description of each.

Criterion	Brief Description
Accuracy	Concern or lack of concern for accuracy
Approximations	Use of approximations or not
Assumptions	Stating or not stating assumptions
Bar Charts	Use or omission of bar charts
Boxing Answer	Boxing or circling answer or lack thereof
Clarity	Ease or difficulty in understanding solution
Confidence	Subject "can tell" the solver knew what they were doing or not
Correctness	Whether answer or portion of solution is correct or not
Cross Outs	Crossing out information
Descriptiveness/ Labeling	Explaining or labeling parts of solution or lack thereof
Detection of Error	Indication that solver found an error in solution
Diagrams	Use or lack of diagrams and/or drawings
Evaluating Answer	Solver did or did not write something about reasonableness of answer
Extra Work/ Info	Solution contains more work than was necessary to solve problem
Formula Choice	Use of a particular formula in solution
Fractionation	Breaking problem into parts or lack thereof
General Appearance	"It looks like" an instructor/student did it
Handwriting	Handwriting
Hurriedness	Subject "can tell" the solver was in a hurry or not
Insertion of Numbers	When numbers were inserted into equations
Length	Whether solution is long or short
List Givens	Listing given information from problem statement
Method	What physics the solver applied to the problem
Multiple Methods	If problem was solved in multiple ways or not
Neat/Messy	Neatness or messiness of solution
Notation	Choice of symbols for physical quantities; use of subscripts
Omissions/Complete	Whether solution is complete or missing steps
Organization	How the solution is organized
Rounding, Digits	Comments on rounding, use of significant figures, or many decimal digits
System ID	Identifying the system
Units Check	Checking units
Use of Units	How units are used or not used in solution

Table 4.2: Criteria Groupings for Card Sorting Task

To compare this work to Chi's study described in Chapter 2, these types were further organized into eight broader categories, as shown in Table 4.3, below.

<b>Superficial:</b> Boxing Answer, Cross Outs, Handwriting, Neatness/Messiness, Confidence, General Appearance, Hurriedness, Length
<b>Presentation:</b> Clarity, Descriptiveness/Labeling, Organization, Omissions/Completeness
<b>Methods:</b> Formula Choice, Notation, Extra Work, Insertion of Numbers, Assumptions, Listing Givens, Method, Multiple Methods, Fractionation
<b>Representational Tools:</b> Bar Charts, Diagrams, Pictures
<b>Error- Related:</b> Correctness, Detection of Error
<b>Checking:</b> Use of Units, Units Check, Evaluating Answer
<b>Accuracy:</b> Rounding, Sig Figs, Approximations

Table 4.3: Broad Categories for Card Sorting Criteria

The first three categories, Superficial, Presentation, and Methods, are, by far, the largest categories. The superficial category is analogous to Chi's surface features, and the methods category is like her deep structure category. Presentation skills were made into a separate category, because while not directly relevant to the deep structure of the problem solution, they are indeed characteristic of expert problem solutions. The



remaining categories contain items which do not as a rule fit into one of these first three categories all the time. For example, a comment on the use of a diagram or the accuracy of a solution could either be a surface feature type of observation or a deeper look into the solution. For example, “There’s a free-body diagram, so it must be an instructor” does not indicate that a subject has looked deeply at the methods employed in solving the problem. However, stating that an extra force is on a diagram and so it probably is by a student indicates that the subject has examined how the problem was solved. There were also a small number of criteria which did not fit well into any of the categories outlined above. It was hypothesized based on Chi’s findings that instructors would tend to base their decisions on the methods used to solve the problem and that students would depend more upon superficial features.

Each subject’s transcript was analyzed to determine how many times a particular criterion was used in arriving at her or his decision. This was somewhat tricky at times, because many subjects made a large number of observations about the solutions before reaching a decision. Only statements which could be directly linked to the final decision for a solution were counted. For instance, when looking at solution O to the spring problem, one of the student transcripts reads like this:

“ ‘There are two parts to this problem, the collision and the compression of the spring. The bullet and block are the system, momentum is conserved, kinetic energy is conserved after the collision.’ Hm. I have to put this on the teacher pile, because of the word ‘thus.’ ”

In this case, the student has noticed the descriptiveness, some of the methods, and system identification. However, the only reason which can be directly linked to his final decision is the use of “thus.” This excerpt was only classified as “other.”

Oftentimes, several reasons were cited for one decision, and each of these were counted. Since each subject made twelve decisions during the task, it was therefore possible for each criterion to be cited twelve times per interview.

Using this information, what has been termed a “decision percentage” for each of the specific criteria, as well as for each of the broad categories, was calculated to indicate how often a criterion or category was used in arriving at a decision. For instance, the total number of reasons given by the novice subjects sorting the spring problem was 276. Of those 276 reasons, 23 fit into the organization category. Taking the ratio of these shows that eight percent of reasons given for a decision were based on the organization of the solution. Figure 4.1, on the next page, shows the decision percentages for each of the broad categories, separated both by expert or novice and by the solution being evaluated. The solid bars represent the experts, while the striped bars are for the students. The black ones indicate the spring problem and the gray ones are for the sliding block.

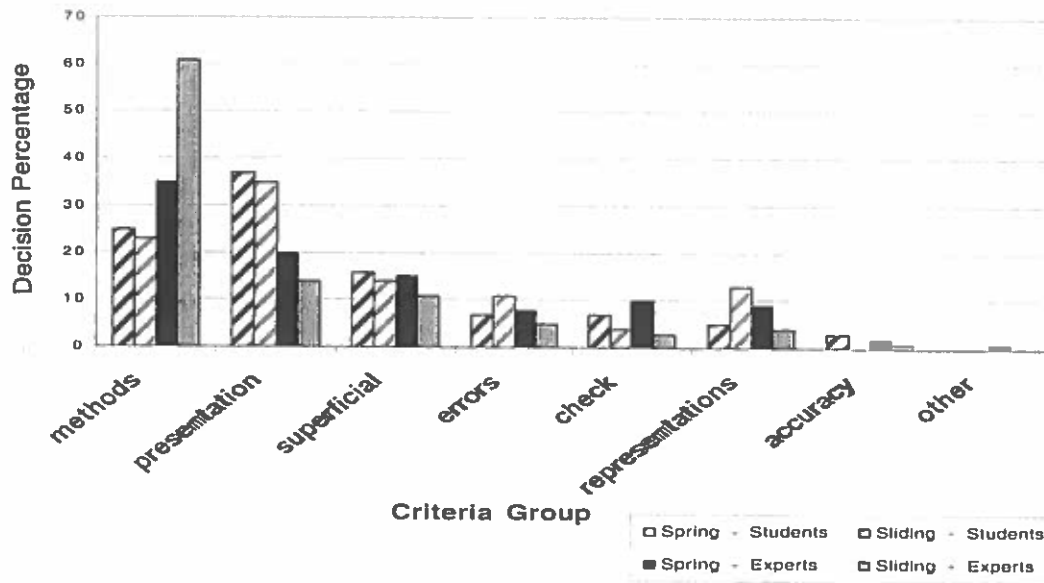


Figure 4.1: Analysis of Student-Instructor Decision Making on Card Sorting Task

Since most of the responses, as expected, were in the first three categories, the analysis focused on these, shown in Figure 4.2, on the next page.

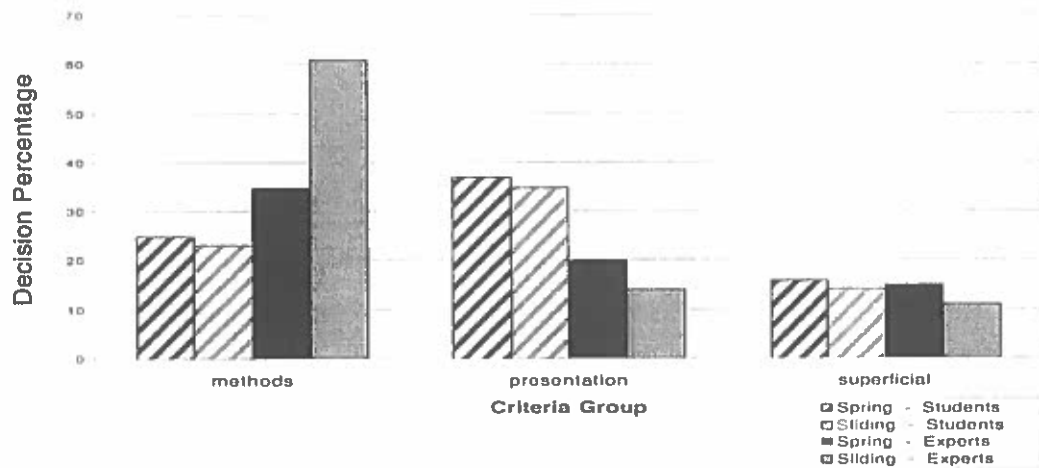


Figure 4.2: Focused Analysis of Student-Instructor Card Sorting Decisions

There are many items to note about this graph. First, as expected, the instructors based their decisions primarily on the methods used in solving the problem. However, the students did not behave as hypothesized – most of their decisions were based upon the presentation features, rather than the superficial details, of the solution. Furthermore, the instructors cited superficial details about the same amount as the students did. There are two possible interpretations of these results. The first is that the students are cueing on the surface features of expert problem solving. [3] When experts solve problems for novices, they tend to concentrate on writing a solution that is clear. This interpretation would be consistent with Chi’s findings. The second possibility is that the students at some earlier time would have cued on the superficial details, but are now in a transitional stage between the superficial area in which they began and the

methods area where they will (hopefully) end. This also fits well with Chi's results, but introduces a new third categorization for subjects.

The second major set of features to discuss is the difference between the two problems as seen by the instructors. Note that the instructors' decision percentage in the methods section is almost thirty percent greater for the sliding block problem than for the spring problem. This is due to the fact, alluded to earlier, that the sliding block problem has two relatively straightforward methods leading to a solution; one is energy conservation, and the other is an application of Newton's second law and kinematics. Both methods were represented in the solutions given to the subjects, and the instructors often used the applied principles as a major criterion in determining whether a solution was by a student or instructor. In the case of the spring problem, dynamics were never attempted, so this was not an issue. However, notice that for the student sample the methods decision percentage is slightly *lower* for the sliding block problem than for the spring problem; the students did not, for the most part, use the differences in physical principles in arriving at their final decisions. Rather, referring back to Figure 4.1, notice the larger representation decision percentage for the sliding block problem than for the spring problem. One consequence of the Newtonian approaches to some of the sliding block solutions was that many more diagrams were drawn. Students tended to decide more based on the presence and quality of diagrams than on the chosen method of solution. By contrast, the instructors used representations as a determining factor about half as much for the sliding block problem as they had for the spring problem. This is due to the presence of the work-energy bar charts on several solutions to the spring problem; it is not surprising that instructors would use this as a cue more than students.

The students were taught to use this representational tool by experts and may assume that everyone uses it, but the instructors know that very few of their colleagues were exposed to bar charts during their formative problem solving years, and view them (usually correctly) as a sign that a student wrote the solution.

With the exception of representations, there were no significant differences in the decision percentages between the two student samples. The reasons behind these differences are made still clearer by looking at the most popular individual criteria for each population on the two problems. As shown in Table 4.4, on the next page, the students based their decisions for the spring problem on descriptiveness and labeling fifteen percent of the time. The other presentation skills (organization, omissions/completeness, and clarity) were used eight, seven, and seven percent of the time, respectively. Number insertion, at eight percent, was the only methods-based skill in the top five. Students also based their decisions on whether the answer was correct or not seven percent of the time. Looking at the instructors on this same problem, they based their decisions on when the numbers were inserted into the problem an overwhelming 25 percent of the time. This was followed by descriptiveness and labeling at eleven percent. The work-energy bar charts discussed above were the third most popular criterion with instructors on this problem at ten percent. Surprisingly, the only place one of the superficial criteria shows up in a top five list is here with the instructors – neatness/messiness tied for the third most popular criterion, also at ten percent. At nine percent, instructors actually cued upon the correctness of the solution more than students did.

Novices			
Spring		Sliding Block	
Criterion	%	Criterion	%
Description/Labeling	15	Digrams/Drawings	13
Organization	8	Omissions/Complete	12
Number Insertion	8	Description/Labeling	9
Correctness	7	Clarity	8
Omissions/Complete	7	Method Used	7
Clarity	7	Correctness	7
Experts			
Spring		Sliding Block	
Criterion	%	Criterion	%
Number Insertion	25	Method Used	20
Description/Labeling	11	Number Insertion	9
Bar Charts	10	Notation	7
Neat/Messy	10	Extra Work/Info	7
Correctness	9	Multiple Methods	6
		Formula Choice	6
		Description/Labeling	6

Table 4.4: Top Five Decision Criteria Used by Each Subject Group (Including Ties)

On the sliding block problem, the top individual criteria further emphasize the difference between the instructors and students. As stated earlier, the students focused much more on diagrams on these solutions than on the spring ones; diagrams and drawings were the primary criterion at twelve percent. The next three criteria were all once again from the presentation category: omissions/completeness, descriptiveness and labeling, and clarity. As will be seen with the instructors, the method used to solve the problem was a fairly popular factor in arriving at a decision. The same percentage of students used the correctness of the answer as a deciding factor as did on the other problem. The instructors on this problem, as described earlier, used methods criteria for

the vast majority of their decision making; six of the top seven criteria (method used, number insertion, notation, extra work or information, multiple methods, and formula choice) were from the methods category. The seventh was descriptiveness and labeling, the only criterion to show up in all four of the top criteria lists.

These results have interesting implications for instruction. It seems that students notice the presentation of an instructor's solution more than the actual methods used in the solution itself. Instructors may want to emphasize the *process* of the solution even more in their courses than they have previously. It should be pointed out that students did not necessarily think that the best presentations were instructors, although this was sometimes the case. One student while looking at a solution said, "It looks very logical to me. Yeah, so I guess a student did it." A very possible interpretation of this comment is that instructors are not logical. Another student, when looking at spring solution O, a very thorough one written by an instructor, said that the person writing it had, "...list[ed] every single step, like they're trying to get more credit for the problem, make sure they show everything that they have. I'm thinking this was a student." This subject has not yet learned the value in writing a complete solution.

Other student reasoning also shed some surprising light on the way pupils view problem solving. Where most instructors would love to see their students exhibit enough knowledge and confidence to check a problem solution by solving it in an alternative way (for example, working a problem with conservation of energy and then checking it with dynamics and kinematics), one student said, "...if I do it a different way after doing it the first way and I come up with the same answer, I usually erase the second way, but I don't know why, I just do." There are two important points to take



from this. The first is that instruction needs to emphasize using multiple methods to check problem solutions as a strength and not a weakness. Second, as discussed earlier, there are some aspects of problem solving which are difficult to measure using merely written exams.

There was a great deal of variation in the student sample, but not very much in the instructor sample. As Table 4.5 shows below, the instructors' primary decision category was methods in 16 of the 24 card sorts in the sample. Presentation skills were the primary category three times, superficial details once, other categories twice, and there were two instances where multiple categories were cited equally much.

Category	Novices	Experts
Method	8	16
Presentation	19	3
Superficial	5	1
Other	2	2
Ties	4	2

Table 4.5: Dominant Decision Making Categories for Novices and Experts

As can be seen in the same table, although 19 of the 38 student subjects focused on presentation features for decision making, there were five students who concentrated primarily on surface features, as well as eight who maturely looked at the methods. Two students had one of the other areas as a primary focus, and four of the students had two or more areas which they accessed equally dominating their decisions. As an example of very superficial reasoning, consider "Fred," who used surface features to

make slightly over 46 percent of his decisions, and only cited presentation or methods criteria about eight percent of the time each. Here is a representative excerpt from his interview:

“I’m guessing this is by a professor. Also another factor coming into it is that generally the more experience people get with writing, the faster they can go and sort of the more, I don’t know, it always seems like the more sort of stylized their numbers are, such as when I write now I have a slight slant and it’s a little bit more intricate and a little bit more neat than it was several years ago, so considering that half of this page seems to be in italics and it’s written extremely smooth, but yet you can tell the person was evidently going really fast, guessing this was done with a person with much faster experience in writing, so therefore, a professor.”

By contrast, Patricia sorted through all the solutions first and placed several of them to the side, stating, “None of those used energy. Those were students.” She then sorted through the solutions again and found more which she thought were by students, based upon the quality of their diagrams. An analysis of all her deciding criteria shows that she used method-based criteria nearly 54 percent of the time, presentation features about 23 percent of the time, and never cited superficial details of the solutions. It is probably not terribly accurate to refer to such a student as a novice, because this student is exhibiting very mature behavior, using the same primary criterion as the instructors.

Other students also made statements indicating that they had moved beyond pure novice reasoning. In contrast to the student described previously, one viewed the multiple solution methods as a desirable behavior: “...the fact that they worked it out using both kinematics and energy leads me to believe it’s a teacher.” Another displayed an awareness of his own development in problem solving: “...that’s a student, plugging in all the numbers [right away] ... at least, that’s what I would have done first quarter, so yeah, that’s a student.”

Another key difference observed between the two populations was that instructors were much more likely to give general statements as part of their criteria. For example, one instructor while sorting said, “One of the major classifications is deciding there are two parts to the problem.... [I’m] drawing the line between the solutions here of whether they essentially split it into two problems and identified what’s the relevant physics of each one.” Another instructor subject, talking about his strategy after he had completed most of the task stated, “lots of students like to take the given information and rewrite it, which isn’t a bad thing, and assign symbols, whereas most of my alleged professor[s] haven’t done that.”

These are classified as general statements, because they were not directed at one solution in particular, but explain part of an overall strategy for accomplishing the task. Ten of the thirteen instructor subjects made at least one of these general statements as part of their final reasoning. Most students also made at least one general comment while observing the solutions, but only nine of them gave general criteria as part of their final decision. If the card-sorting task is viewed as a problem, this difference is consistent with observations from previous research that experts have a general set of problem solving skills which they apply in most contexts, where novices tend to view each problem situation independently. [4, 5, 6]

A related observation centers about the formula choice category. The comments in this category all occurred in interviews with the sliding block problem when subjects were looking at kinematics approaches to the solution. Several instructors and three students cued at one time or another on a form of the equation  $v^2 - v_0^2 = 2a\Delta x$ . Instructors who saw that equation generally thought it indicated a student solution,

reasoning that an instructor would use the more general kinematics equations, rather than this specialized one. (If this reasoning were indeed correct, it would be consistent with previous problem solving research. [7]) Also, the instructors who used this criterion did so several times. By contrast, the three students who based a decision on this equation did so only once each, in spite of the fact that it appeared on six solutions. Furthermore, a student who noticed the equation in solution S, where the equation is written in a nonstandard form, insisted she had never seen that equation before! This difference in behavior is yet another indication of novices viewing each problem uniquely, instead of applying the same basic approach to each situation.

### **Solution-by-Solution Analysis**

Although the subjects' comments while doing this exercise are probably more useful information than how the problems were actually classified, a brief discussion of the sorting results is given here. The discussion will focus on solutions where there was either a strong agreement or a definite disagreement between the two interview samples. For the spring problem, the samples were in good agreement, for the most part, as shown in Figure 4.3 on the next page. Notice that the two groups especially agreed on solutions O, Q, R, T, U, X, and Y. Solution O was almost always classified as an instructor's. Both samples did so primarily because of the presentation features, particularly the descriptiveness.

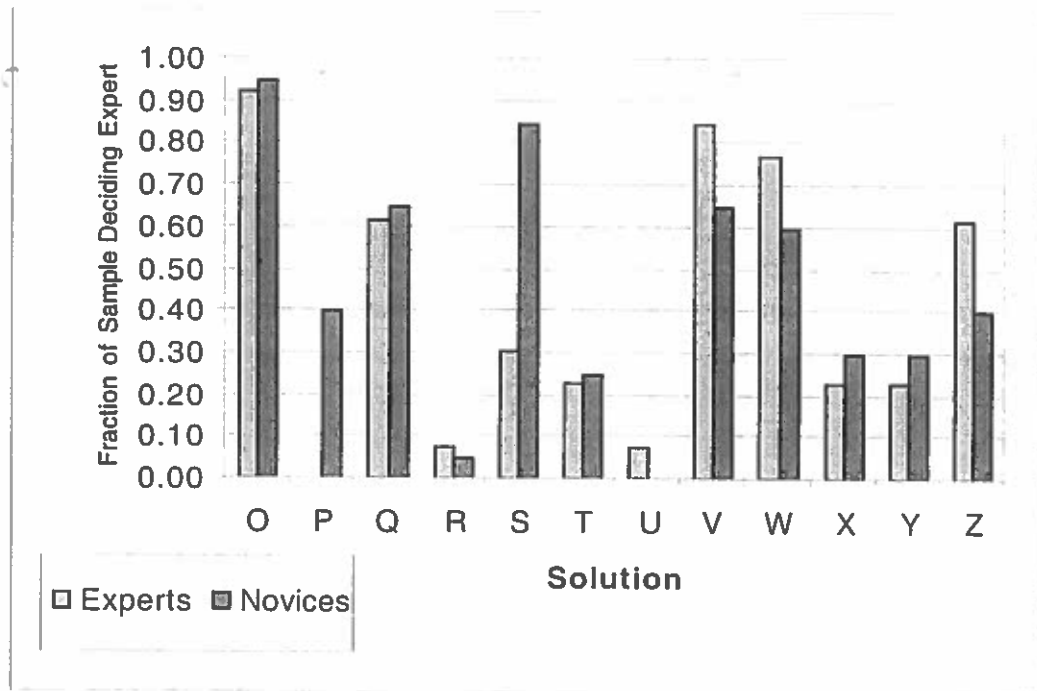


Figure 4.3: Card Sorting Summary for the Spring Problem, Separated by Novices and Experts

Solution Q was classified as an instructor roughly sixty percent of the time by each group. The subjects who said it was an instructor often cited presentation skills again. The most interesting category of comments on this particular solution was the extra work area, sparked by the comprehensive energy equation in it. Of the subjects who cued on this equation, there was roughly a fifty-fifty split as to whether or not this work was by a student. This was true for both samples. One instructor observed,

“...this person’s thought of all forms of energy.... I would think that most students would not write all of that down, but then, I would also think that the teachers would look at a problem like this and know what forms of energy are there and not write them down, so I’m going to say student.”

In contrast, another instructor insisted, “That’s got to be an instructor, absolutely has to be an instructor. There’s no way a student is going to write down, put heat in there.”

The distribution of comments among the student sample was similar.

Solution R was always placed in the student pile, with the exception of one person in each sample. The primary reason for this was the wrong numerical answer. Several instructors based their decision on the early use of numbers in the solution. The one student who said it was an instructor’s solution cited the use of several decimal places in the answer. This once again reveals a student’s false perception of what instructors do when they solve problems.

In each group, solution T was classified as an expert slightly over twenty percent of the time. The most commented-upon feature of this solution was the work-energy bar chart. With the exception of one student, every person who used the bar chart as a criterion decided the solution was a student’s. One interesting difference in reasoning between the instructor and student subjects was in the number insertion category. The person who wrote this solution worked algebraically all the way to the end, but whatever numbers he used to arrive at his answer are not written on the paper. Two out of the three instructors who commented on number insertion felt it indicated a “heavy reliance” on a calculator and suspected the solution was a student’s. By contrast, the two students who commented on number insertion decided the solution was by an instructor. As in the Diana task, a greater percentage of instructors than students noticed specific omissions in the solution. The instructors who noticed this particular omission interpreted it differently than the students who did.

Solution U was categorized as an instructor's solution by only one instructor. Every other subject in both samples felt it was by a student. The primary criterion given by both samples was that the final answer was incorrect. Both populations feel that instructors do not make many mistakes on introductory physics problems.

Solutions X and Y were also considered mostly as student work. The decision criteria for X vary greatly, but there are two trends worth mentioning. First, three instructors made specific comments about the bar charts, all citing them as a reason to think the solution was by a student. Only one student reasoned similarly. Second, this solution, along with solution Y, provoked many more of the superficial comments than the other solutions in the task. Solution Y prompted many comments from both instructors and students about its messiness. All the instructors who mentioned the lack of neatness decided the solution was a student's. Some of the student subjects, however, felt that this indicated the work was by an instructor.

The only solutions where the two groups significantly disagreed were P ( $p < .01$ ) and S ( $p < .005$ ). When looking at solution P, none of the instructors thought it was an instructor solution, but forty percent of the students did. The reasoning for solution P illustrates the global trends seen in the whole analysis. The instructors focussed on the methods, such as number insertion, while the students based their decisions on presentation criteria. Solution S stimulated very different responses from the two groups. The students' reasoning on this one was much more methods-based than on most of the other problems. Over eighty percent of the students thought S was an instructor, primarily because of the initial inclusion of a friction term and then assuming there was no friction. Surprisingly, only three instructors picked up on this aspect of the

solution, and one of them decided it must be by a student. Three of the instructors reached their decision of student because of the bar charts.

With the sliding block problem, there was also considerable agreement, particularly on solutions U, V, W, and Z, as shown in Figure 4.4, below.

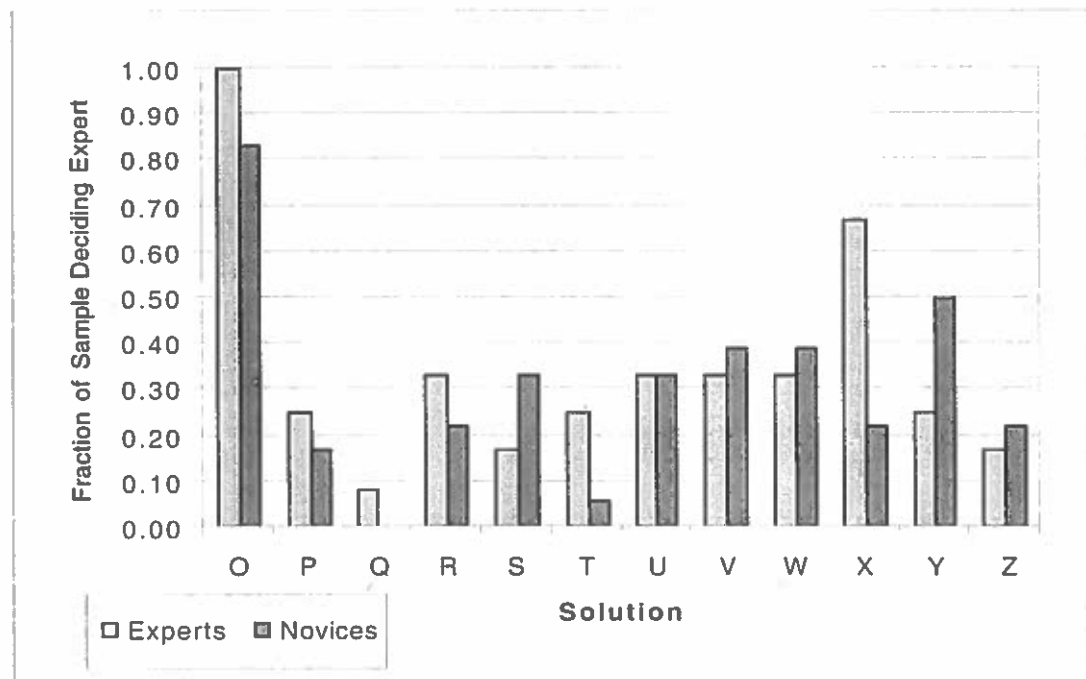


Figure 4.4: Card Sorting Summary for the Sliding Block Problem, Separated by Novices and Experts

Both groups classified solution U as an instructor slightly over thirty percent of the time. Students used the method as a criterion on this solution more than on any other. Solution V was thought to be by an instructor between thirty and forty percent of



the time in each group. In both samples, most of the time a subject gave the detection and correction of V's error as a criterion, he or she determined it was a student solution. The four students who talked about the diagrams all decided an instructor wrote the solution. The expert sample actually used superficial criteria on this solution more than the students and more than all the other solutions except one. These criteria were related to the way the solution indicated the detection of the error.

Solution W had a similar distribution to V as far as the final classification was concerned, but the reasoning behind the decisions was quite different. In the instructor sample, only one subject gave a reason which was not methods-based, and four commented specifically on the method employed. Only one student mentioned the kinematics approach as a reason for making his decision. Students reasoned primarily using the handwriting, diagram, and completeness of the solution.

About twenty percent of each group decided solution Z was by an instructor. The instructors once again based their decisions on the methods criteria, while the students primarily talked about the presentation elements and the diagram. However, it should be noted that three students specifically gave the method as their primary reason.

Solution X was the only one where there was a significant difference between the instructor and student responses, at the  $p < .025$  level. It was viewed as an instructor's by two-thirds of the instructors, but only slightly above twenty percent of the students thought so. A quarter of the expert subjects cited number insertion in their reasoning. The most common reasons for the students were based on the omissions in the work – only one student who commented that work was missing decided it was an instructor.

This solution-by-solution analysis illustrates in a more specific way the general differences seen between the two samples. Some of the reasoning of the students (and sometimes the instructors) reveals perceptions of problem solving which have not been documented before.

### **Summary**

This card sorting exercise shows that experts and students use different criteria when determining whether a problem solution was written by an expert or novice. Experts focus more on method-related criteria, while students focus more on presentation details. The two groups base decisions on superficial details about the same amount. When combined with the fact that some students exhibit behavior very similar to the expert sample and some of the students base their decisions almost exclusively on superficial criteria, this indicates a transitional state in problem solving perception.

The strongest implication this portion of the study has for instruction is that instructors should find ways to emphasize solution methods more when solving problems for their classes while still writing clear, organized solutions. Instruction should place more emphasis on what purpose diagrams serve in problem solving. Most of the students who commented on diagrams merely commented on their presence, without looking at the correctness or apparent function of them. Likewise, some emphasis on the reasons for utilizing certain expert strategies throughout the course is recommended. The apparent mixed student state observed via this task indicates that the transition from novice to expert problem solver does not occur at the same time for

all students. This makes it important to revisit problem solving strategies throughout the term, since a different portion of the class should benefit from the information each time.

## ENDNOTES FOR CHAPTER 4

- [1] Michelene T. H. Chi, Paul J. Feltovich, and Robert Glaser, "Categorization and Representation of Physics Problems by Experts and Novices" *Cognitive Science* 5, 2, 121-152 (1981).
- [2] B. Bloom, editor, *Taxonomy of Educational Objectives: Cognitive Domain* (David McKay, New York, 1956).
- [3] Jerry Touger, private communication.
- [4] Ron Good, "Scientific Problem Solving by Expert Systems," *Journal of Research in Science Teaching* 21, 3, 331-340 (1984).
- [5] Dorteia P. Simon and Herbert A. Simon, "Individual Differences in Solving Physics Problems" in *Children's Thinking: What Develops?* edited by Robert S. Siegler (Lawrence Erlbaum Associates, Hillsdale, NJ, 1978).
- [6] William C. Robertson, "Detection of Cognitive Structure With Protocol Data: Predicting Performance on Physics Transfer Problems," *Cognitive Science* 14, 2, 253-280 (1990).
- [7] Herbert A. Simon, "Problem Solving and Education" in *Problem Solving and Education: Issues in Teaching and Research*, edited by D. T. Tuma and F. Reif (Lawrence Erlbaum Associates, Hillsdale, NJ, 1980).

## CHAPTER 5

### TRENDS FROM EXAM DATA

#### **Context and Overview of the Trends Study**

As stated in the brief description of Chapter 1, one focus of this research effort was to analyze problem solutions from student exams in the 131E and 132E courses to identify ways in which the problem solving behavior of the class changed. The project was not designed to evaluate the effectiveness or ineffectiveness of instructional techniques on problem solving skills, but rather aimed to describe and better understand a portion of the transition from novice to expert problem solver. The students in this particular physics sequence were participants in Ohio State's Freshman Engineering Honors (FEH) program. This population was chosen for several reasons. First, the enrollment of the FEH program is fairly consistent for the entire freshman year, making it possible to carry out a multi-quarter study with a relatively large sample. Second, these students learn quickly, so any effects caused by gaps in physics content knowledge would be minimized.

The format of these courses consisted of three one-hour "lectures," two one-hour recitations, and one two-hour laboratory each week. Each quarter lasted ten weeks. The instruction contained many non-traditional elements. Lectures were extremely interactive; students made observations and predictions, answered qualitative questions,

and solved short numerical exercises, as well as longer multi-part problems, during the class time. Additionally, they were encouraged several times during each class to consult with their peers. Recitations consisted entirely of cooperative work, usually solving context-rich problems similar to those developed at the University of Minnesota [1] or working exercises from the Active Learning Problem Sheets developed by Van Heuvelen. [2] Students worked in the same cooperative groups in the laboratories. This portion of the course consisted entirely of design labs. The TA presented the students with equipment and a question to answer, then the students designed an experiment to answer the question and answered it.

Both the 131E and 132E courses were taught in two sections of roughly 70 students apiece. Although data was initially taken from both classes, some analysis showed that the exam problems from one section of 131E did not contain enough challenging elements to truly be classified as problems for the majority of the class. One of these exercises was even taken verbatim from the course homework. Therefore, the sample was narrowed to students who were in one particular section of 131E and who continued through the 132E course. The exams for both sections of the 132E class were written such that there were problems from the two sections which were conceptually equivalent. The problems selected from the 132E midterms were actually isomorphic. Additionally, the same instructor taught both classes, so the section of 132E was not regarded as a factor which would make a significant difference in problem solving development. The final sample size for this analysis was 67 students.

The instructor of the course wrote the exams, but the researcher would request that certain features sometimes be included in the exam problems. One or two

problems were selected and photocopied from each exam for analysis. These problems can be found in Appendix C. The papers were blinded so that the researcher would not know whose paper was being evaluated. Each solution was evaluated in multiple areas, mostly those with documented differences between experts and novices. The goal was to track these separate problem solving skills for two quarters to see if any significant changes occurred between the beginning and end of instruction.

### **Diagram Usage**

As previous research has shown, there are fundamental differences in the way experts and novices use diagrams and physical representations in their work. Experts use physical representations as a transitional step from the problem statement to the mathematics necessary to solve the problem. Students may not see the need for physical representations, and even if they do use them, often do not use them in this same transitional manner. [3] This analysis will include the usage of both free body diagrams and work-energy bar charts.

It is assumed that the reader is familiar with free-body diagrams, but since the work-energy bar chart was introduced only about ten years ago, a brief description of it is provided here. Work–energy bar charts, developed by Van Heuvelen, are a physical representation useful in solving conservation of energy problems. As with any physical representation, the purpose of them is to reason qualitatively about the process before writing any equations. From the bar chart, one can construct a generalized work-energy equation. A sample of how a bar chart might be used to solve a problem is shown in Figure 5.1, on the next page.

A small ball of mass 1.0 kg is thrown up vertically. The initial speed of the ball is 30 m/s. Ignoring the effects of air resistance, what is the maximum vertical height that the ball achieves?

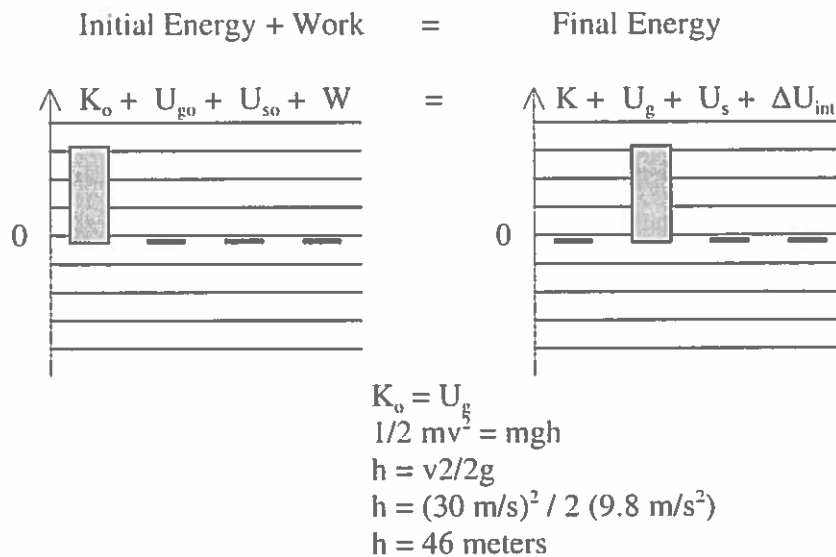


Figure 5.1: Problem Solution Using a Work-Energy Bar Chart

The bar chart here serves as an intermediary between the problem description of the situation and the algebraic equation which begins the mathematical portion of the problem solution. A further similarity between the bar chart and a free-body diagram should be noted. Just as each arrow on the free-body diagram represents a term in a Newton's second law equation, so does each bar on the bar chart represent a term in the energy conservation equation. If done carefully and correctly, the bar chart should make writing the equation easy.

Before any data was taken, it was hypothesized that some fraction, maybe half, of the class would use free-body diagrams on the first exam. However, since some of the students who neglected to draw diagrams would make errors that might have been



avoided had the diagram been present, the hypothesis continued that the number of diagrams would increase on the following exam or two. Then, as some portion of the class became more knowledgeable and expertlike, they would begin to omit the diagrams without penalizing themselves in the process, and so the number of diagrams would begin a gradual decline. The same basic trend was expected for the use of work-energy bar charts.

Reviewing the examination questions showed that there were four exams containing problems which would likely prompt students to draw a free-body diagram. These were the elevator and hiking problems of mechanics midterm 1, the ball-in-a-circle problem of mechanics midterm 2, the motorized cart problem of the mechanics final, and the electromagnetic arcade or pit problem from the electricity and magnetism final. Each student solution was examined and the presence or absence of free-body diagrams recorded. For problems where two free-body diagrams were likely, such as the first hiking problem or the motorized cart problem, the use of each diagram was recorded separately. The fraction of potentially likely diagrams that were actually drawn was computed for each exam. The resulting graph is shown in Figure 5.2, on the next page.

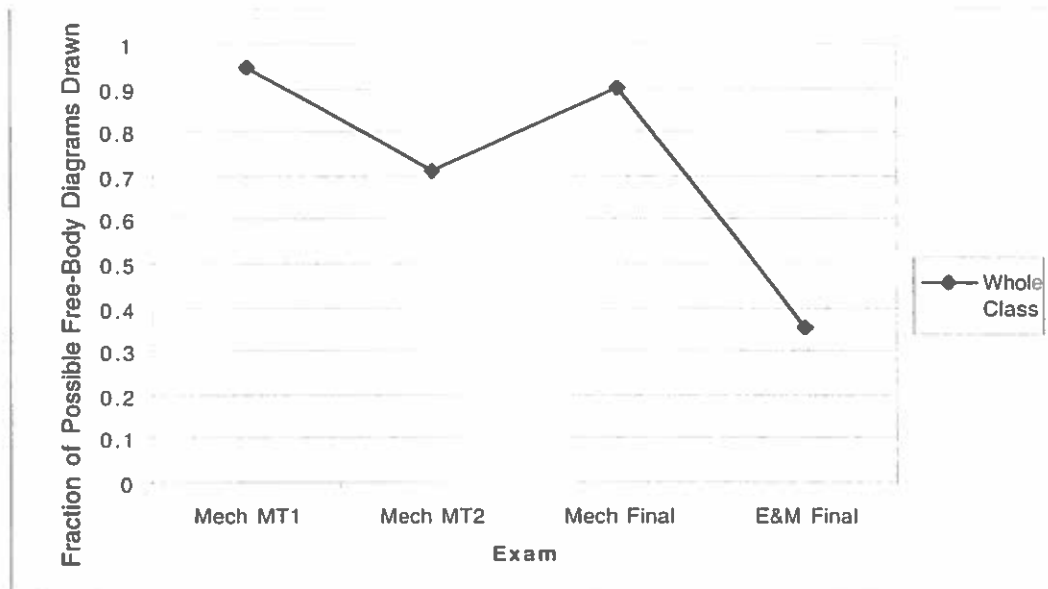


Figure 5.2: Tracking the Number of Free-Body Diagrams During the Two Quarters

Notice that the developmental trend shown by these points bears some resemblance to the hypothesized behavior. The peak and decline are indeed present, but it was not expected that students would draw the greatest number of free body diagrams on the first exam. The development of free-body diagram use is slightly more complicated than hypothesized. It appears that the students started out drawing diagrams on the first mechanics midterm, then some portion of the class felt they had mastered Newton's laws enough that they did not need to draw the diagrams any longer. Then, perhaps due to making errors which could be attributed to omitting the diagram, they drew diagrams again on the final. By the end of the second quarter, the use of

diagrams had dropped significantly below the levels it had been at for the first mechanics midterm ( $p < .001$ ) and the mechanics final ( $p < < .001$ )

In this area, as well as several of the others which follow, an additional stratified analysis was undertaken to better understand the pattern seen in the overall class. For instance, in this case, was the decline an indication of the entire class writing more specific equations or was it due to a change in just a certain portion of the class? Certainly, there are a number of different ways the class could be broken into strata. In this study, it was decided to stratify based upon the performance on the first exam, so that the changes in behavior from the beginning of the sequence could be tracked.

Since there were three potential free-body diagrams on the first midterm, there were four possible strata in which to divide the class (no diagrams, one, two, or three). Fifty-eight of the 67 students drew all three possible diagrams, eight drew two diagrams, and one drew one. Once broken into these three categories, the behavior of each stratum was charted across the two quarters in the same manner as the whole class analysis. The results are shown on the next page in Figure 5.3, with the omission of the one-diagram stratum, since it only contains one student. The numeral in parentheses next to the stratum name is the number of students contained in that stratum. Notice that the two groups start out distinct from each other and gradually come together, finishing at the same level after two quarters. See also that both of the strata curves are in the same basic shape as that for the whole class.

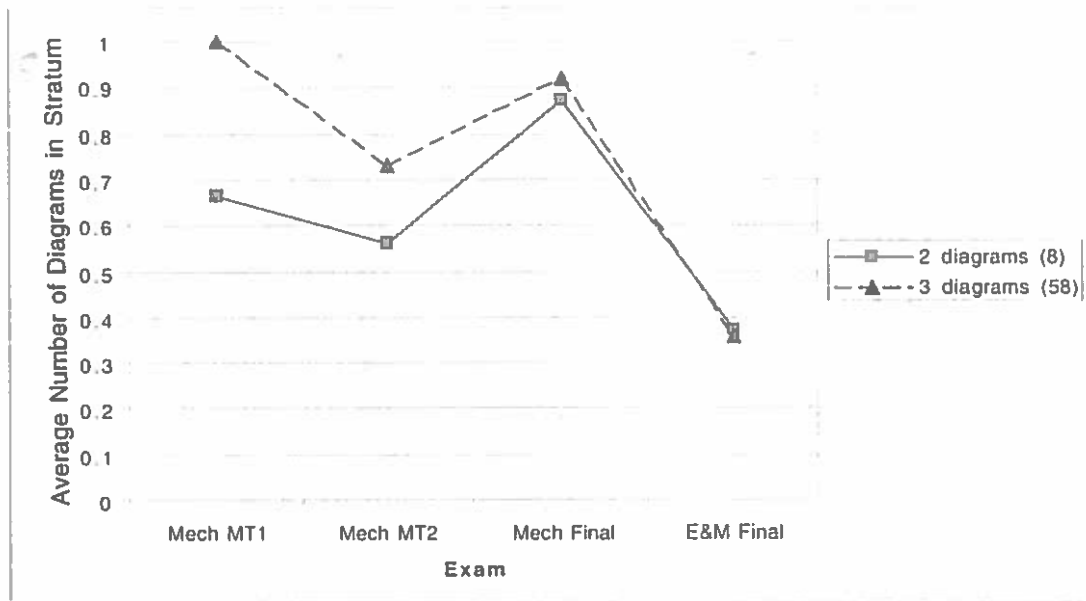


Figure 5.3: Stratified Development of Free-Body Diagram Usage

A similar class analysis was done for the work-energy bar charts. There were four exams with problems appropriate for bar charts. The results of the bar chart analysis are shown on the next page in Figure 5.4.

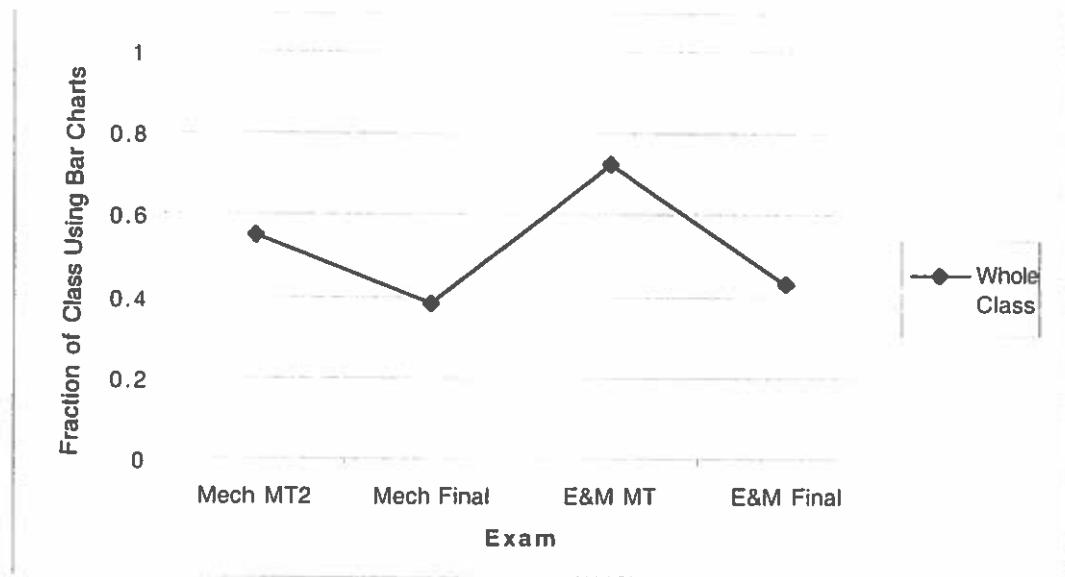


Figure 5.4: Development of Bar Chart Usage During the Two Quarters

Notice that these points exhibit the same basic trend as seen with the free-body diagrams. Carefully comparing the two graphs yields some small but important differences. First, the bar chart usage starts off at a lower level than the free-body diagram usage. This should be expected, since almost all of these students had high school physics, where free-body diagrams are almost certainly part of the curriculum. However, given the recent development of work-energy bar charts, it is highly unlikely that any significant portion of the class entered college with any experience with them. The lower initial usage of the bar charts is thought to be due to less familiarity and comfort with them as a representational tool.

Notice also that the subsequent points for the bar charts are lower than the free-body diagram points. It seems that the students internalized the thought processes

associated with the bar charts more rapidly than they did those for the free body diagrams. This view is supported by two characteristics noted about bar chart usage: almost all of the bar charts drawn were correct, and 98 percent of the energy conservation equations following from a bar chart were consistent with it. This was not so true for the free-body diagrams; there were more errors associated with them, including extraneous forces, missing forces, forces in the wrong direction, and equations following the diagrams which were inconsistent with them. These errors were not wide-spread, but were certainly present. Perhaps this difference should be expected. Forces, after all, are vectors, with all the associated directional issues, and kinesthetically related difficulties with forces are well-documented. [4, 5, 6] Energy, on the other hand, is not a vector quantity, and while one still has to be careful with sign conventions, energy conservation is an easier equation to apply correctly than Newton's second law. Also, student difficulties with energy do not involve adding extra types of energy in the same way that extra forces often appear. The combination of these factors may well have resulted in students appropriately internalizing the issues associated with energy conservation more easily than they did with applications of Newton's second law.

In an initial analysis of the students' free-body diagrams, many of them were seen to differ from the expert-like application of them which would be desired. Remember that the primary function a free-body diagram (or any physical representation) serves to an expert is to help her qualitatively analyze the physics of the situation before beginning the mathematical reasoning process. The diagram then aids in constructing the appropriate equations. A significant fraction of the free-body

diagrams drawn by the students appeared to have been drawn *after* some or all of the mathematical work had been done. Two main factors led to this conclusion. First, the position of many diagrams was near the bottom of the page. Second, a portion of the diagrams had numbers on them which could only be known if the student had done most of the mathematical work necessary to solve the problem.

These observations suggested another analysis of free-body diagram usage, which was to look at the fraction of the diagrams which were purely qualitative. Each diagram drawn was classified as either qualitative or quantitative. It was designated quantitative if it had any numbers on it at all. Once again, the behavior was tracked for the two-quarter sequence. The resulting graph is shown in Figure 5.5, below.

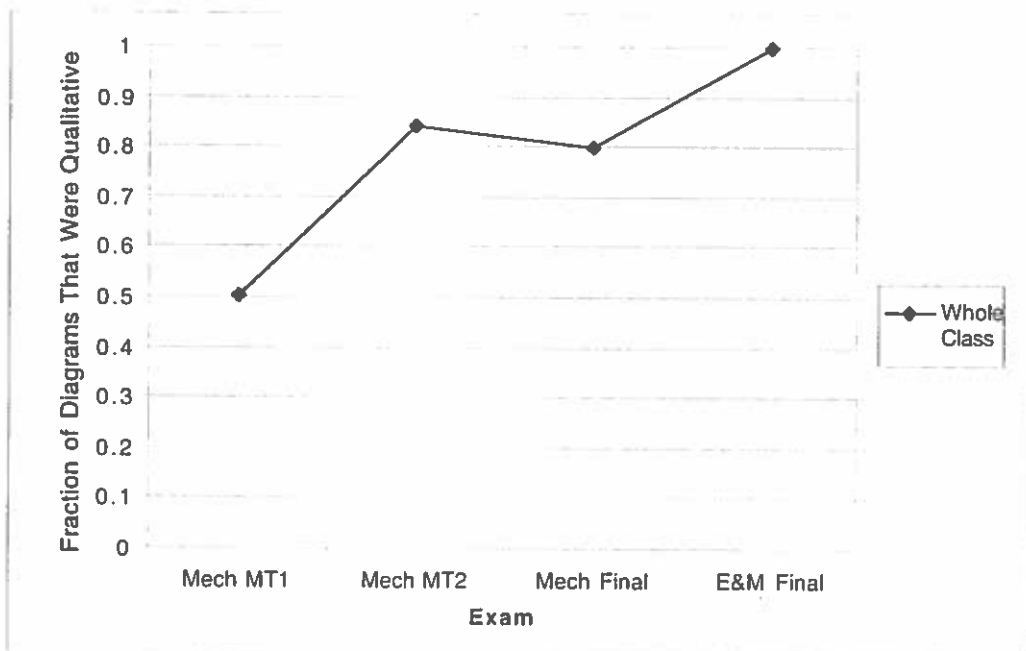


Figure 5.5: Development in Quality of Free Body Diagrams

The difference between the initial and final states is striking, moving from half of the diagrams drawn on the first midterm being qualitative in nature to all of them being qualitative on the electricity and magnetism final. This difference is significant at the  $p \ll .001$  level. Here is one area where the students have adopted expert-like behavior.

### General Versus Specific Initial Equations

A second area where novices and experts differ is in the equations from which they begin reasoning. Recall Simon's finding that novices often start with equations which are only applicable to specific situations, while experts tend to begin with general equations. [7] This area lends itself well to an analysis of written solutions.

Each problem in the sample was analyzed to see whether the solution parts in it began with generalized equations, such as  $\Sigma F = ma$ , or specific equations like  $T = mg$ . This was rated on a five-point scale, explained in Table 5.1, below.

General/Specific Initial Equations	
1	All initial equations are specific (e.g., $T = mg = ma$ ).
2	Some initial equations are specific and some are general, but there are more specific ones.
3	The initial equations are about equally split between specific and general.
4	Some initial equations are specific and some are general, but there are more general ones.
5	All initial equations are general (e.g., $\Sigma F = ma$ ).

Table 5.1: Rating Scale for General/Specific Initial Equations



It was hypothesized that students would begin with fairly specific initial equations and move toward more general ones as they gained experience. The actual behavior was tracked over the course of two quarters, and the resulting data is shown in Figure 5.6, below.

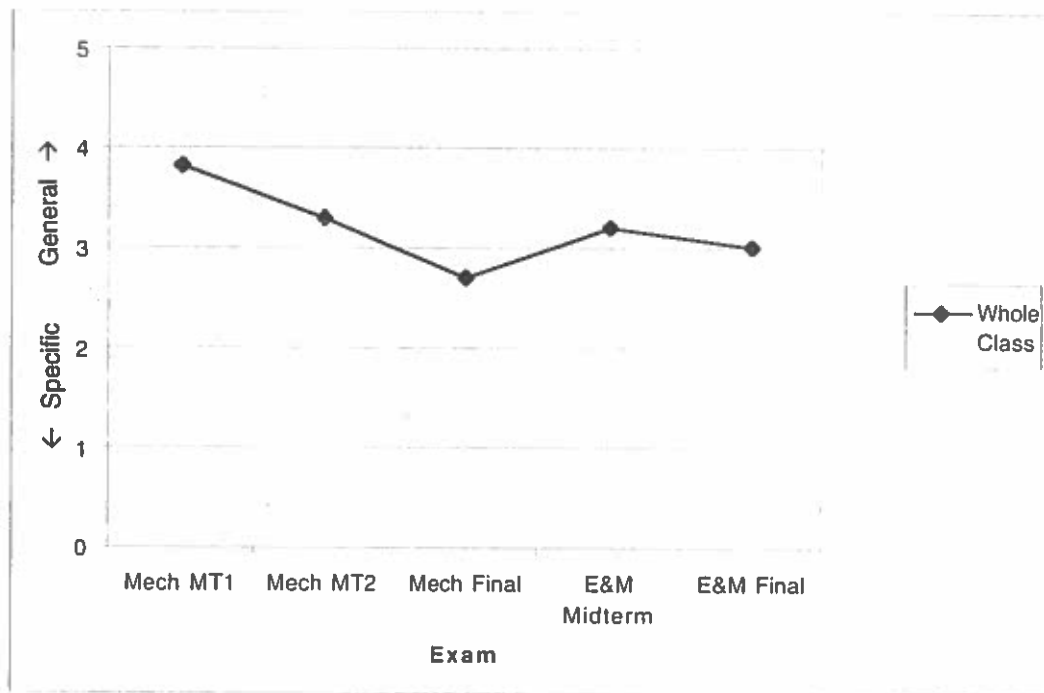


Figure 5.6: General/Specific Initial Equation Behavior Over Two Quarters

The students started with an average general versus specific equation rating of 3.8, which meant that they began their work with a combination of general and specific equations, but that they leaned more toward using general equations. By the end of the

two-quarter sequence, the rating was 3.0, indicating an equal split between general and specific starting equations. The difference between the starting and ending points is statistically significant ( $p < .0003$ ), indicating that the students actually moved away from the desired behavior of beginning problems with general equations.

The class was broken into four strata based upon the general/specific equation usage on the first exam. The development of these strata is shown below in Figure 5.7. Once again, the number of students in each stratum is given in parentheses next to the stratum labels.

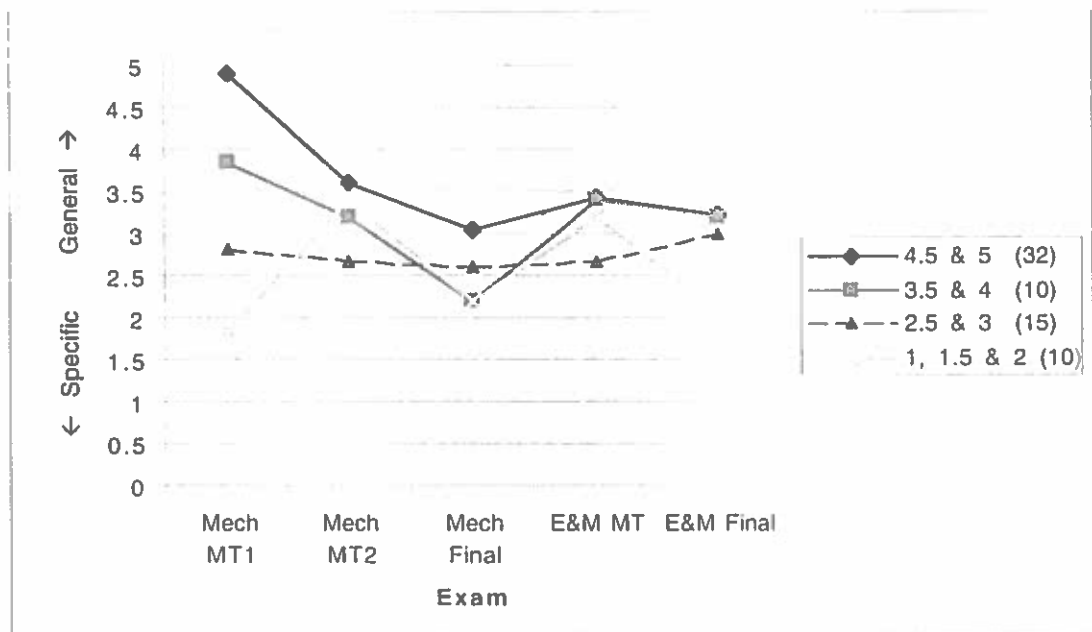


Figure 5.7: Stratified Analysis of General/Specific Equation Behavior

There are a few characteristics of this graph which are especially worth noting. First, the top three strata start out distinct from each other, but by the end of the two quarters have blended together. The bottom stratum shows the greatest variability and ends significantly lower than the other three. The 2.5 and 3 stratum shows not only the greatest stability, but the least variability and improves slightly during the sequence. The top two strata decline over the course of the data collection; the decrease in the highest stratum is the most significant and accounts for most of the global decline observed.

### **Algebraic Versus Numerical Initial Equations**

Another analysis of the initial equations involved whether they were algebraic or numerical. Algebraic here refers to an equation which is purely algebraic, while numerical indicates that the equation contains some numbers. Recall that experts differ from novices in that they reason qualitatively before doing any quantitative problem solving. It was hoped to detect a trend in the student sample toward writing more algebraic initial equations as the instructional sequence progressed. There was no reason to hypothesize that there would be any drop-off once a rise was detected. A five-point scale, shown in Table 5.2 on the next page, was used to rate this problem solving attribute.

Algebraic/Numerical Initial Equations	
1	All the initial equations contain numbers.
2	It's a mix between algebraic and numerical, but there are more numerical equations.
3	Roughly half the initial equations are numerical and half are algebraic.
4	It's a mix between algebraic and numerical, but there are more algebraic equations.
5	All the initial equations are purely algebraic (no numbers at all).

Figure 5.2: Rating Scale for Algebraic/Numerical Initial Equations

At the first midterm, the students were already doing very well in this area, with a rating of almost 4.4 out of 5. This left little room for improvement, but, as Figure 5.8 shows below, the class made a steady climb over the course of two quarters.

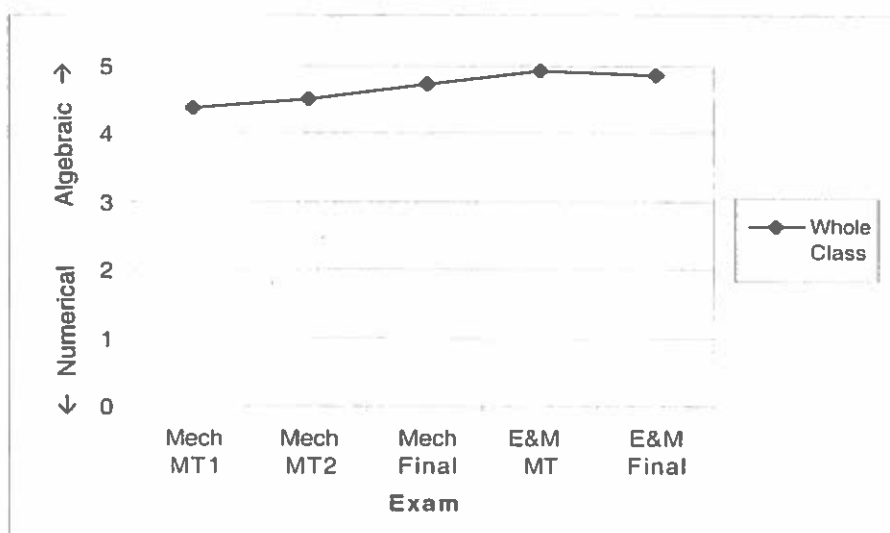


Figure 5.8: Algebraic/Numerical Equation Development for the Two Quarters

There is a small decline from the electricity and magnetism midterm to final, but this is not a statistically significant difference. However, the difference between the initial and final readings is statistically significant at the  $p < .0002$  level.

A stratified analysis is summarized in Figure 5.9, below. Once again, the number in each stratum is shown in parentheses.

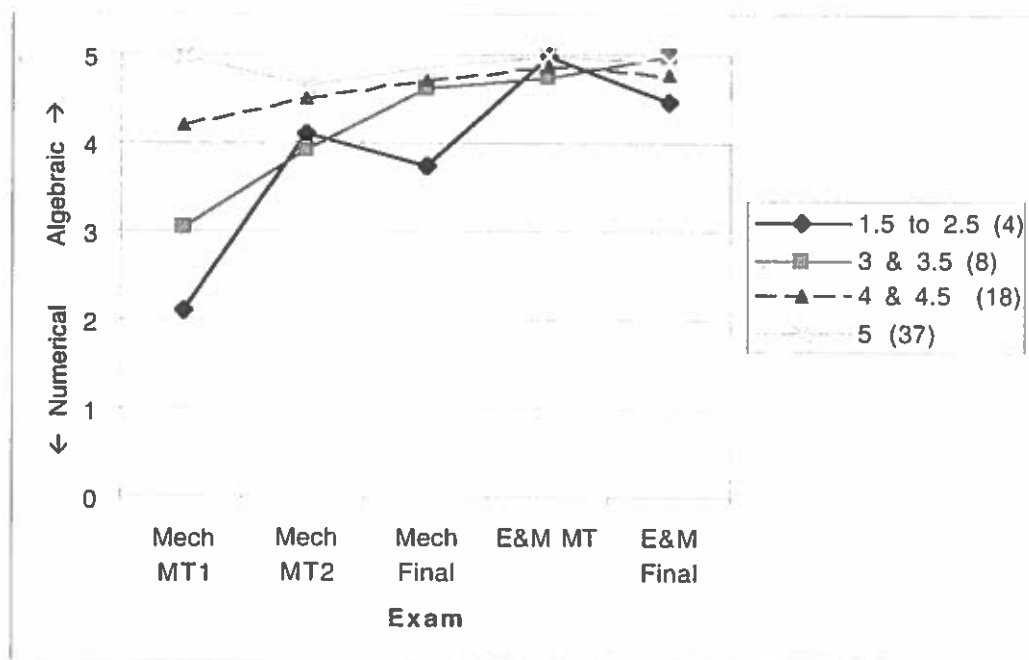


Figure 5.9: Stratified Analysis of Algebraic/Numerical Initial Equations

Note that once again, the greatest instability is in the lowest of the strata. The other strata basically show improvement during the sequence, with the greatest growth shown by those in the 3 and 3.5 stratum.

## Number Insertion

Another area where there are noticeable differences between expert and novice physics problem solvers is when the numbers are inserted into the problem. Novices tend to put the numbers in immediately, while experts will insert the numbers only after all or most of the algebra is done. In the language of representations, experts shift from an algebraic to a numerical representation only when they have done all they can with the algebraic one. By contrast, novices switch to the numerical representation very early in the problem solving process, sometimes omitting the algebraic representation entirely. There are many reasons why, in the majority of cases, it is preferable to refrain from the numerical representation until the end. Factors which seem to be needed but are not available may cancel out, numerical errors are fewer, and, as Zou found in her research, an early switch from the algebraic to numerical representation often causes students to completely lose sight of the physics of the situation.[8]

Based mostly on previous instructional experience and observations, it was hypothesized that students would begin by inserting numbers early in their work, but by the end of two quarters would have shifted to inserting the numbers later. It was not expected that there would be such a dramatic shift that the majority of the students would insert the numbers at the last possible moment, but that the bulk of the students would have at least moved in that direction, with some of them actually exhibiting expert-like behavior.

A seven-point scale for evaluating the number insertion on a given problem solution was developed. This rubric is explained in Table 5.3, on the next page.

Number Insertion	
1	The very first equation of each part contains numbers.
2	On the average, numbers are inserted after the first general algebraic equation.
3	On the average, numbers are inserted after a specific algebraic equation.
4	Number insertion varies along the way, but is early in the algebraic manipulation process.
5	Number insertion varies along the way, but is late in the algebraic manipulation process.
6	Numbers are always inserted at the very end of each part of the problem.
7	Numbers are inserted as the very last step of the entire problem, after all algebraic manipulation.

Table 5.3: Rating Scale for Number Insertion Analysis

Once again, each problem solution in the sample was analyzed for this particular characteristic. An average number insertion score for each exam was calculated, and the trend plotted for the two quarters. This graph is shown on the next page in Figure 5.10.

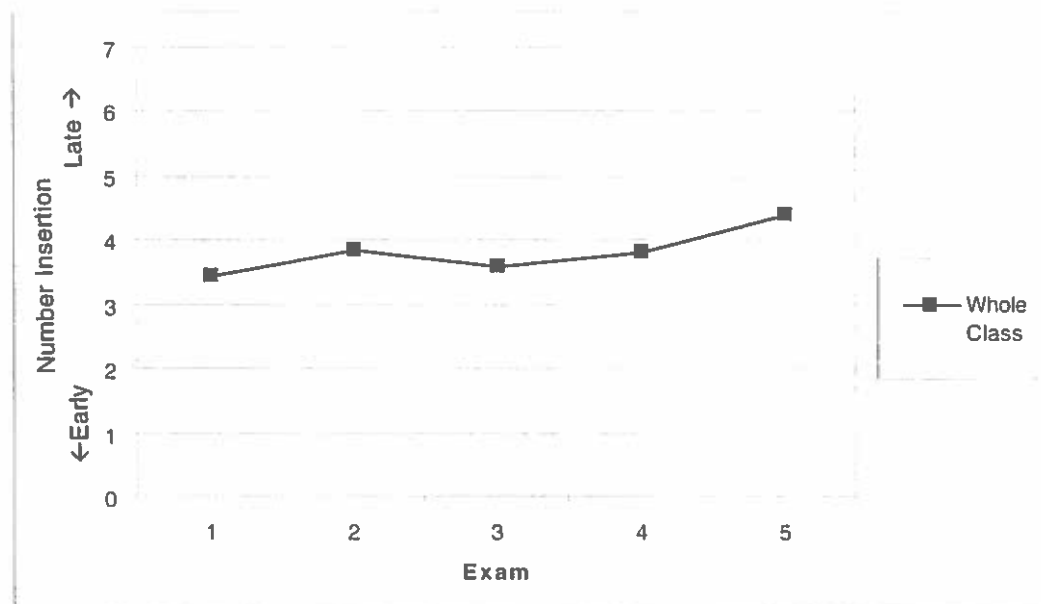


Figure 5.10: Number Insertion Development Over Two Quarters

The beginning average class number insertion score was about 3.5, indicating that the students tended to insert the numbers somewhere between immediately after the first specific qualitative equation and after a small amount of algebra. The final score was 4.4, indicating that students on the average were doing about half the algebra before inserting numbers. The difference between these points is significant at the  $p < .0001$  level, so students made a definite shift from early number insertion to inserting the numbers later in the problem solving process. There certainly was still room for improvement, but the students moved in the right direction, becoming more expert-like in this aspect of physics problem solving.



As before, a stratified analysis was also carried out. Given that the number insertion scale is wider than the other evaluation scales previously discussed, there are more strata in this analysis. The behavior of the number insertion strata is shown in Figure 5.11, below.

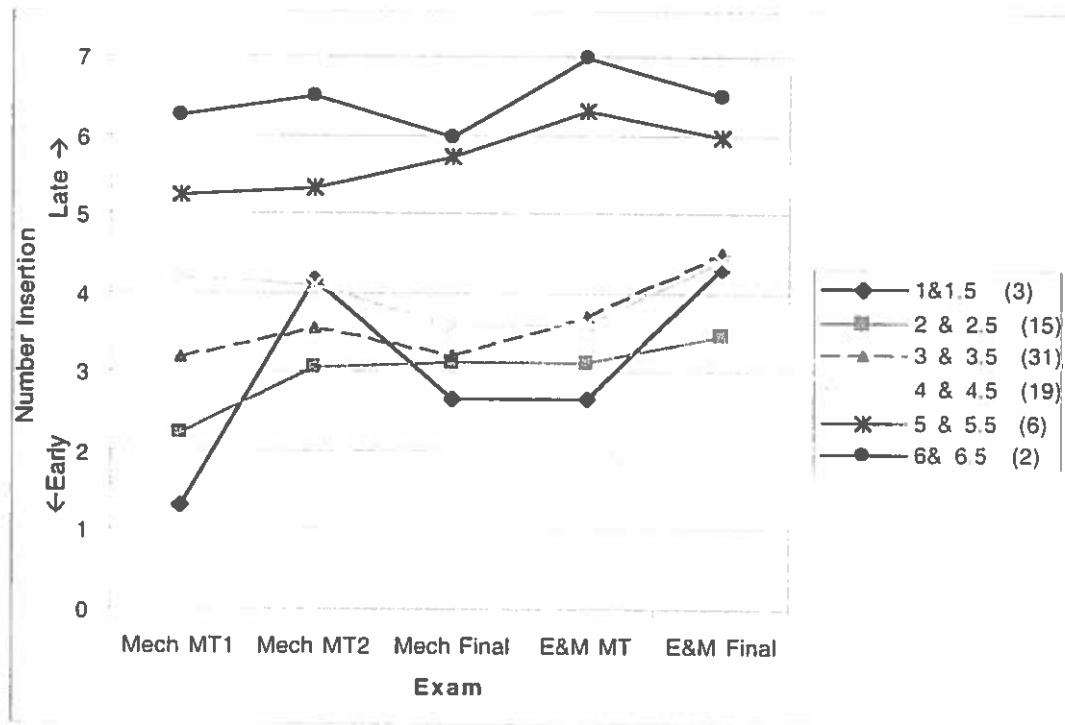


Figure 5.11: Stratified Analysis of Number Insertion

Some of the trends seen before are also observed here, particularly the instability in the lowest stratum. Once again, it seems that the groups, while initially distinct, merge

somewhat by the end of the second quarter. Note also that each stratum has improved in this area, and with the exception of two of them, significantly so.

### **Use of Words**

Another characteristic of expert problem solving that is not present in novice problem solving is the writing of words to explain the process the solver is going through. It was hypothesized that students would use words more often as the physics sequence progressed. A simple analysis was done of each solution, counting not the number of words, but the number of instances where words were written to clarify, explain, or label. Once again, these numbers were averaged over the class and plotted, resulting in Figure 5.12, on the next page.

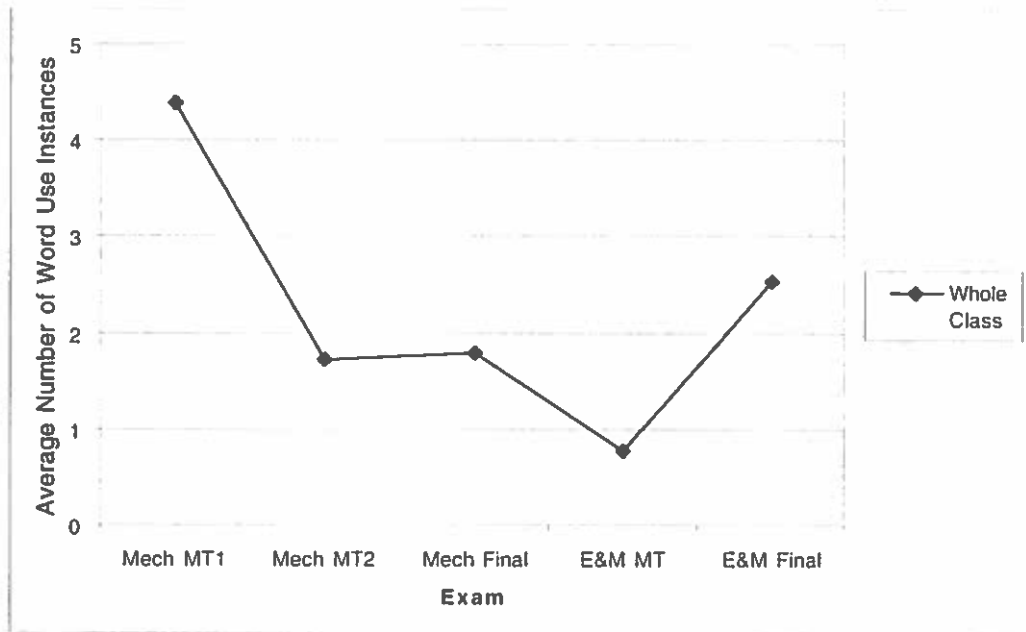


Figure 5.12: Word Usage Development Over Two Quarters

The students began by using words about 4.4 times per solution on the first exam. This number dropped significantly on subsequent exams, reaching a low of less than one instance of words per solution on the electricity and magnetism midterm. On the final exam, the number rose significantly ( $p < .0002$ ) to slightly over 2.5 word instances per solution. The final data point is significantly lower than the initial data point ( $p < .0001$ ), which is not the desired direction for this development trend, but the significant rise from the middle data points suggests a possible turn-around in this characteristic.

The stratified analysis of word usage is shown on the next page in Figure 5. 13.

Note that the top stratum contains only one exceptional student, who is certainly expertlike in this aspect of his problem solving.

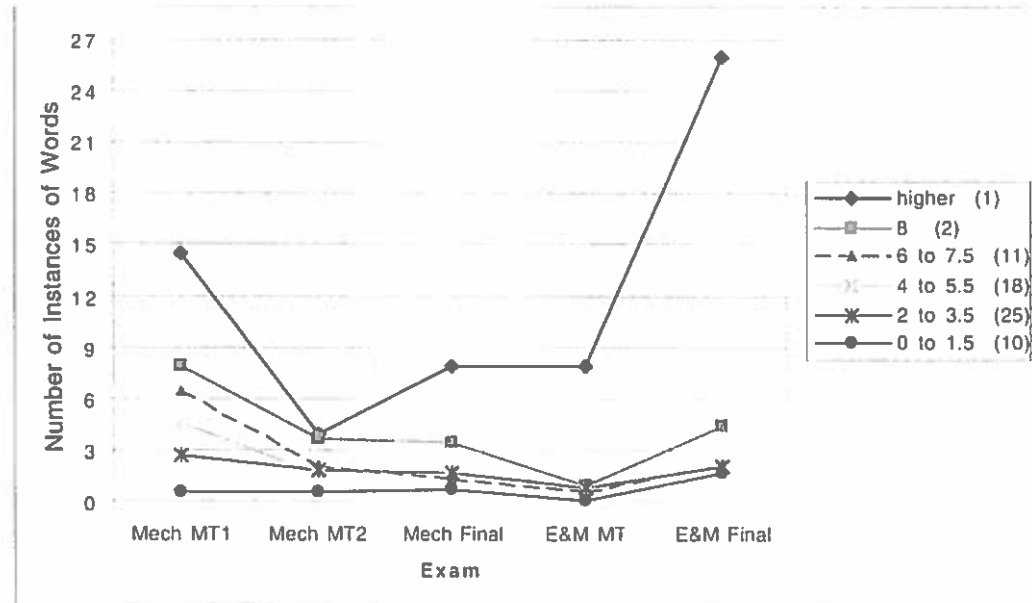


Figure 5.13: Stratified Analysis of Word Usage

Looking at the lower five strata, there is once again a tendency for the class to pull together as the sequence progresses. All of these strata except for the very bottom one significantly decrease in the number of instances where words were employed. The lowest stratum actually shows a significant increase. If the lowest performing group of students in this aspect can make such an improvement in an environment where words

were not emphasized, it may imply that discussing or mentioning this quality during instruction could have a significant impact.

### **Fractionation**

The final skill analyzed was fractionation, that is, how well the students broke problems into smaller pieces to solve. This skill differs from the others discussed so far, with the possible exception of diagram usage, in that it has a stronger tie to content knowledge than the others. If one has a poor understanding of conceptual physics, it will be difficult for him or her to properly identify portions of the problem where particular principles apply. Recall that in the expert-novice research it has been found that experts and novices often cue on different features of the problems, and even when they do cue on the same things, novices do not always access the proper physics principles. For these reasons, it was expected that some greater fluctuation would be seen in this skill development, indicating a problem dependence.

Fractionation was evaluated on a ternary scale. If the solution was split into physically appropriate parts, it received a fractionation score of two. Failure to fractionate at all resulted in a score of zero. One point was awarded for improper fractionation. Two of the problems in the prior analysis could not be fractionated, so they are not considered in this section. They are the ball in a circle problem from the second mechanics midterm and the problem from the electricity and magnetism midterm. The average fractionation scores of the sample for the two quarters are plotted on the next page in Figure 5.14.

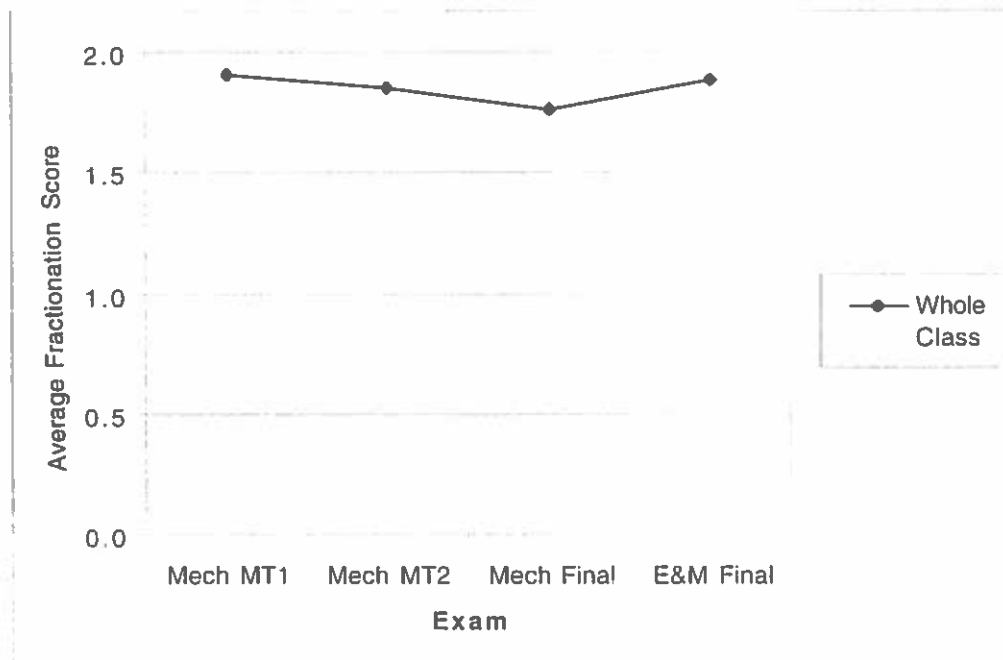


Figure 5.14: Fractionation Development Over the Two Quarters

Notice that the class overall does extremely well in this area, with averages consistently close to 2, and that the scores are quite stable. The high starting average of 1.92 does not leave much room for improvement. The scores for the two mechanics midterms and the second final are not significantly different from each other. However, the dip on the mechanics final is significant when compared to any of the other scores (largest  $p$  value is  $p < .081$  compared to the E&M final). Looking at the problems appearing on that particular exam reveals some potential reasons for this dip.

The first problem was the motorized cart pulling a hanging load. Although the vast majority of the class fractionated this clearly, a non-negligible number of students did not specifically analyze the load. These students did not draw a free-body diagram or write a Newton's second law equation for the load, but usually included some sort of incorrect tension acting on the cart. A few students did not have explicit work for the load, but got the tension correct; these students were counted as having fractionated correctly. The second problem was the bullet imbedding in the wood block and compressing the spring. The scores were slightly lower here due to the well-documented difficulty of students not recognizing that conservation of mechanical energy alone is not an appropriate solution method in this case. As has been seen in other research, students often fail to apply conservation of momentum at the collision.

[8]

The other problems used in this analysis do not contain conceptual elements which make the fractionation difficult. The hanging weight with only a vertical tension causes students to try to lump it in with the analysis of the other object. While the sliding and flying into the mud pit problem of the second mechanics midterm contains a collision where conservation of momentum must be applied, the combination of other elements in this complex problem seems to help all but a few very confused students see it as a fractionation point. In the bullet problem, the relative simplicity of it may cause students to gloss over the implications of the collision. While it is not absolutely certain that the dip in the fractionation score is due entirely to the problems, it is a plausible explanation.

The data do indicate that the fractionation behavior of this class started out extremely expert-like and remained fairly stable throughout the two quarters of measurement. Although it is not the purpose of this study to evaluate the instruction of this physics sequence, it must be noted that splitting a problem into parts is something which the instructor always emphasized and something which the class practiced. Knowing these facts, the high performance in fractionation should have been expected.

### **Summary**

The whole class and stratified analyses of several key physics problem solving skills revealed that there is indeed change in problem solving behavior of students in the Freshman Engineering Honors sequence. Diagram usage, both free-body and energy bar chart, started out at a non-negligible level, dipped, rose again, and then began to drop off. This was fairly close to the predicted behavior. The energy bar chart usage was always lower than the free-body diagram usage. When writing initial equations, the students showed an increase in the number of algebraic equations written, but actually declined in the generality of them. When it came to inserting numbers into algebraic equations, the students improved from doing almost no algebra before insertion to doing about half of the algebra first. Word usage dropped significantly during the two-quarter sequence, which was not the desired behavior. Fractionation was consistently good throughout the two quarters. The stratified studies often showed that even when the class began the first quarter in several distinct groups, they tended to end the second quarter at similar levels. This “clumping” seems to indicate that most of the problem solving behavior observed at the end of the sequence is not purely due to



differing abilities of students upon entering college. It also implies that the instruction, although not specifically targeting most of these skills, had some effect on student problem solving behavior. Some further studies should be done to identify the elements of the instruction that have the greatest impact on student problem solving, so that these can be used as a starting point for developing effective teaching strategies.

## ENDNOTES FOR CHAPTER 5

- [1] Patricia Heller and Mark Hollabaugh, "Teaching Problem Solving Through Cooperative Grouping. Part 2: Designing Problems and Structuring Groups," *American Journal of Physics* **60**, 7, 637-644 (1992).
- [2] Alan Van Heuvelen, *Active Learning Problem Sheets* (Hayden-McNeil Publishing, Plymouth, MI, 1996).
- [3] Alan Van Heuvelen, "Learning to Think Like a Physicist: A Review of Research-Based Strategies," *American Journal of Physics* **59**, 10, 891-897 (1991).
- [4] Jim Minstrell and Virginia Stimpson, "Instruction for Understanding: A Cognitive Process Framework," December 29, 1986.
- [5] Reed Gamble, "Force," *Physics Education* **24**, 79-82 (1989).
- [6] Michael McCloskey, "Naïve Theories of Motion," in *Mental Models*, edited by Dedre Gentner and Albert J. Stevens (Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, 1983), 299-324.
- [7] Herbert A. Simon, "Problem Solving and Education" in *Problem Solving and Education: Issues in Teaching and Research*, edited by D. T. Tuma and F. Reif (Lawrence Erlbaum Associates, Hillsdale, NJ, 1980).
- [8] Xueli Zou, *The Use of Multiple Representations and Visualizations in Student Learning of Introductory Physics: An Example From Work and Energy*, (unpublished doctoral dissertation, The Ohio State University, 2000).

## **CHAPTER 6**

### **APPLICABILITY AND VALIDITY OF EXPERIMENTAL TECHNIQUES**

As stated in the introductory chapter, one of the research questions under investigation was whether looking at copies of student exam solutions was a useful technique for tracking problem solving behavior. In the project's initial stages, a large list of potential areas for exploration was generated. It was found that some aspects of problem solving could be accurately measured using the written exams and others were better explored via the Diana and card sorting-like exercises. Additionally, some areas were not well measured by either of these means. This chapter outlines the findings about the applicability of the research techniques that were used. The validity and reliability of the measures will also be discussed.

#### **Written Exam Analysis**

The written exams provided a vehicle for gaining a large amount of data; without them, tracking the class's behavior would not have been possible. Figure 6.1, on the next page, shows the evaluation sheet originally designed for this analysis.

Analogies	
Diagrams (Count, Type, Agree)	
Checking Answers	
Approach ME vs FW Initial Qual/Quant Blind eqs? Spec vs. gen eq When #'s inserted	
Persistence	
Surf Feat vs. Deep Str	
Incorrect Constraints	
Fractionation	
Will restart/revise?	
Multiple Approaches	
Choose best soln?	
Extra info Logical progression Solving right prob Knowledge needed Making a plan	

Figure 6.1: Evaluation Sheet for Written Exam Analysis

Areas which could be evaluated by written exam analysis were the ones discussed in Chapter 5: diagram usage, algebraic/numerical initial equations, specific/general equations, number insertion, fractionation, and use of words. Note that there is no specific listing for words in Figure 6.1. The idea to analyze words separately grew from the logical progression category.

Other areas listed in Figure 6.1 could have also been analyzed using these means, but the solutions themselves did not provide much opportunity for this. These areas included blind equations (equations which were written on the paper but never used), incorrect constraints, restarting/revising, multiple approaches, choosing the best answer from multiple approaches, using extra information, solving the right problem, and possessing the knowledge needed. The lack of variance in the majority of these areas was most likely due largely to the above-average abilities of these students. There simply was not much evidence of blind equations, incorrect constraints, revision, misunderstanding the question, or not knowing the physics required. If the analysis were repeated for a more general class, it is likely more of these instances would be observed.

After some analysis, it was clear that certain areas were not appropriately measured using this technique. In the areas of analogies, simple checking of answers, and making a plan, there was potential for useful information, but rarely was any seen. Considering specifically the aspects of checking answers and making plans, it is entirely possible and probable that many students did so without writing anything related to this process on the paper. Such is also the case with analogies, although since this has been observed as a primarily expert technique, it is unlikely that many students made them. In the 38 think-aloud interviews conducted to check the validity of these methods, only one student came close to making an analogy, and that was just a statement that he had solved a similar problem on an exam. These think-aloud interviews will be discussed in more detail later. In any case, an analysis of written solutions will not accurately describe behavior in these three areas.

Since the only prior study relying on written exam data was Foster's [1], and he evaluated broader aspects of problem solving, some validity and reliability checks of these techniques are warranted. It is certainly possible that some of the effects described in Chapter 5 have a strong dependence on the actual problems used in the evaluation. Another potential problem could be introduced by the researcher's biases. Several checks were performed to address these issues.

If the methods developed to score the exam papers are indeed valid, experts should outscore students. A group of eight physics instructors (different from those who did the interviews) solved the spring problem used in the interviews; these solutions were evaluated in the same manner as the written exam papers of the students. A comparison of average scores between the two groups is shown below in Table 6.1. The numeric scales here are the same as the ones discussed in the previous chapter.

	Free Bodies (0 or 1)	Bar Charts (0 or 1)	Algeb/ Numeric (1 to 5)	General /Specif (1 to 5)	Number Insert (1 to 7)	Frac (0 to 2)	Word Usage
Experts	0.00	0.13	5.00	2.88	6.25	2.00	4.63
Students	0.00	0.39	4.80	2.11	3.76	1.83	1.28

Table 6.1: Comparison of Expert/Novice Ratings on Spring Problem

The instructors did indeed start with algebraic equations more often (indeed, all the time), start with more general equations, insert numbers much later (mostly at the very end of the algebra), fractionate properly all the time, and write more words. The

students tended to draw bar charts more often, which was expected for the reasons outlined in Chapter 5. The instructors were also asked to solve the related problem involving the sliding block with friction. The ratings of these solutions were comparable to those for the spring problem, as shown below in Table 6.2. These two tests support the validity of the written solution evaluations.

	Free Bodies (0 or 1)	Bar Charts (0 or 1)	Algeb/ Numeric (1 to 5)	General /Specif (1 to 5)	Number Insert (1 to 7)	Frac (0 to 2)	Word Usage
Sliding	0.50	0.00	5.00	2.63	6.25	2.00	5.13
Spring	0.00	0.13	5.00	2.88	6.25	2.00	4.63

Table 6.2: Comparison of Expert Ratings on Two Different Problems

Evaluator reliability was tested by selecting a representative sample of solutions to the spring problem and giving them, along with a copy of the rubrics, to a science education faculty researcher. She evaluated them in the areas of algebraic/numerical initial equations, general/specific initial equations, number insertion, and word usage. The comparisons between the two evaluators are shown in Appendix D. The evaluators' scores were identical in the algebraic/numerical aspect and averaged the same on the general/specific ratings. In the area of number insertion, the researcher's rating was about one third of a point higher than the external evaluator; only twice did the two ratings differ by more than one point on the seven-point scale. When looking at the use of words, the researcher averaged less than a third of a point higher than the

external evaluator. In this category, the two evaluators agreed on 25 of the 28 solutions. The difference between the two scorings came entirely from a difference in counting diagram labels; the external evaluator did not count them, while the researcher did. The results of the evaluator reliability tests show that researcher bias did not play a role in arriving at the conclusions of this study.

A final reliability check was undertaken specifically to address the possibility that the effects seen were due to variations in the problems evaluated. It was conceivable that the students might be more likely to exhibit certain behaviors while solving some problems than others. This effect will be referred to as problem dependence. In his work, Foster concluded that it was essential to use more than one problem from any exam. However, he was evaluating issues with a stronger tie to conceptual understanding. [1] The aspects of problem solving investigated here do not, for the most part, have a specific link to the physics concepts being tested. To examine potential problem dependence, solutions from the same exam were compared to see how consistent the skill scores were. Two problems from each of the mechanics exams were selected for this. The scores were compared for general/specific equations, algebraic/numerical equations, number insertion, and word usage. These results are summarized in Table 6.3 on the next page. The |Diff| columns indicate the differences between the ratings for the two problems on the same exam.



	Mechanics MT1			Mechanics MT2			Mechanics Final		
	Elevator	Hiking	Diff	Slide	Ball	Diff	Cart	Bullet	Diff
Gen/Spec	4.17	3.49	0.68	2.94	3.67	0.73	3.30	2.11	1.19
Numb. Ins	3.33	3.63	0.29	3.66	4.09	0.43	3.45	3.76	0.31
Alg/Num	4.55	4.22	0.32	4.66	4.36	0.30	4.67	4.80	0.13
Words	6.17	2.64	3.53	2.96	0.52	2.43	2.30	1.30	1.00

Table 6.3: Intra-Exam Comparison Summary

It appears that there is some problem dependence in the areas of words used and general/specific equations. A closer look at the word usage shows that students used more words when the problem involved multi-concept fractionation, which should not be too surprising. Words aid in explaining how a problem was split into subproblems and which physics principles were applied to each part. The generality or specificness of the initial equations also appears to depend on the problem itself, and this is more of a surprise. A possible explanation for this finding is that students are more likely to write specific equations for kinematics problems. While this was certainly true for some subset of the class, no consistent behavior was found on a class-wide basis to explain this dependence.

On the other hand, the number insertion and algebraic/numerical scores were quite consistent within exams, always differing by less than half a point and usually differing by a third of a point or less. This agreement certainly should have been expected for the algebraic/numerical measures, as one would not expect a context

dependence for this skill. Initially, the number insertion independence of problem might be surprising; as said earlier, there are cases where problems are solved more easily by inserting numbers. However, recalling that none of the problems in this study fit into that category, these results are exactly what should have been expected. The intra-exam reliability tests indicate some areas where there might be a slight context dependence, but not enough to seriously question the results reported earlier. Specifically, the number insertion and algebraic/numerical areas look very reliable.

To summarize, the written exam analysis techniques are useful in evaluating skills such as diagram usage, whether initial equations are general or specific, whether initial equations are qualitative or quantitative, number insertion, word usage, and fractionation. There are other areas where the exam can provide useful information about student problem solving skills, such as checking an answer or making a plan. However, the absence of written evidence does not confirm that the student did not utilize a particular skill. Evaluation of expert solutions using these written exam techniques showed these methods to be valid. The reliability of them is also fairly good, as the external evaluator test showed. Some of the skills evaluated have a slight problem dependence which may be minimized further in future studies.

### **Think-Aloud Interviews**

As an additional check, 38 students participated in think-aloud problem solving interviews. Each subject was given thirty minutes to solve two problems which were analogous to those from the electricity and magnetism finals. These can be found in Appendix E. Twelve of the interviews were selected as being representative and

transcribed for analysis. Observations of these interviews were compared to interpretations of the written exam solutions to verify the interpretations. Both the results of this validity check and some general findings about the benefits and shortcomings of think-aloud interviews in this kind of research will be addressed.

The think-aloud transcripts contained many of the same observations described in the literature review. The interviews supported the conclusion that these students have not yet progressed to the level where they will make analogies as a solution aid. Recall that Clement found that many experts use analogies when solving problems. [2] As expected, analysis of the interviews yielded more evidence of students making plans and checking answers than was seen on the exams. Seven of the twelve students specifically stated what physics they were applying at least once. At least four of these students started one of the problems by immediately stating how the problem should be attempted. There were five students who indicated at some point that they had reason to doubt their answer. Several of these subjects then repeated their calculations without writing any new information on the paper. There were also a few instances of subjects checking units without writing anything down. These observations confirm that written solutions are not appropriate for evaluating these particular problem solving skills.

The general benefits of think-aloud interviews do not need to be described in detail here, because so many others have done so already. [3] However, the specific benefits and pitfalls of them in the context of this study will be discussed.

The think-aloud interviews were useful because they illustrated the processes students go through as they solve problems. The researcher had the opportunity to see firsthand how certain items on the paper correspond to particular thought processes. A

segment from one transcript will serve as a brief illustration of how these techniques give the researcher a view of the subjects' thought processes. This subject, Priscilla, was solving the mass spectrometer problem and having some difficulty with the portion of the problem where the electron moves in an arc.

“This is the kind of question where it accelerates around a curve... hm... (looks through equations and discusses some of them) ...Force first will be in this direction.....direction of the velocity....and the field is going to bend it down.... (talks about the energy portion of the problem)...I want it to end up eight centimeters in a half-circle, so that it's kind of forced to go this way. Here's my detector. I need to find out what kind of acceleration can make it turn in a half-circle. It's going to go eight centimeters. (Now the light comes on) Oh! It's going to turn in a half-circle of eight centimeters, so the acceleration is  $v^2/r$ .”

Nothing written on her paper indicated the struggle she went through to identify the correct principle. This revelation of the subject's thought process is exactly why think-aloud interviews were included in the study.

There were some unexpected areas where this technique was not helpful. In hindsight, these really should have been expected. For instance, none of the 38 students wrote any words when solving the think-aloud problems, and one of these subjects had the second highest word usage average of the entire class over the two quarters. Once a little thought is given to the experimental situation, this really is not all that surprising – the subject is telling the interviewer *everything* he or she is thinking, including the thoughts that correspond to any words which might be written on an exam page. A further analysis showed that think-aloud techniques are inappropriate for evaluating any presentation skill. Subjects did a much poorer job of writing a clear, easy-to-follow solution in the interviews than they did on actual exams. Once again, this finding

should not be shocking. In the interview, the researcher is watching the entire process, so the subject may not feel the need to organize it the way he would for a graded exam.

### **Solution Evaluation Tasks**

The Diana and card sorting exercises were developed expressly as part of this study to explore differences between students and instructors regarding problem solving perception. These techniques were successful, as described in Chapters 3 and 4. Since the general results from these tasks align well with the previous differences seen between the two groups, these new methods have validity. [4, 5, 6, 7] The Diana exercise brought an added result which was not anticipated at the time of design. When students picked out specific aspects of Diana's solution to criticize, it reliably revealed one of their own strengths. The difference in focus between the two groups when sorting through the instructor and student solutions fits particularly well with Chi's findings. [5]

The conclusion that the student sample was in a mixed state of novice and expert behavior was further supported by the think-aloud interviews. There were instances of extremely novice behavior, such as "Tom" often trying to pump the interviewer for additional information. David was the classic novice blindly using equations: "Oh, I didn't find the velocity, darn it. So.... (writing) it's  $1/2 at^2$ , which will find me the acceleration, one of, it will find me something, hopefully." There also were students who exhibited very mature problem solving methods, like Priscilla, who calculated an intermediate velocity which she suspected was too large: "I'm going to stick with it, just for the efforts of solving the problem. I'm going to pretend like that's the velocity I

need...” Patricia missed the information about the charge on the ion in the mass spectrometer problem, but that did not bother her initially, stating, “I would like to have the charge, but I don’t have that. I’ll just leave it as q.” She then continued with her work. An interesting indication of the mixed student state is the way in which the subjects reread the problems. Almost all of them did refer back to the statements at some point, some more than others. In some cases, the referral was a desperate attempt to find something, anything, that might have been missed earlier. In many cases, though, subjects after rereading part of the problem would indicate they had found what they were looking for and then make rapid progress on the solutions. This collection of instances in the think-aloud interviews confirms the mixed state observations of the Diana and card sorting interviews.

### **Summary**

Each method employed in this problem solving study yielded useful information in domains for which it was appropriate. Think-aloud interviews, long used to probe problem solving processes, are excellent in this respect, but do not give accurate information on presentation skills associated with written problem solutions. Written exam solutions are a practical method for tracking the behavior of many students, but do not give accurate process information. Aspects of problem solving such as checking, making analogies, and making plans are not appropriately evaluated using this technique. The Diana and card sorting exercises probe perceptions of problem solving. All of these methods garner useful information about whether a subject is more novice-like or expert-like in her problem solving.

The validity and reliability of these methods were tested in a variety of ways, particularly cross-checking the new methods with results from think-aloud interviews. Additionally, the written exam method was checked by evaluating expert solutions to make sure experts outperformed the novices, by comparing an external evaluator's results to the researcher's, and by an intra-exam comparison of different problems. Generally, these methods are reliable and valid.

## ENDNOTES FOR CHAPTER 6

- [1] Thomas M. Foster, *The Development of Students' Problem-Solving Skills From Instruction Emphasizing Qualitative Problem-Solving*, unpublished doctoral dissertation (University of Minnesota, 2000).
- [2] John Clement, "Observed Methods for Generating Analogies in Scientific Problem Solving," *Cognitive Science* 12, 4, 563-586 (1988).
- [3] K. Anders Ericsson and Herbert A. Simon, *Protocol Analysis* (Massachusetts Institute of Technology Press, Cambridge, MA, 1996).
- [4] A. Whimbey and J. Lochhead, *Problem Solving and Comprehension - A Short Course in Analytical Reasoning* (Franklin Institute Press, Philadelphia, PA, 1980).
- [5] Michelene T. H. Chi, Paul J. Feltovich, and Robert Glaser, "Categorization and Representation of Physics Problems by Experts and Novices" *Cognitive Science* 5, 2, 121-152 (1981).
- [6] Herbert A. Simon, "Problem Solving and Education" in *Problem Solving and Education: Issues in Teaching and Research*, edited by D. T. Tuma and F. Reif (Lawrence Erlbaum Associates, Hillsdale, NJ, 1980).
- [7] Ray Hyman and Barry Anderson, "Solving Problems" in *The R&D Game: Technical Men, Technical Managers, and Research Productivity*, edited by David Allison (MIT Press, Cambridge, MA, 1969).



## CHAPTER 7

### TRANSFER

#### Context of This Probe

As stated in the literature review, it cannot be certain that any skills have truly developed unless transfer to a new domain is observed. Without confirmed transfer, it is possible that the behaviors under investigation are part of a specific set of techniques only being applied in a limited range of situations. [1] In the context of this study, it was not possible to do a full transfer investigation, but there was an opportunity to do a limited probe into this issue. This involved the analysis of a final exam problem from an engineering statics class. This course, in which seventeen of the original subjects were enrolled, was part of the FEH sequence.

Given the noninvasive nature of this whole study, there were particular restrictions applying to this portion of it. First, the researcher had no input on the exam writing in the engineering course. Second, only the final exam was made available for study. Once the final was inspected, only one problem on it seemed appropriate for this sort of analysis. This problem can be found in Appendix C. Keeping in mind the results of the validity checks, only the algebraic/numerical equations and number insertion areas could be investigated with some amount of confidence in the results. Since the sample size here was much smaller than in the whole class study, and the

problem dependence of word usage and general/specific equations was greater, these areas were omitted from this probe.

## Analysis

Before the engineering solutions were analyzed, the subset of students enrolled in the engineering statics course was compared to the whole sample in the areas under investigation. Figures 7.1 and 7.2 show these comparisons. In each graph, the darker line represents the students in engineering statics, while the lighter line is for the large sample.

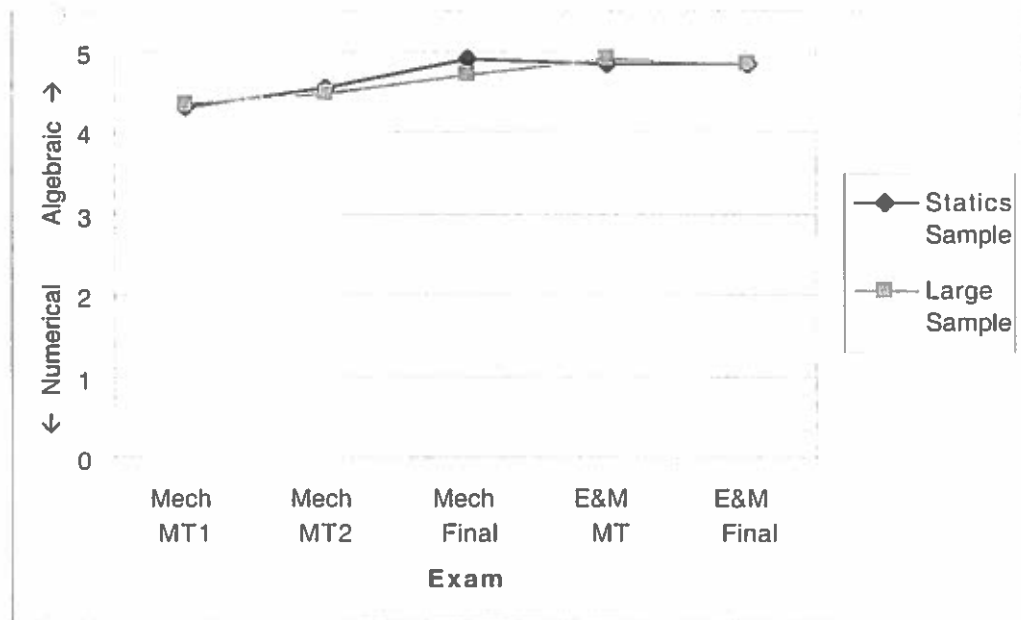


Figure 7.1: Algebraic/Numerical Comparison of Statics Sample to Large Sample

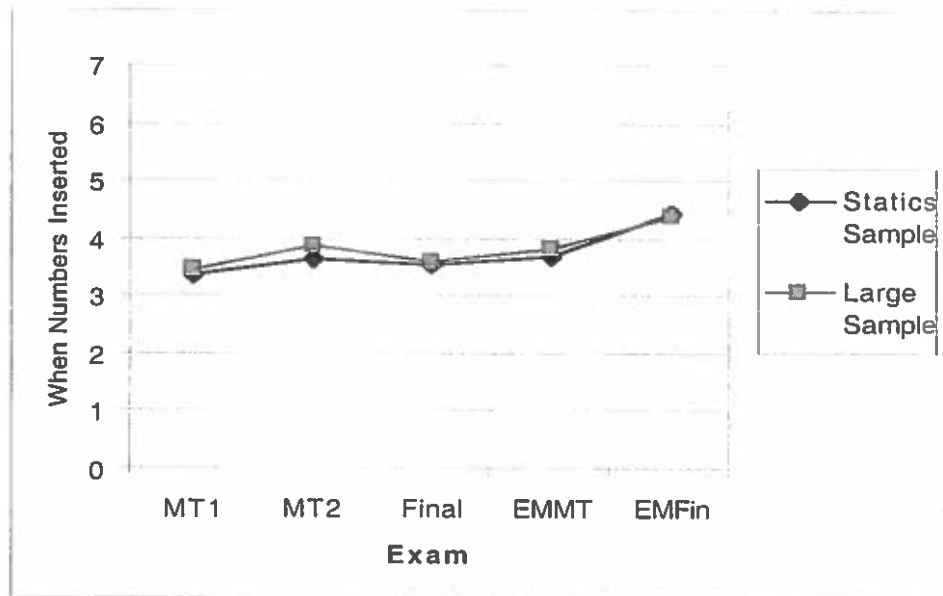


Figure 7.2: Number Insertion Comparison of Statics Sample to Large Sample

In both cases, the trends are fairly similar for both groups. Therefore it is reasonable to assume that the students enrolled in the statics class are representative of the whole class. It is likely that any conclusions reached about this smaller sample would be indicative of the larger one.

Initial results did not show much transfer in the area of algebraic/numerical initial equations, as shown in Figure 7.3, on the next page.

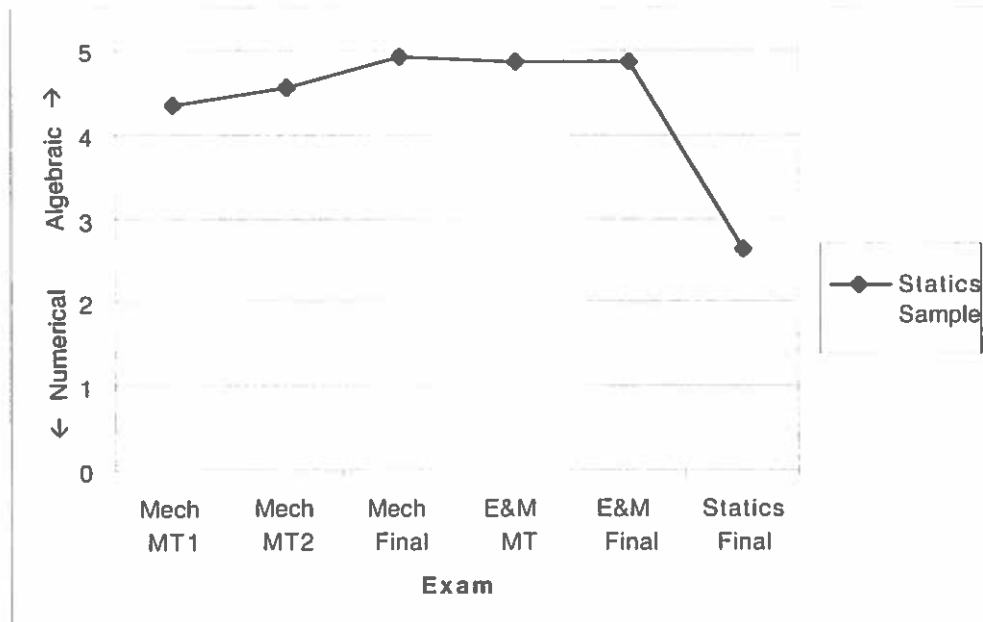


Figure 7.3: Algebraic/Numerical Equation Summary for Statics Sample

This graph indicates that the progress students made during the first two quarters to the point where they consistently began reasoning from algebraic equations was apparently forgotten by the end of the statics course. On the statics final, the students showed a much greater tendency ( $p < .003$  compared to the electricity and magnetism final,  $p < .0023$  compared to the first mechanics midterm) to start with numbers in their equations than they ever had earlier in the year.

Number insertion also showed a significant drop from the end of the second quarter to the end of the third ( $p < .003$ ), as Figure 7.4 shows below.

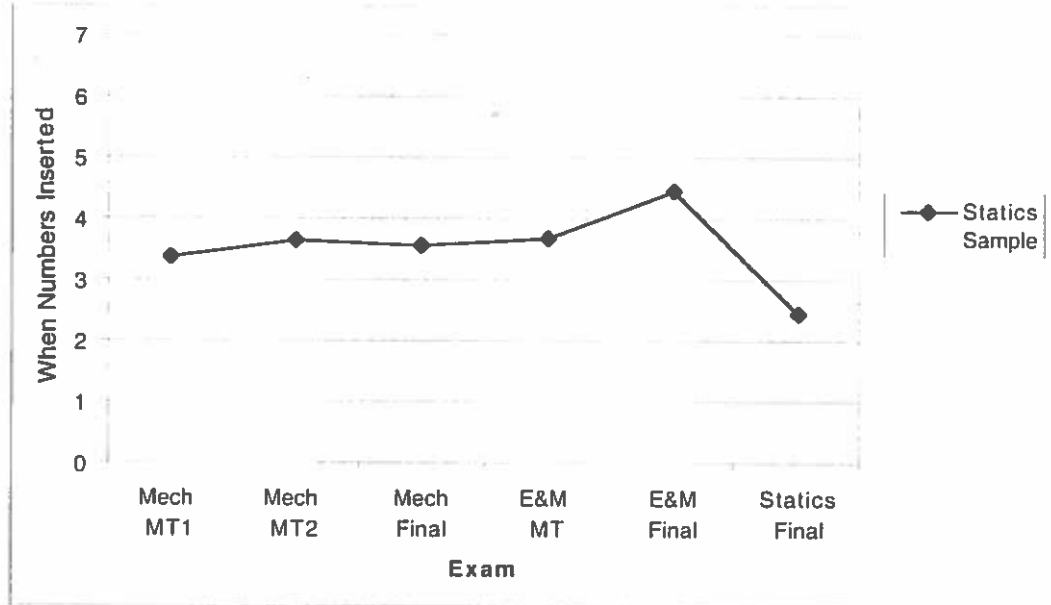


Figure 7.4: Number Insertion Summary for Statics Sample

Initially the results in these two areas appear disappointing. However, viewed together, they may indicate a very positive development. Recall that each of the physics problems was one where it was greatly to the solver's advantage to refrain from inserting numbers until the end. Examining this mechanics problem reveals that it does not fall into this category; it is actually a simpler problem to solve if some numbers are inserted early in the process. It is possible that the significant downward movement in number insertion and algebraic initial equations is a result of students developing an expert-like awareness of when it is preferable to insert the numbers early. However, given the design constraints of this particular study, it is impossible to know which

conclusion is the valid one. Perhaps this can be investigated more fully in a future project.

## **Summary**

The results of the transfer probe in the engineering statics class were mixed. For the most part, the students moved away from the documented expert behaviors in the areas of algebraic/numerical initial equations and number insertion as measured in both classes. It is possible, however, that the students had matured to the point where they could recognize the occasions when altering the usual expert approaches can actually make a problem easier to solve and acted accordingly. Whether this theory is correct or not cannot be determined via this study. Instructional implications from this probe are unclear, because there is still uncertainty in how to interpret these limited results. Additional studies are needed to resolve the transfer issue.

The ambiguity of the results is largely due to the limitations on the probe. The researcher had no input into the writing of the statics exam. Only one problem was selected from the exam for study, and this problem was not ideal for the research being done. Additionally, this problem differed from the other problems used in the first portion of this research in that it was one where it was better to insert numbers early in the solution process. Finally, the sample size was relatively small. In designing future studies, care must be taken to avoid these complications.

## ENDNOTES FOR CHAPTER 7

- [1] John D. Bransford, Ann L. Brown, and Rodney R. Cocking, editors, *How People Learn: Brain, Mind, Experience, and School* (National Academy of Science, 1999).





## CHAPTER 8

### CONCLUSIONS AND FUTURE RESEARCH

#### **Conclusions**

The series of studies described in this dissertation resulted in several new findings in problem solving research. Additionally, most of these results are consistent with those previously published in the literature.

The methods used in this work are reasonably valid and reliable. The Diana and card sorting exercises uncovered fundamental differences between experts and students and the way they view problem solutions. Experts were more aware than students of the complexity of problem solving, realizing that much of the process is not captured by the written solution. When viewing a written solution with deficiencies, the experts had more specific criticisms of these deficiencies than the students. However, any time a student made a specific comment about something lacking in a solution, it reflected a strength in his or her own problem solving. If asked to determine whether a solution was written by a student or instructor, the experts tended to look at how the problem was solved, where the students, on the average, focused more on how the solution was presented. Within the student sample, though, there was some variation. A few accomplished the tasks using methods as sophisticated as the ones employed by the

experts. Others were quite immature, concentrating on superficial characteristics of the solutions.

The analysis of written exam solutions over two quarters revealed several things. First, it is possible to gain useful information about problem solving development via written exams. Second, problem solving skills are not static. The assertion stated in the opening chapters that college has no impact on problem solving skills has been shown to be at least partially incorrect. While it is beyond the scope of this study to see if the development which started in this particular physics sequence carries over to future courses, it is possible to cause at least a temporary change in problem solving behavior. The use of physical representations changed during the year, much as predicted. The students in this study also became more expert in number insertion and their use of initial algebraic equations. Their fractionation was consistently good. The students moved away from expert behavior in the areas of word usage and using initial general equations. The stratified studies showed that the variance in these areas lessened over the course of the two quarters, indicating that the instruction had some sort of effect on problem solving behavior. A limited probe into whether there were indications of potential transfer of some of the more reliably measured skills to another course gave mixed results.

These findings have instructional implications. First, the results of the Diana and card sorting interviews indicate that there is a transitional stage in problem solving development between novice and expert. The student sample in this study was in a mix of novice, expert, and transitional states. Since any given class should contain students in varying stages of problem solving development, instructors should revisit and explain

problem solving strategies throughout the course. As students progress to new levels of maturity in their problem solving ability, they will gain new and different insights due to the instruction. [1] A second strong message from these interviews is that instructors must place more emphasis on explaining the strategies and methods used in solving problems when they work examples in class. When evaluating solutions, the students focused on how a problem was presented more than on how it was actually solved. Perhaps instructors have been doing such a good job of presenting their solutions clearly that students pay more attention to that than the more essential problem solving skills.

Although this study did not attempt to evaluate the problem solving instruction in this particular physics sequence, some inferences can be drawn about that instruction, given the results described in Chapter 5. There were several areas where the class exhibited close-to-expert skill, as well as several where they moved away from the desired behavior. Two particular skills where the students deteriorated were word usage and use of general initial equations. It is recommended that if instructors wish for their students to develop those skills that they develop specific strategies to encourage their growth. The instructor of 131E and 132E does not recall doing anything related to these areas in his instruction. His students did show progress in number insertion, but still needed to make significant strides to reach expert level. This was an area which was emphasized in some of the first-quarter labs and was certainly shown in worked examples, but more explicit attention needs to be given to this skill if it is to develop. The class showed near-expert behavior in the areas of fractionation and using initial algebraic equations. This achievement is probably linked to a couple of specific

instructional elements of the course. The largest factor in the use of algebraic initial equations is most likely homework grading. The grader for 131E and 132E would penalize the students substantially if she could not find an algebraic equation from which the mathematical reasoning began. As stated before, the students solved multi-part problems in lecture, but they also worked on them in recitations, in many laboratories, and on each exam. It is reasonable to assume that this emphasis played a role in the excellent fractionation skills which were observed. The overall message to instructors is that if they want specific problem solving skills to improve during their instruction that they must place specific emphasis on them.

### **Future Research**

As should be expected when undertaking a large exploratory study such as this, many exciting new research questions emerged. Not only are there several new problem solving issues to probe, but there are ideas for improving the techniques developed in this study.

The results of written exam analysis would probably be more reliable if the exam questions were of more uniform complexity. To investigate this claim, one could include two or more problems of roughly equal complexity on an exam or two and do another intra-exam comparison. It is possible that the researcher will be in a situation soon with the kind of control necessary to conduct this experiment.

The card sorting results are fascinating in their own right, but they also suggest several new avenues. To find out whether the average state of the students (focusing on presentation details) was indeed a transitional state or not, a set of interviews should be

done with less experienced students. It is possible that a sample from a non-honors introductory class would suffice. Ideally, the experiment could be done with high school physics students or FEH students very early in their freshman year. Given the nature of the task, the subjects must have a fairly good grasp of energy and momentum concepts.

Additionally, considering the huge amount of data which was accumulated during this research, there are many smaller-scale issues which could be probed. A thorough analysis of the development of some carefully-selected individual students could reveal more about problem solving development. Although the overall class results did not show a significant change during the year, there are certainly students who made significant changes (positive or negative) in the skill areas investigated. It would also be interesting to look at some of these students to see if the development of certain skills are related to each other.

A specific theory regarding diagram usage evolution was raised earlier in this work. The shape of the trend curve for this skill is thought to be due to students beginning with diagrams on the first mechanics midterm, then deciding they had mastered Newton's laws enough that they did not need to draw the diagrams any longer. Then, perhaps due to making errors on the second midterm which could be attributed to omitting the diagram, they drew diagrams again on the final. Another possibility is that some of the problems were simple enough for the students that they did not feel the need to draw a diagram. A specific study to properly explain the observed phenomenon would be beneficial.

Some student interviews with more direct questions about problem solving could be quite revealing. Several different potential formats come to mind. One would be to select a sample of students to interview at several points during the quarter, asking them about what makes for good problem solving and perhaps asking a few questions about some pre-written solutions. This would be different from the Diana and card sorting exercises in that the questions would be more specific. Another possibility would be to copy student solutions to several exams, then have the students come in at a later time during the year to discuss the changes in their own problem solving. Getting the student perspective on why certain strategies change is potentially valuable. Once someone is an expert, he or she is not able to accurately describe the transition. Perhaps talking to students as they are beginning a transition would shed more light on this little-explored stage.

As Reif and Heller cautioned, simply because certain behaviors have been documented to be expert-like does not mean they are optimal. One example is the number insertion issue, discussed previously in this work. A further exploration of how strongly these expert strategies are related to achieving correct answers might revise current views on problem solving.

Ultimately this body of research needs to impact student learning of problem solving in physics. Some controlled studies where specific changes are made in one class to improve problem solving development are a logical step, once some of the issues raised above are more satisfactorily answered. Hopefully, instructional techniques will eventually be developed which will have lasting effects on students and make them proficient problem solvers in their careers.

## ENDNOTES FOR CHAPTER 8

[1] This is certainly true in content areas – think about all the different courses physicists take in mechanics, for example, and how new insights are gained each time a subject is revisited.





**APPENDIX A**

**PROTOCOL AND REPRESENTATIVE TRANSCRIPTS  
FROM THE DIANA TASK**

**Student Protocol:**

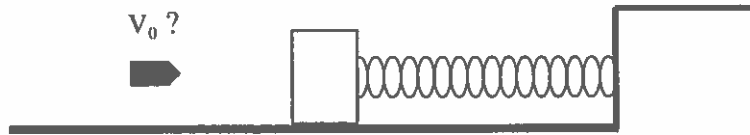
I: This problem solution was written by Diana. Diana is a student just like you, with the same abilities, background knowledge, and time constraints. Her grade in her physics courses didn't matter – she took them pass-fail, so she didn't worry about grades. Her goal was just to understand physics more deeply. She hands this solution to you and asks you if you think it is a good solution to this problem or not. What would you say to her?

**Instructor Protocol:**

I: This problem solution was written by Diana. Diana is a student in your introductory physics class. Her grade in her physics courses didn't matter – she took them pass-fail, so she didn't worry about grades. Her goal was just to understand physics more deeply. She has worked this problem, hands the solution to you and asks you if you think it is a good solution to this problem or not. What would you say to her?

### Spring Problem

A 0.10-kg bullet is fired into a 1.90-kg block. The block is attached to a spring of force constant 1000 N/m. The block slides for 0.40 m while compressing the spring after the bullet runs into the block. Determine the bullet's speed before it hit the block. Assume that the gravitational constant is  $10 \text{ m/s}^2$ . You must show all of the work supporting your answer or no credit will be given.



### Sliding Block Problem

A bullet of mass 50 gram is moving horizontally at a speed of 150 m/s when it strikes a block of wood of mass 1.45 kg at rest on a table and becomes imbedded within it. The coefficient of kinetic friction between the block and the table is 0.25. How far will the block travel after being struck by the bullet before it again comes to rest on the table?

Student #2

S: All right. (reads problem statement aloud) OK. Using energy here, setting the spring constant equal to the kinetic energy and we're trying to find the velocity before. OK. (starts writing on paper) So, um,  $1/2 kx^2$  equal to that,  $1/2 mv^2$ . They [the  $1/2$ 's] cancel. Um, 1000 Newtons times .4 squared is equal to .1... (laughs, pause) .4 is what? That's what I'm doing. Times  $v^2$ , so...all over, all over .1. .4, the root of 1000 over .1 is .4 times the root of 10000. And I don't know if that's 1000 or not, so let's do it. (Gets calculator, mumbles) and it's 100, times .4, 40. That's not what she got. Hm. I don't know if I like her answer. We'll just check mine just to make sure it matches up good. Squared, square root of that, take that out, divide that, do that, plugged it into the calculator right, and right, square root of that. Yeah, I don't think she's right. I don't think we need momentum if you use energy. You don't need to do both; that doesn't make any sense.

I: OK, so that's what you'd tell Diana?

S: Yeah, that's what I'd tell Diana, 'cause once you have this, this gives you your initial velocity of the bullet when it hits this, 'cause you're just converting your kinetic energy of the bullet to spring energy, so it's gonna all drain anyways, so there's no extra energy, unless there's friction, but I don't think there's any friction, 'cause it doesn't tell you anything about that.

Student #5

S: Good solution... Ahh... good solution mean the right answer? (small chuckle, mumbles through problem statement a bit) Looks pretty nice, OK, so  $1/2 kx^2 = 1/2mv^2$ .... We have... OK, hold on, (pause) OK, (mumbles over Diana's work some more, staring intently at page) What's she doing? OK, that's the energy equation. OK. There's energy there. That will be equal to..., you have velocity before you hit the block. No, wait, obviously! First I'd ask her what that velocity is. (points to the 8.95 m/s, pause) 'Cause one...kinetic energy's conserved.... (mumbles over Diana's solution more, pause, grabs paper and works problem with an error, pause) I'd tell her she confused me is what she did. (laughs) She's not displaying her equations or labeling the velocity of the block, mass of what? She has two m's and two v's. (writes some more, pause) It must be part of the box ...I don't know if it's.... How would you get that? I'd tell her I don't understand what she did, and I'm getting a different answer, and I can't follow her work. (laughs)

Student #9

S: (pause while reads over problem) OK, um, I would probably check her math, just to make sure. (calculator, long pause while staring at Diana's solution) Why that .1? (pause) It looks good to me.

Student #30

S: OK. I'll just read over the problem. (reads silently) I think she did work. Let's see.... [unintelligible couple of words] First of all, she didn't show all of her work, so sometimes it's a little bit hard to follow, to see what she's solving for, which masses they are. This one is velocity and...the velocity of the block. Whether or not she, the mass of the, here. (pause) The block and the bullet, and its velocity, must equal the velocity of the bullet times its mass, which I think is right, but you probably want to include units in there and what she did first, I think. (points to an ambiguous  $m$  in an equation) Like  $m$  of the bullet plus  $m$  of the block, multiplying times  $v$  of the total divided by mass of the bullet, and then putting in numbers and the units. Need more work (laughs) to follow it.

Expert #3

S: (pause as S studies question and solution) I don't see, I don't see all the details, so my first question is where did you, how did you find the, uh, speed? Now she wrote down conservation of energy equation, and I guess I'd want to know if that's how she found, uh, the speed. Did she assume that there was no friction? Which is probably a good assumption, depending upon what material has been covered up until that point. Uh, so, I mean, I could, I don't know if you want me to check the numbers in the problem but I would ask her, you know, how she, uh, how she solved it. But if I can assume that she simply, uh, used the maximum compression on the left-hand side of the equation to figure out the initial speed on the right-hand side of the equation, if she did the arithmetic correctly, then, I don't know, and that's OK, uh, see. (pause) Hm. Oh. I don't uh, oh, OK, I have to, I have to back up and, uh, think about this again. OK. (long pause) Well, I'm not, uh, I'm not sure, I mean, do you want me to go, to go through the details?

I: It's up to you.

S: Yeah, OK. I mean, I think in principle she's, uh, addressed the problem correctly. The first thing to do is to find out initially what is the, uh, kinetic energy of the bullet plus the, plus the block. In order to do that, she really had to do this part of the problem first, in some sense. That is, uh, to find the, uh, the speed of the block plus bullet in terms of, uh, the initial, uh, speed of the, of the bullet. Uh, but then, in conserving energy, she knows the compression so that the initial kinetic energy, or the initial speed of the block plus bullet would be this if in fact she used the sum of the two masses, but I don't see that here, so I, you know, that's why I say - Here she's specifically indicating two different masses, uh, but I don't see that over here, so I'm not sure if this number is the correct number. In principle, she's set it up correctly, but she may not have used the correct masses.

#### Expert #7

S: (reads over problem, covering up Diana's solution, then uncovers it.) I see, so she worked the problem in two different ways, or is there some symbols here? I see,  $v$  and  $v_0$ . So she found the speed of the block-bullet system, and then she said that  $mv_0$  had to be the sum of the  $m$ 's  $v$ . (pause) Um, (goes through some more) OK, so you're asking me to tell her what I think of her solution?

I: She asks you, "Is it a good solution?"

S: I think it's a fine solution.

#### Expert #9

S: (pause) OK. "Show *all* supporting work." All the work. All right, she writes down  $1/2 kx^2$  is equal to  $1/2 mv^2$ . So  $mv \dots mv$ . Well, no, it's not a good solution, in my mind. First of all, it's, she's got conservation of momentum, it looks like, in the second part, conservation of energy in the first part, doesn't say anything about that. Comes up with two answers, you don't know which one to take, which is one of my pet peeves, the shotgun approach – write down as many problems as you can, and the right answer is there someplace and leave it for the grader to find it, so no, the direct answer to your question, I don't think this is a good solution.

Expert #10

S: (reads through problem) They don't tell you it's going sideways, but the drawing shows it going sideways. Well, so assuming that, yeah...oh, check numbers. Where did this 8.95 come from? At the very worst, you put in that. Well, so the first answer is (chuckle) um. You want me to solve the problem first?

I: You can do, you can do whatever you want.

S: OK. (solves problem quickly, with an error) So, um, yeah, so how would I answer? Well, OK, I would first ask her, "What the hell is this  $v$ ?" except I wouldn't say, "hell." (laughs) Um, oh, sorry, it was my mistake. Sorry, I made a mistake. Oh, no I didn't. No I didn't. That's fine. Um, oh, yeah, sorry, I made a mistake, I didn't prepare for this question, OK, good. "Um, so what I have to do is say the true answer is, um, OK (resolves problem, compares with Diana's solution a bit) So I would say it's a correct solution, but not particularly clear. Um, and if I were grading it, I would just sort of look and say, "Oh, OK, right answer, check." Um, but you're asking, what, um?

I: Diana said, "Is this a good solution?"

S: Well, OK, I would probably say it's correct, but it could probably use a little explaining. Add some words. Um, in the 260 series, over and over again I say, I almost emphasize the words more than the solution itself. Um, so I would tell her, you know, for the sake of being able to trace your steps earlier on, you know, I would say, um, begin by saying up here, beginning at the top and saying we are equating momentum, and therefore we know how fast this is going in terms of  $v_0$  and then we know the energy and we equate that and so on. As written on here, the first thing it said is we're going to equate some energy. OK, and then the next step is saying there's some velocity, and I'm left saying, "What is the velocity?" So I'm going through sort of forensics here. I'm tempted to say that the student would have started down here, written some of it, and then reached the end of the page and said, "Oh, I think I'll go to the top and continue on writing." Uh, um, because, I mean, as written, it's not clear at all where did this thing, this number, you know. At least it's not clear to me where the number popped out of.

## APPENDIX B

### PROTOCOL, PROBLEM SOLUTIONS, AND REPRESENTATIVE TRANSCRIPTS FROM THE CARD SORTING TASK

Protocol (same for students and instructors):

I: What we have here are 12 solutions to the same problem, which was actually given on a 131E exam. Some of these solutions were written by other FEH students; some are by physics instructors (who were told to solve the problem as if they were taking an exam.) I'd like you to tell me which ones you think are from students and which ones you think are from instructors, and tell me why.

There are two parts to this problem.

- ① the collision
- ② the compression of the spring



- ① Let bullet & block be the system. Then momentum is conserved.

$$m v = (M+m) V$$

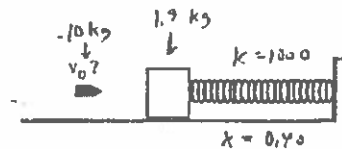
- ② Total energy is conserved... after the collision

$$\frac{1}{2} (M+m) V^2 = \frac{1}{2} k x^2 \quad x = \text{maximum displacement} = 0.40 \text{ m}$$

$$\text{So, } V^2 = \frac{k}{M+m} x^2 = \frac{(1000 \text{ N/m})(0.40 \text{ m})^2}{2.00 \text{ kg}} = 80 \text{ m}^2/\text{s}^2$$

$$\text{Thus } v = \left( \frac{M+m}{m} \right) V = \left( \frac{2.00 \text{ kg}}{0.10 \text{ kg}} \right) (8.94 \text{ m/s}) = 20 \cdot (8.94 \text{ m/s}) = 179 \text{ m/s}$$

Figure B.1: Spring Solution O



$$\frac{kx^2}{2} = K_{\text{ref}}$$

$$\frac{(1000)(0.40)^2}{2} = K_{\text{ref}}$$

$$K_{\text{ref}} = 80$$

$$\frac{m_1 + m_2}{2} v^2 = 80$$

$$(m_1 + m_2) v^2 = 160$$

$$(0.10 + 1.9) v^2 = 160$$

$$v^2 = \frac{160}{2}$$

$$v^2 = 80$$

$$v = 8.94 \text{ m/s}$$

$$0.10(v_0) + 0 = \frac{m_1 + m_2}{2} (8.9)$$

$$0.10 v_0 = (2)(8.9)$$

$$v_0 = \frac{(2)(8.9)}{0.10}$$

$$v_0 = 178 \text{ m/s}$$

Figure B.2: Spring Solution P





FROM 1 TO 2:

$$K_0 + \cancel{v_{0i}^2} + \cancel{v_{0f}^2} + \cancel{K} + \cancel{v_{0f}^2} = \cancel{K} + \cancel{v_{0i}^2} + v_s + \cancel{K} + \cancel{v_{0f}^2}$$

$$K_0 = v_s$$

$$\frac{1}{2} \cdot (m_{20} + m_{2i}) \cdot v_f^2 = \frac{1}{2} \cdot K \cdot d^2$$

$$v_f = \sqrt{\frac{K \cdot d^2}{(m_{20} + m_{2i})}}$$

$$m_{20} = 0.1 \text{ kg}$$

$$m_{2i} = 1.7 \text{ kg}$$

$$K = 1000 \text{ N/m}$$

$$g = 10 \text{ m/s}^2$$

$$d = 0.4 \text{ m}$$

$$v_f = ?$$

$$v_0 = ?$$

COLLISION AT 1:

$$P_i = P_f$$

$$m_{10} \cdot v_0 = (m_{10} + m_{2L}) \cdot v_f$$

$$v_0 = \frac{(m_{10} + m_{2L}) \cdot v_f}{m_{10}}$$

$$v_0 = \frac{(m_{10} + m_{2L})}{m_{10}} \cdot \sqrt{\frac{K \cdot d^2}{(m_{20} + m_{2i})}}$$

$$v_0 = \frac{(0.10 \text{ kg} + 1.70 \text{ kg})}{0.10 \text{ kg}} \cdot \sqrt{\frac{1000 \text{ N/m} \cdot 0.40 \text{ m}^2}{(0.10 \text{ kg} + 1.70 \text{ kg})}}$$

$$v_0 = 178.7 \text{ m/s}$$

Figure B.3: Spring Solution Q

$$\frac{1}{2} k x^2 = \frac{1}{2} m v^2$$

$$1000 (0.4)^2 = 1.90 v^2$$

$$160 = 1.90 v^2$$

$$84.21 = v^2$$

$$v = 9.18 \text{ m/s}$$

$$mV = Mv$$

$$10(V) = 1.90 \text{ kg} (9.18 \text{ m/s})$$

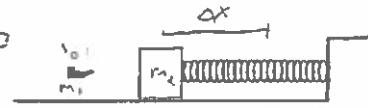
$$10V = 17.43$$

$$V = 174.36 \text{ m/s}$$

Figure B.4: Spring Solution R

Collision }  $m_1 v_1 + m_2 v_2 = (m_1 + m_2) v_3$

$$v_3 = \frac{m_1 + m_2}{m_1} v_1$$



$\mu$  not given, assume zero

Side }  $\frac{1}{2} k x^2 = \frac{1}{2} k x^2 + \frac{1}{2} k x^2$

$$\frac{1}{2} (m_1 + m_2) v_3^2 = \frac{1}{2} k x^2 + \frac{1}{2} k x^2$$

$$v_3 = \sqrt{\frac{k x^2}{m_1 + m_2}}$$

put together

$$v_3 = \frac{m_1 + m_2}{m_1} \sqrt{\frac{k x^2}{m_1 + m_2}}$$

$$= \frac{2.0 \text{ kg}}{1.0 \text{ kg}} \sqrt{\frac{1000 \text{ N/m} (0.4 \text{ m})^2}{2.0 \text{ kg}}}$$

$$= 179 \text{ m/s}$$

Figure B.5: Spring Solution S

$$\frac{1}{2} m v^2 = \frac{1}{2} k x^2$$

$$\frac{1}{2} (m_1 + m_2) v^2 = \frac{1}{2} k x^2$$

$$v = \sqrt{\frac{k x^2}{m_1 + m_2}}$$

$$P_1 + P_2 = P_f$$

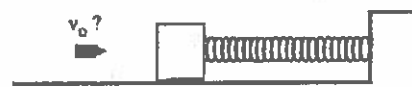
$$m_1 v_1 + m_2 v_2 = (m_1 + m_2) v$$

$$v_2 = 0$$

$$m_1 v_1 = (m_1 + m_2) v$$

$$v_1 = \frac{(m_1 + m_2) v}{m_1}$$

$$v_1 = 180 \text{ m/s}$$



$m_1 = 1$   
 $m_2 = 1.9$   
 $k = 1000$   
 $x = 0.4$

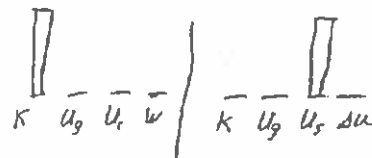


Figure B.6: Spring Solution T

$$\frac{1}{2} m v^2 = \frac{1}{2} k x^2$$

$$\frac{1}{2} \cdot 1 v^2 = \frac{1}{2} 1000 \cdot .4^2$$

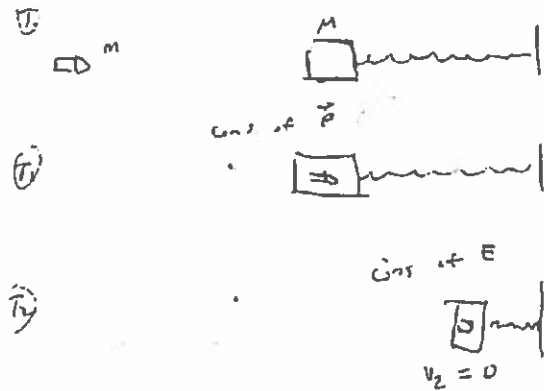
$$v = 40 \text{ m/s}$$

$$| -v + 0 = 2 \cdot 40$$

$$g_0 = .1 \text{ y}$$

$V = 800 \text{ m/s}$

Figure B.7: Spring Solution U



Cons of  $\vec{p}$  since is in first problem: still formula

$$m v_0 = (m+M) v_1$$

$$(I) \quad v_1 = \frac{m v_0}{(m+M)}$$

Now use cons of energy  
 $T_1 + U_1 = T_2 + U_2$

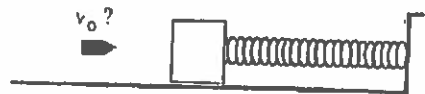
$$T_1 = \frac{1}{2} (m+M) v_1^2$$

$$U_2 = \frac{1}{2} k x^2$$

$$\frac{1}{2} (m+M) v_1^2 = \frac{1}{2} k x^2$$

plus in (I)

$$\cancel{(m+M)} \frac{m^2 v_0^2}{(m+M)^2} = k x^2$$



$$v_0^2 = \frac{(m+M) k x^2}{m^2}$$

units check ✓

$$v_0 = \sqrt{(m+M) k} \frac{x}{m}$$

plus in

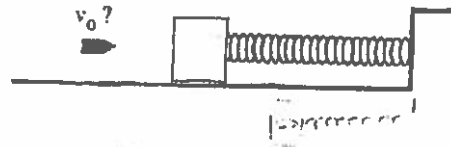
$$v_0 = \sqrt{(20 \text{ kg})(1000 \text{ N/m})} \frac{.4 \text{ m}}{.1 \text{ kg}}$$

$$= 179 \text{ m/s}$$

seems slow for a bullet, but ok.

Figure B.8: Spring Solution V

$$\begin{aligned}
 U_b &= ? \\
 m_b &= 0.10 \text{ kg} \\
 m_2 &= 1.90 \text{ kg} \\
 k &= 1000 \text{ N/m} \\
 x_m &= 0.4 \text{ m}
 \end{aligned}$$

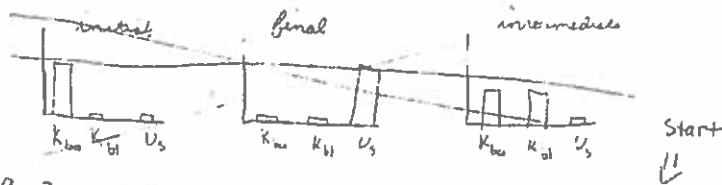
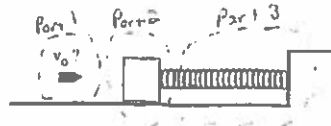


$$\begin{aligned}
 \Delta E_s &= K E_s \\
 \frac{1}{2} (m_b + m_2) v^2 &= \frac{1}{2} k x_m^2 \\
 v^2 &= \frac{k x_m^2}{m_b + m_2} \\
 v &= \sqrt{\frac{k x_m^2}{m_b + m_2}} \\
 &= \sqrt{\frac{(1000 \text{ N/m})(0.4 \text{ m})^2}{0.10 \text{ kg} + 1.90 \text{ kg}}} \\
 &= 8.94 \text{ m/s}
 \end{aligned}$$

$$\begin{aligned}
 E_{2i} &= p_{2i} \\
 m_2 v_2 &= U_s (m_2 - m_b) \\
 v_2 &= \frac{U_s (m_2 - m_b)}{m_2} \\
 &= \frac{8.94 \text{ m/s} (2.0 \text{ kg})}{1.90 \text{ kg}} \\
 &= 178.9 \text{ m/s}
 \end{aligned}$$

Figure B.9: Spring Solution W

$$\begin{aligned}
 m_1 &= 0.1 \text{ kg} \\
 m_2 &= 1.9 \text{ kg} \\
 k &= 1000 \text{ N/m} \\
 d &= 0.40 \text{ m}
 \end{aligned}$$



Part 3: Inelastic collision, conservation of momentum

$$m_1 v_1 = (m_1 + m_2) v$$

$$(1) v_1 = (1 + 1.9)(8.94)$$

$$\boxed{v_1 = 178.8 \text{ m/s}}$$

Part 2: Kinetic Energy of bullet/block system

$$\frac{1}{2} M v^2 = U_s$$

$$\frac{1}{2} (m_1 + m_2) v^2 = U_s$$

$$\frac{1}{2} (1 + 1.9) v^2 = 80$$

$$v = 8.94 \text{ m/s}$$

Part 1: Energy of the Spring

$$U_s = \frac{1}{2} k x^2$$

$$= \frac{1}{2} (1,000) (0.4)^2$$

$$\boxed{U_s = 80 \text{ J}}$$

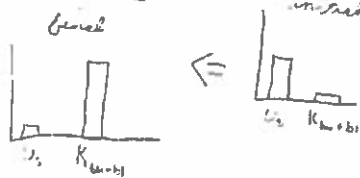


Figure B.10: Spring Solution X

$M_1 v_1 = (m_1 + m_2) v$       $v_2 = \frac{M_1 v_1}{(m_1 + m_2)}$   
 conservation of momentum     elastic collision

$\frac{1}{2} m v_0^2 = \frac{1}{2} k x^2 + \frac{1}{2} (m_1 + m_2) v^2$

elastic collision  $\rightarrow \frac{1}{2} \left( \frac{M_1 v_1^2}{m_1 + m_1} \right) = \frac{1}{2} k x^2$

$\frac{M^2 v^2}{k_2 + m_1} = k x^2$

$.005 v^2 = k x^2$

$.005 v^2 = 160$   
 $v^2 = 32000$   
 $v = 178.9 \text{ m/s}$




Figure B.11: Spring Solution Y

Elastic Collision  
 $m_1 v_1 = (m_1 + m_2) v$   
 $.1 v_1 = (.1 + 1.90) v$   
 $.1 v_1 = (.1 + 1.90)(8.94)$   
 $v = \frac{.2 \cdot 8.94}{.1}$   
 $v \approx 178.89 \text{ m/s} \approx 179 \text{ m/s}$

$\frac{1}{2} M V^2 = \frac{1}{2} k x^2$   
 $\frac{1}{2} (2) v^2 = \frac{1}{2} (10000) (.4)^2$   
 $v^2 = 80$   
 $v = \sqrt{80} \approx 8.9 \text{ m/s} \approx 9 \text{ m/s}$




Figure B.12: Spring Solution Z

This problem has 2 parts: get  $V$  from collisions; find out how it stops

1) bullet & block - momentum of system is conserved.

$$P_i = m v_i = P_f = (m+M)V$$

$$\text{so } V = \frac{m}{m+M} v_i = \frac{50 \text{ g}}{1.45 \text{ kg} + 0.05 \text{ kg}} (150 \text{ m/s}) = \frac{0.05}{1.5} (150 \text{ m/s}) = 5.0 \text{ m/s}$$

2) stopping block - friction

$$\begin{array}{l} \vec{N} \uparrow \\ \vec{F} \leftarrow \\ \vec{W} \downarrow \end{array} \quad \begin{array}{l} N = W \quad (\text{not force vertically}) \\ f = \mu_k N = \mu_k W \quad (\text{net horizontal force}) \end{array}$$

$$-\mu_k m g = m a_x \Rightarrow a_x = -\mu_k g$$

$$v_x(t) = v_{x0} + a_x t \quad (v_{x0} = v)$$

$$\Delta d = v_{x0} t + \frac{1}{2} a_x t^2$$

$$v_x(t) = 0 \Rightarrow t = -v_{x0}/a_x = v/\mu_k g$$

$$\begin{aligned} \Rightarrow \Delta d &= v^2/\mu_k g - \frac{1}{2} v^2/\mu_k g = \frac{v^2}{2\mu_k g} \\ &= \frac{(5.0 \text{ m/s})^2}{2(0.25)(10 \text{ m/s}^2)} = 5.0 \text{ m} \end{aligned}$$

or use conservation of energy

orig KE = thermal energy from friction

$$\frac{1}{2} (m+M) v^2 = \mu_k (M+m) g \Delta d$$

$$\Delta d = \frac{v^2}{2\mu_k g} \quad (\text{same as before})$$

Figure B.13: Sliding Block Solution O

$$m_1 v_1 + m_2 v_2 = (m_1 + m_2) v_f$$

$$0.050 \text{ kg} (150) + 1.45 \text{ kg} (0) = (1.50) v_f$$

$$0.50 (150) = 1.50 v_f$$

$$v_f = 5 \text{ m/s}$$

$$v_f^2 = v_i^2 + 2ad$$

$$0 = 5^2 + 2(-2.5)d$$

$$-25 = -5d$$

$$d = 5 \text{ m}$$

$$m_1 + m_2 = 1.50 \text{ kg}$$

$$F_f = \mu F_N$$

$$F_f = 0.25(15)$$

$$F_f = 3.75 \text{ N}$$

$$F_f = \mu F_N$$

$$3.75 = 1.5 a$$

$$a = 2.5$$

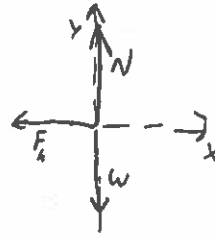
$$F = 0.25(10) = 2.5$$

Figure B.14: Sliding Block Solution P

$$P_b + P_{L1} = P_t$$

$$.05 \cdot 150 + 1.45 \cdot 0 = (1.95 + .05) \cdot v_t$$

$$v_t = 5 \text{ m/s}$$



$$W = 150 \text{ N}$$

$$N = 150 \text{ N}$$

$$F_f = 37.5 \text{ N}$$

$$\sum F_x = -37.5 \text{ N} = 1.5 \cdot a$$

$$a = -25 \text{ m/s}^2$$

$$v^2 - v_0^2 = 2a(x - x_0)$$

$$0 - 5^2 = 2 \cdot (-25)(x - 0)$$

$$x = 1.5 \text{ m}$$

$$v = v_0 + at$$

$$0 = 5 - 25t$$

$$t = .2 \text{ s}$$

$$x = x_0 + v_0 t + \frac{1}{2} a t^2$$

$$x = 0 + 5 \cdot .2 + \frac{1}{2} \cdot (-25) \cdot (.2)^2$$

$$x = .5 \text{ m}$$

Figure B.15: Sliding Block Solution Q

$$m_c = .05$$

$$v = 150$$

$$m_B = 1.45$$

$$\mu = .25$$

$$\textcircled{2} m_{BB} = 1.5$$

$$W = 15$$

$$N = 15$$

$$F = N\mu = 3.75$$

$$F = ma = 3.75 = 1.5(a)$$

$$a = \frac{3.75}{1.5}$$

$$a = 2.5$$

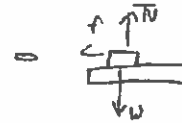
$$\textcircled{1} m_1 v_1 + m_2 v_2 = m_1 v_1' + m_2 v_2'$$

$$m_1 v_1 + m_2 v_2 = v'(m_1 + m_2)$$

$$.05(150) + 0 = v'(1.5)$$

$$v = \frac{.05(150)}{1.5}$$

$$v = 5 \text{ m/s}$$



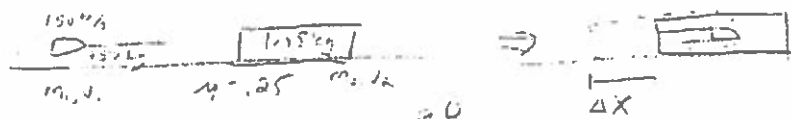
$$\textcircled{3} 2a(x - x_0) = v^2 - v_0^2$$

$$2(2.5)(x) = 0 - (15)^2$$

$$-5x = -225$$

$$x = 45 \text{ m}$$

Figure B.16: Sliding Block Solution R



Collision)  $m_1 v_1 + m_2 v_2 = (m_1 + m_2) v_3$  note:  $v_2 = 0$

$$v_3 = \frac{m_1 v_1}{m_1 + m_2} = \frac{(0.50 \text{ kg})(150 \text{ m/s})}{(1.50 \text{ kg})} = 50 \text{ m/s}$$

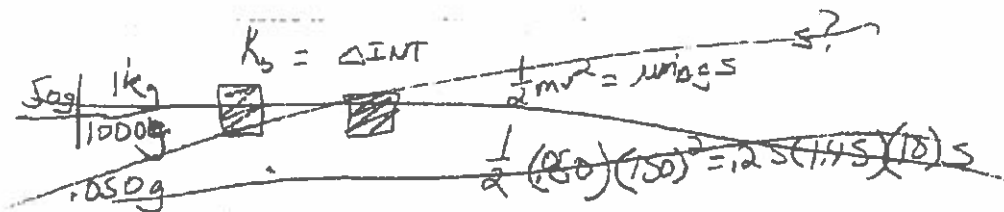
Side)  $v_0 = 50 \text{ m/s}$   $N = (m_1 + m_2)g = (1.50 \text{ kg})(10 \text{ m/s}^2) = 15.0 \text{ N}$

$$F_k = \mu N = (0.25)(15.0 \text{ N}) = 3.75 \text{ N}$$

$$a_k = \frac{F_k}{(m_1 + m_2)} = \frac{3.75 \text{ N}}{1.50 \text{ kg}} = 2.5 \text{ m/s}^2$$

$$\Delta X = \frac{v_0^2 - v_f^2}{2a} = \frac{2500 \text{ m}^2/\text{s}^2 - 0}{2(2.5 \text{ m/s}^2)} = 500 \text{ m} = \boxed{6.82 \text{ m}}$$

Figure B.17: Sliding Block Solution S



$$mv + mv = mv$$

$$0.50(150) + 0 = 1.5v$$

$$v = 50$$

$$K_0 = \Delta int$$


$$\frac{1}{2} mv^2 = mgh$$

$$\frac{1}{2} (50)^2 = 1.5(10)h$$

$$h = 5 \text{ m}$$

Figure B. 18: Sliding Block Solution T





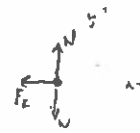
$$P_{\text{inc}} + P_{\text{refl}} = P_{\text{trans}} + P_{\text{abs}}$$

$$0 + M_{\text{air}}(v_{\text{air}}) = (M_{\text{air}} + M_{\text{block}})(v_0)$$

$$.05 \text{ kg} (150 \text{ m/s}) = (.05 \text{ kg} + 1.75 \text{ kg})(v_0)$$

$$7.5 \text{ kg m/s} = 1.8 \text{ kg}(v_0)$$

$$5 \text{ m/s} = v_0$$



$$\sum F_y = N - W = 0$$

$$N = W$$

$$N = (1.75 \text{ kg} + .05 \text{ kg})(g)$$

$$N = 1.8 \text{ kg} (10 \text{ m/s}^2)$$

$$N = 18 \text{ N}$$

$$F_k = \mu N$$

$$F_k = .25 (18 \text{ N})$$

$$F_k = 4.5 \text{ N}$$

$$\sum F_x = -F_k = ma$$

$$-4.5 \text{ N} = (1.8 \text{ kg})a$$

$$-2.5 \text{ m/s}^2 = a$$

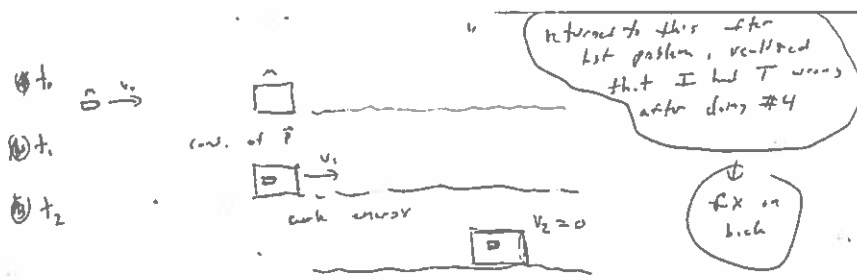
$$2a(\Delta x) = v^2 - v_0^2$$

$$2(-2.5 \text{ m/s}^2)(\Delta x) = 0 - (5 \text{ m/s})^2$$

$$-5 \Delta x = -25 \text{ m}^2/\text{s}^2$$

$$\Delta x = 5 \text{ m}$$

Figure B.19: Sliding Block Solution U



cons of  $\vec{p}$

$$mv_0 = (m+M)V_1$$

$$V_1 = \frac{mv_0}{(m+M)}$$

kinetic energy at  $t_1$ :  $T = \frac{1}{2}mv_1^2 = \frac{1}{2}m\left(\frac{mv_0}{(m+M)}\right)^2 = \frac{1}{2}\frac{m^3v_0^2}{(m+M)^2}$

$T - W = 0$  : Friction  $T = 0$   
 $T - F_f D = 0$   
 $F_f = \mu_k N = \mu_k (m+M)g$

$T - F_f D = 0$   
 $\frac{1}{2}\frac{m^3v_0^2}{(m+M)^2} - \mu_k(m+M)gD = 0$   
 $\frac{1}{2}\frac{m^3v_0^2}{(m+M)^2} = \mu_k(m+M)gD = 0$   
 $D = \frac{1}{2}\frac{m^2v_0^2}{\mu_k g(m+M)^2}$

SCORE

Fix on back

Front

kinetic energy at  $t_1$

$$T = \frac{1}{2}(m+M)v^2 = \frac{1}{2}(m+M)\left(\frac{mv_0}{(m+M)}\right)^2$$

$$= \frac{1}{2}\frac{m^2v_0^2}{(m+M)}$$

$T - W = 0$   
 $T - (m+M)\mu_k g D = 0$   
 $\frac{1}{2}\frac{m^2v_0^2}{(m+M)} - (m+M)\mu_k g D = 0$

$$D = \frac{1}{2}\frac{m^2v_0^2}{(m+M)^2 \mu_k g}$$

$$= \frac{1}{2}\frac{(0.05 \text{ kg})^2 (150 \text{ m/s})^2}{(1.5 \text{ kg})^2 (0.25)(10 \text{ m/s}^2)}$$

unit check ✓

$$D = 5 \text{ m}$$

Back

Figure B.20: Sliding Block Solution V



$u_b = 150 \text{ m/s}$   
 $m_b = 50 \text{ kg}$   
 $u_a = ?$   
 $m_a = 1.15 \text{ kg}$   
 $\mu_k = 0.25$

$$p_i = p_f$$

$$m_b(u_b) + m_a(0) = (m_b + m_a)u_a$$

$$u_a = \frac{m_b u_b}{m_b + m_a}$$

$$= \frac{(0.05 \text{ kg})(150 \text{ m/s})}{0.05 \text{ kg} + 0.05 \text{ kg}}$$

$$= \frac{7.5}{0.1} \text{ m/s}$$

$$= 75 \text{ m/s}$$

$$\sum F_y = 0$$

$$N - mg = 0$$

$$N = mg$$

$$m = (m_b + m_a)$$

$$\sum F_x = ma$$

$$-f = ma$$

$$\mu N = ma$$

$$a = -\frac{\mu N}{m}$$

$$= -\frac{\mu mg}{m}$$

$$= -\mu g$$

$$u_f^2 = u_i^2 + 2ax$$

$$u_a = -ax$$

$$x = -\frac{u_a^2}{2a}$$

$$x = -\frac{u_a^2}{2a}$$

$$= -\frac{(75 \text{ m/s})^2}{2(-0.25g)}$$

$$= \frac{5625}{0.5g}$$

$$= \frac{5625}{0.5 \times 9.8}$$

$$= 1145.8 \text{ m}$$

Figure B.21: Sliding Block Solution W

$$p_b = p_a$$

$$m_b v_b = (m_b + m_w) v_f$$

$$v_f = \frac{m_b v_b}{m_b + m_w}$$

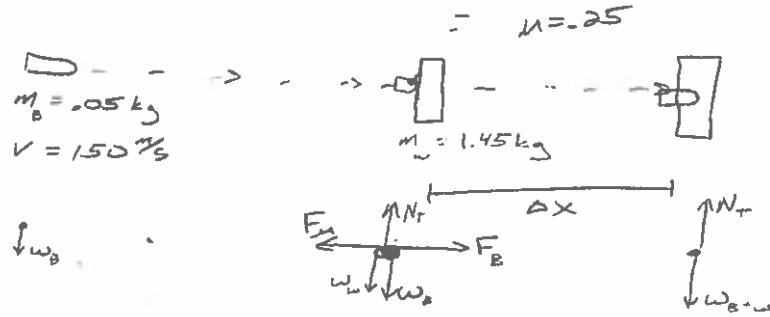
$$v_f = 5 \text{ m/s}$$

$$\frac{1}{2} m v^2 = \mu m g d$$

$$d = \frac{1}{2} \frac{v^2}{\mu g}$$

$$d = 5 \text{ m}$$

Figure B.22: Sliding Block Solution X



$$m_B v_B + m_w v_w = (m_B + m_w) v_{B+w}$$

$$0.05(150) + 0 = (0.05 + 1.45) v_{B+w}$$

$$7.5 = 1.5 v_{B+w}$$

$$5 = v_{B+w}$$

$$\sum F_x = -\mu N = m a \quad \frac{\Delta v}{\Delta t}$$

$$0.25(5) = (0.05) \left( \frac{5}{t} \right)$$

$$-0.5 = \frac{0.25}{t}$$

$$t = 2$$



$$\sum F_y = N - w = 0 \Rightarrow N = 0.5 N$$

$$v^2 - v_0^2 = 2a(\Delta x)$$

$$0 - 25 = 2 \left( \frac{5}{2} \right) (\Delta x)$$

$$\Delta x = 5 \text{ m}$$

SCORE \_\_\_\_\_

Figure B.23: Sliding Block Solution Y

Cons. of Mom

$$(0.05)(150 \text{ m/s}) = (1.45 \text{ kg} + 0.05 \text{ kg}) v$$

$$v = 5 \text{ m/s}$$

X-Diagram

$$v_0 = 0$$

$$x = ?$$

$$v_0 = 5 \text{ m/s}$$

$$v = 0$$

$$a = -7.5 \text{ m/s}^2$$

$$v^2 = v_0^2 + 2ax$$

$$\frac{-v_0^2}{2a} = x$$

$$\frac{-25}{-15} = 5 \text{ m}$$

$$g = 10 \text{ m/s}^2 \quad F_n = 1.5 \cdot 10 = 15 \text{ N}$$

Sum of Forces

$$\sum F = ma$$

$$F_f = ma$$

$$\frac{(0.25)(15)}{1.5} = a$$

$$a = -2.5 \text{ m/s}^2$$

Figure B.24: Sliding Block Solution Z

Student #2 (Spring Problem)

S: All right. [Z] (studies and places to side) [Y] (pause) This is probably a professor 'cause it's all scribbly and his mind's all scattered.

I: OK.

S: Or her mind, whatever. Mind scattered...[X] This is a student. I can tell by the little drawings. I can tell by the drawings, that's a student.

I: OK

S: Nobody got the same answer as me. I must not know my momentum very well. Let's see...

[W] (pause) Hmm...I think it's a student, but I'm not sure why. OK? I can't, I don't know, just the way it's organized, like writing all that stuff down with a little question [mark], that just looks like something a student would write on a piece of paper.

[V] (pause) That's a teacher, cause he's trying to, like, teach you along through the problem. That's what it looks like to me.

[U] That's a student, 'cause they got it wrong. (chuckles) They did what I did, which was not right.

[T] (pause) Probably a student, with those little graphs there.

[S] (pause) That's a teacher, 'cause no student would ever use delta x. I don't even remember where I'm putting my teachers – I'm putting my teachers here, right?

I: Mmm Hmm.

S: OK. [R] (puts to side a bit) [Q] (looks at R & Q together)

[P] (pause) A student, a teacher I don't think would use that (points to  $K_{et}$ ), but they did use K with a little E. I don't think so.

[O] OK. That's probably another teacher. It's probably a trick, but that definitely looks like a teacher's work.

I: How does it look like a teacher?

S: All the notes and that cute cursive. Most students just scribble on some, math there.

[R & Q] These two are getting....

[Q] I think this, (points to list of givens) somebody, an organized student writing all that stuff out, though.

[R] And that's a teacher with all the decimal point stuff.

[Z] And this...we'll give...to the students, this one, cause I'm not sure. That one's...

I: OK. Why did you decide a student?

S: I don't know, the handwriting maybe. I don't know. I got a student vibe from that one.

Student #28 (Spring Problem)

S: [Z] (pause) OK, this one's done the same way. Uh, I don't know. It looks pretty neat. I don't see a lot of teachers writing this neatly. Um, at least [my physics professor] never teaches like that. Um, but they do round, and I know physics teachers like to round. They round it just to 9 meters per second, and I'm going to say – the choices are student or instructor?

I: Mm hm.

S: I'm going to say instructor.

[Y] Um, pretty... (pause) This one looks really messy. Looks messy, but I could see an instructor doing this one, too, but it also looks like something I might do, so I'll say student.

[X] (pause) This one just looks too organized to be a teacher. I think it's, it looks like it's too much work for a very simple problem. Um, it looks like somebody wants to be, I don't know, just clear. I wouldn't say concise; they didn't do it in a short manner, but... The person also just lets you know exactly what, you know, so it'll probably be easy to grade. I think it's a student.

[W] (pause) I'm going to say instructor on this one, because they use the rho and they also drew the after effect of the bullet. I think that looks like something an instructor would do.

[V] This also looks messy, but... I think it looks like [my physics professor's] handwriting, to tell the truth. I'm going to say instructor on this one.

[U] Well, this one's done wrong. I hope it's not an instructor. Uh, hm... OK, they... didn't account for the, they didn't account for the whole problem when doing energy. Uh, I'm going to say student, just 'cause I hope a student.

[T] (pause) This one I don't know, uh. They've got weird looking m's, I can tell you that! And the answer's very rounded at the end, um, which it probably should be, considering the problem statement. I'll say an instructor.

[S] (pause) This looks like an instructor's, because it looks like they were doing it all separate and then it says, "put together," and that's when they put together everything. Plus they solved it in a bunch of variables and then plugged in the numbers.

[R] (pause) This person forgot to take into account the mass of the bullet once it, once the collision.... Other than that, it's done correctly. That's a pretty big mistake. He - The answer is off. I'm going to say student.

[Q] (pause) I'm going to say an instructor, just because it looks like it's so laid out and organized, maybe in a way that an instructor would do it to teach a student.

[P] (long pause) I would say instructor on this one, too. No definite reason why.

[O] (pause) My first thought, this definitely looks like an instructor, 'cause first it's set, they set it up to tell you what parts to consider in the problem, and I just don't see a student taking the time to do that.

Student #11 (Sliding Block Problem)

S: (sorts through them a bit, spreads them out)

[W] [V] [X]

[Y] That one was a student, because, because they didn't use energy, and that looks like a much better way of doing it.

[U] Ooh, that one doesn't have energy, either. (puts in student pile)

[S] (silently puts in student pile)

[R] (silently puts in student pile)

[Q] (silently puts in student pile)

[P] (silently puts in student pile)

None of those used energy. Those were students.

[O] (long pause) That one's a student, 'cause one of the steps, by solving them was a waste of time. It was, it was pointless to calculate it another way. It must be for checking.

[V] [Z] (pause) That one's a student. The first part of the equations weren't there, for the momentum equals momentum.

[V] This one has most of the first equations, except for  $p = p$ , but it says "conservation of momentum," and there's lots of pictures, and a force diagram. Teacher.

[W] This was a student, because the, on the force diagram, body diagram, all the points aren't from the same dot. The block can be treated as a single point.

[T] (long pause) Teacher, because there's a picture, and there's the first equations and, that's about all you have.

[X] No picture. Student.

Student #17 (Sliding Block Problem)

S: OK. [Z] That's some sloppy handwriting, I think that's a student, but also because all the instructors I've ever had have done, just really really focused on explaining things and had pictures, like, like this (pulls out V)

[V] OK, I think this is a student because she fixed it on the back, but I think the instructors probably liked it. (chuckles) Um, unless....this could easily be an instructor because of how well they drew their pictures, and labeled what was going on. And better, my instructors tend to draw better pictures than me.

[Y] Free body diagram. They (V) don't have a free body diagram. I think... (moves on)

[X] If an instructor did this, I'd be surprised. OK, let's see...

[W] Hm. (flips through some more)

[T] OK, this is definitely a student. Another thing that's needed overall that this one is graph. (flips through some more)

[Q] Um, I don't know, I don't know what's going on with this diagram they drew. It doesn't seem like it has anything to do with the problem, at all. West, 150 Newtons, north, 150 Newtons....I'd say that's a student.

[O] This looks like a student's handwriting. (pause) OK, *this* looks like an instructor's solution to me, because it says, "This problem has two parts. Get  $v$  from collision. Find out how it stops," at the top. This one I say is an instructor. This (Q) I say is a student, 'cause that diagram has nothing to do with that question. This (X) I say is a student, because they drew nothing besides, absolutely necessary work.

[Z] Um, conservation of momentum. This person labeled what they were doing, but at the same time, they didn't explain the solution very well. I'm also going to say they were a student.

[P] If the instructor was trying to be cool, this would be an instructor, 'cause there are no questions. You can tell how they just went through the work, and it's pretty much correct, but (mumbles). I think that's a student

[S] I hope this is a student, because they're wrong. But they have a picture with it, but... "collision, slide." They knew what was going on....They use an equation I've never seen before in my life.  $\Delta x$  equals  $v_0$ , squared. I don't know what that is. I'm also going to say they were a student.

[Y] Um, student.

[U] Student.

[V] that could be an instructor, even though they fixed it on the back.

[T] Student.

And Diana I don't have to worry about. She's already a student (puts R in student pile, W in instructor pile)



### Expert #3 (Spring Problem)

S: [Z] (pause) I'm supposed to say which is an instructor, which is a student?

I: Mm hm.

S: Well, this is somewhat closer to the way I would have written it out and so I would put that in the instructor category.

[Y] (pause) This also seems to have the right ingredients, although it seems a bit confused. As an instructor, I would.... Although the answer at the bottom appears to be correct, it looks too complicated, and I would have to make sure that all the algebra was all done correctly. It looks like too much was done. Uh, I might guess that this was a student.

[X] (pause) Well, this is a very systematic way of doing the problem, and, as such, I might guess it was done by a student who wouldn't easily combine steps along the way, so I would guess that this was perhaps a student.

[V] (pause) Uh, I would, I would guess that this might have been done by a faculty member for a couple of reasons. First of all, for the most part, I think for the entire problem, no numbers are plugged in until the end, which is the way I would normally work through the problem until I had the final expression I wanted. Uh, the other thing is they noted that the units were correct, or dimensions, I suspect they meant, uh, were correct. This is something I urge on all the students, but very few of them ever do it. (short pause) Although maybe it was a good student. (little smile)

[W] (pause) I suppose by looking at more and more of these it's getting difficult to really distinguish them. I would put this one almost in the same category, in the sense that it was done systematically, that, uh, numbers weren't plugged in until almost the end, uh, so I could guess that this was uh a faculty member, but it also could have been done by a good student.

[U] (pause) Well, this looks like an incorrect answer. I think the reason is that, uh, in the conservation of energy part, they forgot to put the two masses together. Looks that way to me, so uh, although, I might have the first time around done something similar, I may have been suspicious of the result and checked it. I would call this a student solution.

[T] (pause) Again this is systematic. Numbers aren't plugged in until the very end, uh, so I would, I would be tempted to think of this as a faculty solution. Uh, but I, I mean one of, one of the reasons I may doubt this is that unless I were teaching the course, or a faculty member teaching the course, who emphasized drawing diagrams like this (points to bar charts), I would have never drawn such a diagram, and I doubt very much that a faculty member familiar with the material would actually draw such a diagram. It's very useful for those starting out to think in these terms, but I doubt very much that I would have approached the problem, or used this in the approach to a problem, uh, but if this were a faculty member who was teaching the course, then that would be natural. The way it's been set up systematically and numbers only at the end, I would guess it was a faculty member.

[S] (pause) The same thing, same thing here, maybe another clue is that when I first looked at this problem was is there any friction involved. This person has asked the same question and assumed it was zero. I wonder if a student would have even

thought of it although again I've never taught the E course, and I know the students from the normal 130 sequence probably wouldn't have thought of this. Student.

[R] OK, this is, this is wrong, simply because of the wrong mass used in the conservation of energy, because it was two masses together. The difference isn't very large, because one mass is what, about 20 times the other. Nonetheless, although the approach is correct, there was some carelessness. Maybe a student. Maybe a faculty member. (smiles)

[Q] (pause) Again this is another systematic approach. Looks like almost too much algebra was put in, more so than perhaps someone experienced might have done. Uh, in that some of the terms are zero, which have been crossed out afterwards. Aside from that, uh, I mean, that might lead me to believe it was a student, but that's only a question of experience, when you know some terms aren't going to be contributing and that you don't need them to begin with, so I might guess that that was a student.

[P] (pause) And again, this is a similarly systematic approach to the problem. Just because numbers are put in earlier than I would normally, I might guess that this were a student.

[O] (pause) Well, if I were writing out a solution to post on the bulletin board or a web page, I would probably do it exactly this way, so I'd guess this were a faculty member.

Expert #9 (Spring Problem)

S: Interesting.

[Z] “Inelastic collision” (pause) Actually I would say, this has got some clarity that you might think from an instructor, but given as neatly as it’s written, it’s probably a student.

[Y] (laughs) Oh, I like this one. (laughs) This is probably, given, let’s see. This is a faculty, This is an instructor, I’d say a faculty, not even an instructor. That looks like a faculty type solution, because it’s all done symbolically. Yeah, that’s one thing that stu- most students will plug in numbers as soon as they can, and this, this it looks to me at least, like it remains abstract all the way down, and then start plugging in numbers. That’s got to, that’s got to be an instructor.

[X] Mm. OK, this is a student because of those standard representation things that I know Alan has introduced over the years. Did they get the right answer? It’s something like 179, yeah, right. That’s uh, that’s nice, at least breaking it down into pieces, I like that. Boy, if only students would do that all the time. That’s a student.

Do you want more detail on these than I’m giving?

I: You’re doing fine.

S: [W] All right, let’s see, this one. This is uh, this is a student. Students would be far neater, I would guess, than instructors. (laughs) That’s a student. (pause) Although... I’m not so sure on this one, but I’d, my guess would be a student.

[V] This one... “conservation of momentum,” hm, “use conservation of energy.” Oh, that’s nice, I’d say that’s an instructor.

Are you going to tell me at the ends which ones?

I: Sure.

S: OK, good. (laughs)

[U] OK. That’s a student. One of those who just pulls out a couple of equations, sort of gets it right, but doesn’t really know how it all fits together.

[T] OK, this’ll have to be a student. Instructors, I would guess, don’t draw the representations, even though that’s what we teach. If Alan teaches those in class, it’d be interesting to know if the TAs or instructors actually draw those sorts of things as if they’re taking tests. I imagine on a solution they might, but whether or not they would do if they were taking this thing. So they wrote down the energy, and momentum, and do a bunch of, just plugging in to get to a numerical answer. Student. My guess, they get down, they get to a certain point, even those that will carry along the symbolic stuff, they’ll get to a certain point and then they’ll pull out their graphing calculator and just start plugging in all the numbers and you don’t, so they will calculate this, and they’ll plug it in there, but they’ll never actually write numbers in, so they don’t, those students, I always wonder if they have any idea if they’re close, cause they just plug all the numbers in. That’s got to be a student.

[S] “Collision.” Next they’ve got momentum. “Slide,” and then “slide,” and delta x. “ $\mu$  not given; assume zero.” Oh, somebody thinking about friction. “Put together.” Get that and then plug in numbers. I’d say this is an instructor. Where they put the numbers, and because of  $\mu$ . I don’t think most students, well, maybe some students are anal enough to think about that, even bring up the friction, but that’s nice, that’s nice. I’m going to say instructor.

[R] Let's see, all right. So they plug in numbers – boom! And then do momentum, so I would say that this is a student. They plug in the numbers as soon as they can. They don't really give much explanation of saying "conservation of energy," "conservation of momentum," which is always a tough thing if you're on an exam anyway.

[Q] Let's see, OK. "From 1 to 2." I don't know, the original problem, OK, does not have that labeled. "1 to 2."  $K_0$  plus  $U_{g0}$  plus  $U$ , oh, OK, so writing out all the potential energies, and a  $Q$ . They mean heat? Probably, and work. Wow! There's somebody who's very thorough with the- and  $\Delta U$  internal. (laughs) That's got to be an instructor, absolutely has to be an instructor. There's no way a student is going to write down, put heat in there. And that. Instructors also have a tendency to draw an arrow to where something then plugs in. I don't think students do that. And they have the momentum. So that's got to be an instructor.

[P] OK, so labels the masses. I'm saying this is a student.  $1/2 kx^2$  is  $K$ ,  $K_c$ ,  $K_{el}$ ? So 80, no units, so it's definitely a student. Although, I would guess that many instructors, if you're working through it quickly, you know, we know all about the units, but sometimes we're a bit sloppy in writing them down when we're solving a problem. They plugged in, I would say that's got to be a student. It's not bad, but it's uh...

[O] "There are two parts to this problem: collision and compression of the spring." That's not bad. "Let bullet and block be the system, then momentum is conserved." Wow. "Total energy is conserved at collision." Mm hm. "Maximum displacement." I'm going to guess that's an instructor, because I don't think students would have used "displacement." But they might, 'cause she, whoever this is, why do I say she, almost looks like a she handwriting maybe.  $v^2 =$ , "Then plug in numbers" That's, that's a nice, that's nice and clear. It's either a very good student or an instructor. I'm going to say instructor.

Expert #5 (Sliding Block Problem)

S: (works through Diana's solution.) Are these instructors that were actually working on this class, or?

I: They were um, a variety of instructors.

S: Right. It would probably have taken me about 10 minutes to refresh my memory to remember that this was the way to do it. I would have tried to work out some differential equations, get an exponential decay and find the answer that way.

[Z] Um, let's see, most people, I'm sure, get the first bit, except here's somebody went the route that I was going to go. Were these instructors given a fixed amount of time to do the exam?

I: I trusted them on their own. I just said, "Go take it like you're taking a test." Some of them came back and told me they had imposed time constraints on themselves, and some of them, you know, just....

S: Oh, OK, right. I would probably not let this rest until I found the answer. Here's a person- therefore, anybody who gets it wrong, obviously wrong sort of got stuck because they didn't really know what they were doing, so I'll put them on one side. Um, although if you would had given me a time constraint, I might very well have run out of time on this exam.

[Y] Um, (pause) See here's somebody who worked it out the way I was going to try it. They got the right answer, even though it was a lot of work. I'm probably going to say that was an instructor.

[X] (pause) I'm guessing that's a student.

[W] (pause) I'm just going to sort them by technique

[V] Whoa! What was this?

I: Oh, this is the front and back of the page. (laughs)

S: (pause) I would say that's an instructor. Anybody who does a unit check is not going to be a student.

[U] (pause) Guessing that's a student. This little formula here, I remember working it out every time I teach a class like this, but it's too obscure for me to actually bother to me remember. That would be my guess.

[T] That's a student. Something that was probably covered in class, and if they remember how to do it, it's very quick.

[S] This is wrong, I guess that's a student.

[R] The obscure formula again. (student pile)

[Q] Same with that. (student pile)

[P] (pause, student pile)

[O] I'll guess that's an instructor

### Expert # 10 (Sliding Block Problem)

S: (sifts through a little bit) [V] Wow! How big.

I: That would be the front and back of the page. (laughs)

S: Um, and, oh, I see, it was, oh, he started up here and then went over to the back. So again, that's an immediate clue. (student)

[U] Here's diagrams. (student)

[T] More diagrams. (student)

[Z] [X] [W] (all to side, near instructor pile)

[Y] (instructor)

[O] All right, that same person again. Um... yeah.

(sifts a little) Hm.

[S] (looks at for a long time) I'm tempted to punch in the numbers and figure out why this one got it wrong. Well, I'll put it here (student), just because...

[X & Y] This is tough, just not enough clues. (pause, Y student)

[P] (student)

[R] (student)

[Q] [X] [W] Well, well. (W student) When I'm left with the, I'm tempted to sort them into three categories, basically saying instructors, you can't tell- you know definitely an instructor, well, not definitely, but most likely instructor, most likely a student, and then just can't tell. Um, so, I'm sure, you know I'm mixing some of the instructors here, but I mean, I'm only, how do I say?

[O] This one, this one, um, lays it out, just like the previous one. You know, it even says, "Or use conservation of energy," It's solved for in multiple ways. Um, that's the only one that springs out as definitely I would put in that category.

[X] Um, here's another one that sort of, only because there must be more. (both laugh) Here's one that's to the point and does it in symbols until the very end, OK, which I'm taking as another signature, OK. These other ones, you can't tell.

[V] Again, one that takes so long has got to be [a student].

[U] My previous criterion, um, the fact that, well, um, there's a clue here that um, student way of solving it, writes equation like this:  $F_k = \mu N$ ,  $F_k = .25$  times this,  $F_k =$  this, and somehow, you know, why all those steps, just for the sake of, because it clutters up the paper? Um, it's easy to lose track of - "what was the thing I was looking for? Oh, OK, there it is."

[T] Uh, um, this one looks a little bit messy and um, they started and crossed out, and has this diagram thing. Well, actually, is that? maybe that's not, what's happening, um, but, anyhow... And then these other ones, I can't really tell.

[S] This one somehow got the wrong answer, although, well, yeah, OK, so, um. (sifts a little)

[R] You know, OK, among these, I'll put this one over here (instructors), because it's, well, with mixed emotions (chuckles), because he divides things up into portions which are sensible, encapsulate the parameters in one place, says what you do, solve one part then this part this part, and has a nice drawing, um. But somehow, the only sort of clue on the page to me is that it's laid out in, you know, a slight strange order. Still, I'll put it up there. These others I can't tell.

[P] I mean for example, this could go either way. It's mostly sort of scratching. Um, I'm going to put this here (student), again because they wrote out this  $F_{\text{friction}}$  equals, and they wrote it out in several steps.

[Z] And this one (student) just because someone who knows what they're doing doesn't bother to write out the sum of the forces, sum of this stuff, equals ma.

[Y] And... (student)

[W & Q] I'm not sure. (pause) Well, these could go either way. I mean, I was chuckling because this one, I recognize the handwriting of the previous one that I put in the student pile, so to be consistent, I'd better put it here. (both chuckle) Although, you know, again, it's on the edge, just because the, um, you know. I had a hard time with the other one, too. The thing about this one, is it lays everything out. It's either a student who's parroted, that learned to parrot things very well, because they have all these equations, you know, or how do I say it? It's neatly organized. That's all. My only hesitation is that it's, um, that it's so detailed that it's, you know, almost, verbose and uh, yeah, I'll still go with that.

Expert #12 (Sliding Block Problem)

S: (spreads all the solutions out)

[O] OK, so, up to, again, explaining why things are conserved and why you can use these things, this is a good solution. That's my benchmark there. In fact that looks very much like the one I liked on the other one. (smiles)

[U] And I suspect I'll put this in the same category if I have to lump it into two piles, which you're going to make me lump it into two piles. (middle pile)

[T] And for symbols alone, we'll put this over here. (student)

[V] Uh, let's put this over here. (expert)

(very long pause)

[S] I can't figure out why this one has the wrong answer here. Looks OK. (mumbles something, puts in expert pile)

Do you want to hear a commentary on this, or?

I: Sure, a commentary is always helpful.

S: Yeah. [Q] OK, so this one, we do have – everybody seems to have – momentum conservation, which is nice.

So basically, everyone's coming out with the 5 m/s for the block plus bullet, and... This one is much more continuous than the other one, as far as from the good solutions to the bad solutions, not so good solutions. Essentially everybody gets the, uh, gets the answer here, except for one person, and I don't actually see offhand where they made a mistake on that, so.... Must have evaluated the, uh, acceleration wrong. So...

So I may end up doing the reverse thing I did on the other one; put most of them in the lesser pile and just keep a couple in the good one, if I have to divide into two.

[Q] This one is really a mess. (student)

[X] They didn't explain anything, so put it over there. (student)

Amazingly few words on most of these here, which is a little disturbing, but maybe time pressure.

[W] Looks OK. (expert)

[S] (gets out calculator and punches a few numbers) Oh, OK. So this person was all fine, they just don't know how to use a calculator, which is OK. Put that one there (expert) even though it's the wrong answer.

[U] (into expert pile)

Again, the lack of wording, words on explaining why they can use these constant acceleration formulas and things is a bit disturbing, but it doesn't seem like anybody's done that, so.... (Z goes in expert) Let me try to be consistent here.

[R] (expert)

I think I'm going to end up doing exactly what I did before.

[Y] (expert)

[P] (expert)

OK, so on these ones, I could, I would either make a cut, in which I would maybe keep this one (O) (laughs) and put all in the other category, or roughly lump these together and these ones over here. Um, I think more generally, I'd want to put them into three categories, rather than two, but to be consistent with what I did before, I'd lump, leave the benefit of the doubt to a lot of these things over here, and then put those ones which are not acceptable for various reasons. There'd be lots of things I



would take off on these ones, things that have no words on them and stuff, and sort of structurally broken into different problems. There's only- This one's (V) not bad, this one (O) is clearly the best, and it sort of continuously declines from there. But there are nice elements of a lot of these. Again, it's hard to draw the line, you know, looking between these ones and here, just the little labels they put on things help on that, whereas there's no distinction between things here, other than some spatially. I could divide it a different way, but that's the way I did the first one, so....



## APPENDIX C

### PROBLEMS FROM PHYSICS 131E, 132E, AND ENGINEERING MECHANICS 210 EXAMS USED IN THE STUDY

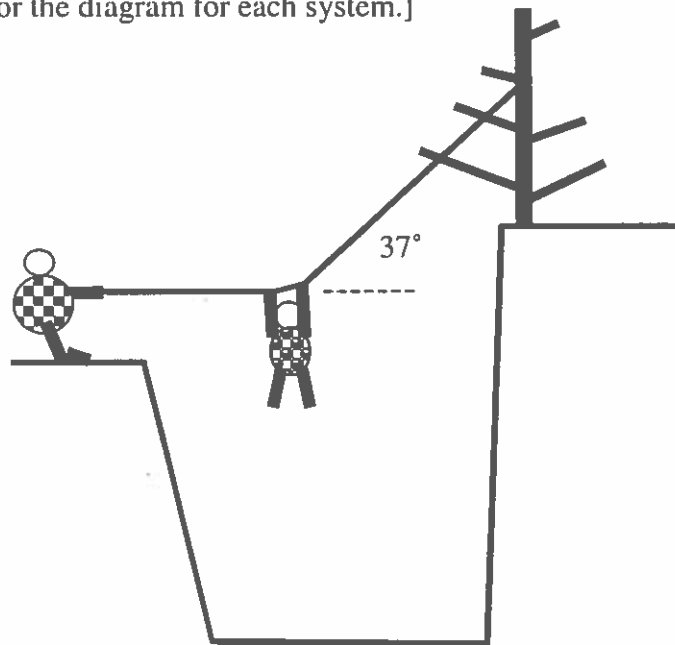
#### Mechanics Midterm 1

6. You are asked to help in the planning for a 500-kg express elevator that is to travel 100 m from the restaurant on the top floor of a building to the ground floor. The elevator starts at rest and stops at the end of the trip. The magnitude of the maximum acceleration of the elevator is  $1.0 \text{ m/s}^2$  both when starting down and when stopping. The elevator's maximum speed is 8 m/s. Determine:

1. the shortest time interval needed for the entire trip
2. the tension in the elevator cable for each part of the trip.

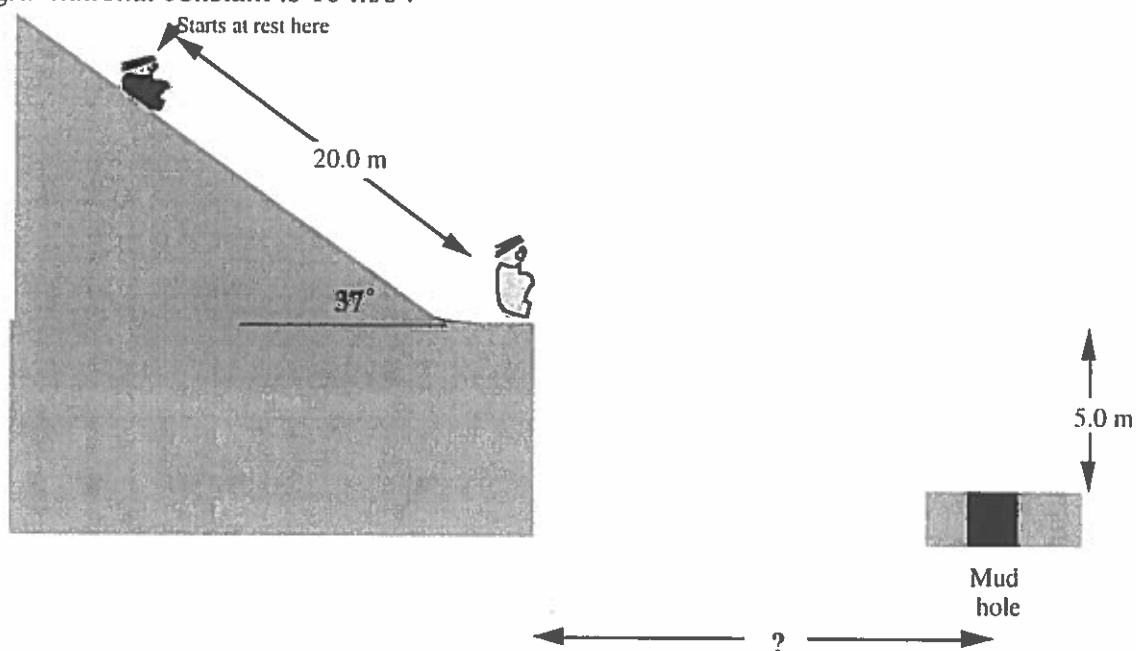
Assume that  $g = 10 \text{ N/kg}$ . Show all of your work or no credit will be given.

7. You and a friend while hiking in the mountains find that you must cross a large crevice—see below. You manage by a miracle to get to the other side with a rope in your hand. Your partner decides to cross by moving hand over hand along the rope. You hold the free end and the other end is tied to a tree. For the situation shown in the sketch, will you be able to remain on the ledge or does the rope tension cause your feet to slide? Your mass is 80 kg and that of your hanging partner is 60 kg. The coefficient of static friction between your feet and the ledge is 0.80. Because your hanging partner is holding the rope with her hands, the rope tension is different on each side of her. Assume that  $g = 10 \text{ N/kg}$ . Be sure to show all of your work or no credit will be given. [Hint: If you are not sure what to do, choose different systems, make force diagrams, and apply Newton's second law for the diagram for each system.]

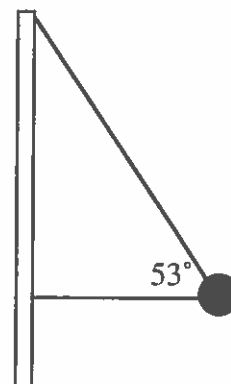


## Mechanics Midterm 2

4. You are involved in the design of a new amusement park activity. An 80.0-kg person slides down an inclined surface and then along a short horizontal section at the bottom. The person then runs into a second 40.0-kg person. They join together and fly off the end of the slide to hopefully land in a mud hole on the other side of a pond. What fun. The dimensions are shown below. Where should the mud hole be placed? The coefficient of friction between the person's bottom and the incline is 0.20. Friction is negligible on the horizontal part at the bottom of the slide. Assume that the gravitational constant is  $10 \text{ m/s}^2$ .

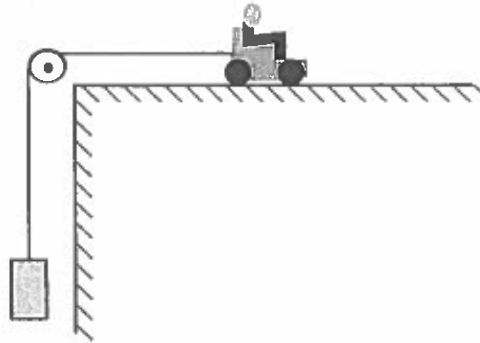


5. A 6.0-kg ball swings in a horizontal circle at the end of two strings. The tension in the slanted string is twice that in the 2.0-m long horizontal string. Determine the speed of the ball. Assume that the gravitational constant is  $10 \text{ N/kg}$ . Show all of your work or no credit will be given.

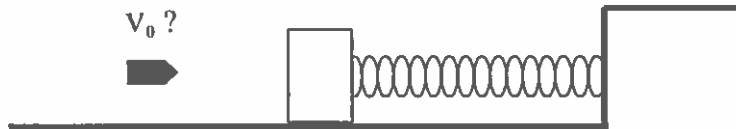


Mechanics Final

27. A 500-kg motorized cart, including the driver, pulls a rope that passes over a very light frictionless pulley down to a 300-kg load that hangs at the end of the rope. The coefficient of static friction between the cart wheels and the ground is 0.80 and the coefficient of kinetic friction is 0.70. Determine the maximum acceleration of the cart and the load. Assume that the gravitational constant is  $10 \text{ N/kg} = 10 \text{ m/s}^2$ . You must show all of the work supporting your answer or no credit will be given.



31. A 0.10-kg bullet is fired into at 1.90-kg block. The block is attached to a spring of force constant  $1000 \text{ N/m}$ . The block slides for 0.40 m while compressing the spring after the bullet runs into the block. Determine the bullet's speed before it hit the block. Assume that the gravitational constant is  $10 \text{ m/s}^2$ . You must show all of the work supporting your answer or no credit will be given.



Electricity and Magnetism Midterm

9. A 0.20-kg cart with a  $2.0 \times 10^{-5} \text{ C}$  charged dome on top starts at rest 8.0 m from a stationary  $-8.0 \times 10^{-3} \text{ C}$  charged dome. The cart moves toward the stationary dome and stops when it runs into and causes a spring to compress 0.50 m. The cart's dome is now 2.0 m from the stationary charged dome. Determine the force constant of the spring.



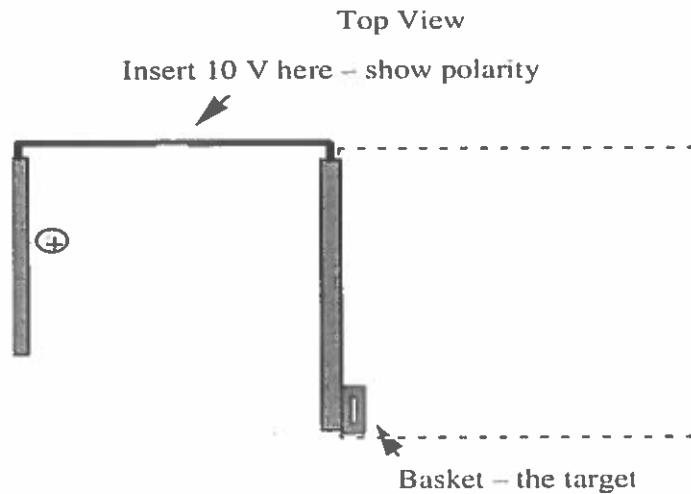
Electricity and Magnetism Final – Section A

31. An 0.080-kg cart with a  $+4.0 \times 10^{-4} \text{ C}$  electric charge starts at rest 0.50 m from a stationary  $6.0 \times 10^{-3} \text{ C}$  charge. After the cart has moved a very very long distance on a horizontal frictionless surface, the cart comes to a 12-m wide crevice. Determine the magnetic field (magnitude and direction) in the crevice region that must exist so that the cart moves straight across the crevice to the other side. Show all of your work.



Electricity and Magnetism Final – Section B

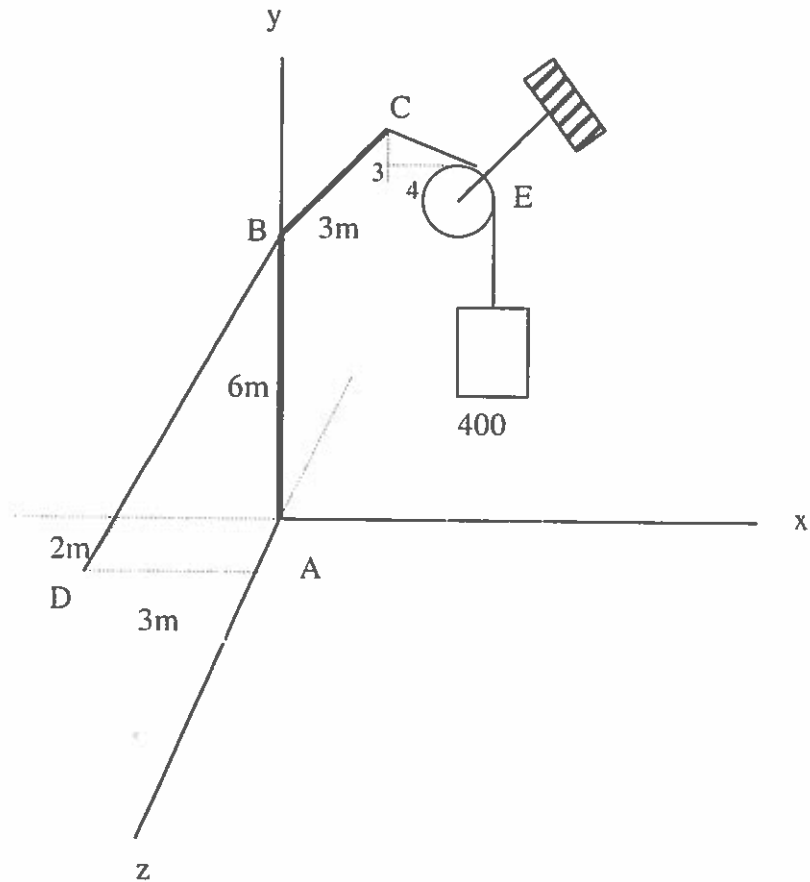
31. You are asked to help design a new arcade game. A  $0.0010\text{-kg}$  ball with charge  $+0.020\text{ C}$  is accelerated across a  $10\text{-V}$  potential difference between two large plates. The ball passes through a slit and then enters a magnetic field. The contestant is to vary the magnetic field so that the ball lands in a basket  $1.0\text{ m}$  to the side of the slit. Determine the magnetic field that will land the ball in the basket and show how to insert the  $10\text{ V}$  so the ball moves from the left plate to the slit with increasing speed. Ignore gravity. Show all of your work.





### Engineering Statics Final

1. Member ABC lies in the  $y$ - $z$  plane and is supported by a pin-and-bracket at A and cable BD. A cable attached to the member at C passes over a frictionless pulley at E and supports a 400 kg mass. Knowing that cable CE lies in the  $x$ - $y$  plane, and that the pin and bracket at A supports 5 components of the reaction at A (3 forces and 2 moments), determine each component of the reaction at A and the tension in cable BD.





**APPENDIX D**  
**COMPARISON OF RESEARCHER AND EXTERNAL EVALUATIONS**

Student	External Evaluator				Researcher				Differences			
	A/N	G/S	# Ins	W	A/N	G/S	# Ins	W	A/N	G/S	# Ins	W
1	5	1	2	0	5	1	3	0	0	0	1	0
2	5	3	5	3	5	3	4	3	0	0	-1	0
3	5	1	2	0	5	1	3	0	0	0	1	0
4	5	1	2	0	5	1	3	0	0	0	1	0
5	5	5	3	0	5	5	3	0	0	0	0	0
6	5	1	2	0	5	1	3	0	0	0	1	0
7	5	1	2	1	5	1	3	1	0	0	1	0
8	5	1	5	2	5	3	4	2	0	2	-1	0
9	5	1	2	0	5	3	3	0	0	2	1	0
10	5	5	7	0	5	5	7	0	0	0	0	0
11	5	1	5	2	5	1	7	2	0	0	2	0
12	5	3	7	1	5	3	7	1	0	0	0	0
13	5	3	2	0	5	1	3	0	0	-2	1	0
14	5	1	2	0	5	1	3	0	0	0	1	0
15	5	1	4	1	5	3	3	1	0	2	-1	0
16	5	1	4	0	5	1	6	0	0	0	2	0
17	5	1	2	5	5	1	3	5	0	0	1	0
18	5	1	4	1	5	1	3	1	0	0	-1	0
19	5	1	5	0	5	1	5	0	0	0	0	0
20	5	5	3	0	5	3	3	0	0	-2	0	0
21	5	3	7	0	5	1	7	0	0	-2	0	0
22		1	2	0		1	3	0		0	1	0
23	5	1	7	2	5	1	6	4	0	0	-1	2
24	2	1	2	0	2	1	2	0	0	0	0	0
25	5	1	2	0	5	1	3	0	0	0	1	0
26	5	1	4	3	5	1	4	6	0	0	0	3
27	5	3	3	1	5	3	3	4	0	0	0	3
28	5	5	2	1	5	5	2	1	0	0	0	0

Table D.1: Comparison of Researcher and External Evaluator Ratings

## APPENDIX E

### THINK-ALoud INTERVIEW PROTOCOL AND REPRESENTATIVE TRANSCRIPTS

Protocol:

I: I'm going to ask you to solve a couple of physics problems and I'd like you to think out loud while you're doing them. I want you to say everything that's going through your mind – it doesn't have to be complete sentences, good grammar, anything like that – just whatever you're thinking. Since you probably haven't done this before, we're going to start with a couple of practice questions.

**Please say everything that is going through your mind as you answer the following question. Say what you are thinking and remember that you are not trying to explain the solution.**

Practice Question 1: Estimate the time you would need to walk in order to go from Smith Lab to German Village, which is south of downtown Columbus.

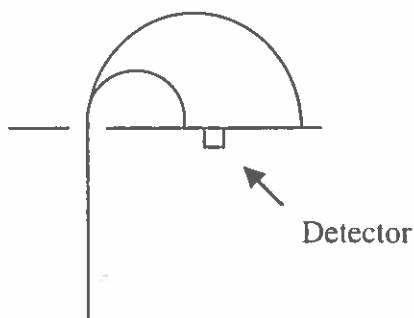
**REMEMBER: Just keep talking to say everything that goes through your mind as you answer the question below.**

Practice Question 2: You earned the only A+ on your physics midterm. A jealous classmate nudges you off a bridge and 3 seconds later you reach the water below. What is your speed when you hit the water?

I: OK, here are the two questions. You can solve them in any order. I'm going to stay out of your way – all I'll do is remind you to keep talking. I'll keep an eye on the time, and if it gets up to about half an hour or so, I'll stop you.

**Please say everything that is going through your mind as you answer the following question. Say what you are thinking and remember that you are not trying to explain the solution.**

1. The mass spectrometer, shown below, is an instrument used to identify molecules in a sample by measuring the charge-to-mass ratio,  $e/m$ . Ionized molecules are accelerated (from rest) through a potential difference  $V$  and then enter a region of uniform magnetic field, in this case of .2 T. The voltage difference can be adjusted so that specific ions can be collected by a detector, which is located 8 cm from the initial opening. What does the potential difference between the plates need to be so that carbon monoxide molecules of mass  $4.68 \times 10^{-26}$  kg with one excess electron are collected?



**REMEMBER: Just keep talking to say everything that goes through your mind as you answer the question below.**

2. Your spacecraft has been on a long mission, and while traveling through space has picked up an excess charge of  $-7$  mC. After a brief stop at Geigel 7, the next task on your mission is an investigation of a nebulous region  $3 \times 10^5$  kilometers away which is known to have a constant and uniform electric field of  $20$  N/C. To travel from Geigel 7 to the nebula, you fire your thrusters at one-quarter impulse, which provides a constant thrust of  $2$  N to your  $26000$  kg craft. Upon arriving at the edge of the nebula, you turn off all propulsion systems. You find that your craft drifts straight ahead at constant velocity through the nebula. If you are drifting perpendicularly to the direction of the electric field, what can you conclude about the magnetic field in this region of space?

### Potentially useful information

$$v = v_0 + at$$

$$x - x_0 = v_0 t + (1/2) at^2$$

$$2(x - x_0)a = v^2 - v_0^2$$

$$F = ma$$

$$F_g = (GMm)/r^2$$

$$F_q = (k q_1 q_2)/r^2$$

$$F = qE$$

$$a_c = v^2/r$$

$$K = (1/2)mv^2$$

$$U_g = mgy$$

$$U_s = (1/2)kx^2$$

$$U_q = (kQq)/r$$

$$U_q = qV$$

$$F_m = qv \times B$$

$$F_m = iL \times B$$

$$B = (\mu_0 I)/(2\pi r)$$

$$k = 1/(4\pi\epsilon_0) = 9 \times 10^9 \text{ N m}^2/\text{C}^2$$

$$e = 1.6 \times 10^{-19} \text{ C}$$

$$m_e = 9.1 \times 10^{-31} \text{ kg}$$

$$\mu_0 = 1.26 \times 10^{-6} \text{ T m/A}$$

Student #17

Problem Solving – mass spectrometer

S: (reads problem aloud, pauses a bit at  $e/m$  part to consider it, then goes on reading) OK, I feel like I need to collect all the little things they gave me. (starts writing down givens) I know the magnetic field is .2. I know the potential difference, oh, I don't know – I'm trying to find it. The distance is 8 centimeters. Um, the mass of a molecule is  $4.68 \times 10^{-26}$  kilograms.

Then I need to look and see what kinds of equations I have to help me out here. This is the kind of question where it accelerates around a curve...hm. I know my initial velocity is zero, 'cause they're accelerated from rest. Hm. (looks through equations) I don't remember what  $\mu_0$  is in this equation. I don't think....  $qv$  cross  $B$ ... I know velocity is zero and it seems like if it's zero and cancels out the charge, then somehow that shouldn't work. OK, so, all I know is FBI. (Looks at diagram with the problem some more)

Force first will be in this direction. I don't know which direction the magnetic field is. (RHR) ...direction of the velocity...hm. Oh, it's going to just accelerate through this part and the field is going to bend it down. So it starts with a potential difference. It's going to be, let's see. One excess electron – it's a negatively charged molecule (starts writing again). It's  $1.61 \times 10^{-19}$  Coulombs then, negative, so I want it to be repelled so it accelerates forward, so I want this to be the negative side. So I want negative, or lower, potential difference over here so it's attracted over here, so I want this to be positive and this to be negative. And the electron's going to start here - well, my negatively charged, molecule? - and then it's going to enter this region. "What does the potential difference between the plates need to be so that carbon monoxide molecules of mass  $4.68 \times 10^{-26}$  kg with one excess electron are collected?" (rereads rest of problem also aloud.)

So I need to figure out what kind of speed I want it to go at over here. And the speed is going to go in this equation,  $F_m$ ,  $qv$  cross  $B$ . I know the charge, I know the magnetic field, and force equals mass times acceleration and I want it [to] end up 8 centimeters in a half-circle, so that it's kind of forced to go this way. Here's my detector. I need to find out what kind of acceleration can make it turn in a half-circle. It's going to go 8 centimeters. Oh, it's going to turn in a half-circle of 8 centimeters, so the acceleration is  $v^2/r$ . And my radius is 4 centimeters, so  $qv$  cross  $B$  equals  $mv^2$  over  $r$ .  $q$  cross  $B$ ... The magnitude of  $q$  cross  $B$  is the same as the magnitude of  $q$  times  $B$ , I think, even though I know it's in a perpendicular direction. I think I can just multiply them together. (small laugh) Um, I figure (unintelligible) the charge is...and solve for velocity, so  $qB$  over  $m$  times  $r$  equals  $v$ .

The charge is  $1.61 \times 10^{-19}$  Coulombs. Oh, yeah, I want to make it into electrons for, I think. (erases) Oh, no I don't. I think I have to use SI units here to get out what I want. .2 T, radius 4 centimeters. I think I should change 4 centimeters to meters... 100 centimeters in a meter...100, not 1000, good job! (laughs) .04 meters. I have the right units to figure out what I want, I think.  $4.68 \times 10^{-19}$  kilograms... And now I multiply them together. (calculator, muttering numbers as uses)  $1.288 \times 10^{-21}$ . (calculates some more) [Get a velocity of] 2752.14. That answer makes *absolutely no sense*. I guarantee that.



OK, what did I even solve for? Velocity, oh, OK, that's what I think the velocity needs to be going, which is probably way too big. I'm going to stick with it, just for the efforts of solving the problem. I'm going to pretend like that's the velocity I need, and we'll round it to 2800. It looks like too big of a number to me. So if that's the velocity that I need, then – I need to know what this distance is right here, how long I have to accelerate the thing. (rereads most of problem again aloud).

OK, this is energy then, because it's going to be the change in electric potential is going to be the same as the change in velocity. I really don't need a distance for this. So  $U_q$  equals  $qV$ , that's... I'm going to start out with all kinds of electric potential energy and make it positive, even though it's probably – oh yeah, it is positive, 'cause of this negative thing, whatever, negative voltage. All kinds of electrical potential, no kinetic, no any other kind of energy in my initial situation. And then my final situation is going to be right here, when it goes through that little hole, and that's going to be, well, I mean, there's still going to be some, oh there's not going to be any electric potential due to that area, so I'm going to have none of that, there'll only be kinetic energy. So the kinetic energy must equal the electric potential I started with.

Electric potential is  $qV$ . Kinetic is  $1/2 mv^2$ . This is big  $V$ , don't get them confused, so I can solve [for] my potential. So  $q$ ,  $1/2$  times the mass, which is  $4.68 \times 10^{26}$  kilograms and velocity squared that I calculated, that huge number, and charge is  $1.61 \times 10^{-19}$ . And I'm going to get what the difference needs to be between them, and I think it's going to be a giant, giant number 'cause I have a negative- no, oh, it should be OK; I thought it was going to be ridiculous, but maybe not. (calculator)

I got the top to be  $1.83456 \times 10^{-19}$  over  $1.61 \times 10^{-19}$ , so I should get a number in the single digits. (calculator) That wasn't too bad. (laughs) Now I'd better check it. (redoes last calculation) And I got an answer close to one, so that's what I was expecting, 1.139, so about 1.14 Volts is my potential difference, and that seems small to me, for having to accelerate it that much, so I would say my answer's not that reasonable, but, um, I would put, you have to go from lower to higher voltage, and the potential difference is 1.14 Volts, according to what I calculated, which I don't think is exactly right. (laughs)

## Spaceship

S: OK (reads problem aloud, writing down info as goes) So I think what I have going on is I have an electric field that is putting a force on my craft one direction and the magnetic field the other direction and so that's cancel each other out and you go at constant velocity. So I'm going to draw a free-body diagram of the situation. There's my spacecraft - a big dot – and, um, I'm just going to arbitrarily say this is the electric field line and this is the magnetic field force, and that they're equal, and since I'm in space, I don't have gravity or anything like that around, and we know everything else, so I'm going to say the electric field has to equal the magnetic field – I mean the force due to each. And, the force due to the electric field is  $qE$ , and the force due to the magnetic field is  $qv$  cross  $B$ . The charge on my spacecraft, excess charge on it, is actually completely irrelevant to the problem. So the electric field, OK, so,  $E/v$ , this is going to give the magnitude of my magnetic field and then I'll have to try and find out what direction it is also.

(unintelligible, rereads part of problem) OK, I just used an equation that I don't have all the information for. (laughs) Hm. (looks at equation sheet, pause) Force due to an electric field,  $qE$ , that makes sense, but I have to find the magnetic field. The force of the magnetic field is completely going to depend on the velocity, so let me reread the problem. (rereads)

This is the key, I have a force on my spacecraft, and I have the mass of it, and I have... so I should be able to find out what the speed is. Oh, I have enough information to calculate the speed when I get there, so I have to calculate the velocity. So, since I'm out in space, I think I just have, I think my force equals my mass times my acceleration. I don't think there's gravity or anything else to deal with. I mean, I have the distance. So, my speed should be...  $v$  of  $t$  – well, acceleration should be 2 times, oh (erases). Force equals mass times acceleration. I have a force of 2 Newtons and a mass of... 26000. 2 over 26000 is the acceleration, should be able to find, so it's 1 over 13000 – that's my acceleration, so my velocity is 1 over 13000 times  $t$ . My position is  $1/2$  of that, so  $1/2$  over 26000  $t^2$  plus 0. That's 0 and nothing else.

So I know that I go  $3 \times 10^5$  kilometers. I need to change that to meters (does conversion, first in head, then checks on calculator) So after I have gone  $3 \times 10^8$  meters, how much time has gone by? I need to find out, so that I can plug time back into the velocity equation. I think I'm doing this the really *long* way, but maybe not. All right,  $3 \times 10^8$ ,  $t^2$ , so then I have  $t^2$  is 26000 times  $3 \times 10^8$ , and then multiply those together and take the square root and see what the time is. It seems like it's going to take an awfully long time. (calculator) 27928.48. That's too many digits. 28, 1, 2, 3, 4, 5 [zeros] seconds. I guess the seconds aren't going to make a difference. I need to put my seconds back in my velocity equation, so that really shouldn't bother me, oh OK, velocity is divide that number by 13000.

214.8 m/s. That's my velocity when I get there, so I can use that equation, so my electric field over velocity is going to be the magnetic field in the region, so (does calculation) .093 T, and I guess I should say which direction it is. I guess if my spacecraft were drifting this way, and the electric field would have to be pointing in the same direction as my spacecraft. It doesn't matter. The electric field is pointing this way, and if my spacecraft is negatively charged, then it's going to have the opposite effect, so it's like my spacecraft is positively charged and the electric field's going in the opposite direction. Magnetic field and the electric field should be pointing in the same direction. (RHR) I have to have the magnetic field into the page. So if my electric field is up, it has an effect of a downward force, so I need my magnetic field to go into the page to have the effect of an upward force. So .093 into the page, according to the little picture I just drew.

### Student #23

#### Problem Solving – mass spectrometer

S: (silently reads problem) OK, so we have... (draws a few arrows) this stuff coming in. I don't think it says its speed. (refers back to problem statement) Oh, they're accelerated through a potential difference  $V$ , so their speed after is just  $qV = 1/2mv^2$ .  $v = 2qV/m$  (makes  $V$  distinct from  $v$ ) square root. So that's the velocity when you come out.

(reads some of problem again, pause) .2 Tesla, so 'cause a magnetic force, depending on this velocity.  $qv \text{ cross } B$ . The force would be towards the center of this. (adds arrow to previous arrow) So this will be  $qvB$  times  $-$  I'm substituting all this stuff [the expression for  $v$ ] that... and that's the magnetic force, which is the centrifugal force...  $mv^2/r$ . That times  $v$  which, or  $v^2$ , which is  $2qV/r$ . Equals... the magnetic force, which is  $qB$  this. Cancel these out. Divide through by this and get the square root on the other side. (draws arrow, continues algebra silently for a bit) And solve for  $v$ . So multiply by  $r$  over... and square root. You can cancel out these  $q$ 's. So I have voltage equals  $qB^2r^2/2m^2$ , which is probably wrong, since I had to go through too many equations, but, still...

Now I'll substitute in what I know. I know mass, know  $B$  (draws arrow). I don't think I know  $q$ . I know  $r$ . So I'll write down what I know, and I'll substitute those in. (calculator) And I get a really big number, of course, because  $m^2$  on the bottom was really big.  $q$  times  $5.8 \times 10^{46}$ . Hm. Seems like it's too big. (pause) I'll check my equations. (reviews work) That seems OK.... That seems OK....  $qB$  times all of that.... Equals  $mv^2/r$ .... Looks all right to me. Should I just do the next one?

I: Sure.

S: That's a big answer.

#### Spaceship

S: (silently reads problem) So we're... (starts drawing) drifting into... an electric field of  $20 \text{ N/C}$ . Um, excess charge  $7 \times 10^{-6}$ , (scratches out)  $^{-3}$  Coulombs. (pause) We need to find the velocity we're at. (writing) Velocity is acceleration times time. Acceleration equals force over mass.  $2 \text{ Newtons over } 26000 \text{ kilograms}$ .  $1 \text{ over } 13000 \text{ m/s}^2$ . So every second we speed up  $1/13000 \text{ m/s}$ . Going  $300000 \text{ kilometers}$ . I'll just add 3 zeros and change it to meters. That's equal to acceleration times  $1/2$  times time squared.

(calculator)  $2,800,000 \text{ seconds}$ . Multiply that by the acceleration. That's  $215 \text{ m/s}$ . (pause) OK. So when we reach there we're going at  $215 \text{ m/s}$ . A charge of  $-7 \times 10^3 \text{ Coulombs}$ . Since it drifts straight ahead, that means that magnetic force is equal to the electric force, so  $20 \text{ N/C times } 7 \times 10^3 \text{ Coulombs}$  is going to be equal to  $7 \times 10^3 \text{ Coulombs times velocity}$ , which is  $215 \text{ m/s}$ , times, oh, and the unknown,  $B \text{ Tesla}$ . So cancel those [ $q$ 's] and the answer should be  $20$  divided by  $215$ . The answer is  $.093 \text{ Tesla}$ . (looks at calculator, looks over work) OK, I'm done.

Student #30

Problem Solving – mass spectrometer

S: (silently reads problem) OK, I always just usually redraw the, what's going on, even if they already state what it is. (draws picture, labels) And putting in what we know. (pause)

I: Keep talking.

S: OK, ah, just trying to find out which direction the field is in, but I don't think it's too important, since we already know which direction it's going to go. Um, check equations to see, what we're looking for is voltage. Um, (pause) Just trying to find equations that will work for me. (pause) You might use the... (starts writing B field due to wire) use that, but, solve for I later, but I can't find voltage in there. (pause)

I: Keep talking.

S: Just trying to find something, trying to remember whether or not work equations can be used for this.... And what I can use I for, I mean, well, if that would work. (pause) Think I'm going to read over the next one.

Spaceship

S: (silently reads problem) Draw what they're talking about, (draws) nebula, you're coming in here. Acceleration would be zero. Charge of  $-7$  milliCoulombs. First have to determine how fast you're going at. The distance... (writes, pause, refers back to problem) Find the acceleration through this distance  $d$ .  $F = ma$ . (puts numbers in and solves for  $a$ ) We start at zero velocity. (calculator) The velocity's  $7.8 \times 10^{12}$  m/s, when you get to the nebula. You stay at that velocity as we go through the nebula. So the force must be equal to zero and that equal to the charge on the ship times the velocity cross product with the magnetic field. That's  $qvB \dots \cosine \theta$ . Cosine of that equals zero, so that's... (pause, looking over work, says something unintelligible and then writes that  $B = 20$  N/C) If there's already a magnetic field, it should drift the spaceship in some direction, so there's probably something else counteracting it. (long pause) Also could be passing through, like, passing in through, like, the same direction as the thing that's causing the magnetic field. If it's in the same direction, apply forces on all directions of the ship. (pause)

I: Keep talking.

S: Just trying to think what's going on here and what it's really asking for. 'Cause we're drifting perpendicular, there's got to be some force on the ship. (pause) Another possibility is the ship is changing the magnetic field by its movement, causing the ship... there's got to be something else acting on it. [If the] Ship was traveling into the paper, and the magnetic field's going around it, then they would cancel each other out and you'd be traveling, like, in the same direction. That's why I think it's donut shaped, something like that. Could also be the spherical movement when it's traveling right through it. (pause) [Unintelligible] what the velocity is, when you're at the two ends. (pause)

I: Keep talking.

S: It's like the magnetic field you might get from a coil or something. Of course.... It's like there's a current going through, so the direction that the ship is traveling in.... create a magnetic field all around the ship.... I don't know what else to do with this problem.

## Mass Spectrometer

Let me look the first one over again, and see if I can get anything else out of it. (rereads problem, looks over work a bit) Well, with this equation (B due to straight wire), um, [unintelligible]. (refers to equation sheet, pause) I think it has something to do with the radius. (writes down potential energy of two point charges, pause)

I: (calls time)



## BIBLIOGRAPHY

ABET criteria can be found at <http://www.abet.org/cac/2000.htm>.

James L. Adams, *Conceptual Blockbusting* (W. W. Norton and Company, New York, NY, 1979).

D. Rosil, "What are Masters Doing?" (AIP Statistics Division, College Park, MD, 1996).

John R. Anderson, James G. Greeno, Paul J. Kline, and David M. Neves, "Acquisition of Problem Solving Skill" in *Cognitive Skills and Their Acquisition*, edited by John R. Anderson (Lawrence Erlbaum Associates, Incorporated, Hillsdale, NJ, 1981).

Yuichiro Anzai and Tohru Yokoyama, "Internal Models in Physics Problem Solving," *Cognition and Instruction* 1, 4, 397-450 (1984).

E. L. Baker and R. E. Mayer, "Computer-based Assessment of Problem Solving," *Computers in Human Behavior* 15, 3, 269-282 (1999).

Miriam Bassok and Keith J. Holyoak, "Interdomain Transfer Between Isomorphic Topics in Algebra and Physics," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15, 1, 153-166 (1989).

Miriam Bassok, "Transfer of Domain-specific Problem-solving Protocols," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16, 3, 522-533 (1990).

R. Bhaskar and H. A. Simon, "Problem Solving in Semantically Rich Domains: An Example From Engineering Thermodynamics," *Cognitive Science* 1, 2, 193-215 (1977).

B. Bloom, editor, *Taxonomy of Educational Objectives: Cognitive Domain* (David McKay, New York, 1956).

Robert C. Bogdan and Sari Knopp Biklen, *Qualitative Research for Education* (Allyn and Bacon, Needham Heights, MA, 1998).

John D. Bransford, Ann L. Brown, and Rodney R. Cocking, editors, *How People Learn: Brain, Mind, Experience, and School* (National Academy of Science, 1999).

Diane J. Briars and Jill H. Larkin, "An Integrated Model of Skill in Solving Elementary Word Problems" *Cognition and Instruction* 1, 3, 245-296 (1984).

Micheline T. H. Chi, Paul J. Feltovich, and Robert Glaser, "Categorization and Representation of Physics Problems by Experts and Novices" *Cognitive Science* 5, 2, 121-152 (1981).

Micheline T. H. Chi, Miriam Bassok, Matthew W. Lewis, Peter Reimann, and Robert Glaser, "Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems," *Cognitive Science* **13**, 2, 145-182 (1989).

Micheline T. H. Chi and Robert Glaser, "Problem Solving Ability" in *Human Abilities: An Information Processing Approach*, edited by Robert J. Sternberg (W. H. Freeman and Company, New York, NY, 1985).

John Clement, "Observed Methods for Generating Analogies in Scientific Problem Solving," *Cognitive Science* **12**, 4, 563-586 (1988).

Ton de Jong and Monica G. M. Ferguson-Hessler, "Cognitive Structures of Good and Poor Novice Problem Solvers in Physics," *Journal of Educational Psychology* **78**, 4, 279-288 (1986).

Robert J. Dufresne, William J. Gerace, and William J. Leonard, "Solving Physics Problems With Multiple Representations," *The Physics Teacher* **35**, 5, 270-275 (1997).

Avish and Nancy Dworkin, *Problem Solving Assessment* (Academic Therapy Publications, Novato, CA, 1988).

Andrew Elby, "Another Reason that Physics Students Learn by Rote," *Physics Education Research Supplement to American Journal of Physics* **67**, 7, S52-S57 (1999).

K. Anders Ericsson and Herbert A. Simon, *Protocol Analysis* (Massachusetts Institute of Technology Press, Cambridge, MA, 1996).

Bat-Sheva Eylon and F. Reif, "Effects of Knowledge Organization on Task Performance," *Cognition and Instruction* **1**, 1, 5-44 (1984).

Monica G. M. Ferguson-Hessler and Ton de Jong, "On the Quality of Knowledge in the Field of Electricity and Magnetism," *American Journal of Physics* **5**, 6, 492-497 (1987).

Monica G. M. Ferguson-Hessler and Ton de Jong, "Studying Physics Texts: Differences in Study Processes Between Good and Poor Performers," *Cognition and Instruction* **7**, 1, 41-54 (1990).

Garlie A. Forehand, *Individual Differences in Problem Solving Processes of College Students* (Carnegie-Mellon University, Pittsburgh, PA, 1967).

Thomas M. Foster, *The Development of Students' Problem-Solving Skills From Instruction Emphasizing Qualitative Problem-Solving*, unpublished doctoral dissertation (University of Minnesota, 2000).

Norman Fredericksen, "Implications of Cognitive Theory for Instruction in Problem Solving," *Review of Educational Research* **54**, 3, 363-407 (1984).

Raymond Friedman, *Problem Solving for Engineers and Scientists: A Creative Approach* (Van Nostrand Reinhold, New York, NY, 1991).

Robert G. Fuller, "Solving Physics Problems - How Do We Do It?" *Physics Today* **35**, 9, 43-47 (1982).



- Robert M. Gagne, "Human Problem Solving: Internal and External Events" in *Problem Solving: Research, Method, and Theory*, edited by Benjamin Kleinmuntz (John Wiley and Sons, Inc, New York, NY, 1966).
- Reed Gamble, "Force," *Physics Education* 24, 79-82 (1989).
- Mary L. Gick, "Problem-Solving Strategies," *Educational Psychologist* 21, 1 & 2, 99-120 (1986).
- John K. Gilbert, Michael D. Watts, and Roger J. Osborne, "Students' Conceptions of Ideas in Mechanics," *Physics Education* 17, 62-66 (1982).
- Ron Good, "Scientific Problem Solving by Expert Systems," *Journal of Research in Science Teaching* 21, 3, 331-340 (1984).
- Bert F. Green, Jr., "Current Trends in Problem Solving" in *Problem Solving: Research, Method, and Theory*, edited by Benjamin Kleinmuntz (Wiley and Sons, Incorporated, New York, NY, 1966).
- Ibrahim Abou Halloun and David Hestenes, "Common Sense Concepts About Motion," *American Journal of Physics* 53, 1056-1065 (1985).
- David Hammer, "Two Approaches to Learning Physics," *The Physics Teacher* 27, 12, 664-670 (1989).
- John R. Hayes, *The Complete Problem Solver* (Franklin Institute Press, Philadelphia, PA, 1981).
- Joan I. Heller and F. Reif, "Prescribing Effective Human Problem-Solving Processes: Problem Description in Physics," *Cognition and Instruction* 1, 2, 177-216 (1984).
- Patricia Heller, Ronald Keith, and Scott Anderson, "Teaching Problem Solving Through Cooperative Grouping. Part 1: Group Versus Individual Problem Solving," *American Journal of Physics* 60, 7, 627-636 (1992).
- Patricia Heller and Mark Hollabaugh, "Teaching Problem Solving Through Cooperative Grouping. Part 2: Designing Problems and Structuring Groups," *American Journal of Physics* 60, 7, 637-644 (1992).
- Peter W. Hewson, "Diagnosis and Remediation of an Alternative Conception of Velocity Using a Microcomputer Program," *American Journal of Physics* 53, 648-690 (1985).
- David C. Howell, *Statistical Methods for Psychology* (Wadsworth Publishing Company, Belmont, CA, 1997).
- Ray Hyman and Barry Anderson, "Solving Problems" in *The R&D Game: Technical Men, Technical Managers, and Research Productivity*, edited by David Allison (MIT Press, Cambridge, MA, 1969).

Christina Karrquist and Bjorn Anderson, "How Swedish Pupils, Age 12-15, Understand Light and its Properties," in *Proceedings of the International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*, edited by Hugh Helm and Joseph D. Novak (Cornell University Press, Ithica, NY, 1983), 380-392.

Lorraine Lee Knefelkamp, *Developmental Instruction: Fostering Intellectual and Personal Growth of College Students*, unpublished doctoral dissertation (University of Minnesota, 1974).

J. H. Larkin and H. A. Simon, "Why a Diagram is (Sometimes) Worth 10,000 Words," in *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, edited by Janice Glasgow, N. Hari Narayanan, and B. Chandrasekaran (AAAI Press/MIT Press, Menlo Park, CA, 1995).

Jill H. Larkin, "Processing Information for Effective Problem Solving," *Engineering Education* **70**, 3, 285-288 (1979).

Jill H. Larkin, John McDermott, Dortha P. Simon, and Herbert A. Simon, "Models of Competence in Solving Problems," *Cognitive Science* **4**, 4, 317-345 (1980).

Jill H. Larkin, "Skilled Problem Solving in Physics: A Hierarchical Planning Model," *Journal of Structural Learning* **6**, 3, 271-297 (1980).

Jill H. Larkin, "Enriching Formal Knowledge: A Model for Learning to Solve Textbook Physics Problems" in *Cognitive Skills and Their Acquisition*, edited by John R. Anderson (Lawrence Erlbaum Associates, Incorporated, Hillsdale, NJ, 1981).

Jill H. Larkin, "Understanding Problem Representations and Skill in Physics," in *Thinking and Learning Skills, Vol. 2*, edited by Susan F. Chipman, Judith W. Segal, and Robert Glaser (Lawrence Erlbaum Associates, Incorporated, Hillsdale, NJ, 1985).

Norman R. F. Maier, *Problem Solving and Creativity in Individuals and Groups* (Brooks/Cole Publishing Company, Belmont, CA, 1970).

David P. Maloney, "Charged Poles?" *Physics Education* **20**, 310-316 (1985).

Michael McCloskey, "Naïve Theories of Motion," in *Mental Models*, edited by Dedre Gentner and Albert J. Stevens (Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, 1983), 299-324.

Jim Minstrell, "Explaining the 'At Rest' Condition of an Object," *the physics Teacher* **20**, 10-14 (1982).

Jim Minstrell and Virginia Stimpson, "Instruction for Understanding: A Cognitive Process Framework," December 29, 1986.

"Shaping the Future: New Expectations for Undergraduate Education in Science, Mathematics, Engineering, and Technology," (NSF Directorate for Education and Human Resources Review of Undergraduate Education, 1996).

H. F. O'Neil, Jr., "Perspectives on Computer-Based Performance Assessment of Problem Solving," *Computers in Human Behavior* **15**, 3, 255-268 (1999).

William G. Perry, Jr., *Forms of Intellectual and Ethical Development in the College Years* (Holt, Rineholt, and Winston, Inc., New York, NY, 1968).

Stephen K. Reed, *Word Problems: Research and Curriculum Reform* (Lawrence Erlbaum Associates, Mahwah, NJ, 1999).

F. Reif, J. H. Larkin, and G. C. Brackett, "Teaching General Learning and Problem Solving Skills," *American Journal of Physics* **44**, 3, 212-217 (1976).

Frederick Reif, "Teaching Problem Solving - A Scientific Approach," *The Physics Teacher* **19**, 5, 310-316 (1981).

Kurt Reusser, "Problem Solving Beyond the Logic of Things: Contextual Effects on Understanding and Solving Word Problems," *Instructional Science* **17**, 4, 309-338 (1988).

H. J. A. Rimoldi, Hermelinda M. Fogliatto, James B. Erdmann, and Michael B. Donnelly, *Problem Solving in High School and College Students* (Loyola University Press, Chicago, IL 1964).

William C. Robertson, "Detection of Cognitive Structure With Protocol Data: Predicting Performance on Physics Transfer Problems," *Cognitive Science* **14**, 2, 253-280 (1990).

Moshe F. Rubenstein, *Patterns of Problems Solving* (Prentice Hall, Incorporated, Englewood Cliffs, NJ, 1975).

J. Schacter, H. E. Herl, G. K. W. K. Chung, R. A. Dennis, and H. F. O'Neil, Jr., "Computer-based Performance Assessments: A Solution to the Narrow Measurement and Reporting of Problem Solving," *Computers in Human Behavior* **15**, 3, 403-418 (1999).

Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences* (McGraw-Hill Book Company, New York, NY, 1956).

Dorthea P. Simon and Herbert A. Simon, "Individual Differences in Solving Physics Problems" in *Children's Thinking: What Develops?* edited by Robert S. Siegler (Lawrence Erlbaum Associates, Hillsdale, NJ, 1978).

Herbert A. Simon, "Problem Solving and Education" in *Problem Solving and Education: Issues in Teaching and Research*, edited by D. T. Tuma and F. Reif (Lawrence Erlbaum Associates, Hillsdale, NJ, 1980).

John Sweller, "Cognitive Load During Problem Solving: Effects on Learning," *Cognitive Science* **12**, 2, 257-285 (1988).

John Sweller, Robert F. Mawer, and Mack R. Ward, "Development of Expertise in Mathematical Problem Solving," *Journal of Experimental Psychology: General* **112**, 4, 639-661 (1983).

David E. Trowbridge and Lillian C. McDermott, "Investigation of Student Understanding of the Concept of Velocity in One Dimension," *American Journal of Physics* **48**, 1020-1028 (1980).

A. P. Carneval, L. J. Gainer, and A. S. Meltzer, *Workplace Basics: The Essential Skills Employers Want* (Jossey-Bass Publishers, San Francisco, CA, 1990).

Alan Van Heuvelen, *Active Learning Problem Sheets* (Hayden-McNeil Publishing, Plymouth, MI, 1996).

Alan Van Heuvelen, "Learning to Think Like a Physicist: A Review of Research-Based Strategies," *American Journal of Physics* **59**, 10, 891-897 (1991).

J. H. P. van Weeren, F. F. M. de Mul, M. J. Peters, H. Kramers-Pals, H. J. Roosink, "Teaching Problem-solving in Physics: A Course in Electromagnetism," *American Journal of Physics* **50**, 8, 725-732 (1982).

Phillip C. Wankat, "Analysis of Student Mistakes, and Improvement of Problem Solving on McCabe-Thiele Binary Distillation Tests" in *Problem Solving*, edited by J. T. Sears, D. R. Woods, and R. D. Noble (American Institute of Chemical Engineers, New York, NY, 1983).

Mark Wood and John Sweller, "Structuring Effective Worked Examples," *Cognition and Instruction* **7**, 1, 1-39 (1990).

J. L. Wheeler and J. W. Regian, "The Use of a Cognitive Tutoring System in the Improvement of the Abstract Reasoning Component of Word Problem Solving," *Computers in Human Behavior* **15**, 3, 243-254 (1999).

A. Whimbey and J. Lochhead, *Problem Solving and Comprehension - A Short Course in Analytical Reasoning* (Franklin Institute Press, Philadelphia, PA, 1980).

Wayne A. Wickelgren, *How to Solve Problems* (W. H. Freeman and Company, San Francisco, CA, 1974).

Donald R. Woods, "PS Corner," *Journal of College Science Teaching*. This column appears in four or five of the six issues printed each year, starting with Volume 12 in 1984.

Richard Zajchowski and Jack Martin, "Differences in the Problem Solving of Stronger and Weaker Novices in Physics: Knowledge, Strategies, or Knowledge Structure?" *Journal of Research in Science Teaching* **30**, 5, 459-470 (1993).

Xueli Zou, "The Use of Multiple Representations and Visualizations in Student Learning of Introductory Physics: An Example From Work and Energy," (unpublished doctoral dissertation, The Ohio State University, 2000).