Development and Validation of a Physics Problem-Solving Assessment Rubric

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Jennifer Lynn Docktor

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Kenneth J. Heller

September 2009

Acknowledgements

First, I am grateful to my adviser Ken Heller whose direction and constructive criticism not only strengthened this research study but also prepared me for future academic pursuits. This document benefited greatly from the feedback of my thesis reviewing committee: Charles Campbell, Cynthia Cattell, and Fred Finley. This dissertation topic is based on previous problem-solving research by the University of Minnesota Physics Education Group, including: Pat Heller, Ronald Keith, Scott Anderson, Mark Hollabaugh, Tom Foster, Jennifer Blue, Tom Thaden-Koch, Charles Henderson, Vince Kuo, and Edit Yerushalmi. I am also grateful to Leon Hsu, Yuichi Kubota, Paul Knutson, Brita Nellermoe, Qing Xu, Mandy Pihlaja, Addishiwot Woldesenbet, Dick Schwarz, and Leonardo Santiago for useful office and seminar conversations (and dark chocolate). I am particularly thankful for Qing's assistance in coding interview transcripts, the eight introductory physics student volunteers who participated in problem-solving interviews, and the many graduate students who tested rubric training materials. I would also like to acknowledge participants in the Research Experiences for Teachers program who assisted in scoring written solutions during summer months: Jay Dornfield and David Parent. In addition I acknowledge Frances Lawrenz for suggesting key references on validity and reliability in education measurement research. Finally, I would like to acknowledge my husband Ben for providing me with a mathematics perspective on problem solving that didn't always agree with my own.

Dedication

This dissertation is dedicated to the family, friends, and teachers who persuaded me to attend graduate school and provided encouragement throughout the process.

Abstract

Problem solving is a complex process that is important for everyday life and crucial for learning physics. Although there is a great deal of effort to improve student problem solving throughout the educational system, there is no standard way to evaluate written problem solving that is valid, reliable, and easy to use. Most tests of problem solving performance given in the classroom focus on the correctness of the end result or partial results rather than the quality of the procedures and reasoning leading to the result, which gives an inadequate description of a student's skills. A more detailed and meaningful measure is necessary if different curricular materials or pedagogies are to be compared. This measurement tool could also allow instructors to diagnose student difficulties and focus their coaching. It is important that the instrument be applicable to any problem solving format used by a student and to a range of problem types and topics typically used by instructors. Typically complex processes such as problem solving are assessed by using a rubric, which divides a skill into multiple quasi-independent categories and defines criteria to attain a score in each. This dissertation describes the development of a problem solving rubric for the purpose of assessing written solutions to physics problems and presents evidence for the validity, reliability, and utility of score interpretations on the instrument.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

Research Motivation

Every day we make decisions, from what to eat and what shoes to wear, to what route to take while running errands. We also make more complex decisions pertaining to education and career paths, or purchasing a home or car. Some of these decisions are made almost automatically without much conscious thought, whereas others require significant time and deliberation. In the latter case, having effective strategies for dealing with everyday problems can reduce the time involved in making them and help to reach a satisfactory solution.

The capacity to approach and solve complex problems is considered an essential skill in today's information-technology-driven society (Jonassen, 2007; Martinez, 1998). Global advances in communication and technology require an ability to adapt to changing circumstances by using knowledge flexibly (Arons, 1990; Reif, 1995; Schwartz, Bransford, & Seers, 2005). Surveys of employers echo this view. In 1991 the U.S. Secretary of Labor and the Secretary's Commission on Achieving Necessary Skills (SCANS) published a report that identified *Thinking Skills* as foundational skills for all competent workers, including creative thinking, decision making, problem solving, visualizing, knowing how to learn, and reasoning. The need for a workforce skilled in technology and scientific problem solving is also reiterated in a recent report by The National Academies Press *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future* (2007).

In response to employer demands, problem solving is often recognized as an important goal for undergraduate education. A list of undergraduate learning outcomes at the University of Minnesota identifies seven such goals, the first of which is “At the time of receiving a bachelor’s degree, students will demonstrate the ability to identify, define, and solve problems” (Carney, 2006). This is especially true for those who earn physics bachelor’s degrees. The American Institute of Physics Statistical Research Center surveyed people who earned physics bachelor’s degrees (with no additional degrees) and work in science-related jobs, including software, engineering, teaching high school, and mangers in technical fields (Ivie & Stowe, 2002). Consistently, scientific problem-solving skills were rated as “essential” to their job, outranking knowledge of physics, lab skills, and computer skills (Ivie & Stowe, 2002).

The physics classroom is one place students can learn the problem-solving skills necessary for their future careers. Problem solving is widely recognized as a primary goal, teaching tool, and evaluation technique of physics courses (Heller, Keith, & Anderson, 1992; McDermott, 1981; Reif, Larkin, & Brackett, 1976). However, many physics instructors find that their students do not solve problems at the desired level of proficiency (Redish, Scherr, & Tuminaro, 2006; Reif, 1995; Van Heuvelen, 1991a). Researchers in Physics Education have responded with alternate pedagogies for teaching problem solving (Heller et al., 1992; Leonard, Dufresne, & Mestre, 1996; Taconis, Ferguson-Hessler, & Broekkamp, 2001; Van Heuvelen, 1991a; 1991b) and alternate types of problems (Heller & Hollabaugh, 1992; Mestre, 2002; O’Kuma, Maloney, & Hieggelke, 2000; Van Heuvelen, 1999; Van Heuvelen & Maloney, 1999). Developing and testing these innovative methods for teaching problem solving requires an agreed-upon definition of what problem solving is, and a satisfactory assessment tool so that student progress in this domain can be assessed. Currently such a tool does not exist (Adams & Wieman, 2006).

Although scoring criteria and rubrics have been used in problem solving research and instruction (Adams & Wieman, 2006; Harper, 2001; Huffman, 1997; Murthy, 2007; Ogilvie, 2007; Reif & Heller, 1982) these instruments are often difficult to use and have not been extensively tested. This study builds upon research at the University of Minnesota (Blue, 1997; Foster, 2000; Heller et al., 1992) to develop an easy to use problem solving assessment instrument for written solutions to physics problems and obtain evidence for validity, reliability, and utility. *Validity* refers to the degree to which score interpretations are supported by empirical evidence and theory. *Reliability* in this context refers to the agreement of scores from multiple raters.

Research Questions

The primary questions that this dissertation addresses include to what extent the scores on a problem-solving rubric are valid, reliable, and useful:

1. To what extent are scores on the problem solving assessment **valid**?
    a. To what extent are rubric categories consistent with descriptions of physics problem solving processes? (*content relevance & representativeness*)
    b. To what extent do scores on the rubric reflect the problem-solving processes undertaken by a solver? (*response processes*)

c. To what extent do scores on the rubric support inferences about students' problem-solving skills from other measures of their performance? (*external structure*)

d. To what extent are the rubric categories independent? (*internal structure*)

e. To what extent is the rubric applicable to multiple populations and contexts, including different student populations, physics topics, and problem features? (*generalizability*)

f. To what extent does the rubric documentation address potential positive and negative consequences of the proposed test use? (*consequences*)

2. To what extent are scores on the problem solving assessment **reliable**?

a. To what extent do multiple raters' scores and score interpretations agree on the same problem solution? (*inter-rater agreement*)

b. What scorer training is necessary to achieve a desired level of rater agreement?

3. To what extent is the problem solving assessment **useful** for evaluating written solutions to physics problems?

a. To what extent can the rubric distinguish between more- and less- skilled problem solvers?

b. How authentic are the assessment's goals, tasks, and constraints?

c. To what extent is the assessment independent of the specific format in which students are taught to express their solutions?

Overview of the Dissertation

Chapter 1 outlines the motivation and need for a quantitative problem-solving assessment tool in physics, lists research questions addressed in this dissertation, and overviews each chapter of the dissertation.

Chapter 2 outlines important definitions relevant for problem-solving, including: problem, problem solving, ill-structured versus well-structured problems, and algorithms and heuristics. Chapter 2 also provides a brief background of problem solving research starting from Aritstotle but primarily reviewing behaviorist research in the early 1900's, information processing theories (including types of knowledge and schema theory), and research with experienced and inexperienced problem solvers in puzzles, games, mathematics, and physics. The chapter concludes with a review of research on problem-solving assessments in physics.

Chapter 3 presents the methodology framework for this study based on the ideas of validity, reliability, and utility for evaluating educational tests. The chapter presents definitions, a brief overview of the concept of validity, and outlines five sources of validity evidence framing this study (content, response processes, internal and external structure, generalizability, and consequences of testing). The chapter concludes by linking the methodology framework and sources of evidence to each part of the study.

Chapter 4 begins with a description of the rubric categories and scores, and how the rubric format and language progressed throughout the study. This chapter also describes data collection procedures and presents data analyses for each part of the study. The parts include a preliminary study with two raters, two studies with training raters, an analysis of written solutions to exams, and student problem-solving interviews.

Chapter 5 discusses the results from each part of the study with respect to the sources of validity, reliability, and utility evidence from the methodology framework. It outlines recommendations for training raters, using of the rubric, and future research.

CHAPTER 2: LITERATURE REVIEW

Introduction

Although there are many descriptions of the important features of problem solving in the literature, there is an agreement that problem solving is a process of decision making. This chapter briefly reviews common definitions and theories of problem solving from cognitive science, mathematics, and physics. It also summarizes research studies on the processes used by experienced and inexperienced solvers on puzzles and games, and within the domains of physics and mathematics. These definitions and processes form the basis for the rubric's development, which will be further clarified in the descriptions of its categories in Chapter 4. The final section of this literature review summarizes other research studies on problem-solving assessment instruments in physics. For a more comprehensive review of problem-solving research, see Mayer (1992); Ormrod (2004); Hsu, Brewe, Foster, & Harper, (2004); and Maloney (1994).

Definitions

Definitions of Problem and Problem Solving

Descriptions of problem-solving emphasize that it is a decision-making process that occurs when a solver is presented with a task for which they have no specific set of actions they can use to reach a solution (Newell & Simon, 1972). For example, Hayes (1989) defines the problem solving process in the following way:

> Whenever there is a gap between where you are now and where you want to be, and you don't know how to find a way to cross the gap, you have a problem. Solving a problem means finding an appropriate way to cross a gap. (p. xii)

Similarly, Martinez (1998) describes problem solving as "the process of moving toward a goal when the path to that goal is uncertain" (p. 605) and Simon (1981) likens problem solving to traversing a maze, in which there is a general sense of a final goal but one cannot predict the obstacles that lie between. Like an ant traveling across a sand dune, Simon comments a problem solver "must adapt his course repeatedly to the difficulties he encounters and often detour uncrossable barriers" (p.64). Problem solving in the workplace was defined by the U.S. Department of Labor SCANS document as: "Recognizes that a problem exists, identifies possible reasons for the discrepancy, and devises and implements a plan of action to resolve it. Evaluates and monitors progress, and revises [the] plan as indicated by findings" (SCANS 1991, p. 32).

In each of these definitions, problem solving depends on the solver's experience and perception of the task (Martinez, 1998). What is considered a problem for one person may be a routine exercise for another person (Schoenfeld, 1985; Woods, 2000). Newell and Simon (1972) also address the issue of encountering a familiar, previously-solved problem. They note that when a problem is solved by the *Recognition Method* (pp. 94-95) and the result is known immediately, the solver's actions must be interpreted in light of their perceived problem difficulty.

Problem Structure

The basic components of a problem include the initial state (*givens*), a desired end state (*goal*), and means to get from the initial state to the end state, (*operations)* (Chi, Feltovich, & Glaser, 1981; Ormrod, 2004). Problems can differ vastly, however, in their structure. On one extreme of the continuum are problems that are a straightforward application of concepts or principles, that clearly state the givens and desired goal, and for which all information needed to solve the problem "correctly" is presented. These are referred to as *well-structured* problems (Jonassen, 1997; Pretz, Naples, & Sternberg, 2003). On the other extreme are problems for which the desired goal is vague, some necessary information is absent, and for which there might be several viable solution paths. These are called *ill-structured* problems, and the topics for such problems often emerge from real-life situations (Jonassen, 1997).

In educational settings, the types of problems appearing in standard textbooks are relatively well-structured in that they have a correct or "preferred" solution that involves the application of concepts and principles from the subject domain. Within physics, recent efforts to develop alternate types of problems that include some characteristics of ill-structured problems include context-rich problems (Heller & Hollabaugh, 1992), Jeopardy problems (Van Heuvelen & Maloney, 1999), experiment problems (Van Heuvelen, 1999), ranking tasks (O'Kuma, Maloney, & Hieggelke, 2000), or problem posing (Mestre, 2002). Within engineering fields, alternate problem types include project-oriented design problems (Dutson, Todd, Magleby, & Sorensen, 1997; Heywood, 2005) or model-eliciting activities (Diefes-Dux, Moore, Zawojewski, Imbrie, & Follman, 2004).

Shin, Jonassen, and McGee (2003) suggest that the degree of structure in a problem affects the skills required to successfully reach a solution. In their research with astronomy students, they found that domain knowledge and justification skills were necessary for solving both ill-structured and well-structured problems. Success on ill-structured problems, however, was influenced by a student's attitudes toward science and their level of cognitive regulation skills.

Algorithms and Heuristics

Descriptions of methods for solving problems often draw a distinction between *algorithms* and *heuristics* (Pólya, 1945, 1957; Schoenfeld, 1985; Martinez, 1998; Pretz et al., 2003; Ormrod, 2004). The term algorithm refers to step-by-step procedures that will guarantee a correct solution every time, if applied correctly. An example of an algorithm is the mathematical procedure for carrying out long division or the steps for tying a shoe. The term heuristic is used to refer to general strategies or "rules of thumb" for solving problems (Martinez, 1998). Examples of heuristics include combining algorithms, hill climbing (working forward), analogies, successive refinements, means-ends analysis, working backward, and using visual imagery or external representations.

Basic-level mathematics and physics problems are often solved using the heuristic of combining algorithms, by using several algebraic procedures in succession (Ormrod, 2004). In hill climbing, problem solvers work forward by performing actions that bring them closer to the goal. In analogical problem solving a solver refers to a familiar or previously-solved problem that is similar to the presented problem and applies a similar method. Successive refinements is a process common in writing; initially a rough outline or draft is produced, and then it is gradually revised to a finished product. Means-ends analysis is perhaps one of the most well-known heuristics, by which a solver breaks the problem up into subgoals, and then works successively on each of them (Newell & Simon, 1972; Gick, 1986; Martinez, 1998). Another common heuristic is working backward, in which the solver starts at the desired goal and considers the reasonable steps just prior to that goal, and continues in this process until reaching the initial state (Martinez, 1998). The use of external representations or visual imagery is a general strategy to represent the complexity of a problem and free up some space in working memory; it also creates a public document that can be examined by others (Pólya, 1945; 1957).

In summary, problem-solving is recognized as a decision-making process that is influenced by a solver's own experience, knowledge, and interpretation of the task. A problem that is familiar or recognized might actually be an *exercise* for that person. Problems can vary in structure, from *well-structured* problems with clear goals and required operations to *ill-structured* problems that might have multiple reasonable solutions. General problem-solving methods can include both step-by-step *algorithms* and general frameworks or *heuristics*. These definitions give an introduction to some of the important terms in problem solving. The next section will expand on these ideas to explore early theories of problem solving and the role of knowledge content, knowledge organization, and processes involved in problem solving.

General Problem-Solving Theories

Efforts to understand human reasoning and problem solving were made early in human history and are particularly evident in the philosophy of ancient Greece. Experimental problem-solving research, however, is often traced back to animal research of behavioral psychologists and Gestalt psychologists in the early 1900s. This research eventually grew to include human participants and led to cognitive theories including the information processing theory of memory and schema theory. This section gives a brief history of these early theories and research relevant to problem solving up through the mid-1900s.

Philosophy and Problem Solving

Theories of human thought processes including logic and the structure of memory can be easily traced back at least 2,000 years to ancient Greece (Mayer, 1992; Anderson & Bower, 1980). Although there were earlier philosophers in East Asia (Confucius, Laozi) and Persia (Zarathushtra, Mani, Mazdak), their teachings on human thought processes as they pertain to problem solving are not well documented (Blackburn, 1994).

In the fourth century B.C. the Greek philosopher Aristotle wrote about intuition and discursive reasoning, or the way that the human mind orders, argues, and reasons (Kal, 1988). He set forth particular procedures for arguments starting with an assumption or agreed-upon premise and chains of reasoning that lead to conclusions drawn from that premise, a process called syllogistic logic (Kal, 1988; Lear, 1980). Some interpretations of Aristotle's works attribute to him the principles of proof by induction, outlining a formal system for valid inferences used in mathematical arguments (Lear, 1980).

Aristotle also identified two main components important for thinking and learning: basic elements called ideas and the links or "associations" between them (Mayer, 1992). He also set forth three doctrines describing the nature of such links:

events that occur in the same time or space are linked (*association by contiguity*), similar items are linked (*association by similarity*) and opposite items are linked (*association by contrast*). Thinking was viewed as traversing ideas via a chain of such associations.

The nature of thinking and reasoning outlined by the Greek philosopher Aristotle laid a foundation for later theories of problem solving, particularly his notion that the human mind is a linked network of ideas. Much later, investigations into the nature of such links in the mind were studied by experimental psychologists in America and Germany. For brevity, examinations of human thought that occurred after Aristotle and before experimental psychology are not included in this review.

Behaviorist Theories of Association

Experimental psychological research became organized in the late 1800's, with studies of observable behavior in animals interacting with puzzle boxes (Thorndike, 1898, 1911; Skinner, 1938). Thorndike described the processes of animals as essentially "trial-and-error" with initial success occurring accidentally. After repeated exposure to the puzzle, the time for an animal to respond correctly was diminished until correct performance became immediate. However, Thorndike was unwilling to attribute this behavior to reason or learning; he claimed that it was the result of an association between a physical or sense impression of the puzzle situation and the actions that will result in positive consequences, such as food or escape (Thorndike, 1898).

Thorndike set forth two laws describing behavior: the Law of Effect and the Law of Exercise (Thorndike, 1911). The Law of Effect states that when several responses are made to a problem situation, the actions resulting in success will be strengthened and the ones resulting in discomfort will be diminished. Also, the Law of Exercise emphasizes the importance of repetition of a task for strengthening the connection between a situation and response (Thorndike, 1911; Ormrod, 2004). Skinner (1938) built on the ideas of Thorndike in his theories of behavior and conditioning, but focused on increasing the frequency of a response with the use of reinforcement rather than the strengthening of connections between a situation and a response. He distinguished *classical* conditioning as an automatic and involuntary response evoked by a stimulus and *operant* conditioning as a voluntary response of an organism to a particular reinforcing stimulus.

Research experiments by Hull (1943) expanded the stimulus-response approach to consider factors such as motivation and its influence on the strength of a stimulus-response association on behavior. In his theory, an organism learns different responses for the same stimulus, and organizes them in what is called a "response hierarchy" according to habit strength. When an organism is presented with a problem (stimulus) the first response produced will have the strongest association with the problem, perhaps based on the perceived effectiveness of that response for solving similar problems in the past. If that response doesn't work, other responses will be produced in order of their strength until either the problem is solved or the list of known responses is exhausted.

Psychologists later concluded that trial-and-error behavior is a generally ineffective and time consuming approach to problem solving and is only of use for problems in which there are a limited number of possibilities to try. As a result, early theories of learning and problem solving that focused on observable behavioral responses have largely been abandoned in favor of cognitive perspectives that consider internal mental processes. This shift in the field began with what is called Gestalt psychology, and then further developed into information processing theory, constructivism, and situated cognition. Also relevant to problem solving are ideas of "types" of knowledge and schema theory.

Cognitive Theories

Gestalt Psychology

During the Behaviorism movement in America, German psychologists were beginning to focus on mental processes to describe learning and problem solving (Ormrod, 2004). The theory of Gestalt Psychology emerged from the works of Wertheimer (1945), Koffka (1935), and Köhler (1929) in their attempts to understand perception of visual objects and optical illusions. They suggested that the organizational structure of perception, learning, and problem solving must be viewed holistically rather than as consisting of separate elements.

Köhler observed chimpanzees interacting with a problem solving task, but rather than attribute their behavior to trial-and-error he thought that the chimps appeared to carefully examine all parts of a problem and deliberately take action in their attempts to solve the problem. Gestalt psychologists concluded that problem solving is a process of combining and recombining components of a problem mentally (what was called "restructuring") until a point of insight into the problem solution is reached. Wertheimer (1945) termed this kind of intentional problem solving based on knowledge and previous experiences *reproductive* thinking, and distinguished it from *productive* thinking which is a quick, unplanned response to a situation resulting from insight.

Unfortunately, Gestalt theories only describe restructuring and insight, and cannot give an explanation of how such processes occur.

Information Processing Theory of Memory

Contemporary theories of learning and problem solving are based on the information processing theory of memory (Newell, Shaw, & Simon, 1958; Newell & Simon, 1972; Ormrod, 2004; Redish, 2003). In brief, this theory asserts that information in the brain is stored in two primary components of memory: short term or “working” memory and long-term memory. Short-term memory is limited in size and the length of time it can hold information, and is thought to contain distinct verbal and visual parts. In contrast, long-term memory can hold vast quantities of facts and data for long periods of time, but to be accessed this information must be activated by being brought into working memory.

In problem solving, it is thought that working memory is utilized to process information about the problem and maintains its accessibility during the problem solving process. Since working memory has limited storage capacity, it is possible that information in a problem can exceed this limit (called cognitive overload) and interfere with attempts to seek a solution (Sweller, 1988). For this reason, information about a problem is often stored externally (written down) or processed externally (such as with a calculator or computer) in order to free up space in working memory that can be devoted to the task. Also, some skills involved in problem solving can be practiced until they become automatic, which also minimizes the use of working memory capacity.

Types of Knowledge

A cognitive approach to problem solving recognizes that there are several different types of knowledge a solver must possess, including knowledge of situations, concepts, procedures, and strategies (de Jong & Ferguson-Hessler, 1996). *Situational* knowledge refers to an awareness of common problem situations or contexts that occur in a domain, which can help a solver recognize important features and form a representation of the problem. Knowledge of how to recognize relevant problem features and construct visualizations might be called *problem-state* knowledge and *representational* knowledge (Dufresne, Leonard, & Gerace 1995). *Conceptual* or "declarative" knowledge includes the major definitions, concepts, and principles of a subject, whereas *procedural* knowledge includes actions appropriate for solving problems in a particular domain. *Strategic knowledge* can include general or domain-specific heuristics.

Another important type of knowledge is *metacognition*, which refers to an individual's awareness of his or her own thinking processes (Flavell, 1976; 1979; Martinez, 1998). Metacognitively engaged problem solvers have developed skills at planning their problem solving approach, monitoring their progress toward the goal while following their plan, and evaluating the effectiveness of their chosen strategies. (Pretz et al., 2003; Schoenfeld, 1983). Since metacognitive solvers are careful to evaluate their assumptions and are less apt to persevere in unproductive strategies, they are more likely to solve complex problems successfully (Chi, Bassok, Lewis, Reimann, & Glaser, 1989).

In addition to cognitive and metacognitive knowledge, there are also affective characteristics relevant to problem solving including attitude, motivation, interest, confidence, or attribution (Mayer, 1992, 1998; Jonassen, 1997; Schoenfeld, 1983). There are also conceptions about problem solving that can influence performance, such as the belief that problems solved in school do not relate to real life, all problems are solvable in ten minutes, or that subjects like physics and math are for “geniuses” so it is fruitless to attempt to understand their problem solutions (Schoenfeld, 1985).

Some theories of cognitive skill acquisition outline a process by which a solver first obtains factual, declarative knowledge about a topic, and then transforms declarative knowledge into a procedural form by solving problems and viewing worked examples (Anderson, 1987; Anderson, Greeno, Kline, & Neves, 1981). From these initial problem solutions, a solver compiles a set of “production rules” similar to conditional if-then statements useful for future problem solving.

Types of knowledge is not the only language for describing the cognitive requirements of problem solving. Other terms include abilities, skills, or processes. Mayer (1998) uses the term *skill*, distinguishing between three types: skill, metaskill, and will. This idea is similar to Schoenfeld’s categories of knowledge resources, control resources (monitoring and decision-making resources), and belief systems which each influence problem solving behaviors (Schoenfeld, 1983).

Schema Theory

In addition to utilizing short-term memory, problem solving also requires accessing relevant information - the solver’s knowledge base about the problem - from storage in long-term memory (Ormrod, 2004). Critical factors in this retrieval of information include *what* has been stored and *how* it has been stored, and the cues or patterns present in the problem that help the individual perceive what information to access from memory. In order to be retrieved, knowledge from a content domain must be present in memory to begin with; it must also be organized in a way that facilitates its retrieval in an appropriate context.

With experience in a content domain, it is believed that problem solvers develop cognitive structures called "problem schemata" that allow them to recognize a problem as belonging to a particular category (Sweller, 1988). This mental classification of problem types can trigger particular actions for solving the problem, based on the perceived similarity of the presented problem to the same category of others stored in memory.

Marshall (1995) describes a schema as an "organized memory structure" (p. 31) or "basic storage device" (p. 46) that is centered around a general concept. This general concept develops as a person repeatedly encounters instances of that concept and abstracts common features from multiple situations. Schemas are believed to have a network structure, meaning that the subtopics of a schema are connected to each other and linked to other related schemas. These structures can vary in their size and complexity, and a schema is flexible in that its connections can change over time. A schema can be composed of both declarative knowledge (facts) and procedural knowledge of how to use or apply the information. The organization of an individual's schemas in memory and the strength of connections between concepts have a profound influence on the accessibility of information for problem solving, which will be described more in the section on *Expert-Novice Research*.

Problem-Solving in Puzzles and Games

Using games and puzzles to gain insight into human problem solving began with logic puzzles such as the Tower of Hanoi puzzle - moving disks on pegs such that a larger disk is always beneath a smaller disk, and the Missionaries and Cannibals problem (Mayer, 1992; Newell & Simon, 1972; Simon, 1975; 1976). This research grew to include more complex rule-based games such as chess (Chase & Simon, 1973) and eventually led to investigations of problem solving within a subject domain such as mathematics or physics (Bhaskar & Simon, 1977; Simon & Simon, 1978).

Perceptual Patterns in Chess

Chase and Simon (1973) investigated the memory of chess board configurations among subjects with varying levels of experience and expertise. When pieces were placed randomly on the playing board, reconstruction of the piece locations was poor for both experts and novices (and in some cases even worse for experts). When the pieces were placed in patterns from actual game configurations (such as common middle game or end game set-ups), experts did much better. Experts could recognize familiar patterns or perceptual "chunks" of pieces commonly encountered together during play and more accurately reconstruct the location of pieces on a blank playing board.

These researchers concluded that an expert chess player doesn't necessarily have more short-term memory capacity (seven items plus or minus two), but rather that information is organized in a way such that each "chunk" is larger (such as a chess board pattern containing several pieces rather than the location of individual pieces). They also suggested the possibility that expert chess players have automated many decision processes, allowing them to manage more perceptual chunks at once (Chase & Simon, 1973).

A Theory of Human Problem Solving

According to Newell, Shaw, and Simon's theory of human problem solving (1958, 1972), the process of problem solving takes place in a mental state called the "problem space", consisting of a solver's internal representation of the given and target information, along with all knowledge available to the solver. They describe problem solving as an iterative process of representation and search within the problem space: first the solver selects a goal, then selects a method to use (such as a heuristic) and applies it, evaluates the result of this choice, revises the goal and/or selects a subgoal, and proceeds until a satisfactory solution is achieved or the problem is abandoned.

Gick (1986) builds on this "representation and search" model of the problem solving process to consider the role of schema activation. If constructing an internal representation of the problem activates an existing schema or a memory framework of how to solve the problem, the solver implements a solution immediately. If this implementation is successful, the problem solver stops. If not, s/he backtracks to redefine the problem or try a different method. If no schema is activated, the solver must "search" through available information and apply general strategies in an attempt to progress toward the goal.

The Newell & Simon (1972) model of problem solving emphasizes that it is not necessarily a linear process, and highlights the role of a solver's internal representation or perception of a problem for suggesting possible solution methods. Unfortunately this model is a very broad description of cognitive processes and does not identify discrete, observable behaviors. For the purposes of research and instruction, it is useful to consider typical stages of problem solving that are observed within a domain such as science or mathematics.

Problem-Solving in Physics and Mathematics

Polya's Four Problem-Solving Steps

One of the early modern attempts to identify stages involved in the type of quantitative problem solving used in mathematics and science was by the mathematician Pólya (1945; 1957). In his first step *Understanding the Problem*, the solver summarizes known and unknown information, introduces suitable notation, and draws a figure. Next, in *Devising a Plan*, the solver uses their knowledge to plan how to connect the given data to the desired goal. Then in *Carrying out the Plan* the solver implements their plan by carrying out the necessary procedures to reach an answer while checking their work along the way. The final step is *Looking Back* or examining the result to check that it makes sense, and if possible using an alternate procedure to achieve the answer. Hayes (1989) expanded these actions to include a first step of recognizing the existence of a problem and a final step of consolidating gains or explicitly considering what was learned from solving the problem and how it might be useful for solving future problems.

Expert-Novice Research

Information about problem solving processes and knowledge structures have been obtained from research studies comparing experienced or "expert" problem solvers to inexperienced or "novice" problem solvers (Chi et al., 1981; Larkin, 1979; Schoenfeld & Hermann, 1982; Simon & Simon, 1978). Many of these studies focused on the content of knowledge and its mental organization as a basis for explaining observed process differences. In many early studies, the experts were professors and the novices were beginning students. Other studies have compared strong and weak beginning students (successful and unsuccessful novices). Within physics research, think-aloud protocols (Larkin, 1979; Larkin et al., 1980a) and problem sorting tasks (Chi et al., 1981) usually focused on solving standard textbook problems and were limited to a few topics in mechanics such as motion with constant acceleration, Newton's second law, or conservation of energy.

Knowledge Organization Differences

In order to be successful at solving a problem in a particular subject domain, the solver must have a strong conceptual knowledge base and this knowledge must be organized in such a way that it is accessible at the appropriate time (Chi, Glaser, & Rees, 1982; Reif & Heller, 1982; Simon & Simon, 1979). Chi et al. (1981; 1982) asked novices and experts to sort cards containing problem statements on them based on their perceived similarity. Novices tended to sort the problems based on the literal objects stated in the problem such as pulleys and ramps, or what was referred to as "surface similarity". Experts, on the other hand, tended to sort the problems based on major principles of physics they would use to solve the problem, such as conservation of energy or Newton's second law, or what was called "deep structure" of the problem (Chi et al., 1981; 1982).

From their observations of expert and novice physics problem solvers, Larkin (1979) and Chi et al. (1981, 1982) drew conclusions about the content and mental organization of physics knowledge. They found that an expert's memory is structured hierarchically around a small number of fundamental physical principles called "chunks" (Eylon & Reif, 1984). Such principles are considered fundamental because they can be applied to a wide range of physical situations (Larkin, 1981). Accessing a chunk also cues other useful relations and the procedures or actions to successfully apply those principles (Chi et al., 1981; 1982; Larkin et al., 1980a). In contrast, the novice's knowledge structures are disconnected and each relation must be accessed individually. There is no clear link between physics principles and application procedures. This mental organization makes the novice's solution search an inefficient and time-consuming process (Larkin, 1979).

Comparisons have also been made between successful and unsuccessful or strong and weak novice problem solvers. For example, Finegold and Mass (1985) concluded that good novices interpreted the problem statement correctly, spent more time planning their solution, made greater use of physical reasoning, but didn't necessarily evaluate their solution more than poor novices. The research of de Jong and Ferguson-Hessler (1986, 1991) confirmed that good novices have their knowledge organized around problem "types" or categories based on fundamental physics principles, whereas poor novices organized problems by nonessential features. Several such studies with varying levels of expertise have concluded that the "expert-novice" distinction is not entirely dichotomous, and beginning students can exhibit some of the same problem categorization behaviors observed in experts and vice versa (Bédard &

Chi, 1992; Hardiman, Dufresne, & Mestre, 1989; Priest & Lindsay, 1992; Zajchowski & Martin, 1993).

Within mathematics expert-novice research, similar results were found in that knowledge of an expert is organized hierarchically around a small number of key ideas. Rather than physical principles, however, it was found that the key ideas in mathematics problems are methods of solution (such as proof by contradiction) rather than an underlying ‘principle’ (Schoenfeld & Hermann, 1982).

Therefore, an important part of problem solving is developing appropriate knowledge structure or schema in memory that is based on central concepts and solution methods. The ability to generalize such key features and solution methods from example problems is important for developing problem-solving expertise (Anderson et al., 1981; Chi, Bassock, Lewis, Reimann, & Glaser, 1989; Gick, 1986). This ability to generalize from examples also facilitates the *flexible use* of knowledge in new, novel, or complex problem situations (Newell & Simon, 1972; Reif, 1981; Schoenfeld, 1992). Another similar distinction is made by Schwartz, Bransford, and Seers (2005) between *routine* experts who can efficiently solve common problems in a domain and *adaptive* experts who can innovatively apply problem-solving skills to new, unfamiliar problems. Observed differences in the problem-solving processes are related to these differences in knowledge content and organization of experts and novices .

Process Differences

Several researchers observed that experts engage in a low-detail overview of problem features and expectations, called a *qualitative analysis*, before writing down quantitative relationships (Chi et al., 1981; Larkin, 1979; Larkin et al., 1980a; Larkin & Reif, 1979). For an expert, simply reading a problem and representing the important features in memory can evoke an appropriate schema containing the necessary concepts, appropriate equations, and procedures to solve the problem (Schoenfeld & Hermann, 1982). Sometimes this phenomenon is referred to as physical intuition about a problem (Chi et al., 1982; Simon & Simon, 1978; Singh, 2002).

Experts use their qualitative analysis and intuition to consider possible solution approaches or physics principles that might be useful in solving the problem. Novices tend to skip this planning step and jump directly to writing down individual equations or miscellaneous mathematical relationships (Larkin & Reif, 1979; Reif & Heller, 1982). Even though novices skip this qualitative analysis or planning step, it doesn't mean they are not capable of engaging in a qualitative analysis. Bagno and Eylon (1997) found that some novices could produce a qualitative representation of a problem, but usually did not do so spontaneously while solving physics problems.

When an expert problem-solver reads a problem, they generally first represent the relevant features and attempt to "categorize" the problem as a particular type based on similarity to problems they might have solved previously and/or an existing schema structure in memory (Chi et al., 1981; 1982). Activating a schema usually also evokes information about solution procedures. For this reason, the processes of an expert are sometimes referred to as *schema-driven* rather than a novice's *data-driven* approach that focuses on specific objects or data given in the problem (Chi et al., 1982; Gick, 1986). An alternate suggestion is that novices possess schemata, but their knowledge is incomplete and disconnected and this organization hinders retrieval of appropriate information (Bagno & Eylon, 1997; Chi et al., 1982).

Some researchers have assigned other names to the processes used by experts and novices. Walsh, Howard, and Bowe (2007) designate the expert's qualitative analysis and concept-based approach as a "scientific approach". Names for novice approaches include "plug-and-chug" in which the solver selects equations based on the quantities in the problem statement and proceeds to solve for unknown quantities (either systematically or by trial and error), or a "memory-based" approach of mimicking the solution for a similar, previously-solved example problem (Walsh et al., 2007). It is also possible that the novice's reasoning is so haphazard they he or she has no discernable approach.

Redish, Scherr, and Tuminaro (2006) have identified what they call knowledge-building or epistemic games (*E-games*) that students engage in during problem solving. One such game is *Recursive Plug-and-Chug* (similar to plug-and-chug described above) in which students perform calculations without much reasoning or sense-making processes. Students can also reason about the problem's physical situation (*Physical Mechanism*) or draw a picture (*Pictorial Analysis*) to aid them in understanding the problem. Students also must identify relevant physics for the problem (*Mapping mathematics to meaning*) and apply concepts to the conditions in the problem to produce mathematical equations (*Mapping meaning to mathematics*).

For easy problems or "exercises", many of the expert's problem solving processes have become automated, they tend to work forward with little explicit planning whereas novices tend to start from the unknown quantity and work backward (Larkin et al., 1980a; Schoenfeld, 1985; Simon & Simon, 1978). Although problem-solving direction was once thought to be a relevant expert-novice distinction, other researchers have concluded that the forward vs. backward direction of a problem solver's process is irrelevant to expertise because experts only work forward on familiar exercises in which they are confident they can reach a correct solution (Larkin et al. 1980a; Priest & Lindsay, 1992; Schoenfeld, 1985).

Simply having appropriate knowledge schemata is insufficient for successful problem solving; other skills and strategies such as general heuristics are necessary to execute solution procedures (Bagno & Eylon, 1997). For example, experts in physics have strong mathematical skills and strategies for monitoring progress and evaluating their answer (Larkin et al., 1980a; Reif, 1981; Reif & Heller, 1982). Within physics problem-solving, methods for evaluating an answer can include examining limiting or extreme cases for quantities, checking the reasonableness of the value obtained, and checking that all questions were answered sufficiently (Rief & Heller, 1982).

Progression from Novice to Expert

Dreyfus & Dreyfus (1986) describe five stages of skill acquisition: novice, advanced beginner, competence, proficiency, and expertise. The *novice* stage is characterized by learning objective facts and rules independent of context or special case exceptions. Decisions are deliberate and tasks require full concentration. As a novice builds up experience or practice in a domain, they acquire more sophisticated rules and begin to base decisions on similarity with previously encountered situations (a stage referred to as *advanced beginner*.) In the next stage called *competence*, a performer learns to organize situation features into a hierarchy that identifies the most important elements and to follow a decision-making plan based on the presence or absence of particular features. *Proficient* performers have acquired an intuitive ability to recognize salient features of a task and organize information, but still consciously analyze available options. *Experts* (on routine tasks) make automatic, intuitive decisions without time for reflection. There is no need to devise a plan because they "know" what works. As a result of some overlap in stages, Alexander (2003) shortens them from five to three: acclimation, competence, and proficiency/expertise.

Within the domain of physics problem solving, the shift from novice to expert is characterized by changes in the way knowledge is organized in memory and the strategies used while solving a problem (Chi et al., 1981; Elio & Scharf, 1990). As in the Dreyfus & Dreyfus (1986) description, a progressing problem-solver learns to extract relevant information from a problem statement or solution and generalize across multiple different problem types and topics (Chi et al. 1989; Chi, 2006; de Jong & Ferguson-Hessler, 1991; Gick, 1986). Chi, Glaser, and Rees (1982) and Bédard & Chi (1992) describe learning in physics problem-solving as a restructuring of schemata from weakly-linked factual knowledge centered on physical problem situations or objects to a rich network of meaningful memory connections focused around fundamental physics principles. Chi et al. (1982) and Gick (1986) describe the development of problem-solving expertise as a progression from initial search-driven strategies toward schema-driven strategies.

Successful beginning problem-solvers may share some characteristics of both novices and experts, by using surface characteristics to categorize problems in some cases and deep structure in other situations (Hardiman et al., 1989; de Jong & Ferguson-Hessler, 1986). Competent problem solvers may also vary in their skills at qualitatively analyzing a problem, monitoring progress while solving, and evaluating a solution (Chi, 2006). A summary of research on experienced and inexperienced problem-solvers in physics is provided in Table 1: Summary of Expert-Novice Research.

Table 1: Summary of Expert-Novice Research

Experienced Solvers	Inexperienced Solvers	References
Categorize problems based on physics principles (deep structure)	Categorize problems based on objects and features (surface structure)	Chi, Feltovich, & Glaser (1981); de Jong & Ferguson-Hessler (1986); Hardiman, Dufresne, & Mestre (1989)
Have knowledge of physics principles stored as a schema that includes procedures and conditions for their application	Have disconnected physics knowledge with weak or no links to application procedures	Larkin et al. (1980a); Larkin (1979,1981a, 1981b); Eylon & Reif (1984); Schoenfeld & Hermann (1982); Chi, Glaser, & Rees (1982); Bagno & Eylon (1997)
Perform a low-detail qualitative analysis (or basic description) of a problem before writing equations	Start problem-solving by writing down mathematical relationships	Larkin (1979; 1981a); Larkin & Reif (1979); Simon & Simon (1978; 1979); Reif & Heller (1982); Heller & Reif (1984); Finegold & Mass (1985); Redish, Scherr, & Tuminaro (2006)
Have strategies for monitoring progress while solving and evaluating the answer	Often get stuck while working on a problem	Chi (2006); Larkin (1981b); Reif & Heller (1982); Singh (2002)
Can generalize key features and solution methods from problems	Have difficulty abstracting problem similarities	Chi, Bassock, Lewis, Reimann, & Glaser (1989); Gick (1986); de Jong & Ferguson-Hessler (1991); Dreyfus & Dreyfus (1986); Schwartz, Bransford, & Seers (2005); Simon (1975)

Implications of Expert-Novice Research

Several physics educators have adapted problem solving definitions and the processes informed by expert-novice research together with observations of student problem solving actions to develop problem-solving strategies or frameworks for physics instruction (Heller & Heller, 2000; Heller et al., 1992; Reif, Larkin, & Brackett, 1976; Van Heuvelen, 1991b). These frameworks use writing to guide the student's use of an organized problem-solving strategy and make explicit the complex processes done implicitly by experts.

These frameworks typically subdivided the first step of understanding the problem (Pólya, 1945) to highlight the importance of multiple representations or problem descriptions in solving physics problems (Heller & Reif, 1984; Larkin, 1981; Larkin et al. 1980a, 1980b; Reif & Heller, 1982). In particular, Heller and Reif (1984) suggest that effective problem solvers first generate a "basic description" that summarizes the relevant information about the situation in symbolic, pictorial, and verbal forms prior to producing a "theoretical description" that contains abstracted diagrams specific to physics concepts and principles.

Although expert-novice research studies provide useful insight into physics problem-solving processes, they also have limitations (Heller & Reif, 1984). The physics topics used in the studies were not representative of the entire domain of physics, and the tasks were typically standard textbook-style quantitative problems. The expert-novice dichotomy does not consider intermediate stages in problem solving, such as progressing levels of competency (Heller & Heller, 2000). In addition, the problems were often “exercises” for the experts and might not reflect the processes engaged in for more difficult problems (Gick, 1986; Schoenfeld, 1985). Nonetheless, for assessment purposes it is important to consider the expert-like processes of qualitative descriptions, approaches based on fundamental physics principles, procedures for the appropriate application of principles, skilled use of mathematics, and strategies for monitoring progress and evaluating results.

Research on Problem Solving Assessments

Currently, there is no single, standard measure to quantitatively assess problem solving (Adams & Wieman, 2006). In most introductory physics courses, students’ problem solutions on homework or exams are given a score based on the correctness of the algebraic or numerical solution. A common grading practice in physics involves giving students partial credit for particular characteristics of their written solution, as compared to the ideal solution developed by the instructor. Usually partial credit values are based on the problem features and physics topic, and can vary substantially across different problems (Henderson et al., 2004). In some instances, instructors award points based on a problem-solving framework that has been modeled for students during the course.

Research into problem solving has used several different means to measure problem solving performance. One method used by Larkin and Reif (1979) involves measuring the time it takes a problem solver to write down each quantitative expression in their solution, and recording the total time to reach a solution. Some researchers have also investigated problem solving using think-aloud protocols or interviews, in which students engage in conversation to explain their thinking processes as they attempt a problem (van Someren, Barnard, & Sandberg, 1994). A difficulty with these methods is the time involved to prepare and conduct them, the vast amount of data generated from interview transcriptions, and the complicated nature of the data analysis (Harper, 2001). In order to compare problem solving performance for many students, it is desirable to have a quantitative measure that can be determined relatively quickly. Researchers who have attempted to assess problem solutions on the basis of criteria or characteristics include Reif and Heller (1982), Heller, Keith, and Anderson (1992), Huffman (1997), Blue (1997), Foster (2000), Harper (2001), Murthy (2007), and Ogilvie (2007).

Research by Reif and Heller (1982) assessed physics problem solutions on the basis of five criteria. The first criterion of *Clear Interpretation* requires that the parameters of the problem are clearly defined and specifications are provided (such as direction and magnitude of vectors, direction of motion, and reference frames). The criterion of *Completeness* requires that all questions are answered completely and expressed only in terms of known quantities. *Internal Consistency* means that the solution is free of logical errors, and *External Consistency* means that the answers agree with expected relationships between parameters. The final criterion of *Optimality* requires that the solution is simple and easily interpretable.

An investigation described by Heller, Keith, and Anderson (1992) used a rating scale of problem solving performance based on six characteristics of expert-like problem solutions. These characteristics include evidence of conceptual understanding, usefulness of the problem description, consistency of the specific equations with the physics description written, reasonableness of the plan, logical progression of the mathematical solution from physics principles to problem-specific expressions, and the use of appropriate mathematics. The scores for each category were examined independently, and, in the study (Heller et al. 1992) were also weighted equally and normalized to obtain a total problem-solving score out of 100 points.

Similarly, Huffman (1997) outlines the following five criteria used to assess written solutions: quality of the physics representation, completeness of the physics representation, match of equations with the representation, organized progression, and mathematical execution. Each criteria was assigned a numerical score in which zero represents nothing written and a high score represents complete and/or correct. The first and third categories were scored from 0-2, the second category was scored 0-3, and the final two categories had a scoring range of 0-4.

Similar coding rubrics were used to measure problem solving in the unpublished doctoral dissertations of Jennifer Blue (1997) and Tom Foster (2000) at the University of Minnesota. Blue (1997) assessed solutions on four criteria: general approach (physics principles used and understanding of them), specific application of physics (including vector components, defining systems and symbols), logical progression (solution organization), and appropriate mathematics. Each category had seven or eight score criteria where the first "Nothing written" was assigned a score zero and the highest criterion was assigned a score of 10; other categories were normalized for a total score of 40. Foster's (2000) coding scheme had the same four criteria headings, but with slightly different interpretations. In addition to changes in wording, General Approach was modified to include the student's representation of the problem statement, and three half-score criteria were added to the Specific Application of Physics. In Foster's scheme some of the coding scores were subdivided (7a, 7b, etc.), effectively resulting in subcategories.

Within the context of a research study to track the progression of problem-solving skills in an introductory physics course, Harper (2001) evaluated students' written solutions on the usage of diagrams, initial equations (starting from generalized equations or specific equations and algebraic or numeric), number insertion, use of words, and fractionation or subdividing the problem into appropriate pieces. In Harper's analysis, Diagram Usage was assessed by counting the fraction of students in the class that used free-body diagrams and bar charts during the term, and characterizing the diagrams as qualitative or quantitative. The General / Specific Initial Equations category was assessed on a scale of 1-5 from all specific to all general for each problem solution. Algebraic / Numerical Initial Equations was also assessed on a scale of 1-5, with 1 representing all equations contain numbers and 5 all the initial equations are purely algebraic. Number Insertion was assessed on a scale from 1-7 from the first equation containing numbers (1) to numbers inserted at the very end (7). Word usage was assessed by counting the number of instances of written words on each problem solution. Fractionation scores were assigned from 0-2: none=0, improper=1, and appropriate=2. Other aspects of problem-solving that were not observed in the study's written problem solutions included unnecessary equations, incorrect constraints, choosing among multiple approaches, (mis)understanding the question, and restarting / revising.

Sahana Murthy (2007) developed a rubric for the purpose of guiding students through a self-assessment and peer-assessment process. Students were trained to assess written homework solutions on the criteria of physics content, relevant representations, modeling the situation, problem-solving strategy, and reasonableness of the answer. Scores ranged from 0-3 with 0=Missing, 1=Inadequate, 2=Needs Improvement, 3=Adequate, and NA not applicable. Only providing the rubric and training was insufficient, however; students also required a taxonomy that outlined specific expectations of the five criteria for a particular problem.

Craig Ogilvie (2007) developed two problem-solving rubrics, one for "Mileposts" in a written solution and one for "Process". The Mileposts rubric identifies five aspects of the solution: diagram(s), citing key physical principles, basic equations, algebraic manipulations, and conclusions consisting of a numerical calculation and clear statement of an answer. The Process rubric identifies six criteria related to problem-solving processes modeled for students during class, with a focus on metacognitive strategies such as representing the problem, planning, monitoring, and evaluating. The six criteria include focus the problem (identify key issues and a goal), qualitative representation, an explicit plan, quantitative representation (defining quantities and a coordinate system, constraints, and initial conditions), ongoing review of the solution, and verification (checking units and assumptions). For both rubrics, each criterion is scored from 0-4 with 0 being Unacceptable, 1=Marginal, 3=Good, and 4=Excellent (a score of 2 is inferred to be somewhere between Marginal and Good).

A difficulty with many of these rubrics is the different number of score options for each category and lack of consistent language for each score (Blue, 1997; Foster, 2000; Harper, 2001; Huffman, 1997). Difficulties with Ogilvie's rubrics are their length (some cells of the rubrics contain as many as four sentences of description) and an inconsistency in language across categories. Also, the criteria of the Process rubric are not often observed unless students are explicitly prompted to write them down, such as an explicit plan, ongoing review of the solution, and verification of the answer. This observation was made by Harper (2001), who stated that an analysis of written solutions does not accurately describe behavior in three areas: analogies, checking answers, and making plans. Table 2 summarizes the score categories for several of the assessments described above.

Table 2: Comparison of Physics Problem-Solving Assessment Categories

<table>
<tr><th>Heller, Keith, & Anderson (1992)</th><th>Douglas Huffman (1997)</th><th>Jennifer Blue (1997)</th><th>Tom Foster (2000)</th><th>Tom Thaden-Koch (2005)</th><th>Sahana Murthy (2007)</th><th>Craig Ogilvie (2007) Mileposts</th></tr>
<tr><td>Evidence of Conceptual Understanding</td><td></td><td>General Approach</td><td rowspan="2">General Approach</td><td rowspan="2">Physics Approach</td><td>Physics Content</td><td>Citing Key Physical Principles</td></tr>
<tr><td>Usefulness of Description</td><td>Quality & Completeness of the Physics Representation</td><td rowspan="2">Specific Application of Physics</td><td>Relevant Representations</td><td>Diagrams</td></tr>
<tr><td>Match of Equations with Description</td><td>Match of Equations with the Representation</td><td>Specific Application of Physics</td><td>Symbolic Translation of Physics Approach</td><td>Modeling the Situation</td><td>Basic Equations</td></tr>
<tr><td>Reasonable plan</td><td rowspan="2">Organized Progression</td><td rowspan="2">Logical Progression</td><td rowspan="2">Logical Progression</td><td rowspan="2">Logical Progression</td><td rowspan="4">Problem-Solving Strategy & Reasonableness of Answer</td><td></td></tr>
<tr><td>Logical progression</td><td>Clear statement of answer</td></tr>
<tr><td rowspan="2">Appropriate Mathematics</td><td rowspan="2">Mathematical Execution</td><td rowspan="2">Appropriate Mathematics</td><td rowspan="2">Appropriate Mathematics</td><td rowspan="2">Appropriate Mathematics</td><td>Algebraic Manipulations</td></tr>
<tr><td>Numerical calculation</td></tr>
</table>

Summary of Literature Review

This chapter outlined some of the key definitions relevant for research on problem solving, highlighted a sample of early research on problem solving from philosophers and behavioral psychologists, summarized a general theory of human problem solving based on the information-processing approach to memory, described problem-solving research on experienced and inexperienced problem solvers in physics and mathematics, and identified attempts to quantitatively assess problem-solving in physics using scoring criteria and rubrics.

The definition of problem solving emphasizes that it is a "process of moving toward a goal when the path to that goal is uncertain" (Martinez, 1998 p. 605) and it depends on both the task and characteristics of the problem solver (Martinez, 1998; Newell & Simon, 1972). Problem structures can vary substantially from *well-structured* problems that have a clear goal with all information provided to *ill-structured* problems in which the goal may be vague goal and there may be several viable solutions (Jonassen, 1997; Pretz et al. 2003). These definitions highlight that there are differences in how problem solvers understand and approach a task, and differences in problem structures that can influence problem solving behavior. When assessing someone's skill at solving problems, it is important to have tasks of appropriate difficulty in order to obtain meaningful scores. Also, an assessment that claims to be general across problem types and topics should be tested on problems that vary in structure.

A brief review of problem solving theories reveal the rich history of this field. As early as the 4th century B.C., the Greek philosopher Aristotle wrote about the structure of human thought as being composed of ideas and links or "associations" between them

(Mayer, 1992), and set forth a formal system of logic built on chains of reasoning that is still used today in mathematical arguments (Kal, 1988; Lear, 1980). Much later, experimental psychological research in America (Hull, 1943; Skinner, 1938; Thorndike, 1898; 1911) and Germany (Koffka, 1935; Köhler, 1929; Wertheimer, 1945) suggested that repetition of a problem-solving task strengthens associations between a particular problem (stimulus) and an appropriate response. These concepts are relevant for studies of problem-solving because they emphasize the following points: knowledge is organized in memory as a network of linked ideas, links can vary in their strength, and past problem-solving experiences influence future problem solving behaviors.

Newell, Shaw, and Simon built on these theories with their information processing theory of memory and general theory of human problem solving (Newell, Shaw, & Simon, 1958; Newell & Simon, 1972). In their theory the brain is composed of a limited short term or "working" memory and a long-term memory. During problem solving, information is maintained in working memory while subject matter knowledge and strategies are accessed from long-term memory. If reading and representing the problem activates an existing memory framework (called a schema) for the problem, a solution is implemented (Gick, 1986; Marshall, 1995). If not, the solver must search through available information and apply general strategies in an attempt to progress toward the goal (Gick, 1986). Information processing theories are relevant for studies of problem solving because they emphasize the influence of knowledge organization on problem solving, particularly that having memory frameworks called schemas can facilitate successful problem solving. These theories also help to explain why some solvers become overwhelmed by complex tasks (working memory overload).

Research with experienced and inexperienced problem solvers in physics indicates differences in the way knowledge is organized in memory and the processes engaged in during problem solving (Chi et al., 1981, 1982; Larkin, 1979; Larkin & Reif, 1979; Simon & Simon, 1978; 1979). Experienced solvers have a hierarchical knowledge structure centered around fundamental physics principles, and that knowledge is stored with procedures for the appropriate application of those principles (Chi et al., 1981; Gick, 1986). Inexperienced solvers tend to focus on the objects in a problem (called its surface features) rather than the physics principles that apply (Chi et al., 1981). When solving a problem, experts will typically start with a low-detail qualitative overview of the problem before writing down mathematical relationships (Heller & Reif, 1984; Larkin, 1979; Larkin & Reif, 1979). Experts also have strategies for monitoring their progress and evaluating the answer (Bagno & Eylon, 1997; Chi, 2006; Reif & Heller, 1982).

These characteristics of experienced and inexperienced solvers are relevant because a useful assessment instrument for problem solving must be able to distinguish between these two groups, and more skilled individuals should score high on the assessment. The development of an assessment in physics must incorporate the findings that experienced solvers are skilled in their use of descriptions, principle-based approaches, and metacognitive strategies that facilitate the logical progression of their solution. The link between these characteristics and the problem-solving rubric developed in this dissertation is discussed in Chapter 4.

Assessments of problem-solving for written solutions to physics problems have been developed by several researchers, including Heller et al. (1992), Blue (1997), Huffman (1997), Foster (2000), Harper (2001), Murthy (2007), and Ogilvie (2007). Many of them are based on the Pólya's steps in the problem-solving process, including: *Understanding the Problem*, *Devising a Plan*, *Carrying Out the Plan*, and *Looking Back* (Pólya 1945; 1957). A difficulty of many of these tools is the different number of score options for each category, lack of consistent language for each score, length, or they include steps that are not generally observed outside of a particular instructional framework. In addition, most have not been tested for evidence of the reliability and validity of scores. These problem solving assessments indicate aspects of problem solving that researchers in physics consider important, and also form the basis for the categories assessed by the problem-solving rubric described in this dissertation. They also suggest difficulties in scoring formats and motivate the need for a general problem-solving assessment that is valid, reliable, and easy to use.

Chapters 3 and 4 further clarify what is meant by reliability and validity of assessment scores, how the categories and scores of the problem-solving rubric in this study relate to research on experienced and inexperienced problem solvers, and how this rubric is similar to or different from other measures of problem-solving.

CHAPTER 3: METHODOLOGY FRAMEWORK

Validity *Reliability* *Utility* *Overview Parts of Study*

Introduction

This dissertation focuses on the development of an assessment rubric to measure problem solving processes observed in written solutions to physics problems independent of the style in which the students are taught to express those solutions. Developing this type of assessment can be thought of broadly as similar to the process for developing an educational or psychological test and evaluating the validity, reliability, and utility of test score interpretations (Messick, 1989b).

This chapter provides definitions pertinent to validity theory, briefly reviews the history of validity theory up to the present, and describes the validity framework adopted for this thesis study. It also describes each source of validity, reliability, and utility evidence and outlines how specific data and research methods from this study fit within the validation framework.

Validity

Brief History of Validity

One of the most prominent resources for validity information is *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council for Measurement in Education [NCME], 1999). There have been at least five editions of this publication, in 1954/1955, 1966, 1974, 1985, and 1999 (Moss, 2007). Revisions to the 1999 edition of the *Standards* are currently underway (http://www.teststandards.org). Another common source is from chapters on validity in different editions of the text *Educational Measurement* (see for example Messick, 1989b). These sources give an indication of the ways in which validity theory has shifted from an early psychometric focus on "types" of validity, each specific to a particular testing aim, toward a more holistic view applicable to all forms of educational and psychological assessment.

This brief history will review some of the primary perspectives of validity as they developed. It begins with validity "types" in the early publications of the *Standards* and moves to Messick's six facets of construct validity, the validity perspective in the most recent *Standards* (AERA et al., 1999), and newer views including Kane's argument approach (Kane, 1992) and a change in terminology recently proposed by Lissitz and Samuelson (2007). Validity considerations for more complex tasks such as performance assessments will also be reviewed (Linn, Baker, & Dunbar, 1991; Messick, 1994; Moss, 1992).

Definitions

Historically from the early 1900s until the 1950s, validity was defined as the degree to which a test measures what it claims to measure or "is capable of achieving certain aims" (APA, 1954, p.213) and a test's validity was determined primarily by statistical measures (Shepard, 1993; Kane, 2001). An important distinction was made by some researchers that it is not the *test* or *scores* that are validated, but the inferences and decisions based on those scores for an assessment's intended purpose (Cronbach & Meehl, 1955; Messick, 1989b). This purpose of an assessment can vary substantially from tests used for college and program admissions, job selection, curricular evaluations, psychological traits, or content knowledge. A test that is valid for one use may not be valid for another use. This is why each use of a test must be validated independently (Shepard, 1993; Moss, 2007). This perspective is reflected in the 1999 edition of the *Standards*:

> Validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests…The process of validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations. It is the interpretations of test scores required by proposed uses that are evaluated, not the test itself. (p. 9)

Threats to validity include what are referred to as construct underrepresentation and construct-irrelevant variance (AERA et al., 1999; Messick, 1995b). Construct underrepresentation refers to a situation in which the assessment is too narrow and does not adequately measure the intended domain. Construct -irrelevant variance occurs when the test is measuring extra characteristics not important to the intended construct, such as the testing methods or situation. Examples of construct-irrelevant variance include issues that can unfairly influence test scores including anxiety, low motivation, illness, fatigue, limited English proficiency, time limits, or bias from familiarity of items. (Messick, 1989b). Test developers and users must pay attention to possible construct underrepresentation and construct-irrelevant variance arising from an inadequate description of the construct, test format, or testing conditions (AERA et al., 1999).

From "Types" of Validity to a Unified Validity Concept

In their 1954 *Technical recommendations for psychological tests and diagnostic techniques*, the American Psychological Association [APA] identified four "types" of validity, each appropriate for different testing aims: content, predictive, concurrent, and construct validity. Content validity was seen as appropriate for achievement and proficiency tests, and was evaluated by professional judgment of the sampling adequacy and representativeness of subject matter items on the test (APA, 1954). Predictive and concurrent validity were appropriate for predicting an individual's future performance or current status, respectively, on a measure external to the test, and were evaluated with correlational measures. Construct validity was appropriate for tests of psychological traits or qualities, and required both logical arguments and empirical evidence. The test recommendations in this publication were stated separately for each testing aim (APA, 1954). Later versions of the publication combined predictive and concurrent validity into one category called criterion-related validity (APA, 1966; Shepard, 1993).

It should be noted that prior to the 1954 APA publication, the concept of "construct" validity had not been formally introduced. Most validity studies were based on content evaluations or criterion measures. Two of the standards committee members, Cronbach and Meehl (1955) followed up the *Recommendations* with a paper highlighting construct validity as "preferable" to the other types of validity for some testing situations, and suggested that "many types of evidence are relevant to construct validity, including content validity, interitem correlations, intertest correlations, test "criterion" correlations..." (p. 300). This statement foreshadows the shift in perceiving validity more holistically as construct validity with evidence based on multiple sources, which wouldn't occur until decades later (Cronbach, 1971).

The 1974 *Standards for educational and psychological tests* explicitly recognized interrelationships among content, criterion-related, and construct validity, yet maintained the language of distinct validity "types" (APA, AERA, & NCME, 1974; Moss, 2007; Shepard, 1993). It wasn't until the 1985 *Standards for educational and psychological testing* that a "unified" concept of validity was presented, with construct validity at its center (AERA, APA, & NCME, 1985; Moss, 2007). Samuel Messick (1989b) explains this unification as,

> One or another of these forms of evidence, or combinations thereof, have in the past been accorded special status as a so-called "type of validity." But because all of these forms of evidence fundamentally bear on the valid interpretation and use of scores, it is not a type of validity but the relation between the evidence and the inferences drawn that should determine the validation focus. The varieties of evidence are not alternatives but rather supplements to one another. This is the main reason that validity is now recognized as a unitary concept. (p. 16)

Messick has played an instrumental role in articulating validity as a unified, non-prescriptive concept, and extended existing theories to consider value implications and social consequences of test use (Shepard, 1993; Messick, 1989a, 1989b, 1995a, 1995b).

He perceived validity as "…a unified though faceted concept" (Messick, 1989b, p.14) and outlined six such "facets" or categories of evidence that contribute to a validity argument. Since his categories are reflected in the most recent *Standards* (AERA, APA, NCME, 1999), it is worthwhile to take a closer look at Messick's validity theory.

Messick's Six Facets of Construct Validity

In his chapter on Validity in the third edition of *Educational Measurement*, Messick (1989b) comprehensively describes a unified theory of validity that considers the proposed use of a test, and outlines six sources of validity evidence including social consequences. He defines validity in the following way:

> Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment….broadly speaking, then, validity is an inductive summary of both the existing evidence for and the potential consequences of score interpretation and use. (p.13)

Messick highlights the importance of determining the appropriateness, meaningfulness, and usefulness of scores (Messick, 1989a). Similar to Cronbach (1988), he also highlights the importance of both *convergent* evidence (which supports the proposed test use and interpretation) and *discriminant* evidence that considers alternate, disconfirming interpretations (Messick, 1989b). His six aspects of construct validity include content, substantive, structural, generalizability, external, and consequential (Messick, 1989b, 1995b).

The *content* aspect addresses selecting tasks that are representative and relevant to the domain theory of the assessment, such as the knowledge, skills, abilities, processes, motives, attitudes, etc. being elicited (Messick, 1995a, 1995b). The *substantive* aspect considers the consistency of an assessment's intended processes with the processes actually engaged in by respondents during tasks. The *structural* aspect compares interrelationships among parts of a test (its internal structure) with expected relationships derived from the structure of the test domain. *Generalizability* refers to applicability of the assessment across different tasks, populations, situations, or times. The *external* aspect considers the relationships of an assessment's scores with other, external measures. Finally, the *consequential* aspect refers to weighing potential positive and negative consequences of the proposed test use.

Standards for Educational and Psychological Testing

The most recent edition of the *Standards* (AERA et al., 1999) incorporates some of Messick's ideas in its terminology and definitions. It cites five primary sources of evidence for construct validity: evidence based on test content, response processes, internal structure, relations to other variables, and consequences of testing. The primary differences with Messick's theory are that *generalizability* is not explicitly listed as a category, the term "response processes" is used in place of *substantive*, "internal structure" is used in place of *structural*, and "relations with other variables" replaces the term *external*. Aside from these label differences, the general meanings of the categories are relatively consistent.

Performance Assessments

Some researchers have questioned the appropriateness of these validation criteria for the assessment of more complex tasks, called performance assessments (Moss, 1992; Messick, 1994). Examples of performance or "alternative" assessment tasks could include solving open-ended problems, writing tasks, portfolios, or performances in artistic or athletic-based subjects. Linn, Baker, & Dunbar (1991) propose a set of eight criteria especially relevant for performance assessments: intended and unintended consequences, fairness of the assessment, the degree to which specific tasks transfer or "generalize", cognitive complexity of the processes required of students, meaningfulness of tasks for students and teachers, content quality, content coverage, and the efficiency of data collection and scoring. It is unclear whether these criteria were intended to be in place of or in addition to other existing validity categories.

Messick (1994) explicitly addresses validity issues related to performance assessments, but does not propose modifying his six aspects of construct validity. He draws a distinction between "task-driven" and "construct-driven" assessments, which differ in their testing focus. He warns that having scoring criteria that are too specific to a task are in danger of limiting generalizability, whereas construct-centered scoring criteria are in danger of being too general and not meaningful. He suggests you "…aim for scoring rubrics that are neither specific to the task nor generic to the construct but are in some middle ground reflective of the classes of tasks that the construct empirically generalizes or transfers to" (Messick, 1994, p. 17). He advocates, when possible, adopting a construct-centered approach over a task-centered approach.

Messick suggests making clear whether the focus of the assessment is a performance or product, or an underlying construct such as knowledge, abilities, skills, or processes (Messick, 1994). He notes that assessing a *performance* is appropriate when a particular procedure has been taught, and the aim of the assessment is to measure deviations from an "ideal" or expected procedure. The assessment of a *product* is appropriate when such a procedure does not exist or has not been explicitly taught, and diverse responses are expected. He also explicitly addresses the trade-offs of performance assessments and more structured tests, recognizing that open-ended tasks require more time intensive scoring procedures, suggesting a mixture of both, "…assessment batteries ought to represent a mix of efficient structured exercises and less structured open-ended tasks" (Messick, 1994, p.22).

Ongoing Validity Discussions

Another important contribution to the ongoing discussion of validity theory among the education measurement community is an argument-based approach (Cronbach, 1988; Kane, 1992, 2001). Kane describes the validation process in the language of developing an argument to persuade an audience, including developing multiple independent sources of evidence (often referred to as "triangulation" of data) and considering possible counterarguments. As Kane (1992) explains:

> A test-score interpretation always involves an *interpretive argument*, with the test scores as a premise and the statements and decisions involved in the interpretation as conclusions….The best that can be done is to show that the interpretive argument is highly plausible, given all evidence. To *validate a test-score interpretation* is to support the plausibility of the corresponding interpretive argument with appropriate evidence. (p.527)

Kane goes on to outline evaluation criteria for a validity argument, including the argument's clarity, coherence, and the plausibility of assumptions (Kane, 1992). He suggests that this evaluation will aid in determining the weakest points in the argument that deserve the most attention.

Kane's approach is in response to dissatisfaction with the current unified construct validity theory, as echoed by other researchers (Lissitz & Samuelson, 2007; Moss, 1992, 2007; Shepard, 1993). They argue that the standards are too complex and lack guidance of how to implement them in practice. There are also very few documented examples of complete validity studies (Shepard, 1993). The idea that validity is an "evolving property" and validation a "continuing process" (Messick, 1989b, p.13) has resulted in some confusion about how much evidence is sufficient for a validity argument, often resulting in inadequate consideration of the issue.

Lissitz and Samuelson (2007) dislike the construct-validity emphasis of recent theories and propose yet another shift in terminology and focus. They suggest that the most central concern is an assessment's internal characteristics, including *content*, *reliability*, and *latent processes*. They also suggest a shift away from correlation-based decisions of criterion-related validity toward other external test factors called *utility, impact*, and *nomological network* (a historical term referring to theoretical psychological traits). This system of internal and external factors is also organized by perspective, noting that content, reliability, utility, and impact are primarily practical concerns, whereas latent process and nomological network are theoretical in nature. Different from other publications, Lissitz and Samuelson list sample questions to guide a validity investigation and potential sources of evidence for each of the six factors.

Validity Framework of this Dissertation

The validity framework used for this dissertation combines five sources of evidence from the *Standards* (AERA et al., 1999) with Messick's idea of Generalizability (Messick, 1995b), and adopts language from each approach. The concepts of validity and reliability are addressed separately, as consistent with the 1999 *Standards* publication (AERA et al., 1999). Reliability in this context refers to the consistency or agreement of scores and score interpretations on the assessment. Utility is also treated separately from validity with the explicit consideration of potential uses of score interpretations for researchers and instructors.

Table 3 outlines five primary perspectives of validity theory reviewed in this brief overview, with the last column representing the validity framework and language adopted for this dissertation. The similarity of categories is designated across the table horizontally, and reflects this author's personal interpretation from available category descriptions. For an alternate and broader meta-analysis of validity frameworks, see Figure 1 in Moss (1992).

The next section explores each source of validity evidence in greater detail and lists the specific data source(s) from this study that correspond to each. The five source categories include validity evidence based on content, response processes, internal and external structure, generalizability, and the consequences of testing. Following the validity source descriptions, reliability and utility will be similarly addressed. The procedures for determining each source in this study is given in section *Methodology Framework*. For a holistic view of the validity, reliability, and utility sources of evidence as they correspond to each stage in the study, refer to Table 11.

Table 3: Multiple Perspectives of Validity

Traditional Validity "Types" (APA, 1954; 1966; APA, AERA, & NCME, 1974)	**Messick's six facets of construct validity** (1989a, 1989b, 1995a, 1995b)	**Linn, Baker, & Dunbar (1991)** Performance Assssessment Criteria	**Standards' sources of construct validity evidence** (AERA, APA, & NCME, 1999)	**Lissitz & Samuelson (2007)** internal and external system of test evaluation	**This study** (sources of validity evidence)
Content	Content relevance and representativeness	Content Quality; Content Coverage; Meaningfulness	Test content	Content	Content
Construct	Substantive	Cognitive complexity	Response processes	Latent processes	Response processes
Criterion-related (Concurrent & Predictive)	Structural		Internal structure	Reliability	Internal & external structure
	External	Cost & Efficiency	Relations to other variables	Nomological network; Utility	
	Generalizability	Transfer & Generalizability	[Reliability & generalizability theory separate from validity]	(Reliability)	Generalizability
	Consequential (social consequences & value implications)	Consequences; Fairness	Consequences of testing	Impact	Consequences of testing

Sources of Validity Evidence

Validity Evidence Based on Content

Content refers to the wording and formatting of items on an assessment, in addition to the documented procedures for scoring. (Messick, 1995b). In this study content is interpreted to mean the process categories being assessed by the rubric and the documentation materials for potential users. Evidence for the relevance and representativeness of content comes from expert judgment, and theoretical descriptions of a domain in the research literature (AERA, et al., 1999).

Table 4: Valildity Evidence Based on Content

Description	Sources of Evidence
a) The extent to which the rubric process categories are consistent with descriptions of problem-solving processes in the research literature.	Descriptions of research literature basis for rubric development, including which problem-solving processes are and are not assessed by rubric
b) The extent to which potential users of the rubric perceive its content (format, categories, and scores) to represent a complete and relevant measure of problem-solving processes in physics.	Comments from expert raters (experienced graduate students / teaching assistants) regarding the extent to which the format, categories, and scores of the rubric represent a complete and relevant measure of problem-solving processes in physics.

Validity Evidence Based on Response Processes

An important validity consideration is the extent to which the assessment represents processes actually engaged in by the person(s) being assessed (AERA et al., 1999). It is also important to consider whether the interpretations of scores by judges or raters are consisted with the developer's intentions. In this study, student response processes are explored using both written work (written physics tests) and verbal interviews. The responses of raters are compared to problem solving grades from instructors and from their feedback while using the rubric to determine the degree of consistency with the rubric developers' intentions.

Table 5: Validity Evidence Based on Response Processes

Description	Sources of Evidence
a) The extent to which the rubric represents problem solving processes actually engaged in by the person(s) being assessed.	i) Problem solving interviews with students to compare written processes assessed by the rubric (5 categories) to verbal evidence of problem-solving processes ii) Scoring students' solutions to exam problems with the rubric for evidence of the written problem-solving processes assessed by the rubric
b) The extent to which the interpretations of the rubric by raters are consistent with the developer's intentions.	i) Scores and comments from graduate students using the rubric to assess examples of written problem solutions

Validity Evidence Based on Internal and External Structure

Internal structure refers to the extent to which relationships among parts of the instrument agree with expectations (AERA et al., 1999). External structure refers to the extent to which scores are related to other measures of the same construct or other hypothesized relationships. In this study, the degree of independence of the process categories are determined from statistical measures. For example, past research (Foster, 2000) indicated that the approach and application were correlated. The external structure of the rubric is evaluated by comparing rubric scores for written physics tests to scores assigned by a grader. It is also evaluated from comparisons of the verbal responses from problem-solving interviews to solutions written on paper during the interview.

Table 6: Validity Evidence Based on Internal and External Structure

Description	Sources of Evidence
a) The extent to which relationships among parts of the rubric (its *internal* structure) agree with expectations.	Statistical measures (e.g. correlations) of the relationships between category scores to assess the level of category independence.
b) The extent to which scores on the rubric agree with other, *external* measures of students' problem solving processes	i) Statistical measures (e.g. correlations) of the relationships between rubric scores and scores assigned by a course grader to exam problem solutions ii) Problem solving interviews with students to measure how students' written problem solutions are similar to or different from their verbal responses during a problem-solving interview.

Validity Evidence for Generalizability

Although not explicitly included in all descriptions of validity evidence, Messick (1995b) highlights the importance of an assessment being general across different populations and contexts. In this study, the rubric is tested on a variety of physics problem solutions that span different topics in standard introductory university physics courses from both mid-term tests and final exams. It is also tested on different types of problems, including those that are similar to traditional textbook problems and those that are context-rich (Heller et al., 1992).

Table 7: Validity Evidence for Generlizability

Description	Sources of Evidence
The extent to which the rubric is applicable to multiple populations and contexts, including different student	i) Rubric scores on archived final exam problem solutions from different introductory physics courses and

populations, physics topics, and problem features.	different problem types ii) Rubric scores of instructor homework solutions iii) Rubric scores of students' written solutions in a semester-long introductory course, spanning multiple physics topics in mechanics

Validity Evidence Based on Consequences of Testing

Descriptions of this source of validity highlight the importance of considering both intended and unintended consequences of score interpretations (AERA et al., 1999). In this study, the purposes of the rubric will be clearly outlined in the documentation materials and training. For example, in validity studies the rubric scores might only be meaningful to assess the performance of a class, and might not be meaningful or valid for diagnosing an individual student. A full study of the consequences of using this rubric, once developed, will be the subject of further work.

Table 8: Validity Evidence Based on Consequences of Testing

Description	Sources of Evidence
The extent to which the rubric and documentation materials consider the intended and unintended consequences of score interpretations.	i) Comments from graduate student teaching assistants regarding the extent to which documentation in training materials adequately outlines the purpose of the rubric ii) Use rubric scores from class-level data to describe intended score interpretation; i.e., not for diagnosing a class pedagogy

Reliability & Utility

In addition to validity evidence, important assessment considerations include the concepts of reliability and utility. As stated previously, reliability in this context is defined to be the agreement of scores and score interpretations from a single rater over time and from multiple raters using the rubric. Utility refers to both the proposed uses of the assessment put forth by the rubric developers, and the perceived usefulness of score interpretations as judged by those using the instrument.

Sources of Reliability Evidence

In this study, reliability is measured from a study in which two expert raters score written solutions to final exams, and from two studies with graduate students who undergo a brief written training in use of the rubric. These graduate students are experienced in grading the work of introductory physics students. Their responses are compared to each other and to the two expert raters. A quantitative measure of reliability is obtained from percentage of perfect score agreement, agreement within one score, and Cohen's Weighted Kappa (Cohen, 1968) which accounts for the degree of difference in scores.

Table 9: Evidence for Reliability

Description	Sources of Evidence
a) The agreement of scores from multiple raters or judges (*inter-rater* reliability)	i) Statistical measures of score agreement from two expert raters using the rubric, eight experienced graduate students, and 19 teaching assistants with two semesters of grading experience. ii) Agreement of interview protocol statement categorizations

Sources of Utility Evidence

Evidence for the usefulness of an assessment includes its acceptance by instructors, the extent to which it can distinguish between experienced and inexperienced problem solvers, and the extent to which it can distinguish between different classroom practices. It is important that researchers, curriculum developers, and instructors are interested in the information obtained from administering an assessment. In this study, interpretations of scores from analyses of written work and problem-solving interviews are used to propose uses of the assessment from the perspectives of researchers, curriculum developers, and physics instructors.

Table 10: Evidence for Utility

Description	Source of Evidence
a) The extent to which the rubric provides meaningful and useful information for instructors and researchers.	i) An analysis of rubric scores from students' written solutions to exam problems to provide a model for score interpretations useful for instructors and researchers. ii) A comparison of verbal protocols and written work from problem-solving interviews to delineate the limitations of measuring problem solving processes from written work
b) The extent to which the rubric can distinguish experienced and inexperienced problem-solvers.	Use the rubric to score student solutions and two types of instructor solutions to determine if it can distinguish between these two groups
c) The extent to which rubric scores respond to different instructional practices	*subject of further study*

Summary of the Methodology Framework

The methodology for this dissertation as outlined in previous sections is based on a framework of validity, reliability, and utility. A brief history of validity theories has indicated a shift from “types” of validity appropriate for a particular testing aim to a more holistic view of construct validity with multiple lines of evidence (see Table 3). This dissertation considers five aspects of validity based on currently accepted theories (AERA et al., 1999; Messick, 1989b). These five categories of evidence sources include content, response processes, internal and external structure, generalizability, and consequences. Each of these aspects was examined in greater detail, with a description of the data source(s) and analysis provided by this research study.

To summarize the information another way, a timeline for parts of the study is listed below. The titles for each part of the study are modeled after a similar study by Lavoie (2003): Drafting the assessment, Preliminary testing, Pilot testing, and Field testing. Another chart that explicitly links the parts of the study to the validation framework is provided in Table 11.

Overview of Parts of the Study

1. Drafting the Rubric

    Literature Review

    Drafting the rubric and testing it on exam solutions

2. Preliminary testing

    Scoring final exam solutions with two expert raters

    Scoring instructor homework solutions

3. Pilot testing

    First test with training raters

Second test with training raters

4. Field Testing

Scoring exam solutions from one semester of mechanics

Student problem-solving interviews

Parts of the Study Linked to Framework

Table 11 connects each stage of the study with a specific source of evidence (data and analysis) and its purpose within the validity-reliability-utility framework.

Table 11: Sources of Evidence Linked to Methodology Framework

Stage	Source of Evidence	Validity	Reliability	Utility
Drafting Rubric	Literature review	Content relevance and representativeness		Purpose and content of this rubric in research literature context
	Scoring exam solutions	Response processes of students (written)		Degree to which rubric can be used to distinguish among student written solutions
Preliminary testing	Scoring students' final exam solutions	Generalizability (across course, physics topic, level of expertise, and solution detail); Response processes	Agreement of scores between two raters	Degree to which the rubric can be used consistently
	Scoring instructor homework solutions			Degree to which the rubric distinguishes students and experts
Pilot testing	First study with training raters	Content relevance and representativeness; Response processes of raters; Consequences	Agreement of scores and score interpretations among multiple raters; Agreement with researchers	Perceived utility of rubric scores (comments from raters)
	Second study with training raters			
Field testing	Scoring written tests over one semester of an introductory physics course	Generalizability; Internal and External structure; Response processes		Proposed usefulness of the rubric scores for researchers and instructors
	Student problem-solving interviews	External structure; Response processes (written and verbal)	Agreement of interview codings between two researchers	Limitations of written work as evidence of problem-solving processes

CHAPTER 4: DATA COLLECTION & ANALYSIS

Introduction

This section outlines the development of the rubric's categories and scores, in addition to four major stages in testing validity, reliability, and utility of the rubric categories and scores. After developing a draft instrument based on previous research, preliminary studies with two experienced raters using feedback were used to determine initial reliability measures and modify categories. Utility was also tested by comparing the ratings of instructor and student solutions. Next, two studies with graduate students involving a written training exercise was used to further measure the reliability and validity of the rubric content when accompanied by minimal training materials. An analysis of students' written solutions to physics tests from a semester of introductory physics (mechanics) was used to obtain evidence for response processes, generalizability, internal and external structure, and to propose uses of the rubric. Analysis of student interviews was used to obtain further evidence of response processes and structural measures.

Drafting the Rubric

Introduction

The Literature Review in Chapter 2 outlines important definitions related to problem solving, gives a brief history of research on problem solving, highlights results from expert-novice research studies in mathematics and physics, and surveys different problem solving assessments. The research on process differences for experienced and inexperienced problem solvers in physics (Chi et al., 1981; Heller & Reif, 1984; Larkin, 1979; Larkin et al., 1980a; Schoenfeld, 1985; Simon & Simon, 1978; 1979) is particularly relevant for identifying important characteristics of a problem solution. For example, these studies observed that experienced problem-solvers usually represent important features from the problem statement and consider possible physics concepts and principles applicable to the problem before writing down equations, or what Larkin called a "qualitative analysis" of the problem. Pólya referred to these stages as "Understand the Problem" and "Plan the Solution" (Pólya, 1945; 1957). As explained in the next section, these processes are reflected in the *Useful Description* and *Physics Approach* categories of the rubric.

Successful solvers also have knowledge for how to apply physics concepts and principles to the specific conditions in a problem (Chi et al., 1981; 1982; Eylon & Rief, 1984) and can "Execute the Plan" by carrying out mathematical procedures (Pólya, 1945; 1957). Competent problem solvers have metacognitive skills for monitoring their progress while solving and evaluating their answer (Chi, 2006; Reif & Heller, 1982) which can contribute to the logical progression of their solution. Problem-solving coding schemes used by researchers at the University of Minnesota (including this one) are based on these characteristics of qualitative descriptions, approaches based on fundamental principles, procedures for the appropriate application of principles, skilled use of mathematics, and strategies for monitoring progress and evaluating results.

The purpose of the following sections on problem-solving process categories and scores is to place this rubric within the context of existing research literature on problem solving, and provide evidence to the extent to which the rubric content is relevant to and representative of existing conceptions of problem solving in physics.

Problem-Solving Process Categories

The process categories for the assessment rubric were based on the research literature in cognitive science, mathematics, and physics (Chi et al., 1981; Gick, 1986; Larkin et al., 1980a, 1980b; Larkin & Reif, 1979; Pólya 1945, 1957; Reif & Heller, 1982; Schoenfeld, 1985; Simon & Simon, 1978, 1979). They were developed within the constraints of being easy to use and interpret for physics instructors, independent of pedagogy, generalizable to multiple problem types and topics, and focused on written work. Many other related rubrics that have been developed to assess student problem solving in physics and other disciplines are available from a general search of the Web. Such rubrics are often developed for classroom use to support a specific pedagogy and typically have not been extensively tested for reliability or validity.

The rubric described in this dissertation is based on research on student problem solving at University of Minnesota over many years (Blue, 1997; Foster, 2000; Heller et al., 1992). Although there are many similarities in the problem solving processes assessed by instruments in those studies, the current study differs by attempting to simplify the rubric and adding more extensive tests of validity, reliability, and utility. It explicitly considers applicability to a broad range of problem types and topics in physics and the ease of use for both research and instruction.

Table 12 lists of process categories for past coding rubrics at the University of Minnesota. Although many of the category names are similar, their interpretations evolved with time (Refer to the section in Chapter 2 on Assessments). Drafts of the rubric used in this dissertation are described in Appendix 1.

Table 12: Development of the Rubric by the University of Minnesota Group as Used in Research

Pat Heller, Ronald Keith, & Scott Anderson (1992)	Jennifer Blue (1997)	Tom Foster (2000)	Tom Thaden-Koch (2005)	Jennifer Docktor & Ken Heller (2009)
1. Evidence of conceptual understanding 2. Usefulness of description 3. Match of equations with description 4. Reasonable plan 5. Logical progression 6. Appropriate Mathematics	1. General Approach 2. Specific Application of Physics 3. Logical Progression 4. Appropriate Mathematics	1. General Approach 2. Specific Application of Physics 3. Logical Progression 4. Appropriate Mathematics	1. Physics Approach 2. Symbolic Translation of Physics Approach 3. Appropriate Mathematics 4. Logical Progression	1. Useful Description 2. Physics Approach 3. Specific Application of Physics 4. Appropriate Mathematics 5. Logical Progression

To make the rubric easy to use, it was constructed with as few dimensions as possible to still span most of the space that distinguishes novice and expert problem solving. The Minnesota rubric considers five problem-solving processes: organizing problem information into a *Useful Description*, selecting appropriate physics principles (*Physics Approach*), applying the physics principles to the specific conditions in the problem (*Specific Application of Physics*), using *Mathematical Procedures* appropriately, and the overall communication of an organized reasoning pattern (*Logical Progression*).

Useful Description

Useful Description assesses a solver's process of organizing information from the problem statement into appropriate and useful representations that summarize essential information symbolically, visually, and/or in writing. It is similar to Pólya's (1945) stage of understanding the problem or Hayes' (1989) stage of representing the problem.

A useful problem description could include specifying known and unknown information, assigning appropriate symbols for quantities, stating a goal or target quantity, a sketch or picture of the physical situation, stating qualitative expectations, drawing an abstract physics diagram, drawing a graph, defining coordinate axes, and/or choosing a system. Unlike other models of problem solving (Heller & Reif, 1984; Heller & Heller, 2000; Van Heuvelen, 1991b), this combines both a basic description and a physics-specific description into a single category. The term "description" was chosen to be consistent with other uses of the term (Heller et al., 1992; Reif et al., 1976) and avoid the multiple interpretations of the term "representation" (Hayes, 1989; Larkin et al., 1980a, 1980b). The useful description category differs from other instruments (Foster, 2000) by being assessed separately from the general physics approach.

Physics Approach

The *Physics Approach* assesses a solver's selection of appropriate physics concepts and principles to use in solving the problem. Here the term "concept" is used to mean a general physics idea, such as the general concept of vector or specific concepts such as momentum and velocity. The term "principle" is used to mean a fundamental physics rule used to describe objects and their interactions, such as conservation of energy or Newton's second law. The interpretation of the term *approach* is similar to that used by experts in the Chi et al. (1981) study. Physicists in the Chi et al. (1981) study typically responded to instructions to describe their "basic approach" by stating the major physics laws or principles they would use to solve each problem.

In addition to assessing the selection of a principle, this category also includes its basic understanding, such as the independent treatment of perpendicular components of vectors. This is similar to the evidence of conceptual understanding category outlined by Heller et al. (1992) and the general approach category used by Blue (1997) and Foster (2000).

The *Physics Approach* category reflects the expert-like process of selecting relevant physics principles before applying them to the specific context of the problem (Chi et al., 1981; Larkin et al., 1980b). Although several descriptions of problem-solving emphasize a stage of planning the solution (Hayes, 1989; Heller et al., 1992; Pólya, 1945), selecting important relations is a necessary first step in planning the solution (Leonard, Dufresne, & Mestre, 1996; Reif et al., 1976). In addition, the details of planning are difficult to assess because students often do not write down the steps of their solution plan unless explicitly instructed to do so. The planning process is implicitly addressed by this rubric in its other categories.

Specific Application of Physics

Specific Application of Physics assesses the solver's process of applying physics concepts and principles to the specific conditions in the problem. Specific application often involves connecting the objects and quantities in the problem to the appropriate terms in specific physics relationships. It can include a statement of definitions, relationships between quantities, initial conditions, and consideration of assumptions or constraints in the problem.

This category separates the identification of appropriate principles and concepts in the Physics Approach from the actual application of those principles to the specific conditions in the problem. This is consistent with other descriptions of problem solving strategies (Leonard et al., 1996) and other assessments of problem solving (Blue, 1997; Foster, 2000). Writing down specific physics relationships, typically in the form of equations, can be seen as an aspect of planning the solution (Heller et al., 1992; Reif et al., 1976). This category is similar to the problem-solving model by Larkin et al. (1980b) that designates "connecting symbols in an equation with information in the problem" as a process that follows "selecting relevant physics principles" and "generating the corresponding equation" (p. 323).

Mathematical Procedures

Mathematical Procedures assesses the solver's process of executing the solution with respect to using appropriate mathematical procedures and following mathematical rules to obtain target quantities. Examples of these procedures include: isolate and reduce strategies from algebra, substitution, use of the quadratic formula, matrix operations, or "guess and check" from differential equations. The term mathematical "rules" refers to processes from mathematics, such as the Chain Rule in calculus or appropriate use of parentheses, square roots, logarithms, and trigonometric identities.

This category corresponds to carrying out the plan (Hayes, 1989; Pólya, 1945) or the plan implementation process (Reif et al., 1976). It also corresponds to Van Heuvelen's (1991a) "math representation" (p. 901) and Larkin et al.'s (1981b) "solving equations" function (p. 323). It is consistent with other assessments of appropriate mathematics (Blue, 1997; Foster, 2000; Heller et al., 1992) but differs in that it doesn't require students to solve equations symbolically to receive the highest score.

Logical Progression

Logical Progression assesses the solver's processes of communicating reasoning, staying focused toward a goal, and evaluating the solution for consistency. The category checks whether the overall problem solution is clear, focused, and organized logically. The term "logical" means that the solution is coherent (the solution order and solver's reasoning can be understood from what is written), internally consistent (parts do not contradict), and externally consistent (results agree with qualitative physics expectations). It does not imply a linear or continuous process in the solution.

This category agrees with the problem-solving assessment by Reif and Heller (1982) that includes clear interpretation or specification of parameters, completeness of the answer, internal logical consistency of the argument, external consistency of relationships and the magnitude of values, and optimality or the simplicity of the solution. It also emphasizes the importance of "the ability to provide coherent explanations" in science and engineering careers (Leonard et al., 1996, p. 1502). The term logical progression is taken from earlier assessments of problem solving (Blue, 1997; Foster, 2000; Heller et al., 1992) but it differs from those measures in that it doesn't score the student's process as working forwards or working backwards.

Several models of problem solving emphasize the final stage as looking back (Pólya, 1945) or evaluating the solution to check that it makes sense (Reif et al., 1976; Van Heuvelen, 1991b, Heller et al., 1992; Heller & Heller, 2000). The logical progression does not require an explicit evaluation of the solution because students and experts often do not write down this step unless explicitly instructed to do so, and the rubric is intended to be independent of strategy-modeling instructional techniques. However, steps such as planning and evaluation or checking the result help avoid errors in consistency and coherence, which are scored as part of the logical progression.

Processes Excluded From the Rubric

To make the rubric as independent of specific pedagogy and as easy to use as possible, the metacognitive processes of planning and evaluating the answer are not explicitly assessed by the rubric. Although they are excluded as specific criteria from this rubric, planning and evaluation are implicitly assessed by the several other categories because these processes affect the overall coherence and consistency of the solution. Other aspects of problem-solving not assessed by the rubric include affective qualities such as motivation, interest, and beliefs about physics. These qualities are not usually evident from written work.

Score Range and Descriptions

The current version of the rubric is given on the following page. Scores for each category on the rubric range from 0 (worst) to 5(best) with additional "not applicable" categories for the problem and for the specific solver, NA(Problem) and NA(Solver). The NA(Problem) score means that a particular category was not measured by the problem usually because those decisions were not required for that problem. For example, if an explicit description was provided in the problem statement or was not really necessary to solve the problem, the Useful Description would be scored as NA(Problem). The NA(Solver) score means that based on the overall solution, it was judged that this set of decisions might not be necessary for the solver to write down. This often occurs for experts and for students who were generally successful in solving the problem without writing down all of their internal processes, such as a description or explicitly stating a physics approach. These "not applicable" scores are included because the rubric needs to recognize the possibility that students are beginning to develop some of the automated processes engaged in by experts (Heller & Reif, 1984).

The final version of the rubric used in this dissertation is provided in the following figure (Rubric version 4) and in Appendix 1. To promote ease of use, the language of the score descriptions for each category is consistent. A score of 0 means that there is no evidence of the category and it was necessary for the solver, 1 means the category evidence was entirely inappropriate, 2 means mostly inappropriate or missing, 3 means parts are inappropriate or missing, 4 designates minor omissions or errors, and 5 is complete and appropriate. The numerical score range of 0 to 5 developed from several tests of raters using the rubric to score solutions (see Appendix 1). Several raters found that a narrower score range (0 to 3 or 0 to 4) did not provide sufficient delineation of abilities, or that some solutions fell "between" scores. The score range was modified accordingly in response to these comments and stabilized at the range 0 to 5.

Previous research indicates that when scoring written solutions to physics problems, it is important to consider only what is written and avoid the tendency to assume missing or unclear thought processes are correct (Henderson, Yerushalmi, Kuo, Heller, & Heller, 2004). Similarly, it is important not to overly emphasize the amount of detail in student explanations. Training materials provide examples of scored solutions (Appendix 5), to help raters avoid this tendency to project correct reasoning onto student work.

Drafts of the Rubric

Appendix 1 includes four versions of the rubric representing key stages of the instrument's development throughout this dissertation. Version 1 built on problem solving assessments by Heller, Keith, and Anderson (1992), Jennifer Blue (1997), Tom Foster (2000), and Tom-Thaden Koch (personal communication in 2005) with substantial changes to the formatting of the scores and criteria. As seen in Versions 1-4 in Appendix 1, all drafts described in this dissertation are formatted as a table or grid that lists problem-solving process categories along the left column of the table, and scores along the top row. Scoring criteria are described in each cell of the table. This format necessitates the same score range for each category, which is typical of rubrics used for education purposes (Arter & McTighe, 2001; Mertler, 2001; Montgomery, 2002). Version 1 also differs from most past Minnesota instruments by including a NA or Not Applicable score and separating the description process from selecting physics concepts or principles.

The rubric Version 2 differs from the first version in that the Not Applicable (NA) score was split into two scores: NA(Problem) and NA(Student). The language was also changed to be more consistent across scores for some categories. For example, the scores distinguished between "one part missing and/or incorrect" (score 3) and "more than one part missing and//or incorrect" (score 2). This is the version of the rubric that was used in the Preliminary study with two raters.

The third version of the rubric (Version 3 in Appendix 1) was formatted to fit vertically on one page with NA score descriptions summarized in one line at the bottom of the table. The language was also changed for some scores, such as distinguishing an important part or key feature that is missing or inappropriate (score of 2) rather than the 'one' or 'more than one' counting language used in Version 2. In this version the score of zero represented either all missing or all incorrect, whereas in past versions zero represented all missing. Also, category descriptions were provided on the second page of the instrument. This is the version that was used with the first study with training raters.

The fourth version of the rubric represents changes that were made after the first study with training raters, in preparation for the second study with training raters. Notably, the score range was changed from 0-4 to be 0-5, where the zero score represents no evidence for the category (or all missing) and a score of one represents all inappropriate / incorrect. The NA(Solver) and NA(Problem) scores were returned to their former column positions to be more prominent than in Version 3, and it was formatted to one landscape page. The language was made consistent for each score across categories, particularly 4: minor omissions or errors, 3: parts missing and/or contain errors, 2: most missing and/or contain errors. This is the rubric version that was used for the second study with training raters, scoring written solutions to exams, and scoring written solutions to the interview problems.

Figure 1: Problem-Solving Assessment Rubric (Version 4)

	5	4	3	2	1	0	NA(Problem)	NA(Solver)
USEFUL DESCRIPTION	**The description is useful, appropriate, and complete.**	**The description is useful but contains minor omissions or errors.**	**Parts of the description are not useful, missing, and/or contain errors.**	**Most of the description is not useful, missing, and/or contains errors.**	**The entire description is not useful and/or contains errors.**	**The solution does not include a description and it is necessary for this problem /solver.**	**A description is not necessary for this problem. (i.e., it is given in the problem statement)**	**A description is not necessary for this solver.**
PHYSICS APPROACH	**The physics approach is appropriate and complete.**	**The physics approach contains minor omissions or errors.**	**Some concepts and principles of the physics approach are missing and/or inappropriate.**	**Most of the physics approach is missing and/or inappropriate.**	**All of the chosen concepts and principles are inappropriate.**	**The solution does not indicate an approach, and it is necessary for this problem/ solver.**	**An explicit physics approach is not necessary for this problem. (i.e., it is given in the problem)**	**An explicit physics approach is not necessary for this solver.**
SPECIFIC APPLICATION OF PHYSICS	**The specific application of physics is appropriate and complete.**	**The specific application of physics contains minor omissions or errors.**	**Parts of the specific application of physics are missing and/or contain errors.**	**Most of the specific application of physics is missing and/or contains errors.**	**The entire specific application is inappropriate and/or contains errors.**	**The solution does not indicate an application of physics and it is necessary.**	**Specific application of physics is not necessary for this problem.**	**Specific application of physics is not necessary for this solver.**
MATHE-MATICAL PROCEDURES	**The mathematical procedures are appropriate and complete.**	**Appropriate mathematical procedures are used with minor omissions or errors.**	**Parts of the mathematical procedures are missing and/or contain errors.**	**Most of the mathematical procedures are missing and/or contain errors.**	**All mathematical procedures are inappropriate and/or contain errors.**	**There is no evidence of mathematical procedures, and they are necessary.**	**Mathematical procedures are not necessary for this problem or are very simple.**	**Mathematical procedures are not necessary for this solver.**
LOGICAL PROGRESSION	**The entire problem solution is clear, focused, and logically connected.**	**The solution is clear and focused with minor inconsistencies**	**Parts of the solution are unclear, unfocused, and/or inconsistent.**	**Most of the solution parts are unclear, unfocused, and/or inconsistent.**	**The entire solution is unclear, unfocused, and/or inconsistent.**	**There is no evidence of logical progression, and it is necessary.**	**Logical progression is not necessary for this problem. (i.e., one-step)**	**Logical progression is not necessary for this solver.**

Category Descriptions:

Useful Description assesses a solver's skill at organizing information from the problem statement into an appropriate and useful representation that summarizes essential information symbolically and visually. The description is considered "useful" if it guides further steps in the solution process. A *problem description* could include restating known and unknown information, assigning appropriate symbols for quantities, stating a goal or target quantity, a visualization (sketch or picture), stating qualitative expectations, an abstracted physics diagram (force, energy, motion, momentum, ray, etc.), drawing a graph, stating a coordinate system, and choosing a system.

Physics Approach assesses a solver's skill at selecting appropriate physics concepts and principle(s) to use in solving the problem. Here the term *concept* is defined to be a general physics idea, such as the basic concept of "vector" or specific concepts of "momentum" and "average velocity". The term *principle* is defined to be a fundamental physics rule or law used to describe objects and their interactions, such as the law of conservation of energy, Newton's second law, or Ohm's law.

Specific Application of Physics assesses a solver's skill at applying the physics concepts and principles from their selected approach to the specific conditions in the problem. If necessary, the solver has set up specific equations for the problem that are consistent with the chosen approach. A *specific application of physics* could include a statement of definitions, relationships between the defined quantities, initial conditions, and assumptions or constraints in the problem (i.e., friction negligible, massless spring, massless pulley, inextensible string, etc.)

Mathematical Procedures assesses a solver's skill at following appropriate and correct mathematical rules and procedures during the solution execution. The term *mathematical procedures* refers to techniques that are employed to solve for target quantities from specific equations of physics, such as isolate and reduce strategies from algebra, substitution, use of the quadratic formula, or matrix operations. The term *mathematical rules* refers to conventions from mathematics, such as appropriate use of parentheses, square roots, and trigonometric identities. If the course instructor or researcher using the rubric expects a symbolic answer prior to numerical calculations, this could be considered an appropriate mathematical procedure.

Logical Progression assesses the solver's skills at communicating reasoning, staying focused toward a goal, and evaluating the solution for consistency (implicitly or explicitly). It checks whether the entire problem solution is clear, focused, and organized logically. The term logical means that the solution is coherent (the solution order and solver's reasoning can be understood from what is written), internally consistent (parts do not contradict), and externally consistent (agrees with physics expectations).

Summary of Drafting the Rubric

The first draft of a problem-solving assessment built on scoring criteria developed over time by the University of Minnesota Physics Education Group (Heller et al., 1992; Blue, 1997; Foster, 2000; and Tom-Thaden Koch, 2005) with substantial changes to the formatting of the scores and criteria. The revised rubric also differs from most past Minnesota instruments by including a NA or Not Applicable scores and having the same score range for all categories. The progression of the rubric drafts throughout this study (Appendix 1) indicate changes to the score range and improved consistency in the language for scoring criteria.

The rubric developed in this study considers five problem-solving processes: organizing problem information into a *Useful Description*, selecting appropriate physics principles (*Physics Approach*), applying the physics principles to the specific conditions in the problem (*Specific Application of Physics*), using *Mathematical Procedures* appropriately, and the overall communication of an organized reasoning pattern (*Logical Progression*). These processes are consistent with research on problem-solving in physics (Chi, 2006; Chi et al., 1981; de Jong & Ferguson-Hessler, 1986; Eylon & Reif, 1984; Heller & Reif, 1984; Larkin, 1979; 1981b; Larkin et al., 1980a; Larkin & Reif, 1979; Reif & Heller, 1982; Reif et al., 1976; Singh, 2002; Van Heuvelen, 1991a).

The next sections describe a series of studies conducted to test the validity, reliability, and utility of the rubric scores.

Preliminary Study with Two Expert Raters

Introduction

The goals for this preliminary study with two expert raters include obtaining: validity evidence for the response processes of students on written solutions to physics problems, validity evidence for the rubric's generalizability (applicability across different courses, physics topics, pedagogy, and solution detail), a measure of the reliability for two people using the rubric, and evidence for the utility of the rubric scores, including the degree to which the rubric distinguishes more- and less-skilled problem solvers.

The research questions addressed by this study are listed below, where the number and letter refer to the specific Research Question stated in chapter 1:

- 1b) To what extent do scores on the rubric reflect the problem-solving processes undertaken by a solver? (*response processes*)
- 1e) To what extent is the rubric applicable to multiple populations and contexts, including different student populations, physics topics, and problem features? (*generalizability*)
- 2a) To what extent do multiple raters' scores and score interpretations agree on the same problem solution? (*inter-rater agreement*)
- 3a) To what extent can the rubric distinguish between more- and less- skilled problem solvers?

Data Collection Procedures

Following the rubric's initial development, it was used by two raters (one researcher and one experienced high school physics teacher) to score final exam problem solutions from introductory university physics courses. A total of eight different problems were scored over a time period of one month. During this time, the teacher was in residence with the University of Minnesota physics education group. Five problems were from a calculus-based mechanics course for science and engineering and three problems were from an algebra-based mechanics course. Twenty solutions were randomly selected for each problem (out of approximately 200) that were legible and reflected a range of detail and quality. Interpretation of the rubric was discussed by the raters after independently scoring each problem (after every twenty solutions) and a final consensus was reached on the scores for each solution. The early rubric draft that was used to score these student exam solutions was Version 2 (Appendix 1).

Rubric Scores

The rubric score frequencies for the Algebra-based mechanics solutions (N=60) and the Calculus-based mechanics solutions (N=100) are plotted in Figure 2 and Figure 3, respectively. These graphs represent the final consensus scores that resulted from the discussions following each set of twenty solution ratings. Since the solutions were selected to represent a range of detail and quality, very few blank papers were chosen and subsequently there are very few zero scores. As might be expected, a higher fraction of solutions in the Calculus-based course scored high (4) in Mathematical Procedures. In both groups, there were few students scoring a high (4) score in the

Specific Application of Physics category, indicating difficulty appropriately applying physics concepts and principles to the specific conditions in the problems. For the Calculus-based solutions, there was one problem for which the Useful Description was rated Not Applicable (Problem), whereas no problems met this criteria in the algebra-based course solutions. Overall, the solutions reflected a range of scores across the five aspects scored on the rubric.

Figure 2: Frequency of Rubric Scores for Student Solutions to Exams (Algebra-Based Mechanics)

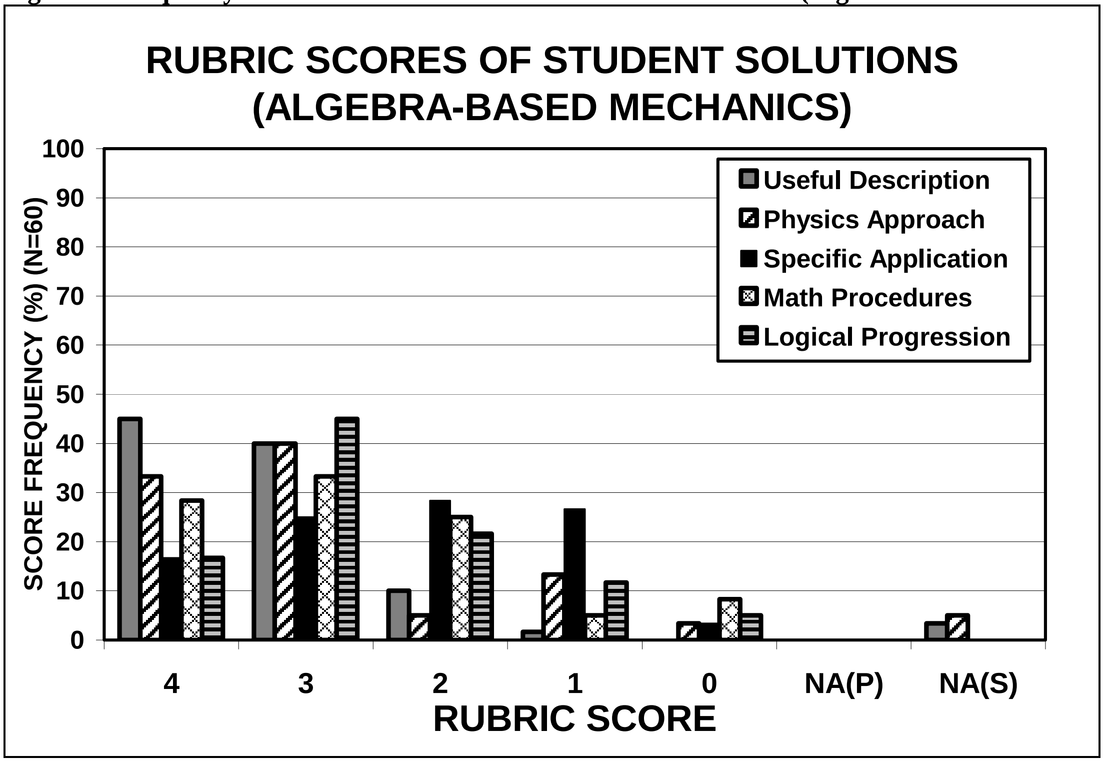
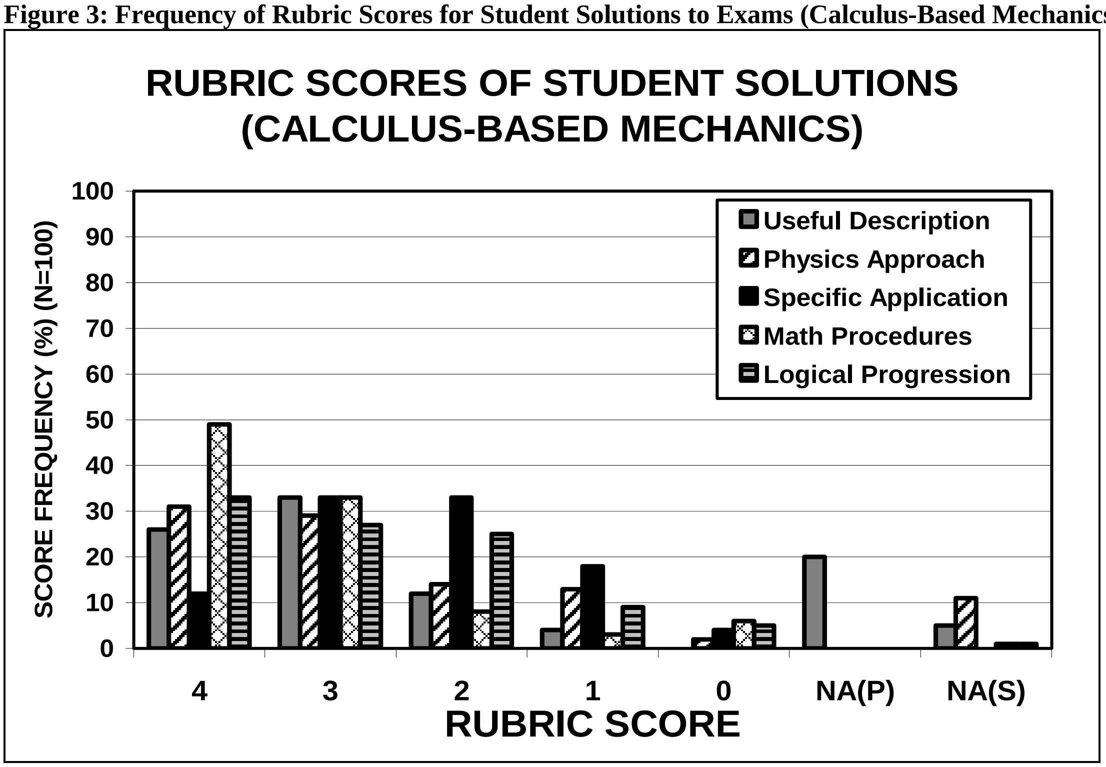


Figure 3: Frequency of Rubric Scores for Student Solutions to Exams (Calculus-Based Mechanics)



Score Agreement

Scores on all 160 solutions were used to determine the agreement of the two raters. The results are reported in Table 13. With discussion following each set of twenty solutions (implicit training), the overall percent exact agreement in each of the five categories ranged from 61±4% to 77±3% with an average of 68±4%. Agreement within one score (excluding NA scores) was above 96±1% in every category. As seen in the following table, the categories with lowest agreement were Logical Progression and Specific Application of Physics and the category with highest agreement was Useful Description.

In addition to percent agreement, a statistical measure of reliability is also reported for each rubric category. Kappa (Cohen, 1960; Howell, 2002) is a measure of raters' exact score agreement after correcting for expected agreement by chance. Weighted kappa is an extension of the kappa measure that considers the degree of difference in raters' scores (Cohen, 1968). Scores that are closer (such as agreement within one score) are given more weight in calculating the kappa agreement score than scores which differ more substantially. A more detailed explanation of kappa and its calculation are provided in Appendix 2, along with the pros and cons of using this particular measure. One limitation to using kappa is in the way it calculates an "expected" level agreement based on the responses of each rater and considers this a "chance" agreement. The meaning of chance in this context is unclear, and by correcting for it kappa is considered by some to be an overly conservative measure. For this reason, it is important to consider both the raw agreement counts (percent agreement) and this statistical measure together.

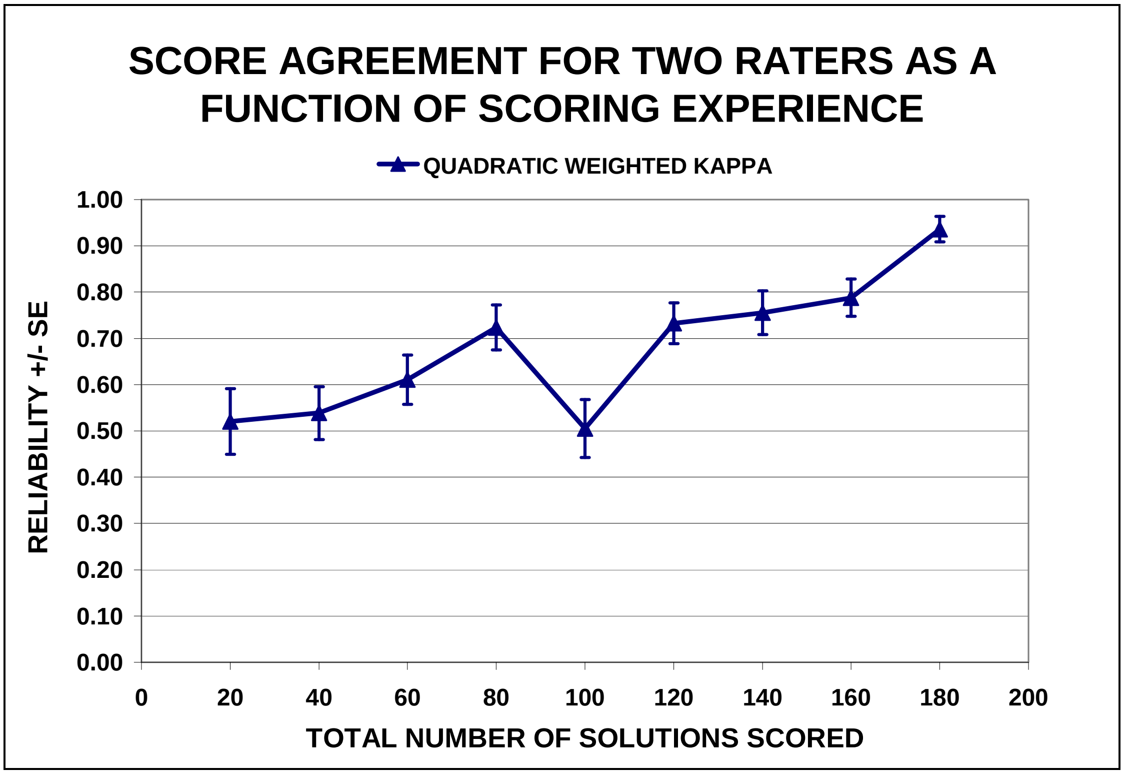
As seen in Table 13, the weighted kappa values for the categories study ranged from 0.59±0.04 to 0.75±0.04 with an overall reliability of 0.66±0.02. A kappa value above 0.60 is considered by some researchers to indicate "substantial agreement" and a value above 0.80 is considered "almost perfect agreement" (Landis & Koch, 1977).

Table 13: Score Agreement for the Preliminary Study with Two Raters

Category	% Perfect Agreement	Agreement Within One	Quadratic weighted kappa	Kappa Sig.
Useful Description	77±3 %	99±1 %	0.71±0.05	$p<0.001$
Physics Approach	76±3 %	97±1 %	0.75±0.04	$p<0.001$
Specific Application	61±4 %	97±1 %	0.59±0.04	$p<0.001$
Math Procedures	66±4 %	99±1 %	0.64±0.03	$p<0.001$
Logical Progression	63±4 %	97±1 %	0.61±0.04	$p<0.001$
Overall	68±4 %	99±1 %	0.66±0.02	$p<0.001$

Although this table examines the agreement by rubric category, it should be interpreted as an "average" level of agreement over the eight problems scored in the month-long time period. As seen in the following plot (Figure 4), the discussions following each problem (every 20 solutions) provided an "implicit" level training that resulted in a general increase in the weighted kappa reliability scores over time. The final and ninth point on the graph is a re-scoring of the first problem on the graph and was not included in the previous table of agreement by category. As seen in the graph, reliability as measured by quadratic weighted kappa was initially around 0.50. After one month of scoring and discussing 180 solutions the reliability rose to be above 0.90, nearly at the maximum possible level of agreement. The problem statements for these eight mechanics problems are listed in Appendix 3.

Figure 4: Graph of Score Agreement for Two Raters as a Function of Time for Eight Problems. Reliability is measured by quadratic weighted kappa. The ninth data point is a re-scoring of the first problem initially scored as the first data point. Discussion occurred after the data for a set of solutions was recorded.



The graph indicates a drop in score agreement for the fifth problem scored. A closer look at the percent agreement by problem and category in Table 14 shows that the Logical Progression and Physics Approach agreement measures for Problem 5 were the lowest of all the problems. The Specific Application of Physics score was also lower than average for that particular problem, whereas the Math Procedures agreement was consistent with other problems.

The reason for this dip in some categories is unclear, however one possible source of disagreement could be the problem topic (Refer to Appendix 3 for the problem statements). This particular problem focused on the topic of simple harmonic motion, or the horizontal oscillation of a mass between two springs. A comparison of the scores assigned by each rater in the Physics Approach and Specific Application of Physics categories indicated there were several instances in which the teacher scored a "zero" for these categories or NA(Solver) when the researcher assigned a numerical score. Zero represented a missing but necessary solution aspect for the rubric version used, suggesting the teacher had different and perhaps narrower criteria for evidence of these categories than the researcher. Graphs of these two physics categories on problem 5 are plotted below. The Logical Progression category scores typically agreed within one score, with the Teacher assigning a higher score than the researcher.

Figure 5: Researcher and Teacher Scores for Physics Approach on Each Solution to Problem Five

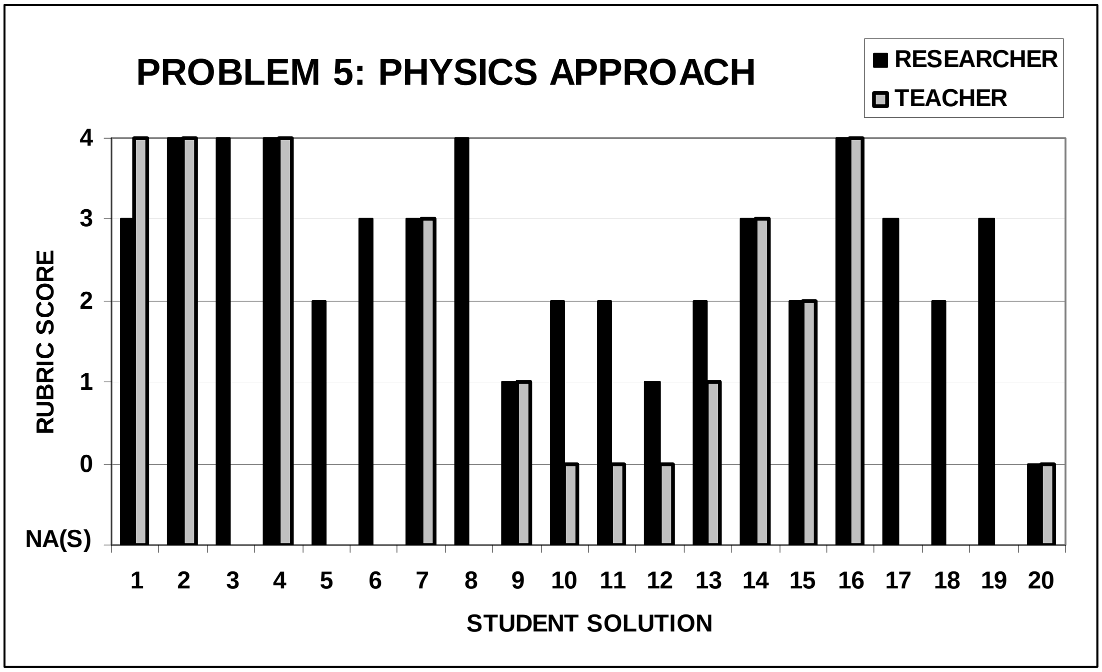


Figure 6: Researcher and Teacher Scores for Specific Application of Physics on Each Solution to Problem Five

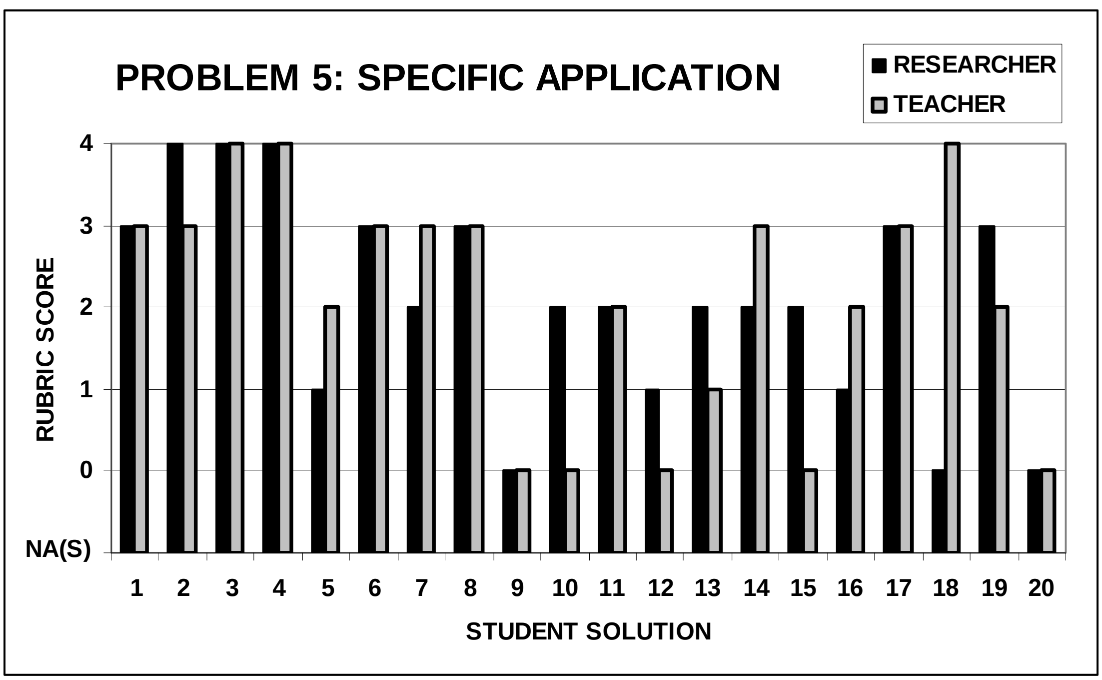


Table 14: Percent Score Agreement for the Preliminary Study by Problem

Percent agreement values include NA(Solver) scores. Solutions to problems 1-5 were from a calculus-based mechanics course and solutions to problems 6-8 were from an algebra-based course.

Category	**P1 Energy (N=20)**	**P2 Forces (N=20)**	**P3 Momen-tum (N=20)**	**P4 Angular (N=20)**	**P5 Oscilla-tions (N=20)**	**P6 Kine-matics (N=20)**	**P7 Forces (N=20)**	**P8 Momen-tum (N=20)**	**P1 Rescore (N=20)**
Useful Description	70±10%	70±10%	85±8%	65±11%	NA(P)	85±8%	85±8%	65±11%	95±5%
Physics Approach	65±11%	45±11%	90±7%	70±10%	40±11%	90±7%	75±10%	95±5%	100±0%
Specific Application	60±11%	55±11%	40±11%	75±10%	45±11%	50±11%	75±10%	90±7%	90±7%
Math Procedures	60±11%	55±11%	45±11%	75±10%	65±11%	80±9%	75±10%	70±10%	95±5%
Logical Progression	65±11%	55±11%	65±11%	75±10%	30±10%	60±11%	85±8%	70±10%	95±5%
Overall	64±5%	56±5%	65±5%	72±4%	45±5%	73±4%	79±4%	78±4%	95±2%
Weighted Kappa	0.52±.07	0.54±.06	0.61±.05	0.77±.04	0.57±.06	0.74±.04	0.76±.05	0.79±.04	0.94±.03

Distinguish Instructor and Student Solutions

In addition, a preliminary study was conducted to determine the rubric's utility for distinguishing instructor or "expert" solutions from student solutions. Two problems were selected randomly from each of 38 chapters in a popular calculus-based physics textbook (Halliday, Resnick, & Walker, 1997) (N=76), and the solutions printed in the instructor solution manual were scored with the rubric (Rubric Version 2 in Appendix 1). The solutions were typically very sparse and did not include much explicit reasoning. Then, homework solutions hand-written by a physics instructor for an entire introductory physics course (N=83) were scored with the rubric. These solutions were more detailed and included steps of the reasoning process.

The frequency of rubric scores was very similar for the instructor solution manual and the instructor, regardless of the level of detail. Most rubric scores for instructors were the highest possible value or a not applicable score (Figure 7 and Figure 8). In comparison, scores of student solutions to different problems on exams spanned the entire range of rubric scores (Figure 2 and Figure 3). From the differences in score frequencies it was easy to distinguish between the sets of instructor and student solutions.

Figure 7: Frequency of Rubric Scores for Randomly Selected Problems in a Textbook Instructor Solution Manual

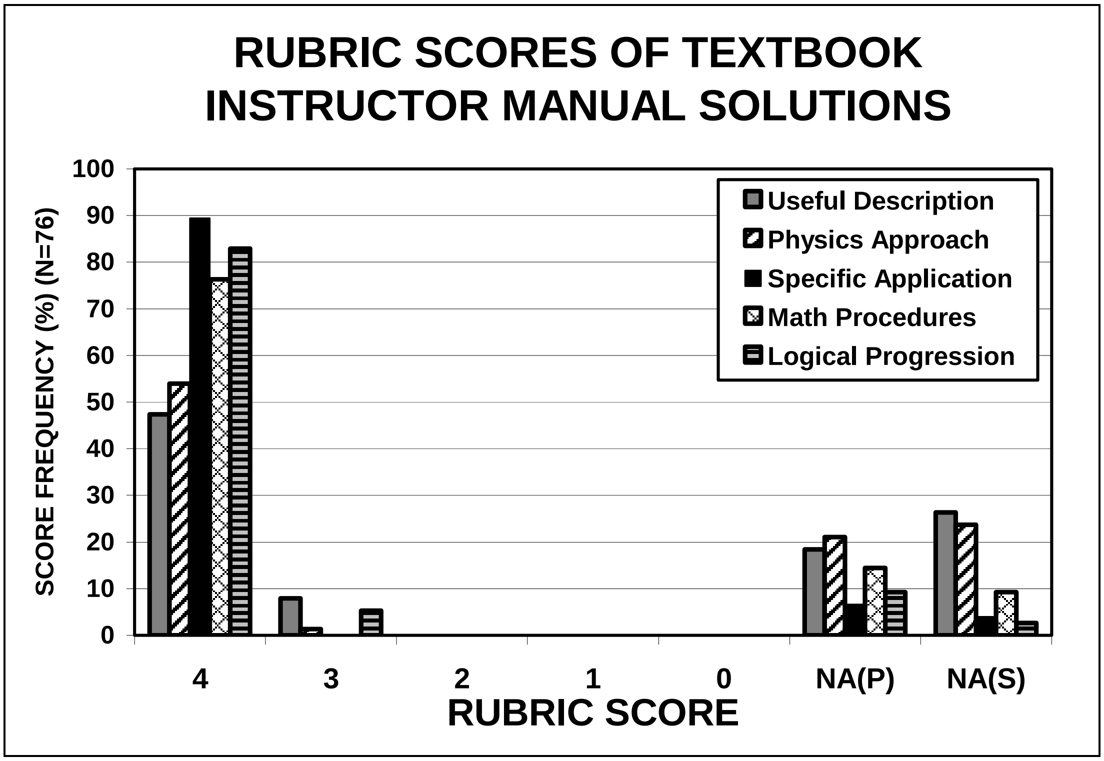
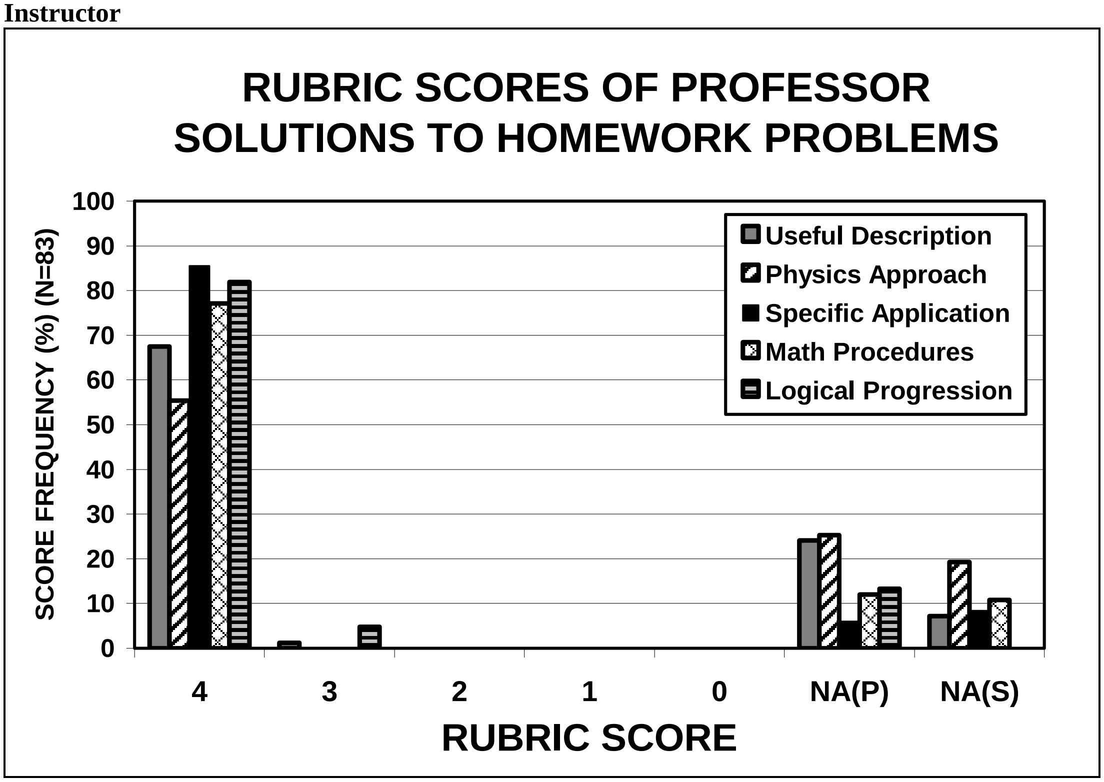


Figure 8: Frequency of Rubric Scores for Detailed Problem Solutions Written by a Course Instructor



Summary of Preliminary Study

In this study, a researcher and an experienced high school teacher scored student solutions to final exam problems (N=180) over the period of one month with periodic feedback as mutual training. The solutions reflected a range of physics topics from mechanics, including forces (Newton's second law), energy conservation, momentum conservation, rotational (angular) motion, and oscillations. These scores provided baseline information for the maximal level of agreement of two people using the instrument and the generalizability of the rubric to multiple courses and physics topics. In addition, detailed homework solutions written by a course instructor and sparse solutions in a textbook's instructor solution manual were also scored to assess the rubric's ability to distinguish more- and less-skilled solvers.

A key goal for this preliminary study with two raters was to measure the reliability or inter-rater agreement for two people using the rubric. With the training of mutual feedback the agreement increased over time. Averaged over the training interval and over the topics in an introductory mechanics course, the percent exact agreement in each of the five categories ranged from 61±4% to 77±3% with an average of 68±4%. Agreement within one score (excluding NA scores) was 97±1% or higher in every category. Weighted kappa values for the categories study ranged from 0.59±0.04 to 0.75±0.04 with an overall reliability of 0.66±0.02 or "substantial agreement" above chance. The increased agreement over time (Figure 4) is characterized by re-scoring the solutions to the first problem rated where the weighted kappa increased from around 0.52±.07 at the beginning to 0.94±.03 at the end of a month. The agreement also showed, with one exception, a steady increasing trend with time for different problem solutions (Table 14).

For the student solutions scored, there was evidence of all of the rubric categories, suggesting a consistency between the response processes measured by the rubric and students' problem-solving processes observed in written solutions. The rubric was also applicable to two different physics courses (algebra-based and calculus-based mechanics) and multiple physics topics in mechanics.

Another goal of this preliminary study was to obtain evidence for the utility of the rubric scores, including the degree to which the rubric distinguishes more- and less-skilled problem solvers. A comparison of the frequency of rubric scores for final exam solutions by students in algebra-based and calculus-based physics courses (Figure 2 and Figure 3) indicate solutions spanned the range of scores, whereas scores of instructor solutions (Figure 7 and Figure 8) were primarily scored high or NA, regardless of the level of detail in the solution. These frequency charts indicate the rubric can distinguish more- and less- skilled problem solvers when several measures are made. The next section describes further study of the reliability of rubric scores for multiple raters and describes a test of minimal written training materials and documentation for the rubric.

First Study with Training Raters

Introduction

After the preliminary study with two raters established a baseline maximal measure of score agreement, the next step was to develop documentation and training materials for the rubric that could be used in less time with a less expert group of users and then subsequently test these materials for the agreement of those users. To make the rubric easy to use, the object of this study was to establish the minimal training necessary for get reasonable reliability among inexperienced raters. The goals for the first study with training raters included: obtaining validity evidence for the rubric content and format (relevance and completeness) as judged by graduate students experienced in traditional grading practices, assessing the adequacy of documentation in outlining the purposes of the rubric, obtaining a reliability measure of score agreement from multiple raters using the rubric, testing the clarity of task instructions in the training materials, and assessing the raters' perceptions of the rubric's usefulness.

The research questions addressed in this study are listed below, where the number and letter refer to the specific Research Question stated in Chapter 1:

- 1a) To what extent are rubric categories consistent with descriptions of physics problem solving processes? (*content relevance & representativeness*)
- 1f) To what extent does the rubric documentation address potential positive and negative consequences of the proposed test use? (*consequences*)
- 2a) To what extent do multiple raters' scores and score interpretations agree on the same problem solution? (*inter-rater agreement*)
- 2b) What scorer training is necessary to achieve a desired level of rater agreement? (*reliability and utility*)
- 3b) How authentic are the assessment's goals, tasks, and constraints? (*utility*)

Data Collection Procedures

Eighteen graduate students who met the criteria for this study (experienced teaching assistants in at least their third year of graduate school) were contacted by e-mail with a brief description of the research study, the task which would be asked of them, and the expected time involved. The eight interested volunteers were randomly assigned to two groups. Four people used the rubric to score student solutions from a mechanics final exam problem and four people scored student solutions from an electricity and magnetism (E&M) final exam problem. The graduate students were provided with an instruction sheet, a copy of the rubric, brief definitions of each category on the rubric, the problem statement, an example instructor solution to the problem, a blank scoring template table, and a set of student solutions. There was no other contact with the researcher and no organized contact among the graduate students. These materials are described in more detail in Appendix 4.

In both groups, the graduate students were asked to use the rubric to score eight student solutions without any explicit training or discussion. After submitting their scores and rationale they received a brief written self-training consisting of example scores and rationale for the first three solutions. Raters were instructed by writing to read the example scores and rationales (written in a table) for these three solutions and compare them to their own scores. They were then instructed to rescore the remaining five solutions from before and score five new solutions. Although the graduate students were not given an explicit deadline for returning materials, most completed the task in one week.

Score Agreement

Reliability was assessed by comparing the graduate students' scores to the consensus scores of two expert raters. Since the reliability values are approximately the same for both the mechanics and E&M problems, the scores for all eight graduate students have been combined into a single analysis. As seen in

Table 15, perfect agreement in scores for each category of the mechanics problem ranged from 13±6% to 38±9% before training with an overall average of 28±4%. The after training agreement on this problem ranged from 25±7 to 63±8% with an average of 44±4%. Agreement within one score above or below was 74±2% before and 85±2% after training, with most of the agreement within one occurring with graduate student scores above the researcher scores. On the E&M problem perfect agreement in scores for each category before training ranged from 19±7% to 56±9% with an average of 41±4% and after training ranged from 38±8% to 50±8% with an average of 45±4%. Agreement within one score above or below was 81±2% before and 88±2% after training, with most of the agreement within one occurring with graduate student scores above the researcher scores.

Table 15: Percent Agreement of Graduate Student Scores with Expert Raters' Scores Before and After Training (Mechanics)

	BEFORE TRAINING			**AFTER TRAINING**		
	Perfect Agreement (N=32)	TAs One Above (N=32)	TAs One Below (N=32)	Perfect Agreement (N=40)	TAs One Above (N=40)	TAs One Below (N=40)
Useful Description	18±7%	32±8%	7±5%	53±8%	40±8%	5±3%
Physics Approach	31±8%	38±9%	9±5%	25±7%	20±6%	20±6%
Specific Application	38±9%	43±9%	9±5%	50±8%	33±7%	8±4%
Math Procedures	13±6%	56±9%	3±3%	30±7%	40±8%	8±4%
Logical Progression	38±9%	28±8%	3±3%	63±8%	10±5%	18±6%
Overall	28±4%	40±4%	6±2%	44±4%	29±3%	12±2%

Table 16: Percent Agreement of Graduate Students' Scores with Expert Raters' Scores Before and After Training (E&M)

	BEFORE TRAINING			AFTER TRAINING		
	Perfect Agreement (N=32)	TAs One Above (N=32)	TAs One Below (N=32)	Perfect Agreement (N=40)	TAs One Above (N=40)	TAs One Below (N=40)
Useful Description	56±9%	31±8%	3±3%	41±8%	25±7%	15±6%
Physics Approach	43±9%	39±9%	4±3%	50±8%	28±7%	18±6%
Specific Application	53±8%	41±9%	6±4%	45±8%	20±6%	30±7%
Math Procedures	29±8%	8±5%	12±6%	50±8%	8±4%	17±6%
Logical Progression	19±7%	41±9%	13±6%	38±8%	35±8%	13±5%
Overall	41±4%	33±4%	7±5%	45±4%	25±3%	18±6%

Scores in the categories Mathematical Procedures and Logical Progression were most affected by the training. These aspects initially had the lowest agreement with the expert raters, indicating differences in interpretations of the categories. Viewing the written examples helped to achieve a closer match with the expert rater scores. An example of a shift in scores for Logical Progression is shown below in Figure 9 (before training) and Figure 10 (after training) for a single student solution on the Mechanics problem. An example of a shift in Approach and Math scores (but not Logic) is shown in Figure 11(before training) and Figure 12 (after training) for a single student solution on the E&M problem.

Figure 9: TA and Expert Rubric Scores on Mechanics Student Solution Five (Before Training)

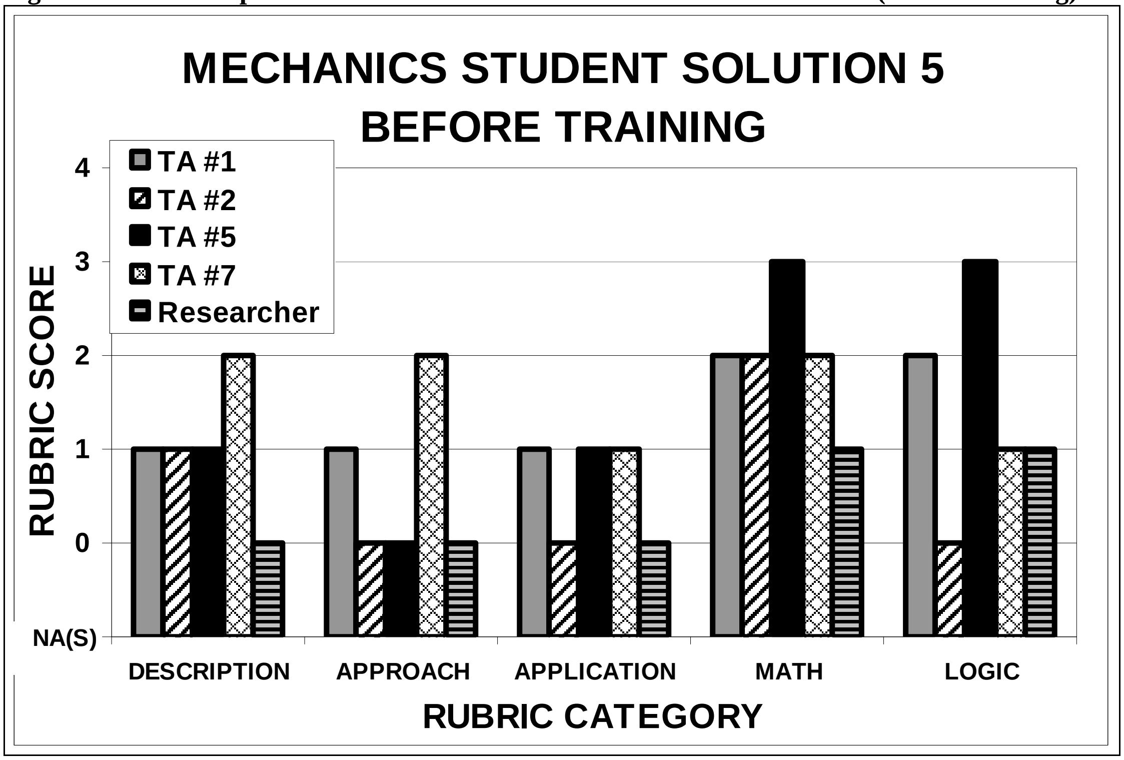


Figure 10: TA and Expert Rubric Scores on Mechanics Student Solution Five (After Training)
Training improved the agreement in Logic scores.

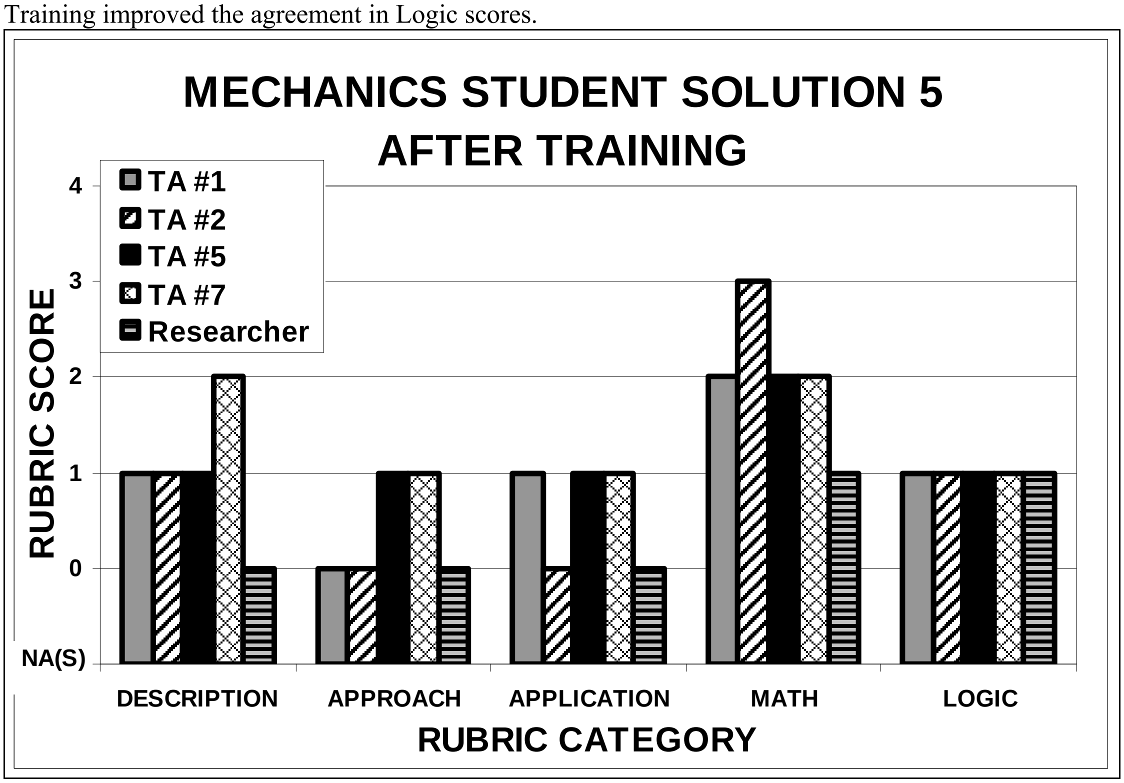


Figure 11: TA and Expert Rubric Scores on E&M Student Solution Six (Before Training)

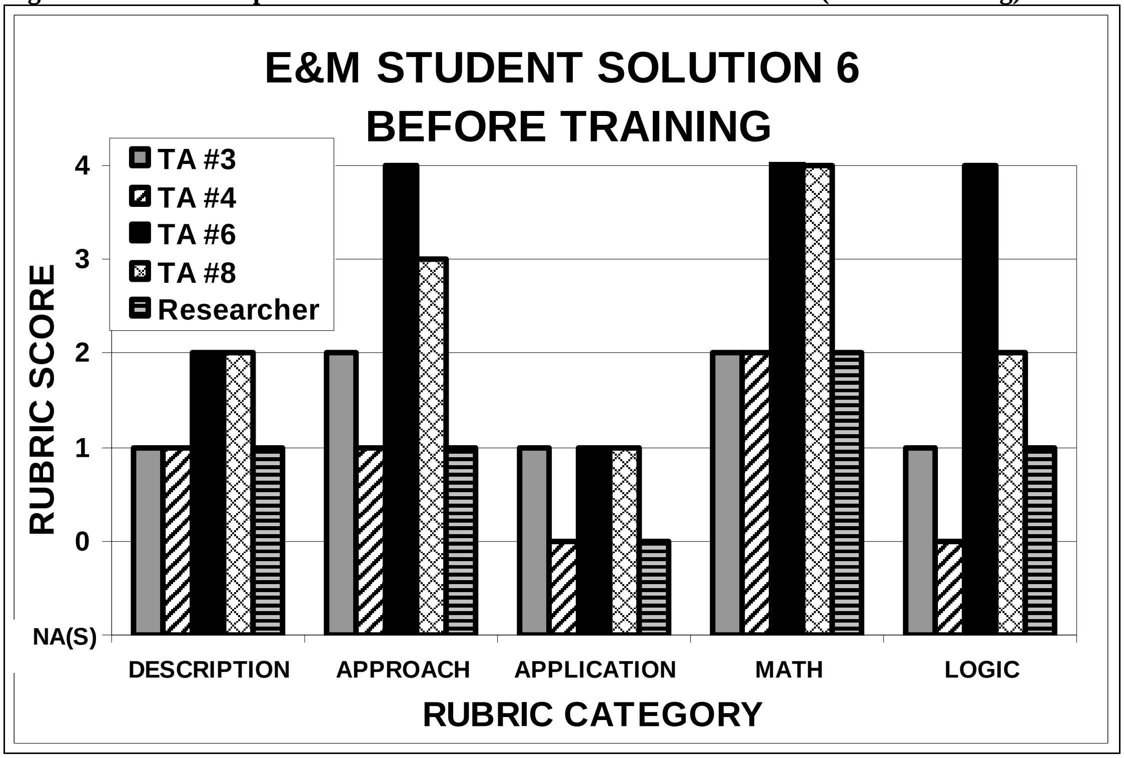
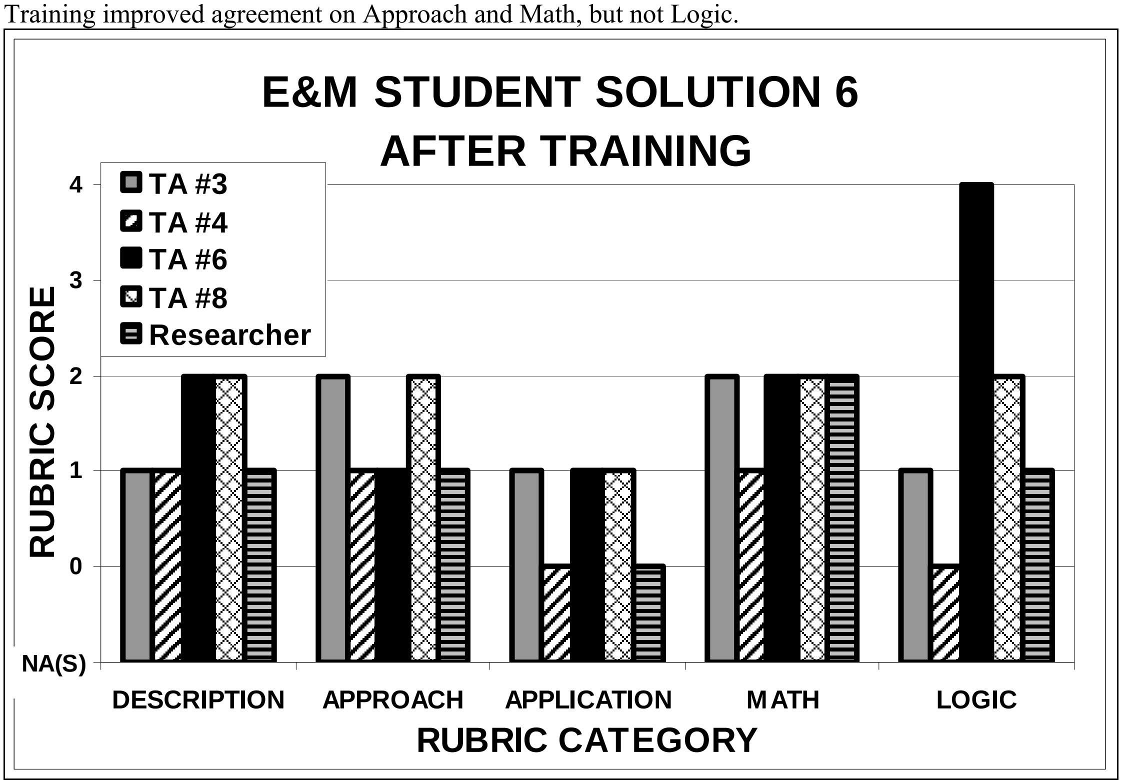


Figure 12: TA and Expert Rubric Scores on E&M Student Solution Six (After Training)
Training improved agreement on Approach and Math, but not Logic.



Useful Description and Specific Application of Physics were not significantly affected by the self-training. Table 17 and Table 18 indicate another reliability measure of rater agreement with the expert raters' scores. Values for agreement above chance as measured statistically by quadratic weighted kappa (Cohen, 1968) are reported for each category before and after training. Appendix 2 lists the equations and values necessary to calculate kappa and quadratic weighted kappa (Cohen, 1960; Howell, 2002).

Kappa is a standard measure used for inter-judge agreement, and its calculation based on probability theory is relatively straightforward (Cohen, 1960; Howell, 2002) although its method of correcting for chance agreement is controversial and leads to an underestimation of agreement (see Appendix 2). Although it is a standard measure, kappa only estimates the level of perfect score agreement above chance, and does not consider other levels of agreement (for example, agreement within one score or two scores). For this reason, the values for weighted kappa are provided in the table instead of kappa (Cohen, 1968). This means that scores are weighted by how close they are to perfect agreement such that scores that are close but not necessarily perfect receive more weight in the calculation than scores that differ substantially. For more information, review Appendix 2. Weighted kappa can vary between -1 and +1 with 0 being consistent with agreement by chance, a negative number indicating disagreement above chance, a positive number indicating agreement above chance, and 1 being perfect agreement.

As indicated by Table 17 the overall weighted kappa before training on the mechanics problem was 0.23 ± 0.04 (often designated fair agreement) and improved to a weighted kappa of 0.41±0.04 (often designated moderate agreement) after a minimal written training exercise (Landis & Koch, 1977). As indicated in Table 18, the overall weighted kappa before training on the E&M problem was 0.31±0.05 and improved to a weighted kappa of 0.43±0.04 after training. Consistent with the percent agreement scores for each category, the categories that were most influenced by the training include the Mathematical Procedures and Logical Progression categories. The reliability measure for Mathematical Procedures was not significant before training on either problem, and was significant at the 0.01 level after training on the E&M problem. The Logical Progression category was significant at the 0.01 level on the mechanics problem before training and increased to significance at the 0.001 level after training. Here the level of significance is the probability that the agreement could occur by chance. The agreement in all categories after training was statistically significant.

Table 17: Reliability of Graduate Students' Scores as Measured by Quadratic Weighted Kappa (Mechanics Problem)

	Before Training (4 raters x 8 papers)		After Training (4 raters x 10 papers)	
Category	Quadratic Weighted Kappa	Kappa Sig.	Quadratic Weighted Kappa	Kappa Sig.
Useful Description	0.07±0.10	Not Sig.	0.11±0.10	p<0.1
Physics Approach	0.30±0.11	p<0.01	0.55±0.08	p<0.001
Specific Application	0.33±0.10	p<0.01	0.47±0.09	p<0.001
Math Procedures	0.04±0.08	Not Sig.	0.19±0.09	p<0.1
Logical Progression	0.30±0.10	p<0.01	0.58±0.11	p<0.001
Overall	0.23±0.04	p<0.001	0.41±0.04	p<0.001

Table 18: Reliability of Graduate Students' Scores as Measured by Quadratic Weighted Kappa (E&M Problem)

	Before Training (4 raters x 8 papers)		After Training (4 raters x 10 papers)	
Category	Quadratic Weighted Kappa	Kappa Sig.	Quadratic Weighted Kappa	Kappa Sig.
Useful Description	0.41±0.12	p<0.01	0.48±0.09	p<0.001
Physics Approach	0.36±0.12	p<0.01	0.42±0.09	p<0.001
Specific Application	0.51±0.09	p<0.001	0.53±0.07	p<0.001
Math Procedures	0.01±0.14	Not Sig.	0.31±0.12	p<0.01
Logical Progression	0.05±0.10	p<0.05	0.25±0.10	p<0.05
Overall	0.31±0.05	p<0.001	0.43±0.04	p<0.001

The frequency of rubric scores for the four raters before and after training on the mechanics problem are shown in Figure 13 and Figure 14. As seen in the graph, the scores for Mathematical Procedures and Logical Progression had a noticeable shift (decrease) in scores after training. Initially, the graduate students rated several of the solutions with a high score (4) for these categories, which was inconsistent with the researchers' scores. This pattern is not observed for the E&M problem (Figure 15 and Figure 16) because one rater considered Math "Not Applicable" to the problem before training and the shift in high scores is not as noticeable.

The training appeared to have more influence on the E&M problem raters than the Mechanics raters. On the Mechanics problem, the researcher had a lower average score than the raters for all categories before training and after training, with the exception of the Logical Progression category. On the E&M problem, the scores and rationale provided in the training resulted in a shifted interpretation of these categories and higher agreement with the researchers' scores. As seen in the Table 20 below, the researcher had a lower average score than the raters for most categories before training, but the rater and researcher averages were closer after the training.

Table 19: Average Rater and Researcher Rubric Scores for Each Category (Mechanics)

Category	Before Training (8 papers)		After Training (10 papers)	
	Average Rater Score	Average Researcher Score	Average Rater Score	Average Researcher Score
Useful Description	2.3±0.2	1.3±0.4	2.2±0.2	1.2±0.3
Physics Approach	2.2±0.2	2.0±0.6	2.3±0.2	2.4±0.4
Specific Application	2.0±0.2	1.5±0.4	2.0±0.2	1.5±0.4
Math Procedures	3.3±0.1	2.3±0.3	3.1±0.2	2.4±0.3
Logical Progression	2.9±0.2	2.0±0.4	2.4±0.2	2.4±0.4

Table 20: Average Rater and Researcher Rubric Scores for Each Category (E&M)

Category	Before Training (8 papers)		After Training (10 papers)	
	Average Rater Score	Average Researcher Score	Average Rater Score	Average Researcher Score
Useful Description	2.2±0.2	1.8±0.3	2.4±0.2	2.2±0.4
Physics Approach	2.8±0.2	2.1±0.2	2.4±0.2	2.5±0.6
Specific Application	1.8±0.2	1.5±0.5	1.9±0.2	2.1±0.5
Math Procedures	3.5±0.2	2.4±0.3	2.5±0.2	2.4±0.3
Logical Progression	2.7±0.2	2.4±0.4	2.4±0.2	2.3±0.4

Figure 13: Frequency of Rubric Scores Before Training (Mechanics)

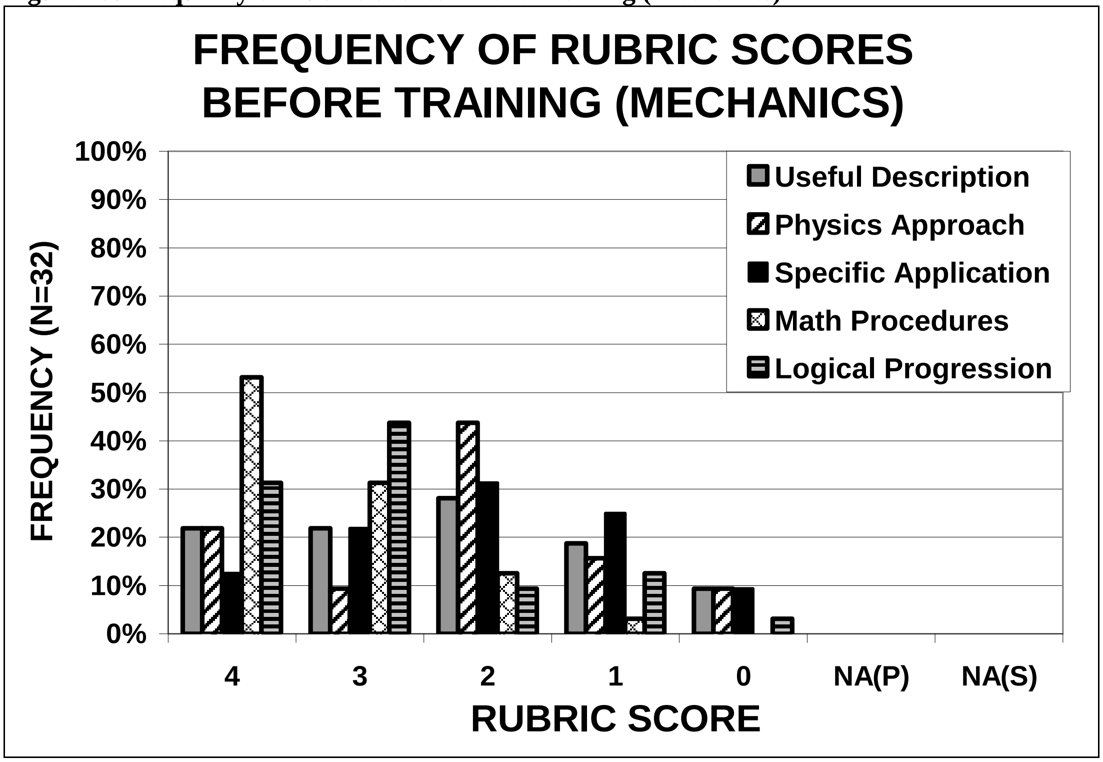


Figure 14: Frequency of Rubric Scores After Training (Mechanics)

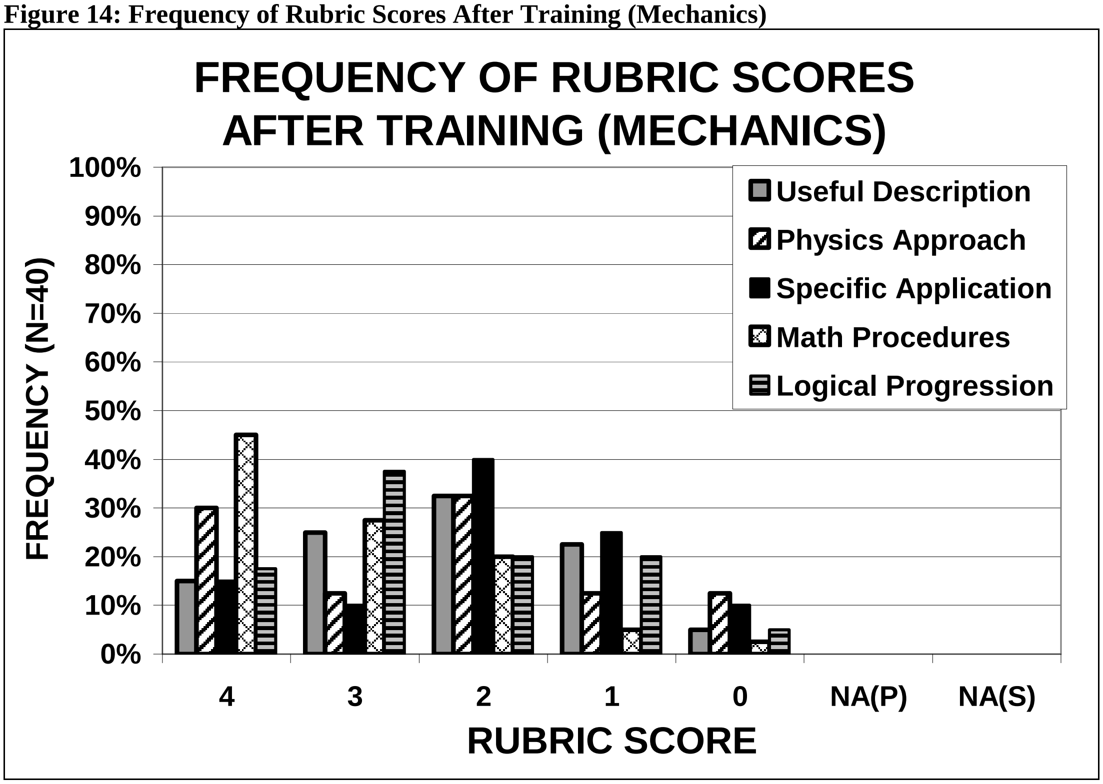


Figure 15: Frequency of Rubric Scores Before Training (E&M)

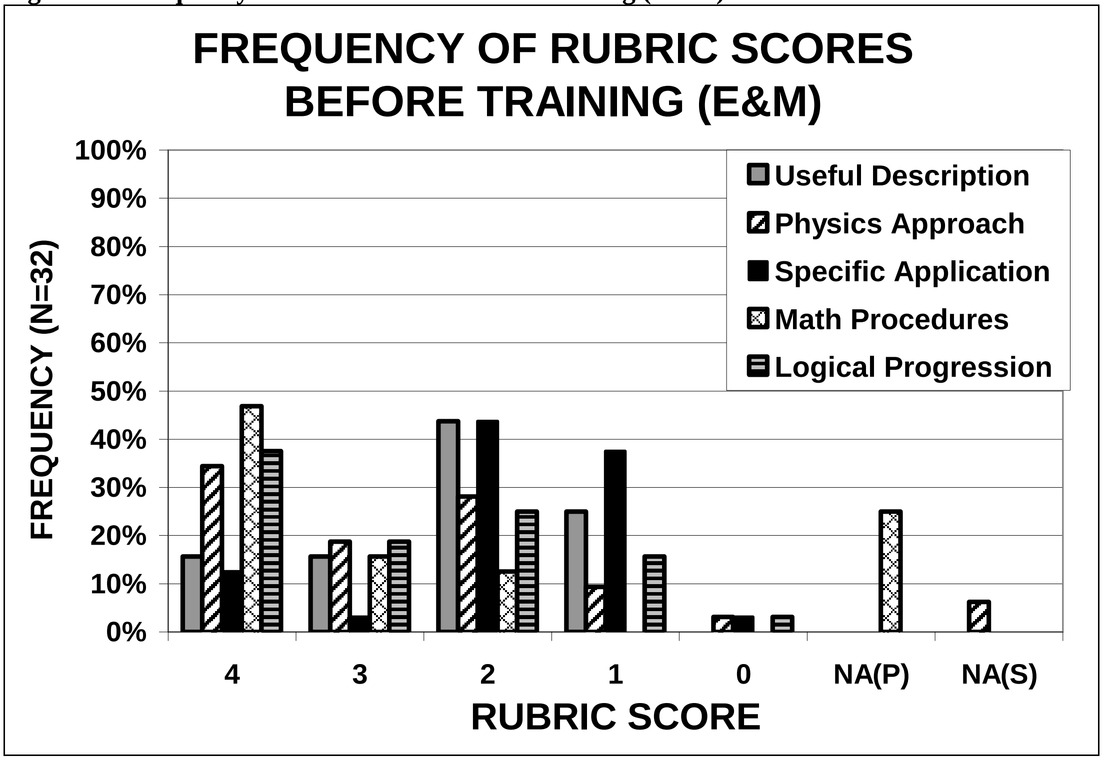
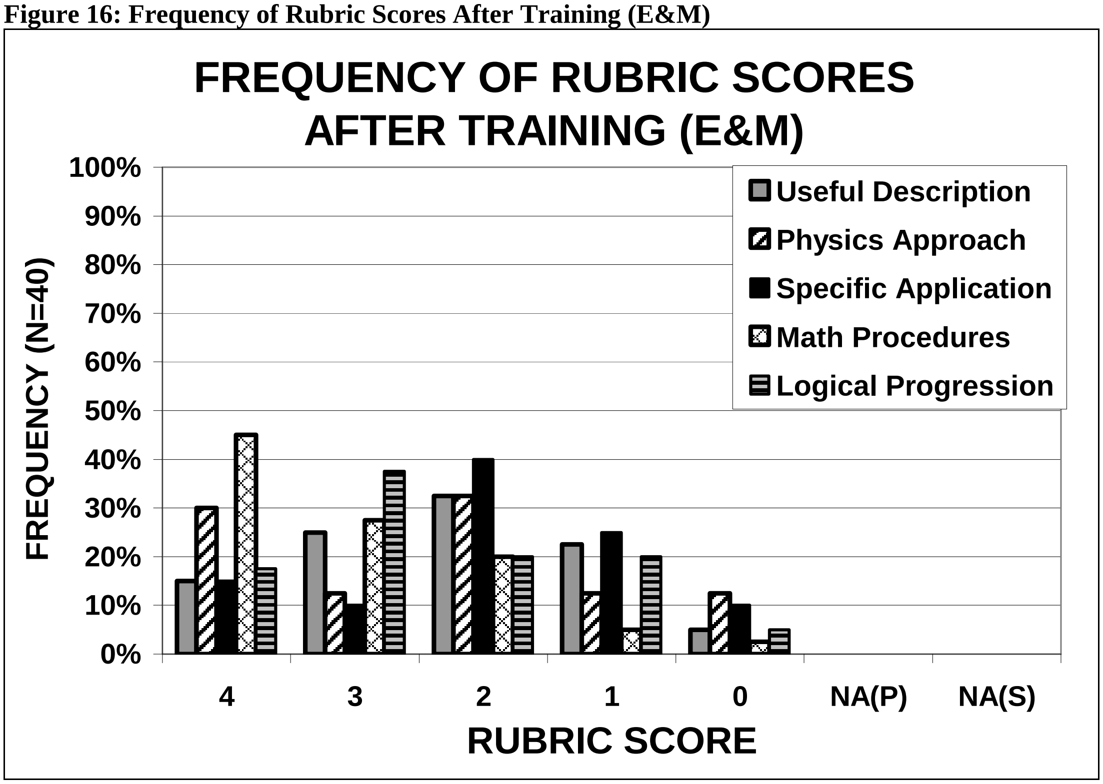


Figure 16: Frequency of Rubric Scores After Training (E&M)



The scores in Table 17 and Table 18 indicated the extent of the agreement of the eight raters' scores with researchers' consensus scores before and after a brief written training exercise. It is also important to consider the agreement of the raters' scores with each other before and after training. Since Cohen's kappa only compares pairs of ratings, a different measure, called Fleiss's kappa (Fleiss, 1971) is necessary to compare the ratings for the eight raters. A description of the equations necessary to calculate Fleiss's kappa is provided in Appendix 2.

In this study, Fleiss's kappa was 0.30±0.04 before training and 0.23±0.03 after training on the Mechanics problem, and 0.15±0.05 before training and 0.20±0.03 after training on the E&M problem. When considering the standard errors, this indicated the graduate students had a similar level of agreement with each other before and after training, but the mechanics raters agreed more with each other than the E&M raters did. Since the agreement with researchers' scores improved after training, this indicates there was a similar spread in scores at both times, but the overall average shifted to be closer to the researchers' scores. Fleiss's kappa should be compared to Cohen's kappa and not the weighted kappa since it is only a measure of perfect agreement for nominal scales, and does not consider the ordering of those scores in the same way weighted kappa does.

Table 21: Reliability of Graduate Students' Scores as Measured by Fleiss's Kappa (Mechanics)

	Before Training (4 raters x 8 papers)		After Training (4 raters x 10 papers)	
Category	Fleiss's Kappa	Kappa Sig.	Fleiss's Kappa	Kappa Sig.
Useful Description	0.01±0.08	Not Sig.	0.03±0.08	Not Sig.
Physics Approach	0.33±0.11	p<0.01	0.30±0.08	p<0.001
Specific Application	0.48±0.08	p<0.001	0.30±0.09	p<0.001
Math Procedures	0.31±0.14	p<0.05	-0.01±0.10	Not Sig.
Logical Progression	0.12±0.09	Not Sig.	0.31±0.07	p<0.001
Overall	0.30±0.04	p<0.001	0.23±0.03	p<0.001

Table 22: Reliability of Graduate Students' Scores as Measured by Fleiss's Kappa (E&M)

	Before Training (4 raters x 8 papers)		After Training (4 raters x 10 papers)	
Category	Fleiss's Kappa	Kappa Sig.	Fleiss's Kappa	Kappa Sig.
Useful Description	0.16±0.11	p<0.20	0.13±0.08	p<0.20
Physics Approach	0.14±0.09	p<0.20	0.16±0.06	p<0.05
Specific Application	0.23±0.12	p<0.20	0.34±0.09	p<0.001
Math Procedures	-0.08±0.12	Not Sig.	0.20±0.08	p<0.05
Logical Progression	-0.08±0.08	Not Sig.	0.14±0.06	p<0.05
Overall	0.15±0.05	p<0.01	0.20±0.03	p<0.001

Comments from Raters

The graduate students also responded to questions about the rubric and suggested changes. Their comments focused on scoring difficulties, difficulties understanding either the category descriptions or the evidence for a category, and the adequacy of the training materials. A list of the questions is below:

(After the first scoring of eight solutions)

1. What difficulties did you encounter while using the scoring rubric?
    a. Which of the five categories was most difficult to score and why?
    b. Which student solutions were the most difficult to score and why?
2. What changes, if any, would you recommend making to the rubric? Why?
3. If you were deciding how to grade these student solutions for an introductory physics course exam, how would you assign points? (out of 20 total points)

(After training and second scoring of ten solutions)

4. What difficulties did you encounter while using the scoring rubric?
5. Were the example scores useful? Why or why not?
6. What further changes, if any, would you recommend making to the rubric?

Some graduate students expressed confusion about the "Not Applicable" scores. These scores and the score zero were largely ignored or avoided, even after training. It should be noted that no examples of the NA(Problem) rating were included in the self-training materials and there were very few examples of NA(Solver). In answering the written questions, several graduate students commented on the need to include examples of the NA scoring in the training materials. The training examples were subsequently modified to include more NA(Solver) scores, as is explained in the Second Study with Training Raters. Examples of NA(Problem) were not included in the training because the examlpe problems were chosen to include all aspects of the rubric, however the NA(Problem) score criteria was modified to include “i.e.” statements giving examples of when that score is appropriate, such as “i.e. a description is provided in the problem statement” for the Useful Description category.

One graduate student expressed difficulty scoring the mechanics problem, which had multiple parts (a and b) that each required a student to solve for a separate physics quantity. This person expressed difficulty deciding whether to assign separate rubric scores for each part of the problem, or to give one overall score for the solution. Although the remaining three raters of the mechanics problem did not express difficulty with scoring a multi-part problem, training materials were modified to exclude the multi-part problem. Although multi-part problems are common in physics, for the purposes of the training materials this difficulty was avoided and addressed in later studies, such as scoring written solutions to exams.

Written comments also indicated the graduate student raters were strongly influenced by their traditional grading experiences. They expressed concerns about scoring math and logical progression when the physics is inappropriate: “I don't think credit should be given for a clear, focused, consistent solution with correct math that uses a totally wrong physics approach” (GS#1). Some also expressed a desire to weight the categories based on their importance to the problem or difficulty level, such as giving a lower weight for simple algebraic math procedures.

The graduate students also perceived substantial overlap in some categories and had difficulty treating some of the categories independently. GS# 1 remarked, “Specific application of physics was most difficult. I find this difficult to untangle from physics approach. Also, how should I score it when the approach is wrong?”

In response to the training materials, GS #6 commented, “They [example scores] helped me understand what someone else thought was important. They did seem a touch harsh. I also think I was a little lax the first time around. Examples help clarify the details.” This is consistent with the shift in scores observed from Figure 13 to Figure 14. One graduate student did not perceive the training example scores as very helpful, because “I did not always agree with them” (GS #2).

In response to the third question about how they would assign grading points on this question for an exam, three of the graduate students mentioned scoring the diagram or description, four mentioned their use of a physics principle (such as conservation of energy), five mentioned correct formulas or how the principle was used (a sixth person just said “physics” rather than separating this aspect into the principle and its application), and four stated they would assign points to the correct answer or correct symbolic and numerical solution. One person explicitly mentioned “math” and two said they would score “logic”. In general, the teaching assistants stated they would assign the most points for the correct formulas and the correct answer, with some also giving a few points for a diagram and use of the physics principle. These aspects and weightings indicate that graduate students with teaching experience are accustomed to scoring solutions on the basis of presence or absence of particular formulas and numerical answers rather than considering the overall process taken by the solver.

Revisions to the Rubric and Training

Based on this data, both the rubric and training materials were modified. The scores were changed to include NA(Problem) and NA(Solver) more prominently in the rubric rather than as a single line description at the bottom, and the 0-4 scale was changed to 0-5. In the previous version, the zero score designated both "all missing" or "all inappropriate", and this score was split into two scores due to the graduate students' tendency to give a score of 1 for showing some work, even if it was all inappropriate. The language was also made more parallel in every category and the category "Logical Organization" was changed back to its original name "Logical Progression". The order of scoring the categories in the rubric was changed with Useful Description placed before Physics Approach, because most students begin their solution by organizing the problem information visually and/or in words. The training materials were revised to include more examples of scored solutions (five student solutions instead of three), to exclude the mechanics multi-part problem solutions, to include more NA(Solver) score examples and a wider range of score examples for most categories, and score rationales written directly on the student solution rather than in a separate table.

Summary of First Study with Training Raters

In this part of the study, eight graduate student volunteers who had experience as teaching assistants used the rubric to score eight student solutions to physics problems before and ten solutions after a minimal written training exercise. Half of the graduate students scored solutions to a mechanics problem and half of the raters scored solutions to an electricity and magnetism problem. Data collected in this study includes rubric scores and rationales for each of the problem solutions, and written comments in response to questions about the rubric and scoring task.

Perfect agreement in scores for each category of the mechanics problem ranged from 13±6% to 38±9% before training with an overall average of 28±4%. The after training agreement on this problem ranged from 25±7 to 63±8% with an average of 44±4%. Agreement within one score above or below was 74±2% before and 85±2% after training, with most of the agreement within one occurring with graduate student scores above the researcher scores. On the E&M problem perfect agreement in scores for each category before training ranged from 19±7% to 56±9% with an average of 41±4% and after training ranged from 38±8% to 50±8% with an average of 45±4%. Agreement within one score above or below was 81±2% before and 88±2% after training, with most of the agreement within one occurring with graduate student scores above the researcher scores.

The overall agreement of the raters' scores with the researchers' scores as measured by quadratic weighted kappa (Cohen, 1968) before training on the mechanics problem was 0.23±0.04 (often designated fair agreement) and improved to a weighted kappa of 0.41±0.04 (often designated moderate agreement) after a minimal written training exercise (Landis & Koch, 1977). The overall weighted kappa before training on the E&M problem was 0.31±0.05 and improved to a weighted kappa of 0.43±0.04 after training. In both cases, the probability was far from chance agreement ($p<0.001$).

The comments about the rubric indicate that the rubric category focusing on use of physics (Specific Application of Physics) is most consistent with the graduate students' existing ideas for scoring students' written work, and the Physics Approach and Useful Description are somewhat consistent with their existing ideas. The Mathematical Procedures category, as it was interpreted for this rubric, differs from the graduate students' ideas of assessment, in that it focuses on the procedures and not just the end result of calculations. Logical Progression seemed to be an aspect of judging a solution that these graduate students did not explicitly considered during grading, but may consider when scoring other features of the solution such as the organization of the formulas or clarity of students' physics reasoning. Overall, the table format seemed to be clear enough for the graduate students to follow, but comments suggest it may have been more complex than they would like.

The written comments indicated the graduate student raters were strongly influenced by their grading experiences. They expressed concerns about scoring math and logical progression when the physics is inappropriate, a desire to weight the categories based on their importance to the problem or level difficulty, and a desire to sum scores to an overall score. Graduate students also perceived substantial overlap in some categories and had difficulty treating some of the categories independently. The rubric score agreement and comments from these eight graduate student raters suggested specific changes to the rubric format and training materials that were implemented for the next round of testing.

Second Study with Training Raters

Introduction

The goals for the second study with training raters are very similar to the first study with training raters, with a modified rubric and training materials. However, both the population of raters and the setting were changed. The raters were graduate student teaching assistants near the end of their first year of graduate school. Both the training and rating processes occurred in a 50 minute classroom setting. The goals for this study were to obtain validity evidence for the rubric content (relevance and completeness) as judged by graduate students with only some experience in grading student work, to assess the adequacy of documentation in outlining the purposes of the rubric, to obtain a reliability measure of score agreement from multiple raters using the rubric with revised training materials, to obtain a controlled measure of the time required to use the rubric and training materials for the first time, and to assess the raters' perception of the rubric's usefulness.

The research questions addressed in this study are listed below, where the number and letter refer to the specific Research Question stated in Chapter 1:

- 1a) To what extent are rubric categories consistent with descriptions of physics problem solving processes? (*content relevance & representativeness*)
- 1f) To what extent does the rubric documentation address potential positive and negative consequences of the proposed test use? (*consequences*)
- 2a) To what extent do multiple raters' scores and score interpretations agree on the same problem solution? (*inter-rater agreement*)
- 2b) What scorer training is necessary to achieve a desired level of rater agreement? (*reliability and utility*)
- 3b) How authentic are the assessment's goals, tasks, and constraints? (*utility)*

Data Collection Procedures

Revisions to the rubric following the first study with training raters included the following: the Useful Description was placed before the Physics Approach because the description usually appears first on a student's paper, the zero score was separated into two scores (1=all inappropriate and 0=all missing) increasing the score range to 0-5, the language was made more consistent, and the Not Applicable scores were made more prominent in the rubric as separate columns because these scores were confusing and largely ignored by the graduate students. The training materials were modified to avoid multi-part problems, to include more examples (five instead of three solutions) and more examples of NA scores, to represent a greater score range for most categories, and to include the example scores and rationale directly on the solution rather than in separate table.

The training exercise instruction sheet and example solutions are in Appendix 5.These materials were used by 19 graduate students who participated in a rubric training activity during a 50-minute seminar for first-year graduate student teaching assistants near the end of their first year of teaching. Each TA scored 2 problem solutions (3 versions = 6 different solutions) and provided written and verbal comments about the rubric scoring task. The number of problems scored with the rubric was reduced substantially from the first study with training raters in order to fit the activity within a limited 30-minute time frame and allow discussion after the activity. The written questions included:

1. What features do you usually look for when scoring a student exam paper?
2. What difficulties did you encounter during this activity?

a. Difficulties understanding the scoring task

b. Difficulties using the scoring rubric

3. Additional comments:

Score Agreement

During the rubric scoring activity, most of the graduate student teaching assistants took approximately 20 minutes to go through the training procedure which consisted of: reading the one page of instructions, solving the problem, reading the example instructor solution, reading the rubric and category descriptions, and going through 5 example solutions scored by an expert rater. Following this minimal training, these graduate students took 10 minutes to score two solutions and write comments. In total the TAs spent 30-35 minutes participating in the rubric scoring task. The remainder of the 50 minute class was spent in small-group discussions with their peers and a brief whole-class discussion. The rubric scores assigned by each TA and the Researcher are plotted separately for each solution (F to K). The solution with the most disagreement was solution I. Table 23 summarizes the agreement of the TAs' rubric scores with two researchers' consensus scores in each category.

Figure 17: TA and Researcher Rubric Scores for Student Solution F

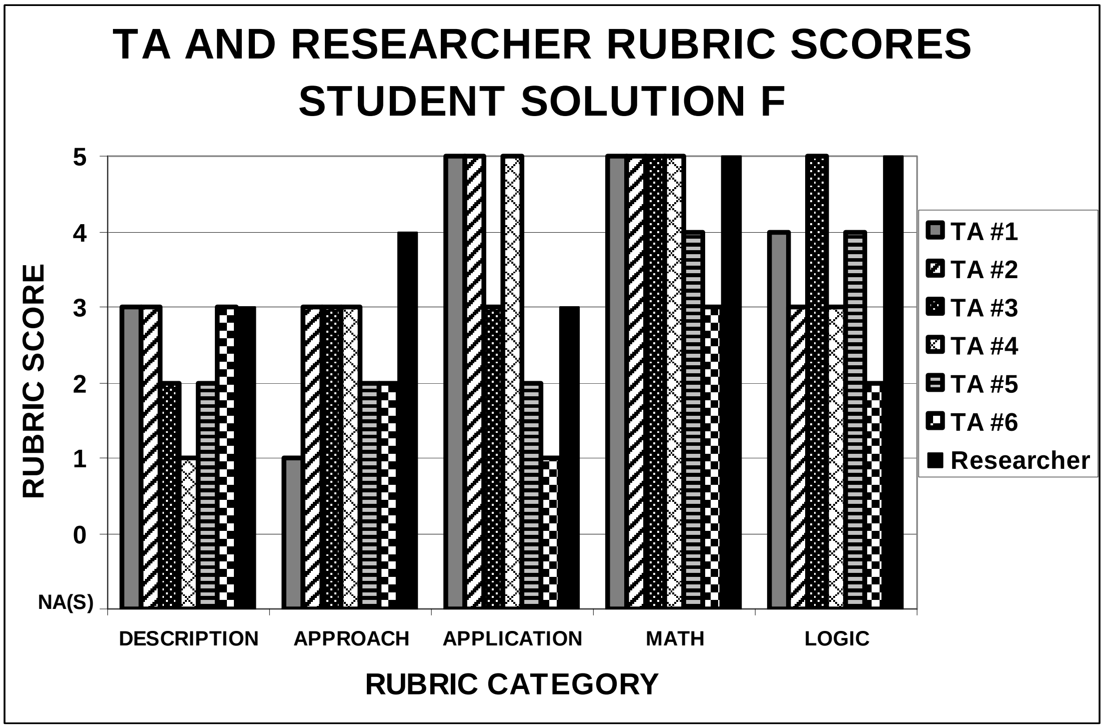


Figure 18: TA and Researcher Rubric Scores for Student Solution G

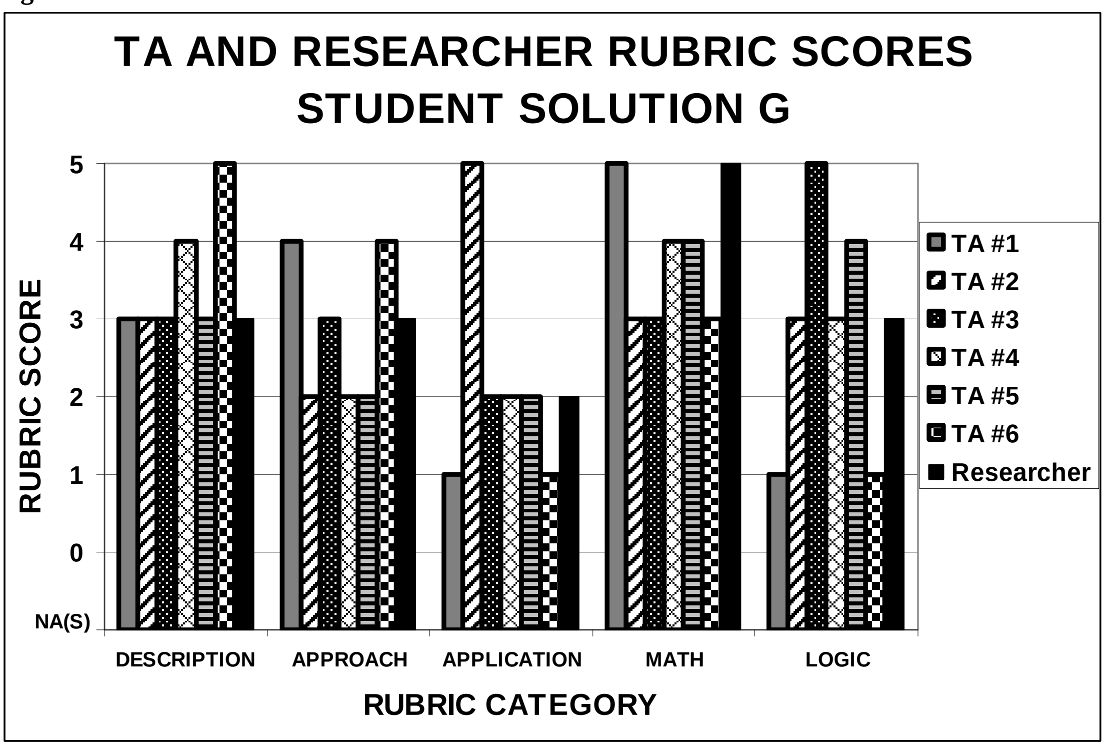


Figure 19: TA and Researcher Rubric Scores for Student Solution H

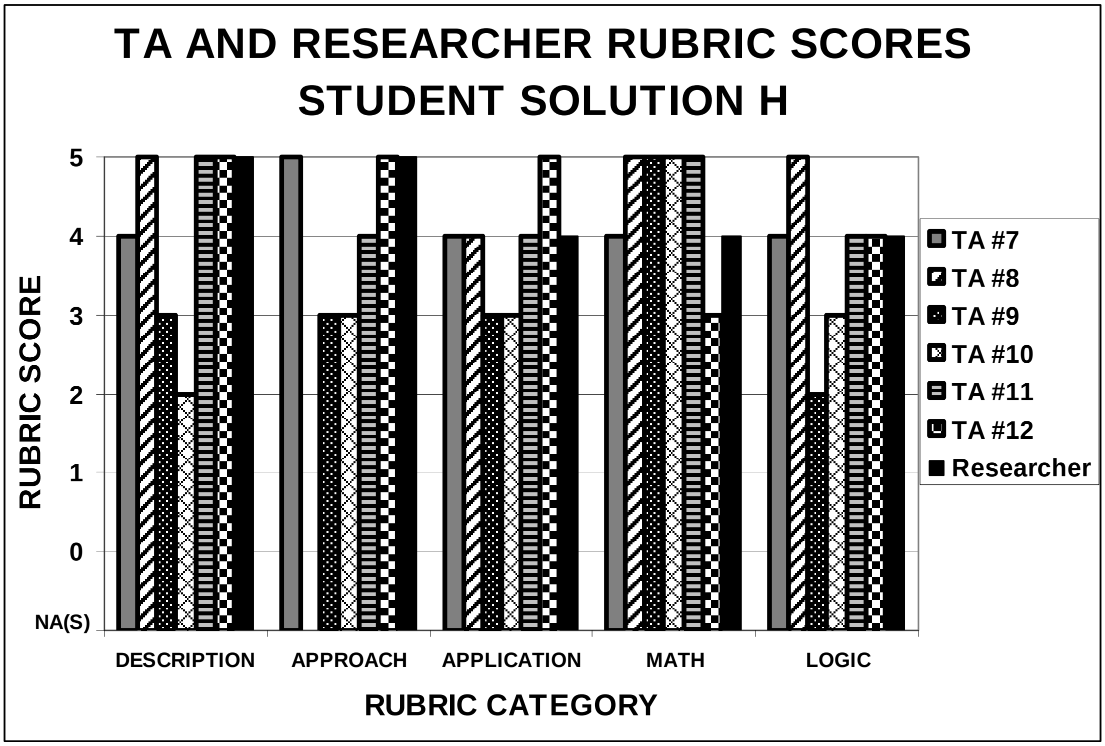


Figure 20: TA and Researcher Rubric Scores for Student Solution I

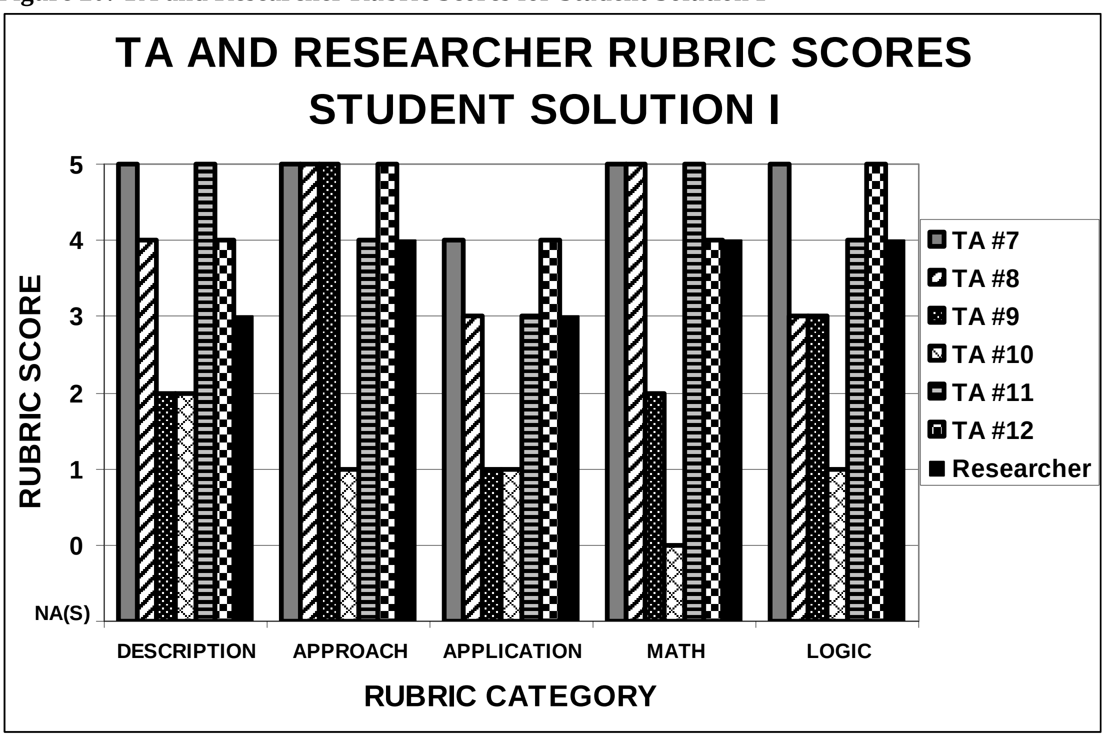


Figure 21: TA and Researcher Rubric Scores for Student Solution J

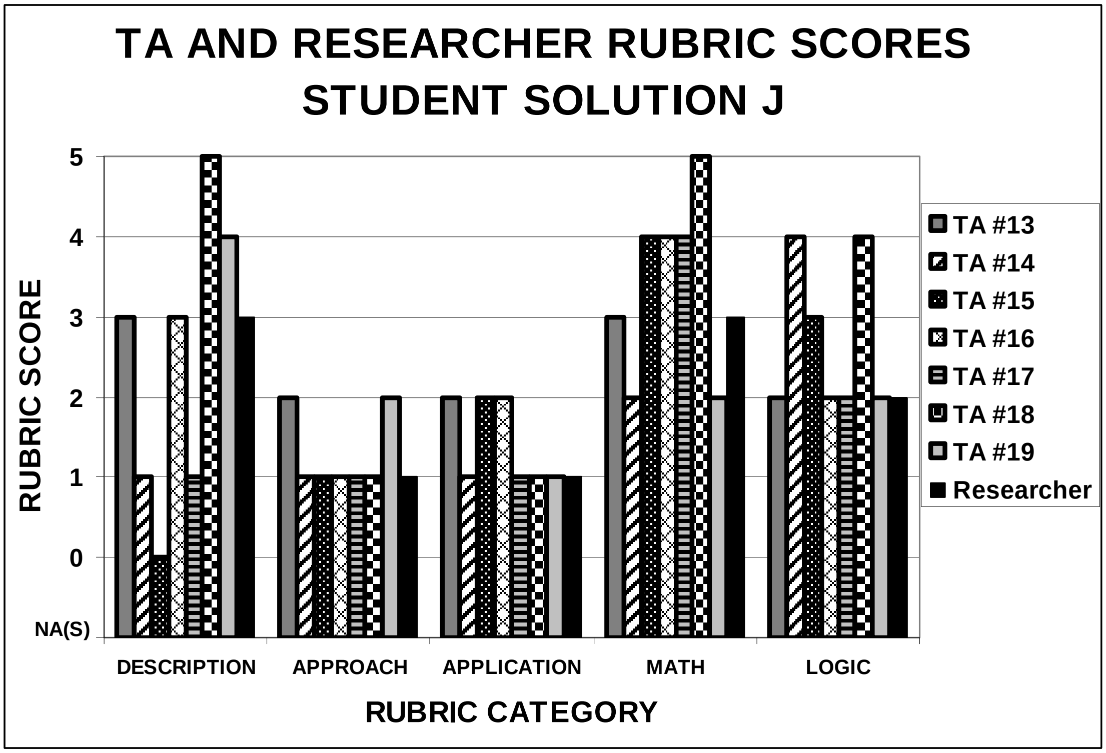


Figure 22: TA and Researcher Rubric Scores for Student Solution K

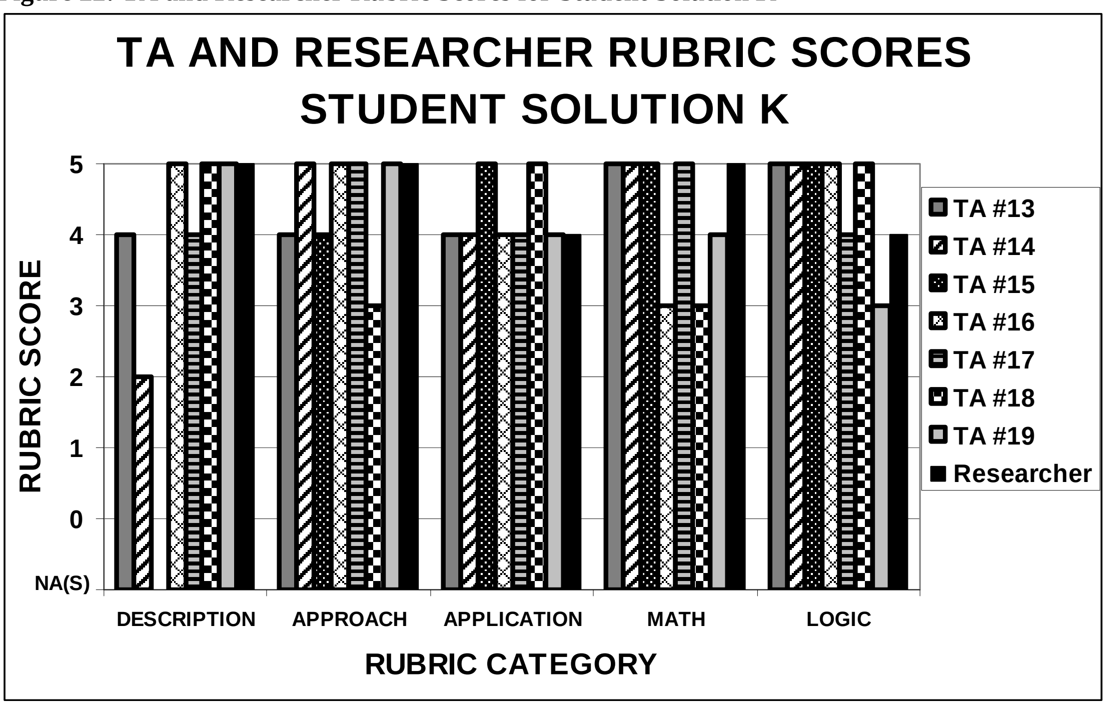


Table 23: Percent Agreement of Graduate Student Scores with Expert Raters' Scores After a Brief Written Training Activity

Category	Perfect Agreement (N = 38)	TAs One Above (N=38)	TAs One Below (N = 38)	Quadratic Weighted Kappa (N = 38)	Kappa Sig. (N = 38)
Useful Description	41±8%	13±5%	18±6%	0.24±0.12	p<0.05
Physics Approach	35±8%	21±7%	24±7%	0.40±0.08	p<0.001
Specific App.	47±8%	21±7%	13±5%	0.46±0.09	p<0.001
Math Procedures	32±8%	26±7%	18±6%	0.04±0.12	Not sig.
Logical Prog.	32±8%	26±7%	16±6%	0.17±0.11	Not sig.
Overall	37±4%	22±3%	18±3%	0.32±0.04	p<0.001

The agreement of these graduate students with the expert rater is very similar to that of the first set of graduate students even though the graduate students had less experience grading and a significantly shorter training. The category with highest agreement is Specific Application of Physics and the categories with least agreement are Math Procedures and Logical Progression. The range of agreement is narrower across the categories (math & logic are higher but the other categories were lower) than for the first training trial. The overall perfect agreement of scores ranged from 32±8% to 47±8% with an average of 37±4% and the agreement within one score (above or below) ranged from 73±7% to 82±6% with an average of 77±3%, both are significantly above chance. The quadratic weighted kappa measure of agreement ,0.32±0.04 gives a probability of chance agreement of less than 10^{-3}. In the previous rubric training activity, in which eight graduate students were allowed unlimited time for a higher number of scoring tasks, the kappa value was 0.27±0.02 before training and 0.42±0.03

after training, the overall percent of perfect agreement was 44±2% after training and agreement with one was 85±2% after training.

In general, the agreement for the second study with training raters was higher than the "Before Training" situation from the first study, but not as high as the "After Training" of the first study. It is possible that the longer time available to complete a scoring task (one week compared to 35 minutes) and a higher number of solutions scored resulted in a more complete training experience. However, another difference was that the first set of graduate students had more experience in graduate school than the second set.

The lower kappa scores for math procedures and logical progression despite percent agreement values that are not too different from the other averages indicates a known difficulty of the weighted kappa measure and is explained in more detail in Appendix 2. In the case of the math category, most of the student solutions that were scored had few math errors (scored 4 or 5) and as a result the "expected" score frequencies were very similar to the "observed" frequencies in the kappa calculation (see Appendix 2), giving an agreement level that is not significantly different from chance using its technique of estimating chance. Because kappa attempts to correct for rater biases in scoring, it does not give meaningful results for narrow distributions of rater scores.

Comments from Raters

More score examples of NA(Solver) in the training materials helped clear up some confusion about these scores. There were still some comments about the NA scores, but in the context of trying to use these scores in grading the problem: "For N/A situation, how do we compare with other students' answers? Should N/A be counted as 5 points?" (TA#12) or "NA→ does this mean they are not docked points?" (TA#9) TAs generally did not express difficulties understanding the training materials or scoring task itself, or a desire to see more score examples. Most comments focused on the content of the rubric itself and its application to student solutions.

Consistent with previous results, these graduate students were influenced by their need to relate the assessment to grading the students even though the instructions did not address grading. In particular, they expressed dislike for assigning separate scores for math and logic when the physics is incorrect, "Seems to give too many points for correctly and logically solving a problem with the completely wrong setup and understanding" (TA #5). Four of the 19 TAs expressed difficulty distinguishing physics approach and application, "Determination between what is physics approach and specific application of physics is not always clear" (TA #2), and "Sometimes hard to differentiate between physics approach and specific application" (TA #12). However, some TAs also expressed that having two categories for physics gave more "weight" to this aspect of the solution which was good. There were essentially no comments about Useful Description, indicating this category was relatively apparent to TAs.

During a small group and whole class discussion which occurred after the rating task, the teaching assistants were asked to state what features they usually look for while grading a student solution. One group said they try to 'see if it looks like the student knows what they're doing'. When asked what specifically they look for on a paper, the group of TAs did not have a response. Other groups stated that they look for the overall pattern of the solution, the way the solution is organized, setting up the problem, important equations, and the final answer. Next, the TAs were asked what features they noticed while using the rubric to score solutions. They listed scoring logic as something unique to the rubric even though they had previously stated that this was one of the features they looked for when grading. They also stated that treating math as an independent measure from physics was unique to the rubric.

Despite this class discussion that explicitly contrasted the rubric with traditional grading practices, the purpose of the rubric scoring scheme was unclear to some teaching assistants. Four of the 19 TAs expressed the rubric scores were “not useful” or “of little use”. They did not see any value in an assessment tool that was not explicitly used for grading in the traditional sense.

Summary of Second Study with Training Raters

Overall, the written rubric training activity that took place in a limited time period (30-35 minutes) produced similar results to earlier training in which graduate teaching assistants were given a week or more to complete tasks. The increased number of examples and a printed rationale directly on the student solutions improved TAs’ interpretation of scores, especially the NA(Solver) score. Comments from TAs indicate they felt the rubric format was too wordy, the rubric inflates scores when the physics is incorrect, and some people experienced difficulty distinguishing a physics approach from the application of physics. Questions from TAs regarding the usefulness of the rubric indicated a description of potential uses of the rubric should be included in the documentation materials with explicit contrast to other, traditional scoring practices.

Scoring Written Solutions on Exams

Introduction

The goals for this part of the study include: obtaining validity evidence for the response processes of students on written solutions to physics test problems, obtaining validity evidence for the rubric's generalizability (applicability to multiple topics in a semester-long mechanics course), obtaining validity evidence of the external and internal rubric structure, and measuring the consistency of a single rater over time. The information in this part of the study will help delineate a possible set of uses of the rubric for an instructor.

The research questions addressed in this study are listed below, where the number and letter refer to the specific Research Question stated in Chapter 1:

- 1b) To what extent do scores on the rubric reflect the problem-solving processes undertaken by a solver? (*response processes*)
- 1c) To what extent do scores on the rubric support inferences about students' problem-solving skills from other measures of their performance? (*external structure*)
- 1d) To what extent are the rubric categories independent? (*internal structure*)
- 1e) To what extent is the rubric applicable to multiple populations and contexts, including different student populations, physics topics, and problem features? (*generalizability*)
- 3b) How authentic are the assessment's goals, tasks, and constraints?

- 3c) To what extent is the assessment independent of the specific format in which students are taught to express their solutions?

Data Collection Procedures

Copies of test papers were collected from two sections of a semester-long introductory calculus-based physics course for science and engineering (mechanics) that had the same lecture instructor. The tests during the term each required a free-response solution to two problems. The problems had a very traditional format similar to the ones in the textbook used for the course. One section of the course had an enrollment of 230 students and the other had an enrollment of 250 students. The teaching assistants made copies of students' papers after they had been graded and before they were returned to students. For some tests the course instructor gave the same problems to both sections, and for some tests the problems differed (on test 2 the problems were different).The tests represented standard physics topics including motion with constant acceleration in one and two dimensions, Newton's second and third laws, rotational motion, and conservation of energy. Available problem solutions from the first three tests were scored using the rubric and compared to the scores assigned by graders.

Rubric and Grader Scores

The following subsections report a problem-by-problem analysis of scores for test problems scored with the rubric in two sections of the course. These subsections also report how course TAs graded each problem and the relationship of grades to rubric scores for low, middle, and high-scoring groups of students (as measured by the grades). For each problem, common student difficulties are identified and a description of how rubric scores responded to each difficulty is provided.

For each paper, a total rubric score was computed by taking the sum of the category scores divided by the number of categories with a numeric score. In this way, a percent score was calculated in a way that excluded NA(Problem) and NA(Solver) scores for that paper. Agreement for each category score and this “total” rubric score with the grader’s score for that same paper are reported in the correlation tables. Adding together all of the rubric scores was done to obtain a convenient single score to compare with the grade and does not indicate that this procedure is either mathematically justifiable or educationally desirable.

Rubric and Grader Scores on Test 1

The first test focused on motion under constant acceleration in one and two dimensions, or what is referred to as “kinematics”. The first problem on this test was problematic for several reasons (see Appendix 6) and was excluded from the study. For example, approximately half of student papers were scored 100% by both the grader and the rubric, indicating the problem was perceived to be an algorithmic “exercise” for students and not a problem. The second problem on the first test focused on two-dimensional constant acceleration kinematics. The problem statement is copied below.

Test 1 Problem 2:

A punter kicks a football during a critical football game. The ball leaves his foot at ground level with velocity 20.0 m/s at an angle 40^o to the horizontal. At the very top of its flight, the ball hits a pigeon. The ball and the pigeon each stop immediately and both fall vertically straight to the ground from the point of collision.

(a) With what speed is the ball moving when it hits the pigeon? [10 points]
(b) How high was the ball when it hit the pigeon? [10 points]
(c) What is the speed of the ball when it hits the ground? [5 points]

The frequency of rubric scores for this problem in the first section (N=65 solutions) and the second section (N= 163) are presented in Figure 23 and Figure 24 below. Numbers are smaller than the class populations because the TAs did not always remember or take the time to make copies before returning the graded solutions to students. This is especially true in section 1 that had an 8:00 a.m. recitation time. The spike in NA(Solver) for Useful Description indicates that several students did not write any description of the problem (visually or in words) but that it didn't substantially affect their capacity to reach a solution. The frequencies for score 5 provide some indication that students experienced more difficulty with categories Specific Application and Logical Progression than the categories Math and Physics Approach.

A Useful Description in this problem meant that students assigned appropriate symbols for quantities in the problem, designating them with subscripts as necessary, and included a picture with vector symbols. A Physics Approach meant that students used the concept of motion with constant acceleration and treated the horizontal and vertical directions independently. Specific Application scored the students' ability to match quantities in the problem with an appropriate kinematics equation, and their use of a coordinate system with appropriate positive and negative signs. Logical Progression meant that the solution was clear, coherent, and consistent, independent of whether the student solved for sub-quantities (like time) or solved the problem more directly.

Figure 23: Frequency of Rubric Scores for Test 1 Problem 2 (Section 1)

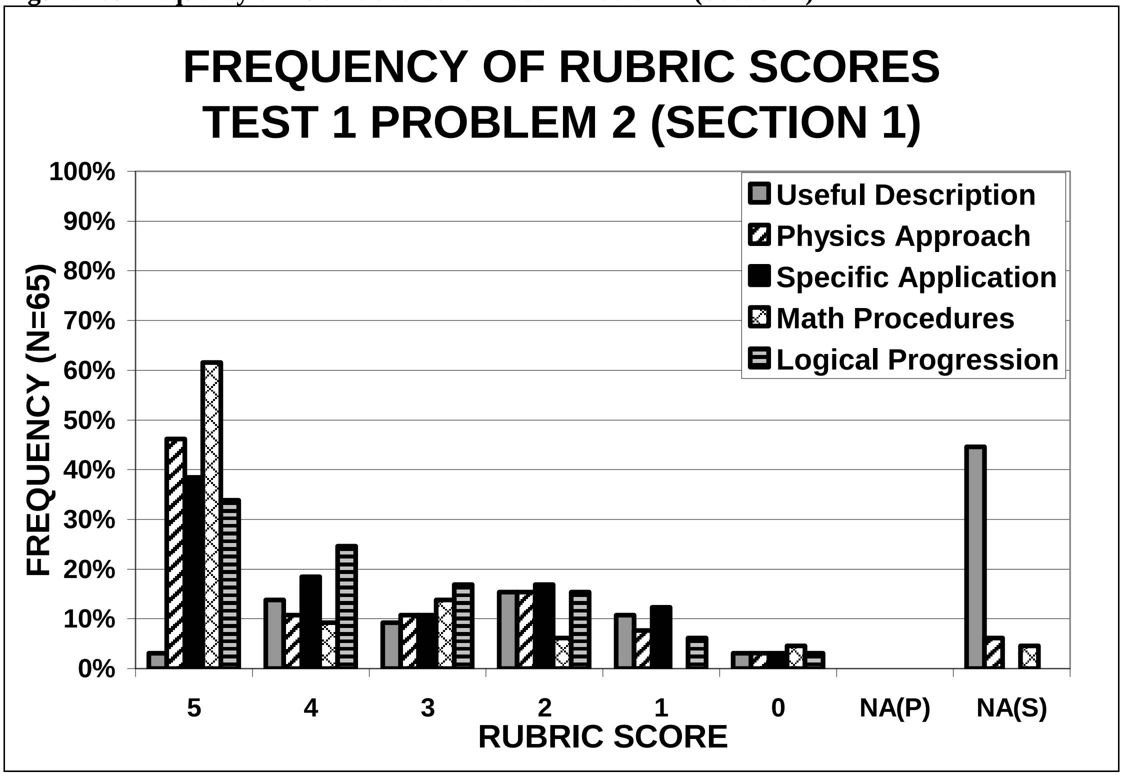


Figure 24: Frequency of Rubric Scores for Test 1 Problem 2 (Section 2)

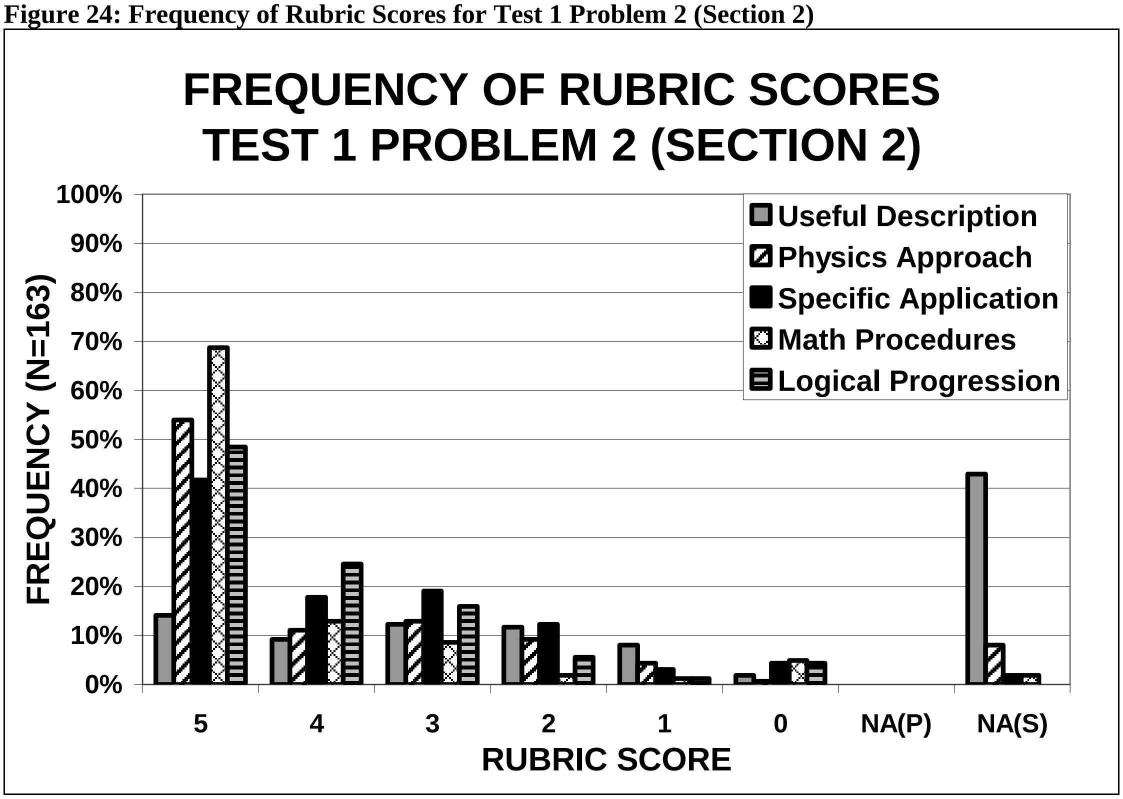


Figure 25: Frequency of Problem Grades for Test 1 Problem 2 (Section 1)

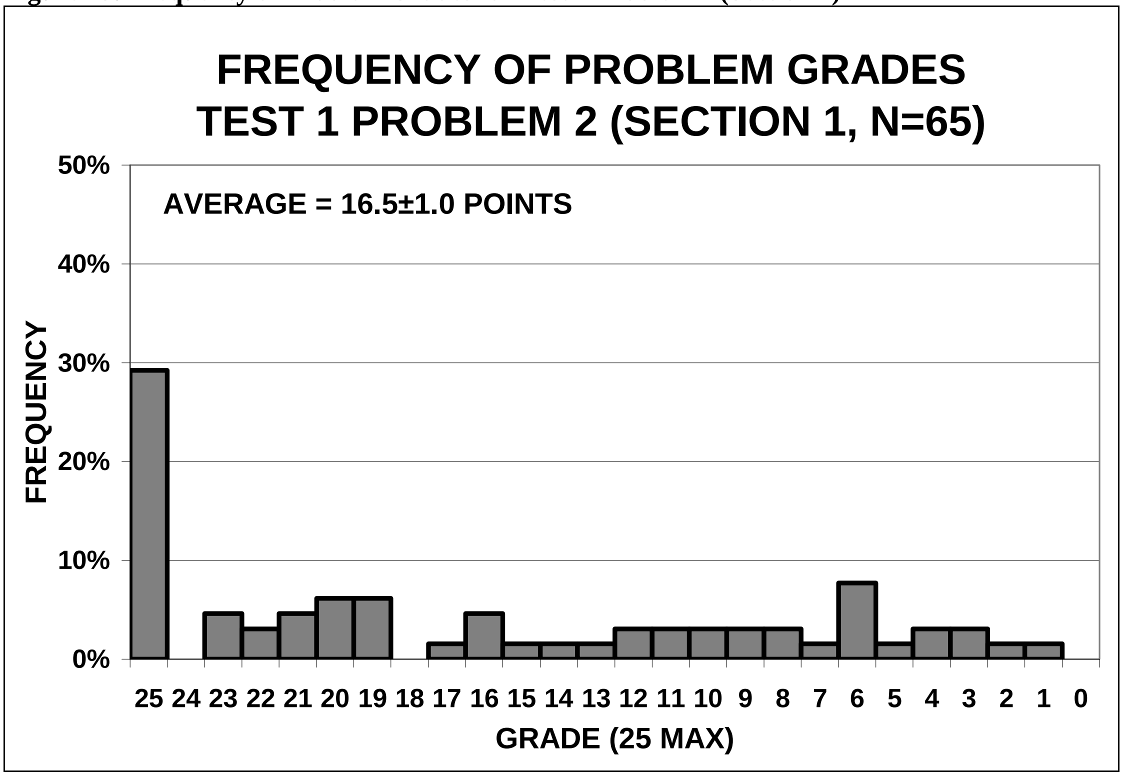
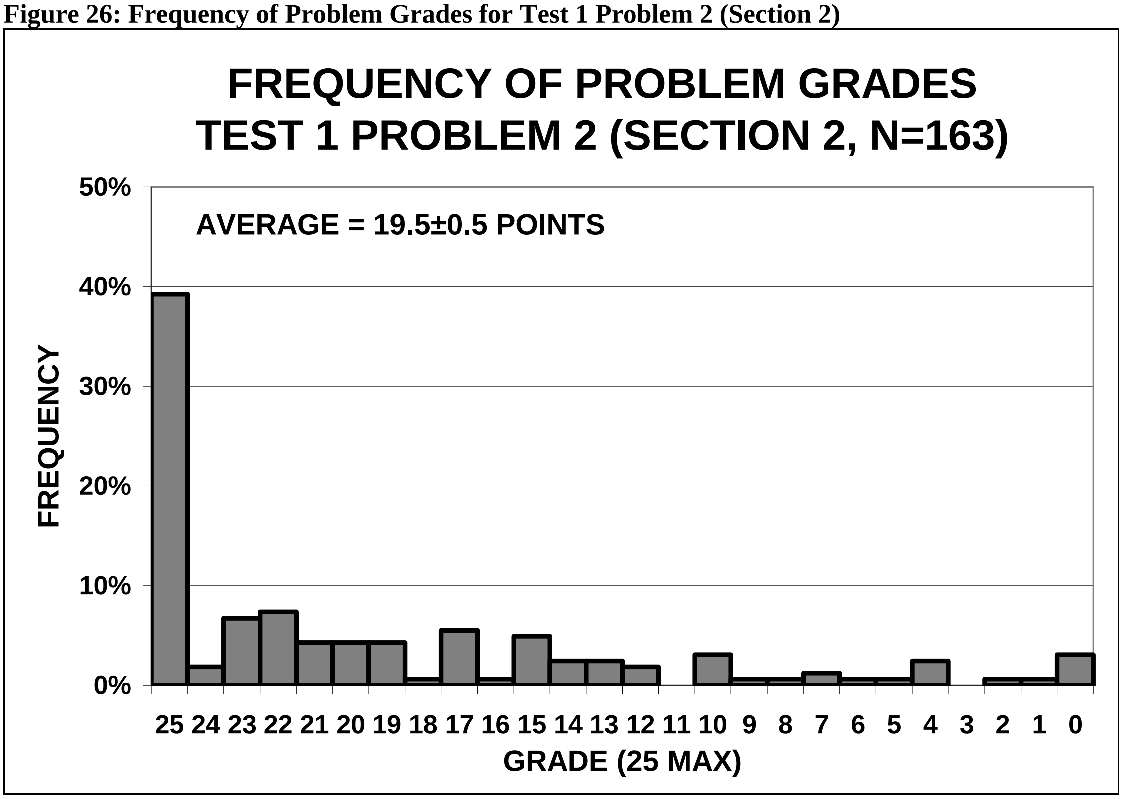


Figure 26: Frequency of Problem Grades for Test 1 Problem 2 (Section 2)



These frequency charts reflect a similar pattern of scores for sections 1 and 2, with a lower fraction of students scoring high (5) in section 1. The average rubric and problem grades are consistent with this observation of higher scores for section 2. The distributions for both sections reveal two distinct populations of students. After excluding students who scored 100% or 0%, the average score assigned by the grader in section 1 of the course was 54±4% and the average score assigned by the grader in section 2 was 67±3%. The average total rubric score for these sections were 65±4% and 70±2%, respectively. Rubric category averages are calculated by taking the average score (from zero to five) and dividing it by the maximum score of 5. The total Rubric Score represents a sum of the category scores. This is done to make a direct comparison with the overall grader score from the TA. As seen in the following table, section 2 scored higher on both the problem grade and the rubric measures, with the largest differences observed in the Physics Approach and Logical Progression categories.

Table 24: Average Rubric Scores and Problem Grades for Test 1 Problem 2

	Averages Section 1 (N=48)	**Averages Section 2 (N=110)**
Useful Description	49±4%	54±3%
Physics Approach	64±5%	73±3%
Specific Application	58±4%	62±3%
Math Procedures	78±4%	80±3%
Logical Progression	61±4%	71±3%
Rubric Score	65±4%	70±2%
Problem Grade	54±4%	67±3%

A detailed analysis of the solutions showed that for students who did experience difficulty with this problem, a common error in their Physics Approach included failing to treat the horizontal and vertical directions independently (resulting in an approach score of 2 or 3 and a problem grade of 12 or lower out of 25). Common errors related to the Specific Application of Physics included failing to recognize that the vertical velocity at the ball's peak height is zero and the horizontal velocity is nonzero (rubric score of 3 and problem grade of 0 out of 5 for part a), using the wrong sign for the gravitational acceleration in their coordinate system (minor error scored a 4 on the rubric or problem grade 8 out of 10 on part b), or confusing some quantities in a kinematics equation such as using a distance instead of a velocity (rubric application score of 2 and problem grade of 2 out of 10 for part a or b). Other students had difficulty with their Logical Progression such as solving for a quantity other than the problem target (logic score of 2 or 3 and problem grade of 10 or lower out of 25). A combination of errors could result in a lower score for that category, such as a 1 for most parts missing and/or contain errors.

Figure 27 and Figure 28 show scatterplots of total rubric scores versus the problem grades on Test 1 Problem 2. Points are shifted by a small random number so that clusters of scores in the distribution will not be masked by being included as a single point [shifted score=score - score*0.05*RAND()]. The Pearson correlation coefficient for this distribution is R=0.93 and R=0.92 showing that the total rubric score accounts for 88% and 84% of the grader score variance in sections 1 and 2, respectively.

Figure 27: Scatterplot of Rubric Scores vs. Problem Grades for Test 1 Problem 2 (Section 1)

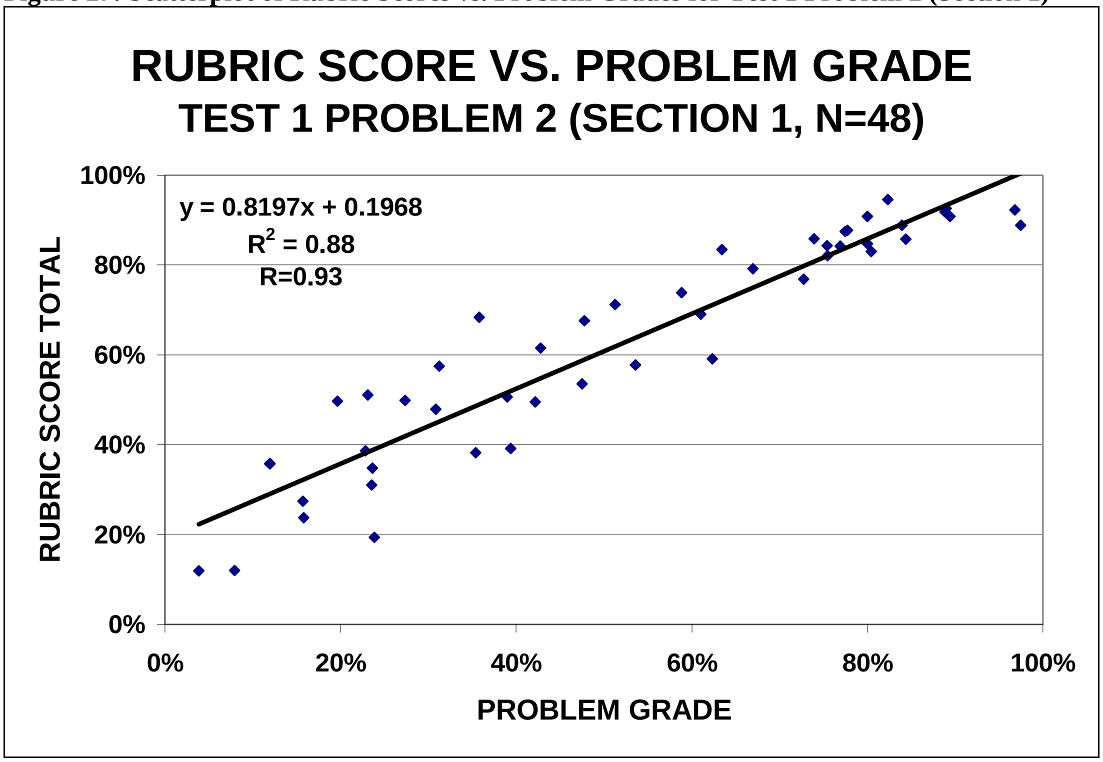
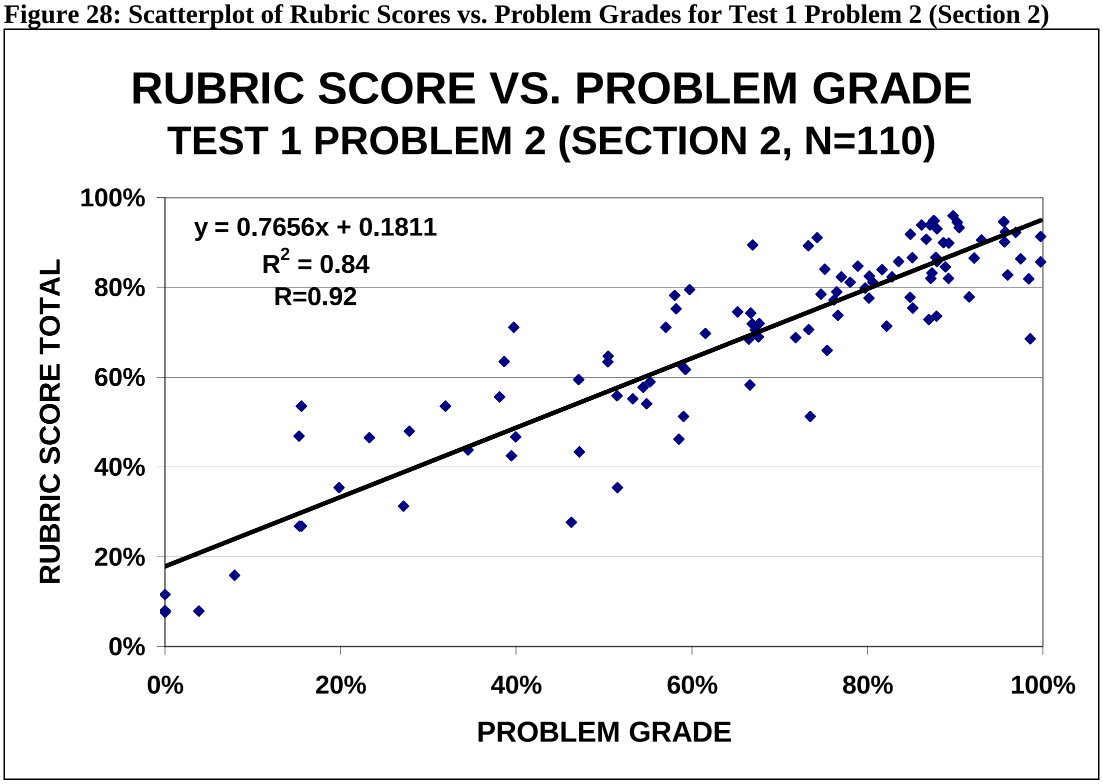


Figure 28: Scatterplot of Rubric Scores vs. Problem Grades for Test 1 Problem 2 (Section 2)



The correlation between the grader's scores and rubric scores for this problem in section 1 are reported in Table 25 and plotted in Figure 29. Correlations are reported for all students (overall) and separately for the bottom third, middle third, and top third of the distribution of grader's scores. The correlations for section 2 are reported in Table 26 and Figure 30. Since the designations of bottom, middle, and top third were determined by the distribution of grader scores in each section, they have different score ranges (the middle range is wider for section 1). This could be a result of the small sample for section 1.

Table 25: Correlations of Rubric Scores with Grader Scores for Test 1 Problem 2 (Section 1)

	Overall (N=48)	**Bottom Third (0-36%) (N=17)**	**Middle Third (37-76%) (N=17)**	**Top Third (77-100%) (N=14)**
Useful Description	0.51	0.20	0.42	NA
Physics Approach	0.90	0.37	0.88	0.21
Specific App.	0.94	0.75	0.83	0.36
Math Procedures	0.69	0.71	0.47	-0.10
Logical Progression	0.82	0.61	0.61	0.48
All Categories	0.94	0.75	0.86	0.62

Table 26: Correlations of Rubric Scores with Grader Scores for Test 1 Problem 2 (Section 2)

	Overall (N=110)	**Bottom Third (0-56%) (N=33)**	**Middle Third (57-84%) (N=40)**	**Top Third (85-100%) (N=37)**
Useful Description	0.56	0.31	0.67	0.23
Physics Approach	0.90	0.80	0.45	0.10
Specific App.	0.86	0.81	0.04	0.34
Math Procedures	0.71	0.66	0.18	0.18
Logical Progression	0.84	0.80	0.30	-0.27
All Categories	0.92	0.81	0.55	0.16

Figure 29: Correlations of Rubric Scores with Grader Scores for Test 1 Problem 2 (Section 1)

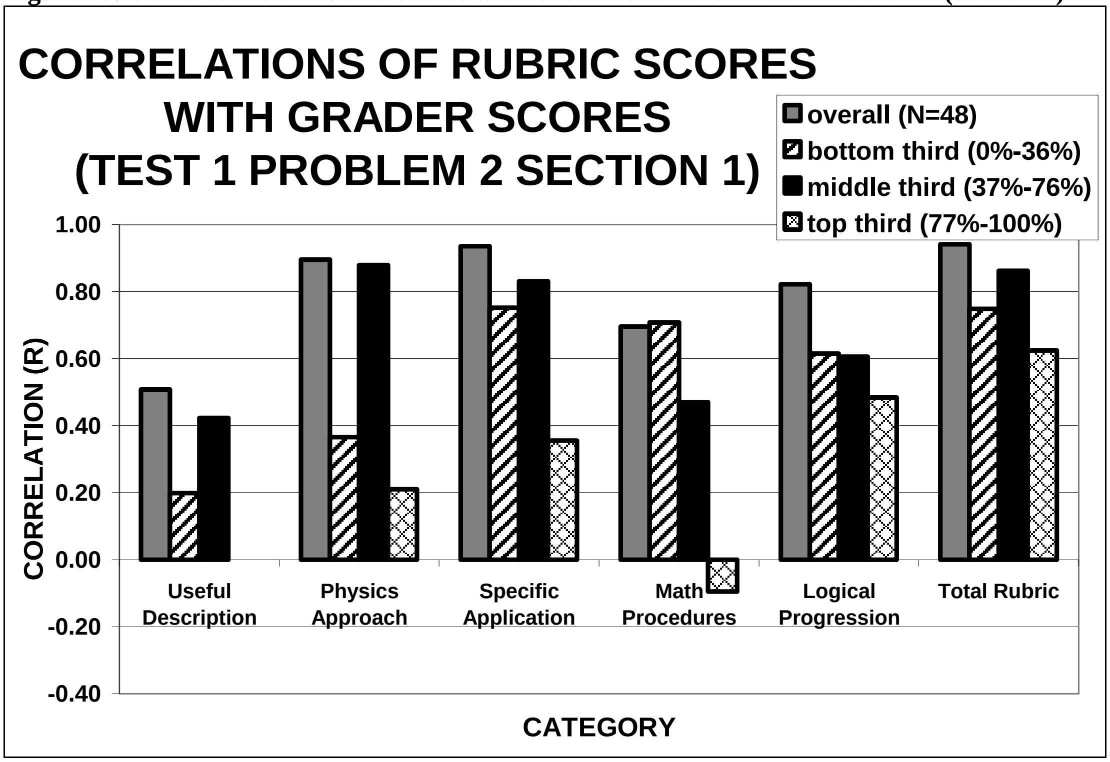
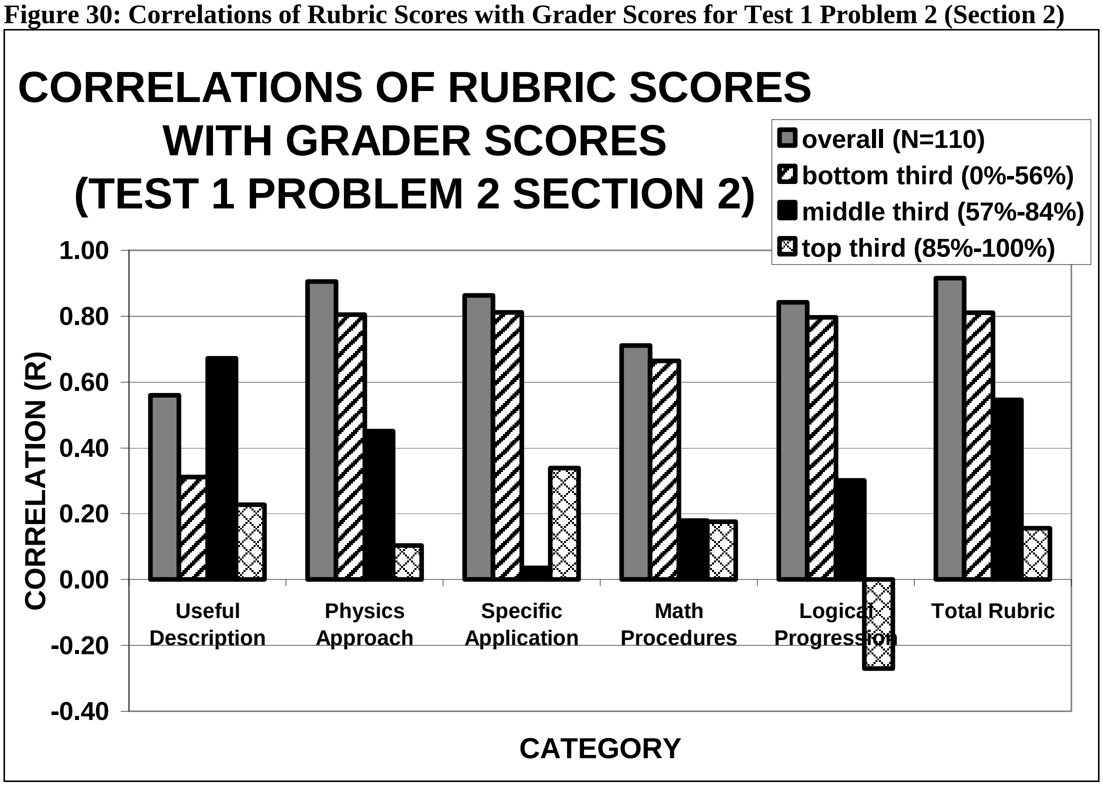


Figure 30: Correlations of Rubric Scores with Grader Scores for Test 1 Problem 2 (Section 2)



The correlation plots in Figure 29 and Figure 30 indicate differences in the relationships between rubric scores and grades for each category, and how these correlations compare for different groups of students (bottom, middle, and top third) and compare across the two course sections. The “overall” correlations of the rubric scores and problem grades indicated that in both sections the Useful Description and Math Procedures category scores had a lower relationship to the problem grades than the Physics Approach, Specific Application of Physics, and Logical Progression scores. This suggests rubric scores for description and math aspects of a physics solution are less consistent with the grading practices of TAs than other aspects, such as physics and logical progression of the solution. The specific grading criteria used was not known, so it is uncertain whether those aspects were not used in the grading at all, were given a lower weight in the overall solution, or if the grading was inconsistent.

In section 1, the middle third group exhibited a stronger correlation between the rubric scores and grader scores than the other groups for the Useful Description, Physics Approach, and Specific Application of Physics categories. The top third group exhibited a negative correlation between Mathematical Procedures rubric scores and the grader scores that was not observed for other categories or the other section (A negative correlation indicates that an increase in the grader score for the solutions corresponded to a decrease in the rubric Math score for this group). In section 2, the bottom third group typically exhibited a stronger correlation between the rubric scores and grader scores than the middle and top groups for all categories except for Useful Description, and the top third group had a negative correlation between the Logical Progression category scores and the grader scores.

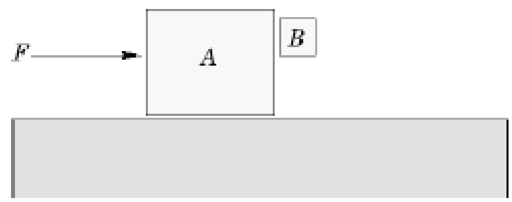
The relationship between the "Total Rubric" score and grader score for each group (bottom, middle, and top third) exhibits a different pattern in each section. In section 1, the "Total Rubric" correlations are above 0.60 for every group. In section 2, however, the "Total Rubric" correlations are highest for the bottom third (R=0.81), lower for the middle third (R=0.55), and even lower for the top third (R=0.16). In addition, there are noticeable section differences in the Physics Approach bottom and middle groups, Math Procedures middle and top groups, and Logical Progression middle and top groups (section 2 had lower correlation values in each of these instances). These differences between sections make it difficult to generalize grader and rubric relationships on this problem.

Rubric and Grader Scores on Test 2

The second test dealt with forces and Newton's Second and Third Laws of Motion. One problem focused on circular motion. Since each lecture section had different test problems, there were four problems scored with the rubric for Test 2. The first problem for one lecture section is stated below.

Test 2 Problem 1 (Section 1)

The mass of block *A* is 75kg and the mass of block *B* is 15kg. The coefficient of static friction between the two blocks is $\mu = 0.45$. The horizontal surface is frictionless. What minimum force *F* must be exerted on block *A* in order to prevent block *B* from falling?



As seen in the rubric score frequency plot for this problem (N=94), students scored lower on this problem, with several students scoring 2 or 1 for the categories Specific Application of Physics and Logical Progression. In this problem, a Useful Description meant that the student represented the forces acting on each block with a picture, free-body diagram, and/or a verbal statement. Physics Approach meant the student used Newton's Third Law to consider the contact force between block A on block B, and Newton's Second Law independently in each direction. Specific Application of Physics measured the specific identification of forces acting on each object and its relationship to the appropriate acceleration, and Logical Progression meant the solution was coherent, consistent, and progressed to an answer for the appropriate quantity.

Figure 31: Frequency of Rubric Scores for Test 2 Problem 1 (Section 1)

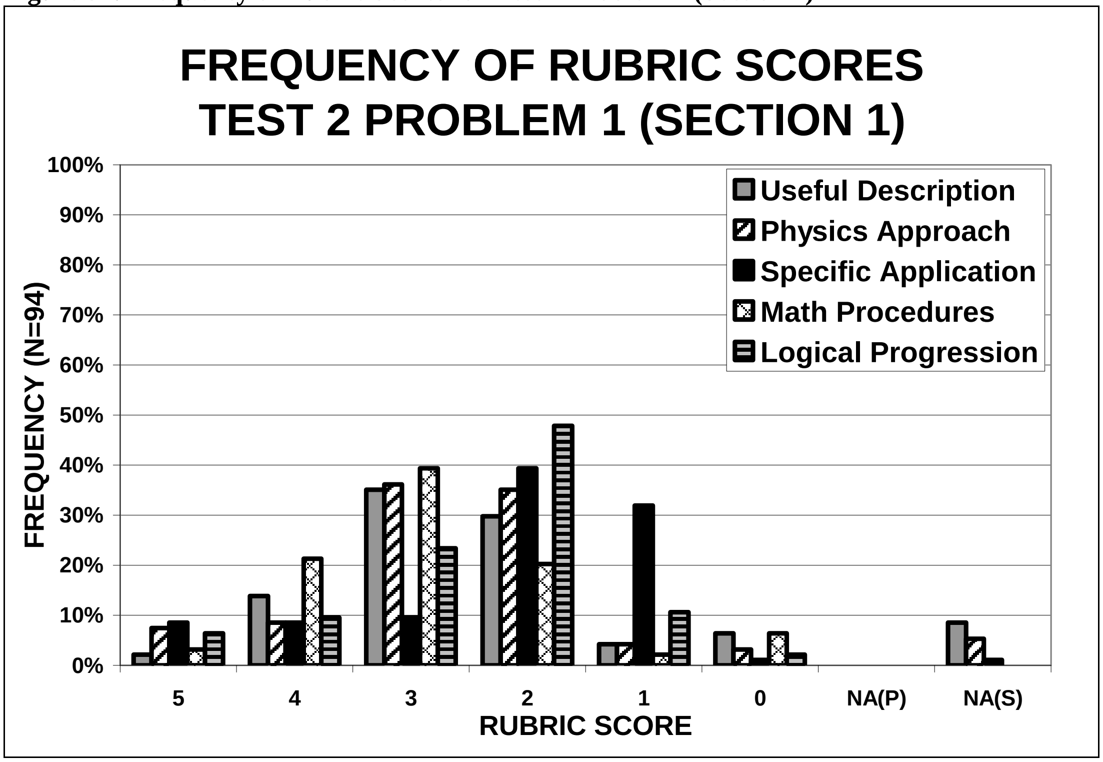
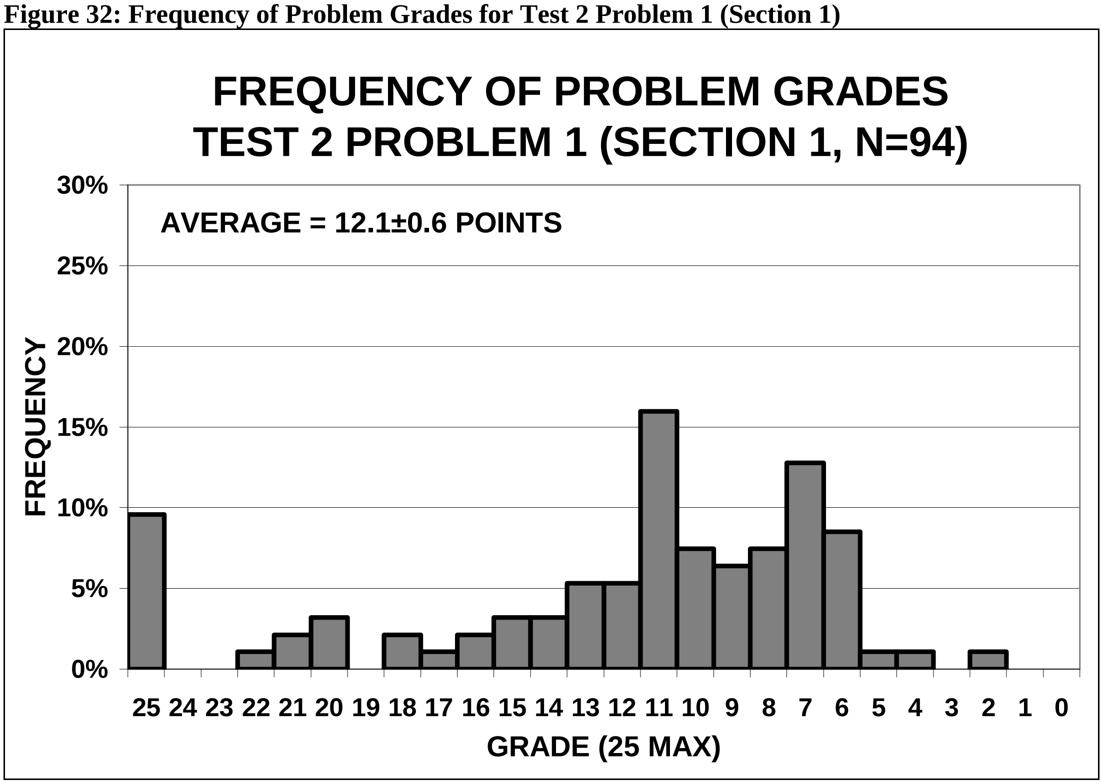


Figure 32: Frequency of Problem Grades for Test 2 Problem 1 (Section 1)



The frequency chart indicates the distribution of rubric scores peaked around 2 or 3 points in all categories, or around 50%. As seen in the table below, the average grade for this problem was 47±2% and the average rubric score was 50±2%. The lowest scoring category was Specific Application of Physics and the highest scoring was Math Procedures.

Table 27: Average Rubric and Grader Scores for Test 2 Problem 1 (Section 1)

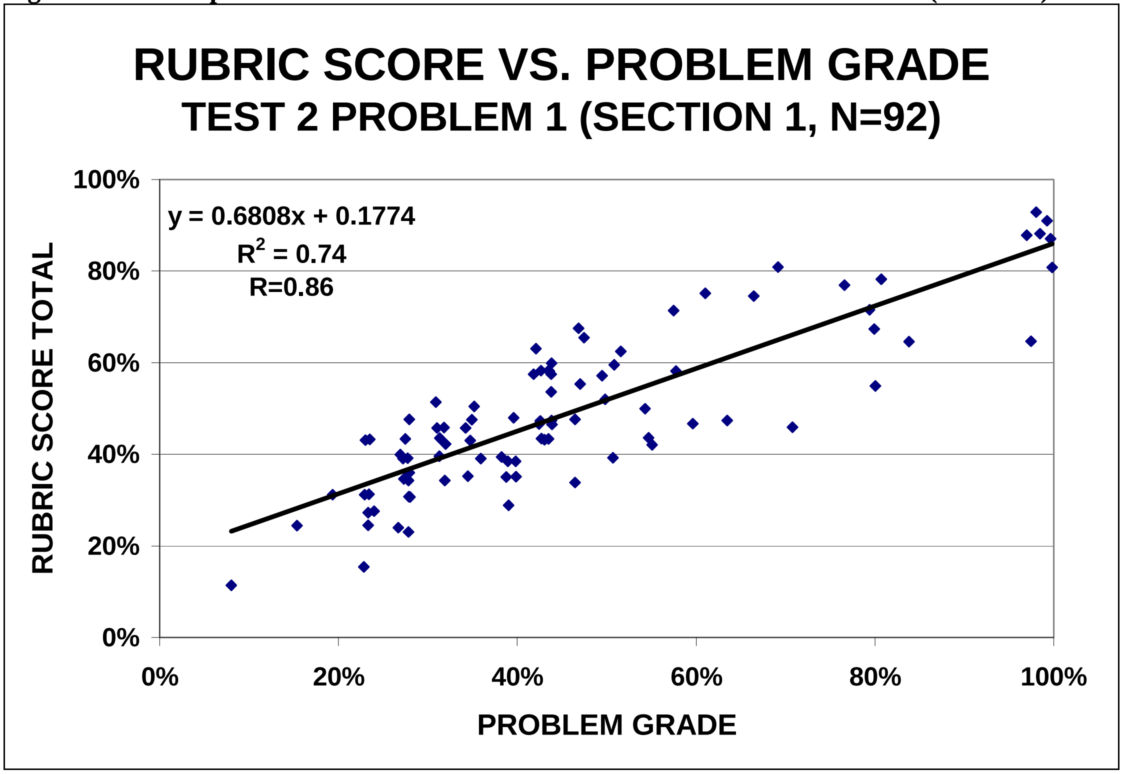
	Averages (N=92)
Useful Description	51±2%
Physics Approach	53±2%
Specific Application	42±3%
Math Procedures	56±2%
Logical Progression	48±2%
Rubric Score	50±2%
Grader Score	47±2%

Only 14% of the student papers collected had a completely correct solution. A closer look at students' solutions revealed that common errors resulted from an incorrect identification of forces acting on each block or inappropriate directions for the forces (both Specific Application of Physics errors scored 2 or 3 and a problem grade of between 9 and 12 points out of 25). For example, several students did not consider the contact force of block B on block A, were missing the friction force acting on block A, drew an incorrect direction of the normal force on block B, or drew the friction force between the two blocks horizontally instead of vertically. Some also didn't distinguish symbols for the two block masses (just used "m") which was an error in the Useful Description category. Some papers considered that any normal force was "mg", which resulted in a lower score for Specific Application of Physics.

The most common error was to assume that the blocks were not accelerating, or that the applied force F was equal to the normal (contact) force between blocks A and B a (29% of papers). Depending on precisely how this was written on the paper, it could be interpreted as a conceptual error for Newton's Second Law (Physics Approach error) or an error in the Specific Application of Physics (explicitly setting acceleration to zero). This typically resulted in a problem grade of between 8 and 12 points. Another common error (16% of student papers) was a failure to treat forces in the horizontal and vertical directions independently, such as stating that the applied force F was equal to the difference of the surface friction and block B's weight. Consistent with the same error in Test 1, this was a Physics Approach error that resulted in a score of 2 or 3 for that category and a problem grade between 7 and 9 points out of 25.

Figure 33 shows a scatterplot of total rubric scores versus the problem grades on Test 2 Problem 1 in section 1. Points are shifted by a small random number so that clusters of scores in the distribution will not be masked by being included as a single point [shifted score=score - score*0.05*RAND()]. The Pearson correlation coefficient for this distribution is R=0.86 showing that the total rubric score accounts for 74% of the grader score variance.

Figure 33: Scatterplot of Rubric Scores vs. Problem Grades for Test 2 Problem 1 (Section 1)

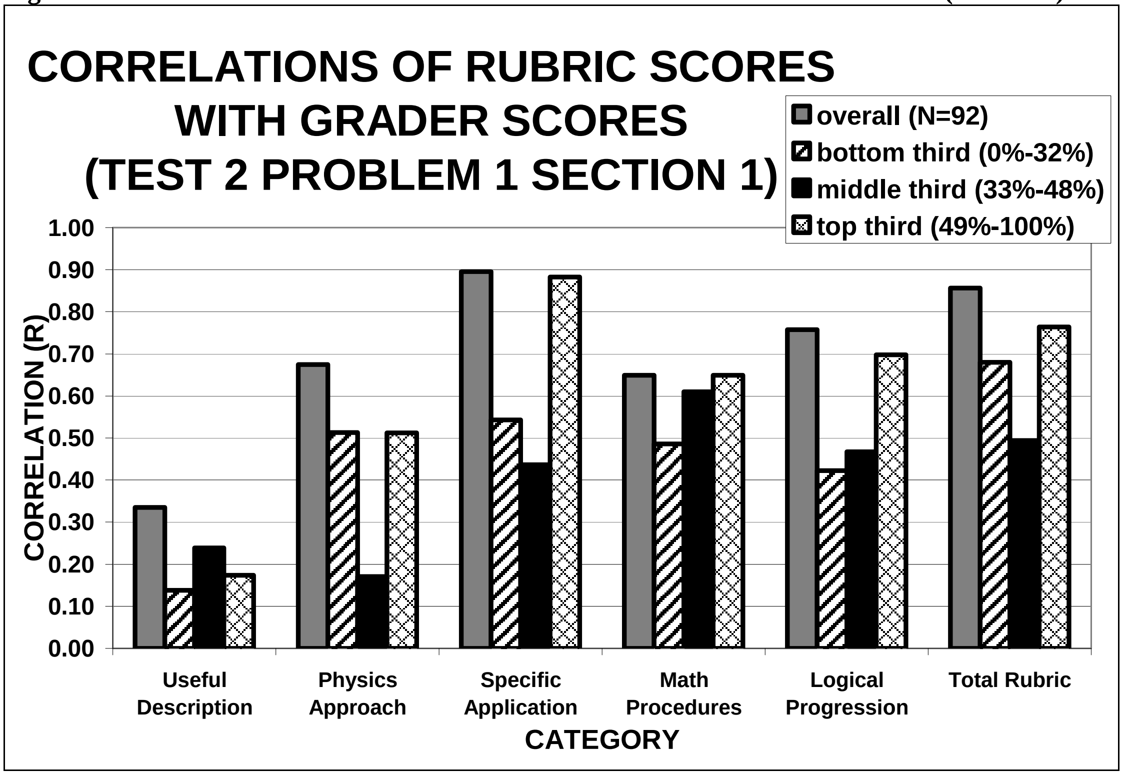


The correlation between the grader's scores and rubric scores for this problem in section 2 are reported in Table 28 and plotted in Figure 34. Correlations are reported for all students (overall) and separately for the bottom third, middle third, and top third of the distribution of grader's scores.

Table 28: Correlations of Rubric Scores with Grader Scores for Test 2 Problem 1 (Section 1)

	Overall (N=92)	Bottom Third (0-32%) (N=30)	Middle Third (33-48%) (N=33)	Top Third (49-100%) (N=29)
Useful Description	0.34	0.14	0.24	0.17
Physics Approach	0.67	0.51	0.17	0.51
Specific App.	0.90	0.54	0.44	0.88
Math Procedures	0.65	0.49	0.61	0.65
Logical Progression	0.76	0.42	0.47	0.70
All Categories	0.86	0.68	0.49	0.76

Figure 34: Correlations of Rubric Scores with Grader Scores for Test 2 Problem 1 (Section 1)



The correlation plot in Figure 34 indicates differences in the relationships between rubric scores and grader scores for each category, and how these correlations compare for different groups of students (bottom, middle, and top third). The "overall" correlations of the rubric scores and grader scores indicate that similar to Test 1, the Useful Description and Math Procedures category scores had a lower relationship to the grader scores than the Physics Approach, Specific Application of Physics, and Logical Progression scores.

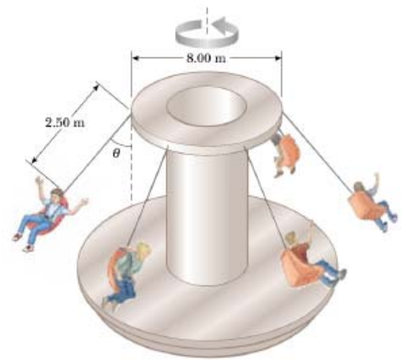
On this problem, the high-scoring group exhibited a stronger correlation between the rubric scores and grader scores for all categories except Useful Description. The relationship between the "Total Rubric" score and grader score is stronger for the bottom third group and the top third group than for the middle group. This pattern is also true for the categories Physics Approach and Specific Application of Physics.

The second section had a different problem on Test 2. While this problem still assessed use of forces in Newton's Second Law, it also required an understanding of circular motion and centripetal acceleration. The problem statement is copied below.

Test 2 Problem 1 (Section 2)

An amusement park ride consists of a rotating circular platform 8.00 m in diameter from which 10.0-kg seats are suspended at the end of 2.50-m massless chains (see figure). When the system rotates, the chains make an angle $\theta = 28.0^{o}$ with the vertical.

a) What is the speed of each seat? [5 pts]
b) Draw a free-body diagram of a 40.0-kg child riding in a seat [5 pts]
c) Find the tension in the chain [15 pts]



As seen in the rubric score frequency plot for this problem (N=162), students had a distributed response to this question. High scores for Useful Description and Math Procedures were observed more frequently than other categories, such as Physics Approach, Specific Application of Physics, or Logical Progression. In this problem a Useful Description involved a visualization of the forces acting on the seat and appropriate symbols for quantities, a Physics Approach included use of Newton's second law independently in perpendicular directions, Specific Application focused on the particular forces identified and an equation for centripetal acceleration, Math Procedures depended on algebraic procedures, and Logical Progression was the extent to which the solution was coherent, consistent, and progresses to an answer for the appropriate target quantity.

Figure 35: Frequency of Rubric Scores for Test 2 Problem 1 (Section 2)

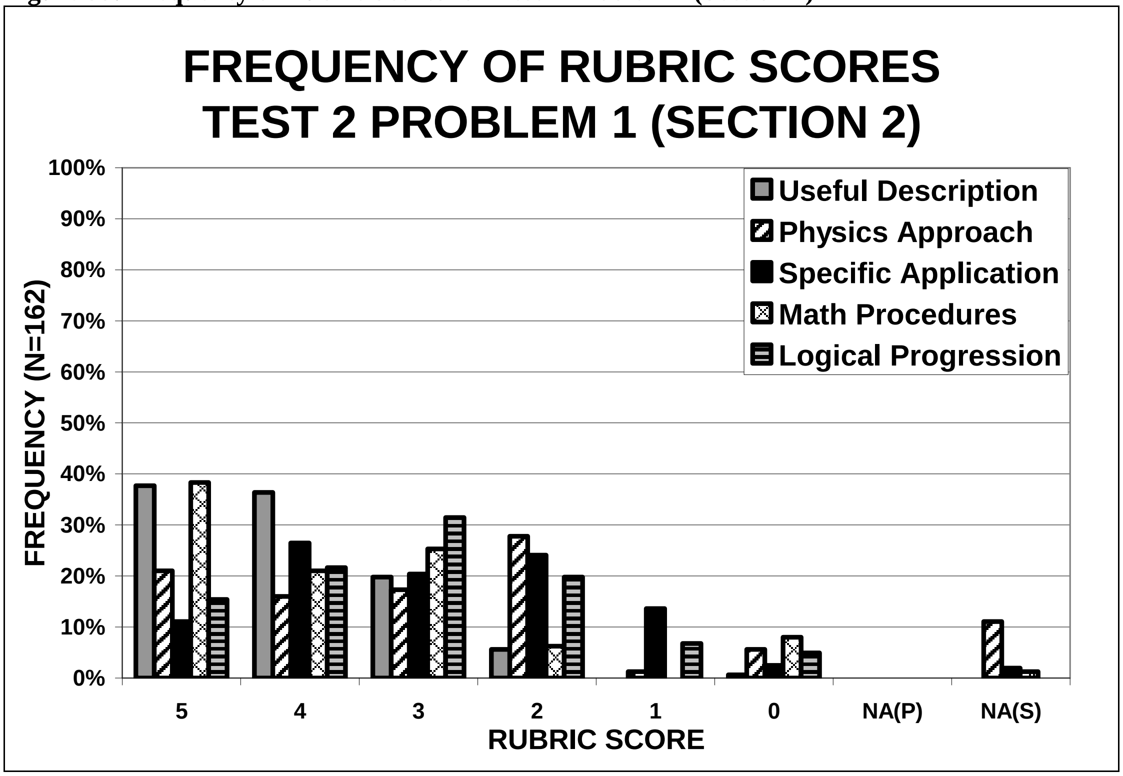
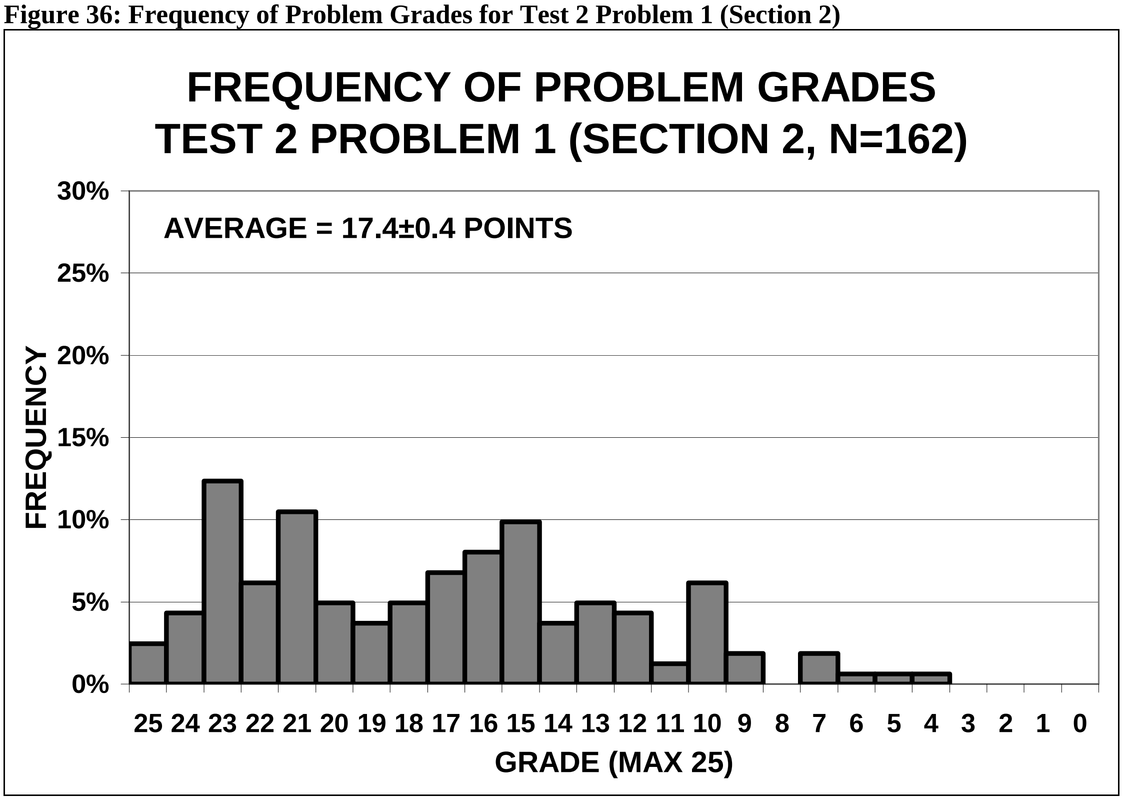


Figure 36: Frequency of Problem Grades for Test 2 Problem 1 (Section 2)



After excluding solutions that scored 100% on both the rubric and grader measures (2 solutions) the average grader score for this problem was 69±2% and the average rubric score was 67±2%. As seen in the frequency chart above and reported in the table below, the Useful Description and Math Procedures categories scored higher on average than the Specific Application of Physics and Logical Progression.

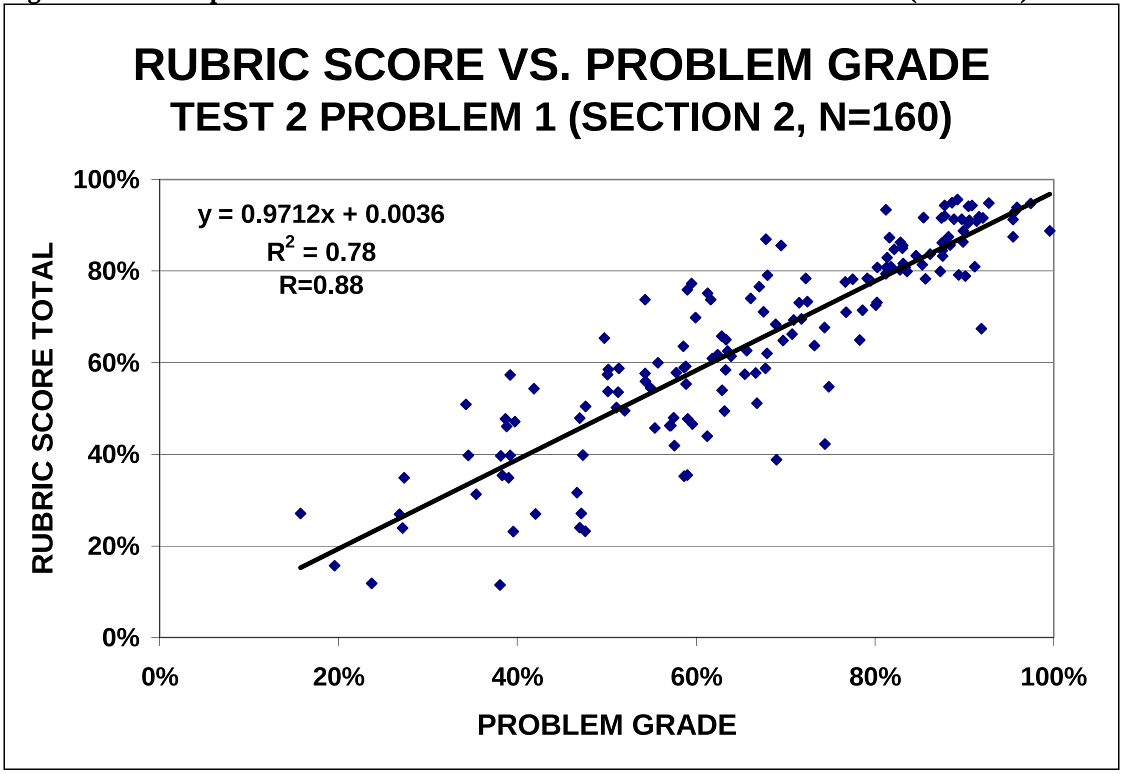
Table 29: Average Rubric and Grader Scores for Test 2 Problem 1 (Section 2)

	Averages (N=160)
Useful Description	81±2%
Physics Approach	62±2%
Specific Application	58±2%
Math Procedures	73±2%
Logical Progression	60±2%
Rubric Score	67±2%
Grader Score	69±2%

Approximately 40% of the student papers reached a correct answer. A common error observed in 25% of papers involved using equations for period and angular velocity and leaving the solution in terms of unknown quantities (Logical Progression score of 3 and problem grade of 2 out of 5 points for part a). Specific Application errors included assuming no acceleration along the direction of the chain or entirely along the chain (20% of papers) which resulted in a rubric score of 2 or 3 and a problem grade of between 6 and 10 points out of 15 for part c). Physics Approach errors included not treating force directions independently and/or not considering force components (15%) which resulted in an approach score of 2 or 3 and a problem grade of 9 or 10 out of 15 for part c). Some students had one of these errors in combination with using an inappropriate value, such as the mass quantity (6%) or not considering the radius of the axle (25%), which were minor errors in Specific Application of Physics scored a 4 on the rubric or a problem grade deduction of two points on part c). Some students reached a correct numerical answer, but had conceptual errors indicated by their labeling of forces, such as labeling a “centripetal force” or a “ma” force on their free-body diagram which resulted in a lower Description score of 2 or 3. Labeling one of these forces resulted in a problem grade of 3 out of 5 for part b.

Figure 37 shows scatterplots of total rubric scores versus the problem grades on Test 2 Problem 1 in section 2. Points are shifted by a small random number so that clusters of scores in the distribution will not be masked by being included as a single point [shifted score=score - score*0.05*RAND()]. The Pearson correlation coefficient for this distribution is R=0.88 and showing that the total rubric score accounts for 70% and 78% of the grader score variance.

Figure 37: Scatterplot of Rubric Score vs. Problem Grade for Test 2 Problem 1 (Section 2)

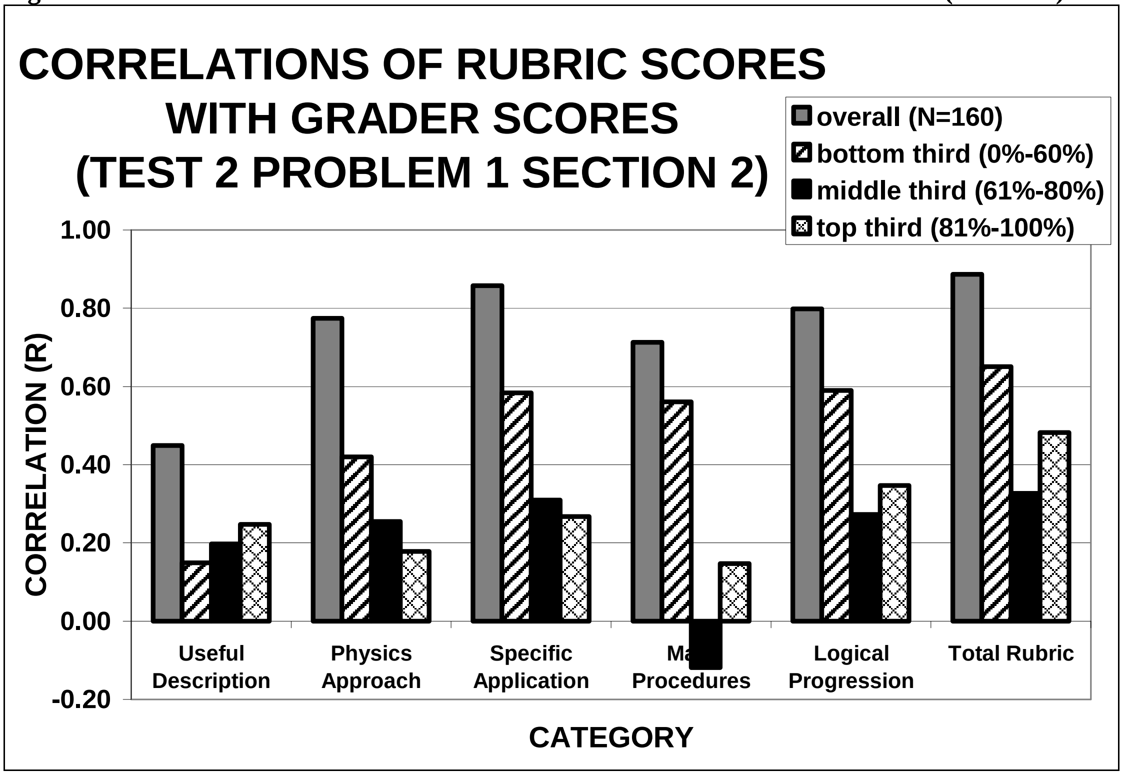


The correlation between the grader's scores and rubric scores for this problem in section 1 are reported in Table 30 and plotted in Figure 38. Correlations are reported for all students (overall) and separately for the bottom third, middle third, and top third of the distribution of grader's scores.

Table 30: Correlations of Rubric Scores with Grader Scores for Test 2 Problem 1 (Section 2)

	Overall (N=160)	Bottom Third (0-60%) (N=58)	Middle Third (61-80%) (N=46)	Top Third (81-100%) (N=56)
Useful Description	0.45	0.15	0.20	0.25
Physics Approach	0.77	0.42	0.25	0.18
Specific App.	0.86	0.58	0.31	0.27
Math Procedures	0.71	0.56	-0.12	0.15
Logical Progression	0.80	0.59	0.27	0.35
All Categories	0.89	0.65	0.33	0.48

Figure 38: Correlations of Rubric Scores with Grader Scores for Test 2 Problem 1 (Section 2)



The correlation plot in Figure 38 indicates differences in the relationships between rubric scores and grader scores for each category, and how these correlations compare for different groups of students (bottom, middle, and top third). The "overall" correlations of the rubric scores and grader scores indicate that similar to previous problems, the Useful Description and Math Procedures category scores had a lower relationship to the grader scores than the Physics Approach, Specific Application of Physics, and Logical Progression scores.

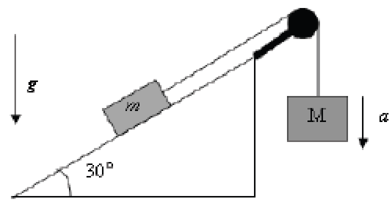
On this problem, the low-scoring group exhibited a stronger correlation between the rubric scores and grader scores for all categories except Useful Description. The relationship between the "Total Rubric" score and grader score is stronger for the bottom third group and the top third group than for the middle group (consistent with section 1's problem). Unlike section 1, however, this pattern does not hold true for the categories Physics Approach and Specific Application of Physics. There is also a negative relationship between the rubric scores and grader scores in Math Procedures for the middle group that was not observed for other problems.

The second problem for each section of the course had the same situation and figure, but asked different questions. Section 1's version explicitly prompted students to draw a free-body diagram in part a) whereas Section 2's problem did not, and the first version asked for the value of the tension force numerically before calculating work done by the tension whereas the second version asked for the work done by the tension force expressed symbolically. Each version is copied below.

Test 2 Problem 2 (Section 1)

A block of mass $m = 3$ kg and a block of unknown mass M are connected by a massless rope over a frictionless pulley, as shown below. The kinetic frictional coefficient between the block m and the inclined plane is $\mu_k = 0.17$. The plane makes an angle 30^o with horizontal. The acceleration, a, of the block M is $1\ m/s^2$ downward.

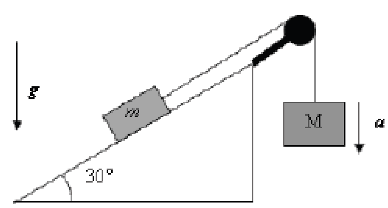
(A) Draw free-body diagrams for both masses. [5 points]
(B) Find the tension in the rope. [5 points]
(C) If the block M drops by 0.5 m, how much work, W,is done on the block m by the tension in the rope? [15 points]



Test 2 Problem 2 (Section 2)

A block of known mass m and a block of unknown mass M are connected by a massless rope over a frictionless pulley, as shown. The kinetic frictional coefficient between the block m and the inclined plane is μ_k. The acceleration, a, of the block M points downward.

a) If the block M drops by a distance h, how much work, W, is done on the block m by the tension in the rope? Answer in terms of known quantities [15 pts]
b) Now let the mass m=3kg, the coefficient of kinetic friction between the block m and the inclined plane be $\mu_k = 0.17$, and the acceleration a, of the block M be $1\ m/s^2$ downward. How much work, W, is done on the block m by the tension in the rope if the block M drops by 0.5m? [5 pts]
c) If the inclined plane were frictionless, would the total work done on both blocks by the tension in the rope increase, decrease, or stay the same? [5pts]



As seen in the frequency plot of rubric scores in

Figure 39, most papers in section 1 included a useful description of the problem and more than half of students did so without errors. This was not surprising because a free-body diagram was prompted in part a). As seen in Figure 40 this was not true of section 2, where fewer than 40% of students had an error-free description. One category with a high frequency of low scores was the Specific Application of Physics, and the Math Procedures were lower than on previous problems. One possible reason for lower math scores is that this problem involved two objects (blocks) that are both accelerating, producing simultaneous equations. In addition, more students in section 2 scored a 2 on Physics Approach and Logical Progression. Since students had to solve symbolically before numerically, a greater fraction of students left the answer in terms of the unknown quantity M than in version 1, or had haphazard, confused reasoning in their solution that impacted their Logical Progression score.

In this problem, a Physics Approach involved applying Newton's second law to each block and treating force directions independently and a basic equation for calculating the work. The Specific Application of Physics assessed the forces identified on each object, appropriate signs of acceleration, and appropriate terms in calculating the work. Math Procedures included algebraic steps to solve simultaneous equations, and Logical Progression assessed the coherence, consistency, and progression of the solution to an answer for the target quantity in terms of known quantities.

Figure 39: Frequency of Rubric Scores for Test 2 Problem 2 (Section 1)

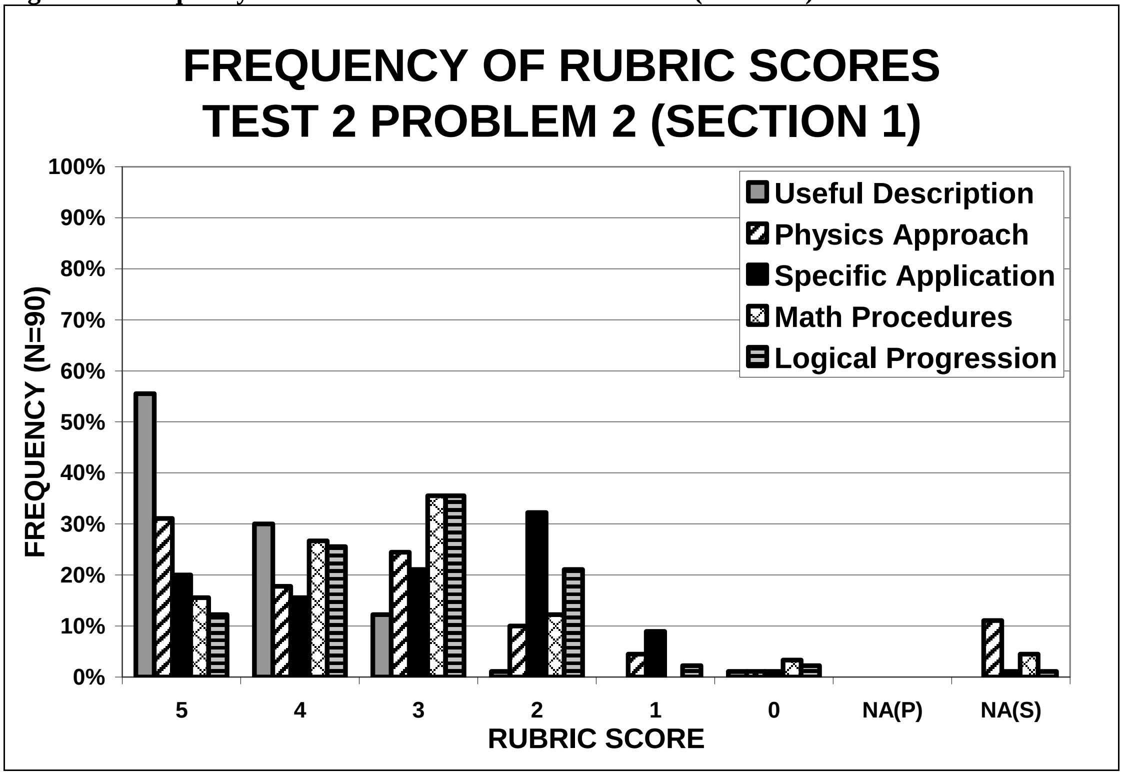


Figure 40: Frequency of Rubric Scores for Test 2 Problem 2 (Section 2)

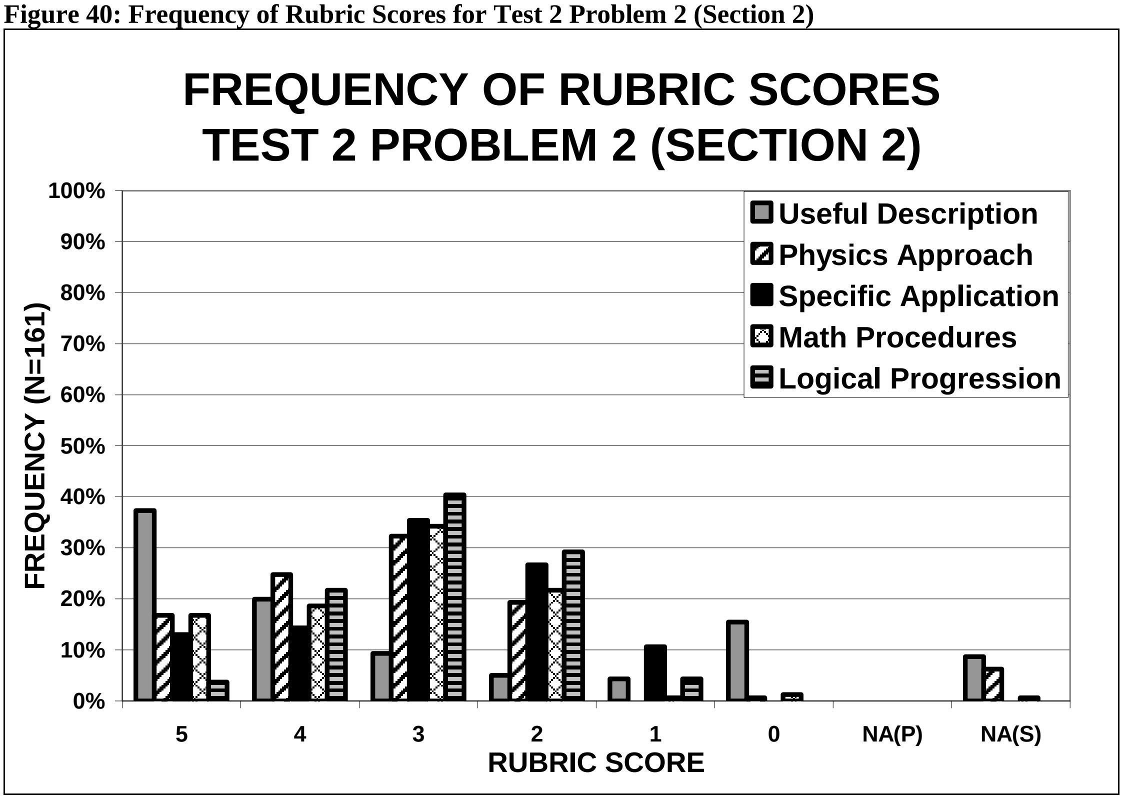


Figure 41: Frequency of Problem Grades for Test 2 Problem 2 (Section 1)

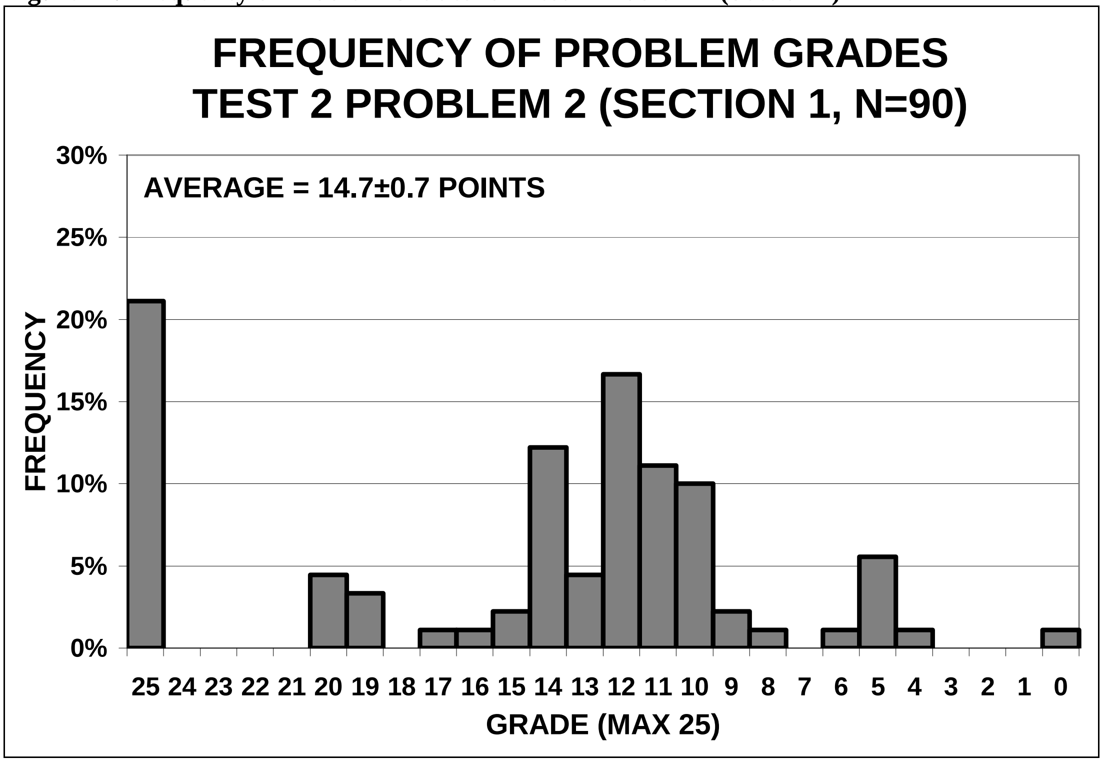
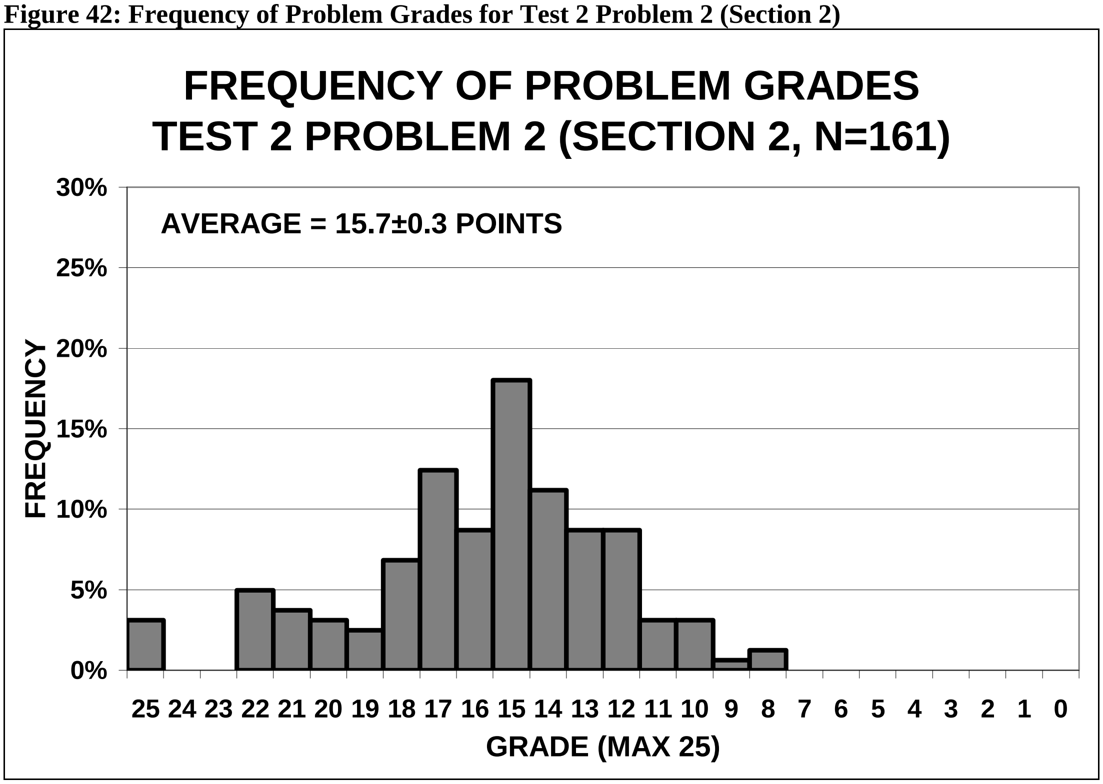


Figure 42: Frequency of Problem Grades for Test 2 Problem 2 (Section 2)



After excluding solutions that scored 100% on both the rubric and grader measures in section 1 and section 2 (9 solutions and 5 solutions) the average score from the grader on problem version 1 was 55±3% and the average grader score on version 2 was 62±1%. The total rubric score was higher in section 1 than section 2, 69±2% compared to 62±2%. As reflected in both the frequency plot above and table of average scores below, the description and physics approach categories scored higher on average than the specific application in section 1. As noted previously, the Useful Description was lower in the non-prompted version, and the Physics Approach and Logical Progression averages were also somewhat lower in section 2. The Specific Application and Math averages were the same in both versions.

Table 31: Average Rubric and Grader Scores for Test 2 Problem 2 (Section 1)

	Average Section 1 (N=81)	**Average Section 2 (N=156)**
Useful Description	87±2%	67±3%
Physics Approach	72±3%	67±2%
Specific Application	57±3%	57±2%
Math Procedures	66±2%	65±2%
Logical Progression	61±2%	57±1%
Rubric Score	69±2%	62±2%
Grader Score	55±3%	62±1%

As seen in the frequency of problem grades for section 1 in Figure 41, approximately 21% of papers received the full 25 points. Approximately 37% of the papers in version 1 scored with the rubric had a correct solution to parts a) and b), but not necessarily to part c) on calculating work. A closer look at students' solutions indicate that common errors included summing forces to zero instead of mass times acceleration (20%), which could be interpreted as a Physics Approach error if Newton's second law was not written in basic form, or a Specific Application error if the acceleration was explicitly set to zero or if the approach was considered Not Applicable for the solver (rubric score 2 or 3). This same error resulted in a problem grade of 2 points out of 5 on part b). Other errors included missing a force such as friction or gravity (14% of papers), resulting in a Specific Application score 2 or 3 and problem grade between 0 and 2 points out of 5 on part b. Another error was leaving the answer in terms of unknown quantity M or substituting a false value for M (14% of papers), resulting in a Logical Progression score of 3 and a problem grade of between 0 and 2 points in part b). Some papers had these errors combined with other minor errors such as Specific Application errors with signs of quantities, or Mathematical procedures errors with trigonometric functions or dropping terms during calculations (19% of papers). Minor errors were typically scored a 4 with the rubric and resulted in a 1 or 2 point deduction by the problem grader. For part c) on calculating the work done by the tension force, approximately 23% of the papers incorrectly assumed that the general equation for work is $W = Fd\cos\theta$ rather than multiplying the tension force by the distance it acts over. This was reflected in the rubric as a lower Physics Approach score (rubric score 4), because the error regarded a general equation and not the specific terms in that equation. The problem grade for this error was 10 points out of 15 on part c). Other work equation errors (such as using $W = mad$) resulted in a problem grade of 5 out of 15 points.

The fraction of students receiving the maximum problem grade on the section 2 problem version was lower, however, at only 3% (compared to 21% on version 1). If you neglect the multiple choice question (part c) approximately 14% of students had a correct solution (some with minor errors). Again, some of the most common student errors included summing forces to zero instead of mass times acceleration (18% of papers, problem grade of 10-12 points of 15 on part a) or missing a force (10% of papers, problem grade of 8-12 points of 15 on part a). The error of leaving the answer in terms of the unknown quantity mass *M* or tension *T* or using a false value was scored by the problem grader between 8 and 10 points out of 15 on part a (20% of papers, Logical Progression error of 3). or minor errors with signs (Specific Application), trigonometric components (Math Procedures), or dropping terms (17%). One notable difference from version 1 is that fewer students were missing a force (10% compared to 14%) but more students left the answer in terms of an unknown quantity (20% compared to 14%). This suggests that the symbolic form of the question gave students more difficulty expressing a final answer in terms of known quantities, or that in the numeric question students were more likely to find this error when performing the calculation.

Another difference was in the students' calculation of work. In version 1, several students (23%) used an incorrect equation for work done one the block *m* by the tension in the rope, using $W = Fd\cos\theta$ where they typically substituted the tension force T in place of the symbol F. Recall this was marked by the grader as 2 out of 5 points on part b. In the second version, there were very few instances of this error (2%) however more students calculated work from the net force: $W = F_{net}d$ or $W = mad$ (22%) rather than the work done by the tension force, resulting in some of the errors for part a) stated above and typically scored between 8-12 points out of 15 on part a). Numerical answers to part b) that were consistent with part a) equation were scored 2 out of 5 points. It is possible that the format of the question for version 1 made the calculation of work from the tension force more explicit, since the step of calculating the tension force was prompted in parts b) in version 1 but was not prompted in version 2.

Figure 43 and Figure 44 show scatterplots of total rubric scores versus the problem grades on Test 2 Problem 2. Points are shifted by a small random number so that clusters of scores in the distribution will not be masked by being included as a single point [shifted score=score - score*0.05*RAND()]. The Pearson correlation coefficient for this distribution is R=0.72 and R=0.76 showing that the total rubric score accounts for 52% and 58% of the grader score variance in sections 1 and 2, respectively.

The correlation between the grader's scores and rubric scores for this problem in section 1 are reported in Table 32 and plotted in Figure 45. Section 2 correlations are reported in Table 33 and plotted in Figure 46. Correlations are reported for all students (overall) and separately for the bottom third, middle third, and top third of the distribution of grader's scores.

Figure 43: Scatterplot of Rubric Scores vs. Problem Grades for Test 2 Problem 2 (Section 1)

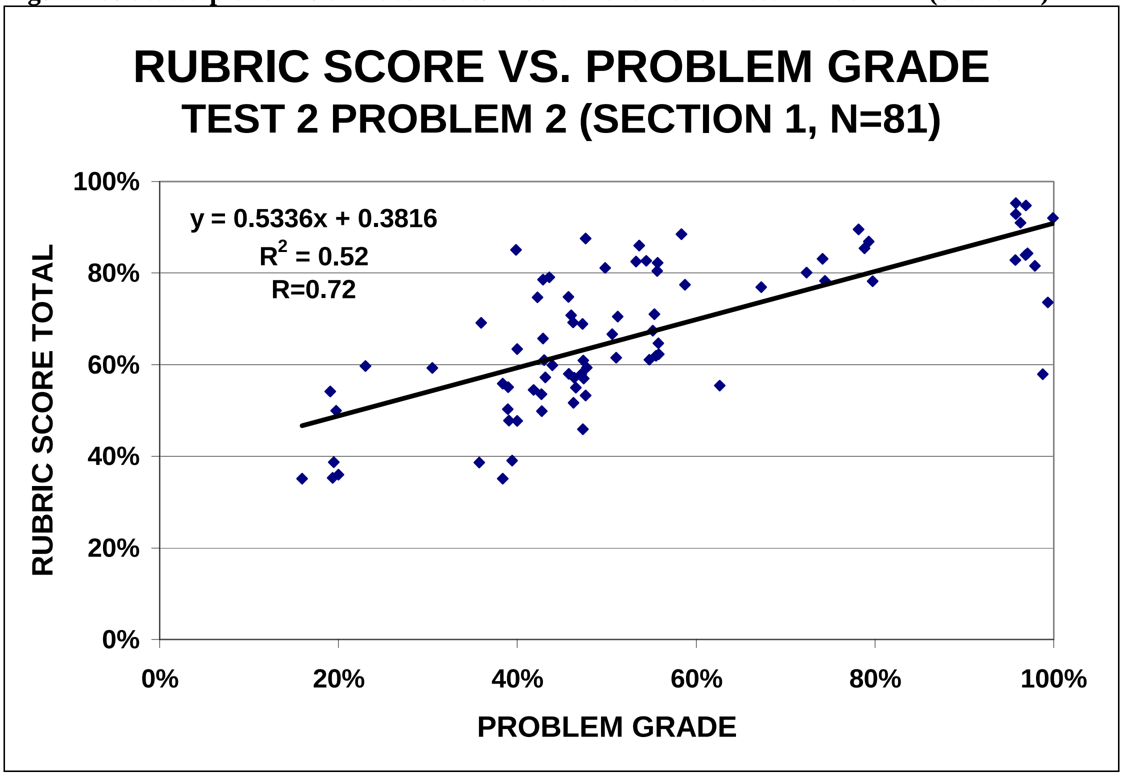


Figure 44: Scatterplot of Rubric Scores vs. Problem Grades for Test 2 Problem 2 (Section 2)

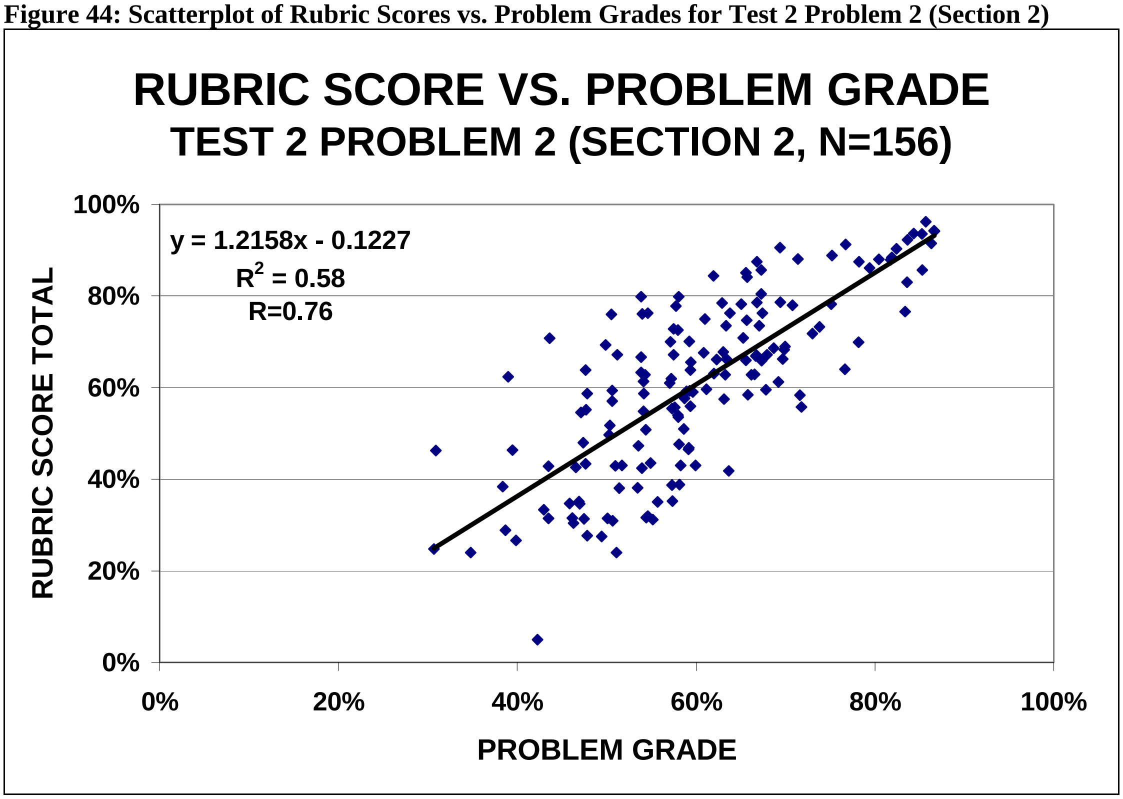


Table 32: Correlations of Rubric Scores with Grader Scores for Test 2 Problem 2 (Section 1)

	Overall (N=81)	Bottom Third (0-44%) (N=29)	Middle Third (45-56%) (N=30)	Top Third (57-100%) (N=22)
Useful Description	0.34	0.30	0.37	-0.20
Physics Approach	0.65	0.46	0.52	0.36
Specific App.	0.81	0.50	0.56	0.67
Math Procedures	0.47	0.30	0.03	0.15
Logical Progression	0.63	0.35	0.13	0.35
All Categories	0.72	0.51	0.49	0.30

Table 33: Correlations of Rubric Scores with Grader Scores for Test 2 Problem 2 (Section 2)

	Overall (N=156)	Bottom Third (0-56%) (N=59)	Middle Third (57-64%) (N=43)	Top Third (65-100%) (N=54)
Useful Description	0.48	0.19	0.25	0.25
Physics Approach	0.69	0.40	0.30	0.48
Specific App.	0.74	0.32	0.37	0.71
Math Procedures	0.70	0.33	0.14	0.45
Logical Progression	0.68	0.21	0.25	0.43
All Categories	0.77	0.35	0.37	0.65

Figure 45: Correlations of Rubric Scores with Grader Scores for Test 2 Problem 2 (Section 1)

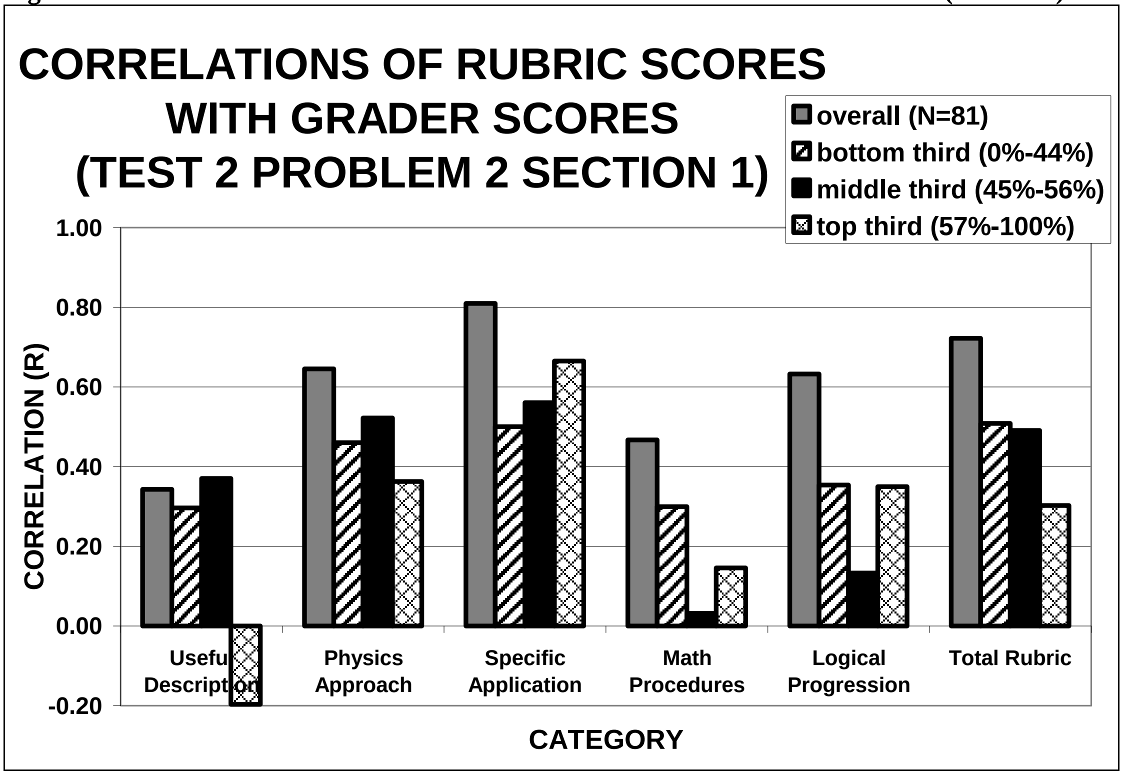
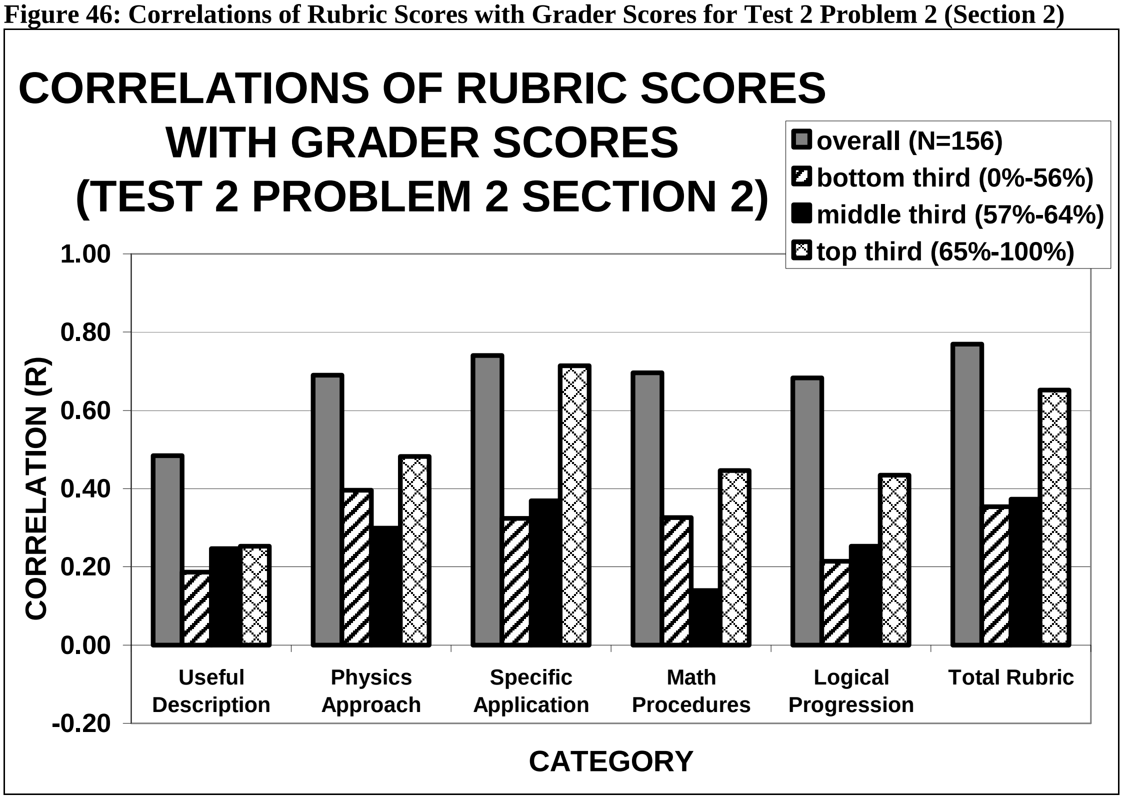


Figure 46: Correlations of Rubric Scores with Grader Scores for Test 2 Problem 2 (Section 2)



The correlation plots in Figure 45 and Figure 46 indicate differences in the relationships between rubric scores and grader scores for each category, and how these correlations compare for different groups of students (bottom, middle, and top third) and across sections. The "overall" correlations of the rubric scores and grader scores for section 1 indicate that similar to previous problems, the Useful Description and Math Procedures category scores had a lower relationship to the grader scores than the Physics Approach, Specific Application of Physics, and Logical Progression scores. In section 2 this was true of the Description but not the Math category; the overall correlation between Math Procedures and grader score for section 2 was higher than on previous problems (R=0.70).

In section 2, the high-scoring group had the strongest correlation between the rubric scores and grader scores for all categories, but this was only true for Specific Application and Logic in section 1. The relationship between the "Total Rubric" score and grader score was weakest for the top group in section 1 but strongest for that same group in section 2. There is also a negative relationship between the rubric scores and grader scores in Description for the high-scoring group that was not observed for other problems or the other section.

Rubric and Grader Scores on Test 3

The first problem on the third test was the same for both sections of the course. Although the instructor intended students to apply the Law of Conservation of Energy to this problem, it could also be solved using Newton's Second Law. Approximately 55% of the papers used an energy approach, 30% used a forces approach with Newton's second law, and the remaining 15% did not have a discernable physics approach. As seen below, the problem statement cued on a particular object in the problem (the middle block M_3) which affected the response processes for some students.

Test 3 Problem 1:

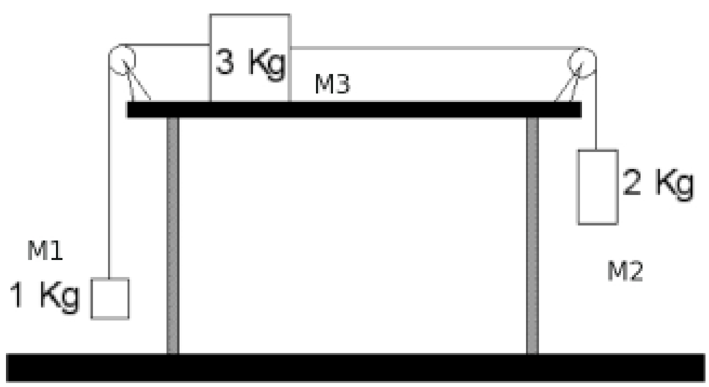
The system of three blocks shown is released from rest. The connecting strings are massless, the pulleys ideal and massless, and there is no friction between the 3kg block and the table.

(A) At the instant M_3 is moving at speed v, how far d has it moved from the point where it was released from rest? (answer in terms of M_1, M_2, M_3, g and v.) [10 pts]

(B) At the instant the 3 kg block is moving with a speed of 0.8 m/s, how far, d, has it moved from the point where it was released from rest? [5 pts]

(C) From the instant when the system was released from rest, to the instant when the 1 kg block has risen a height h, which statement (1, 2 or 3) is true for the three-block system? (1) The total mechanical energy of the system increases. (2) The total potential energy of the system increases. (3) The net work done on the system by the tension forces is 0. [5pts]

(D) Now suppose the table is rough and has a coefficient of kinetic friction μk = 0.1. What is the speed, v, of the 3 kg block after the 2 kg block drops by 0.5 m? (Assume again that the system is released from rest.) [5pts]



In this problem, a description was useful for a Newton's second law approach (analyzing the forces acting on each block) but was not necessary for most students using a conservation of energy physics approach. As seen in

Figure 47, 50% of solutions in section 1 were scored NA(Solver) in the Useful Description category and more than 35% in section 2. Very few categories had papers scored high (5) with the exception of Math Procedures. The scores for Specific Application of Physics were lower than on other problems, with 20% of students scoring a 1 in this category for both sections and nearly half scoring a 1 or 2.

Figure 47: Frequency of Rubric Scores for Test 3 Problem 1 (Section 1)

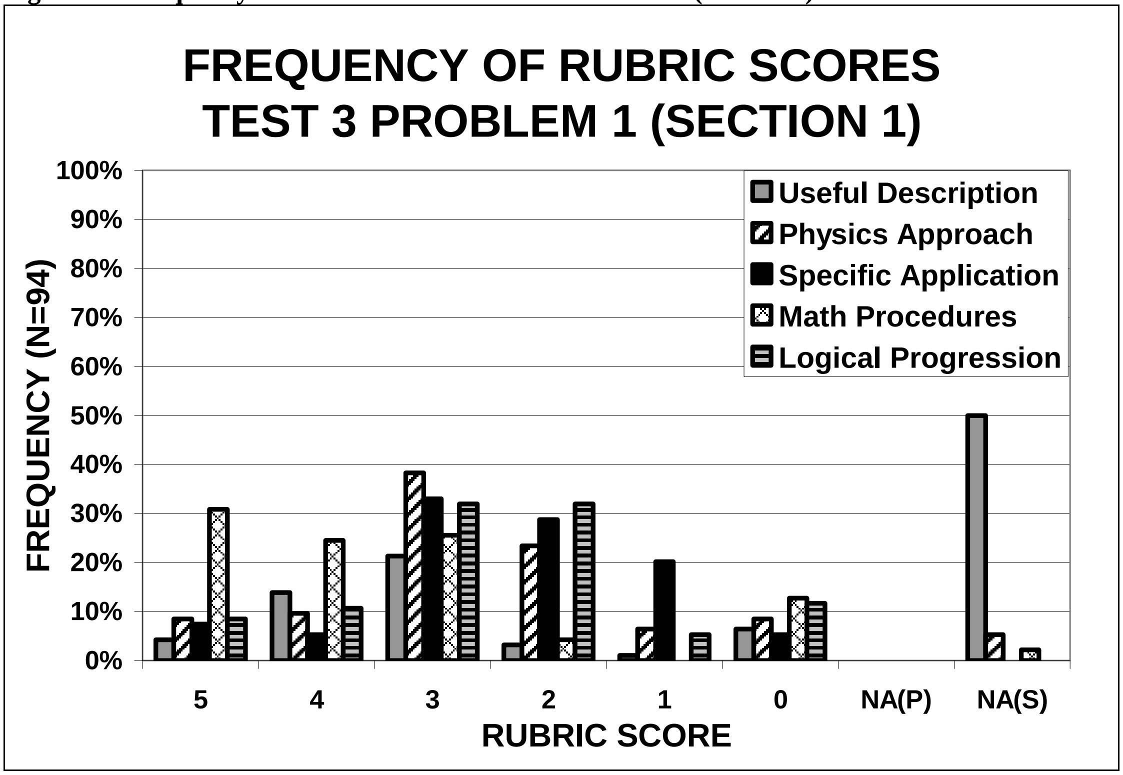


Figure 48: Frequency of Rubric Scores for Test 3 Problem 1 (Section 2)

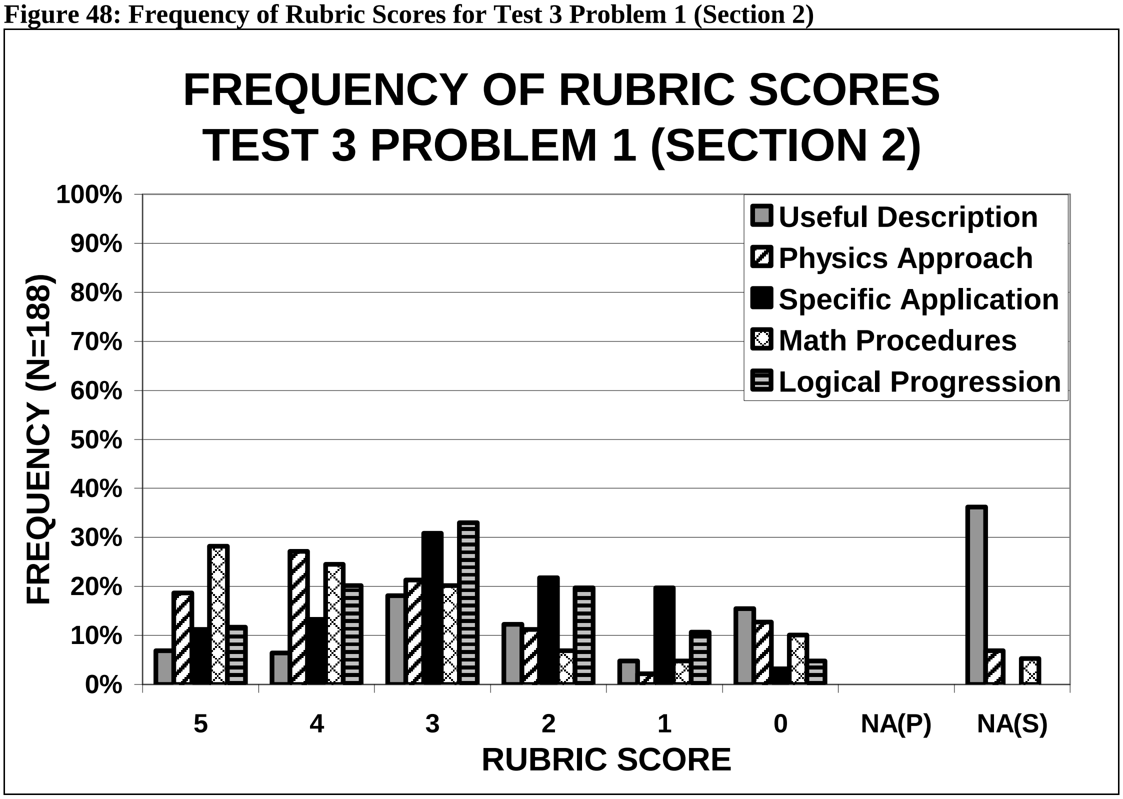


Figure 49: Frequency of Problem Grades for Test 3 Problem 1 (Section 1)

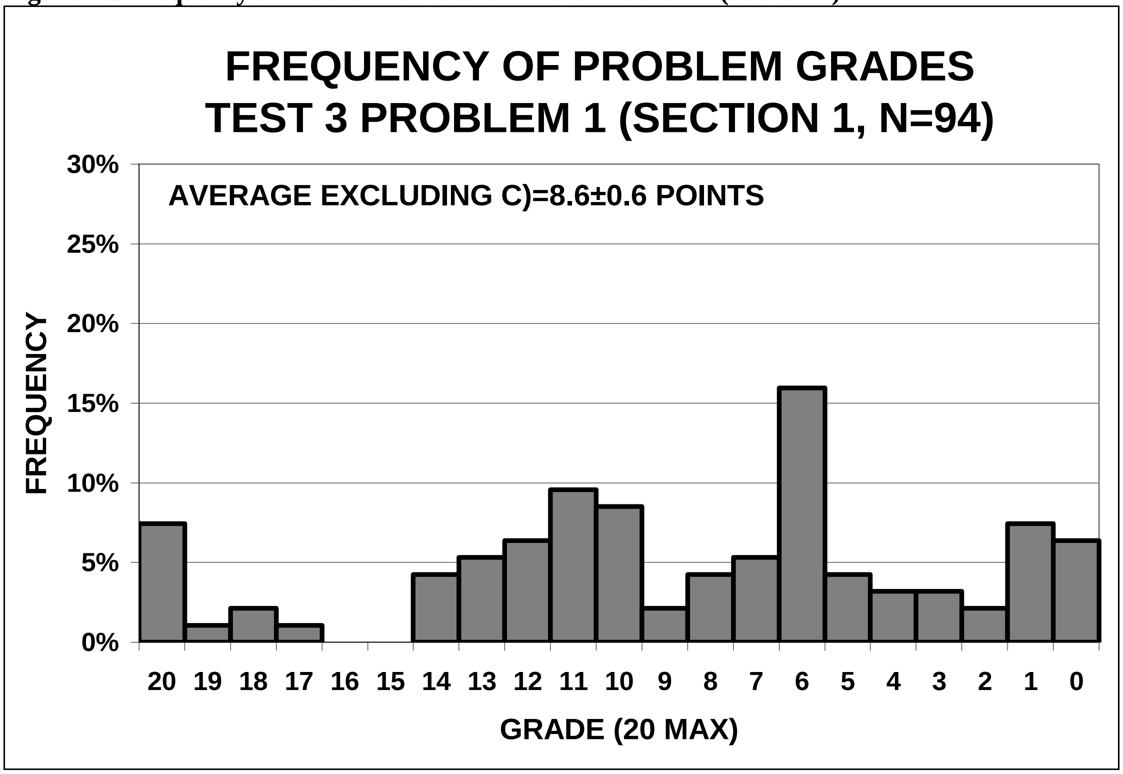
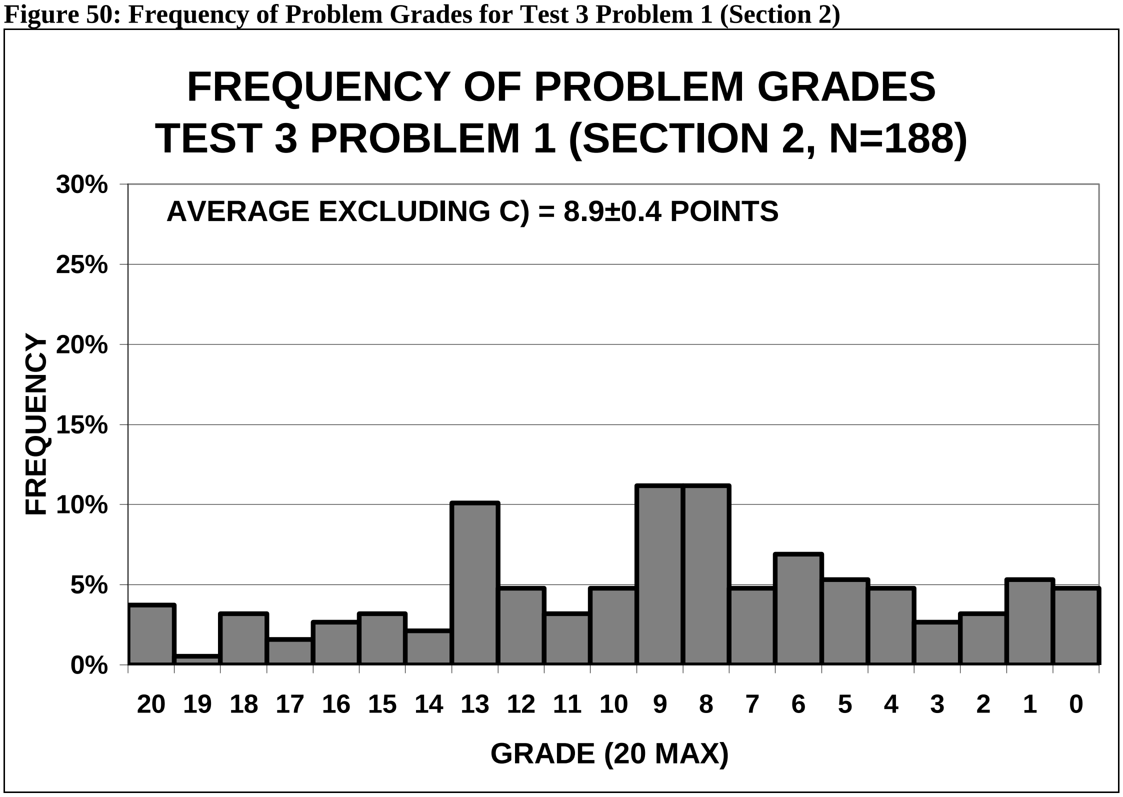


Figure 50: Frequency of Problem Grades for Test 3 Problem 1 (Section 2)



After excluding solutions that scored 100% on both the rubric and grader measures in section 1 and section 2 (2 solutions and 9 solutions) the average score from the grader in section 1 was 44±2% and the average grader score in section 2 was 46±3%. The total rubric score reflected a similar trend with higher scores, 56±2% and 59±3%. As reflected in both the frequency plot above and table of average scores below, the description category scored higher in section 1 than section 2 (the zero score frequency is higher in section 2). The Physics Approach scored higher in section 2, and the Math averages were the same for both groups.

Table 34: Average Rubric and Grader Scores for Test 3 Problem 1

	Averages Section 1 (N=92)	**Averages Section 2 (N=179)**
Useful Description	60±3%	44±3%
Physics Approach	53±3%	62±3%
Specific Application	47±3%	52±3%
Math Procedures	69±3%	67±3%
Logical Progression	50±3%	57±3%
Rubric Score	56±2%	59±3%
Grader Score (w/o C)	43±3%	46±2%

When you neglect the multiple choice question c), approximately 11% of students in section 1 and 16% of students in section 2 answered this problem correctly and/or with minor errors, with most selecting to use the principle of Conservation of Energy. The most common specific application error (25% of students) was to only consider the kinetic energy of block 3, rather than the kinetic energy of all three blocks. This error resulted in a rubric score of 2 or 3 for specific application of physics and a problem grade of 6 out of 10 points on part a). An example of this application error is shown in Appendix 6. Another common error was to apply Newton's Second Law with incorrect reasoning that the tension in each string was equal to the weight of the hanging masses (T=Mg) instead of considering the acceleration of the blocks. As stated in Test 2 problems, neglecting acceleration can be considered a Physics Approach error if Newton's second law was not written in basic form, or an Application error if the acceleration was explicitly set to zero. At least 15% of students inappropriately used Newton's Second Law with this reasoning, resulting in a problem grade of 5 out of 10 points for part a). For some student solutions, the final answer was correct but the reasoning was unclear. A typical procedure for these papers was to calculate acceleration from acceleration equals the force "F" divided by total mass (sum of three masses), where $F = M_2 g - M_1 g$. For these students, it is possible that the answer was obtained using correct reasoning (using "F" to represent net external forces) but it is also possible that the student used false reasoning, such as the T=Mg error. The problem grade for this procedure was 10 out of 10 points for part a), whereas on the rubric it was scored as 4 out of 5 in Logical Progression.

One notable difference between the two graders was in their scoring of part b), the numerical calculation of the distance. The grader in section 1 assigned solutions that had a calculation consistent with part a) as 4 out of 5 points (even if part a was incorrect), whereas the grader for section 2 assigned an incorrect numerical answer as 2 out of 5 points, regardless of its consistency with the equation written in part a). The rubric scored math procedures independent of the particular physics equation, so correct algebra and numerical calculations received 5 out of 5 points for Math on the rubric.

Figure 51 and Figure 52 show scatterplots of total rubric scores versus the problem grades on Test 3 Problem 1. Points are shifted by a small random number so that clusters of scores in the distribution will not be masked by being included as a single point [shifted score=score - score*0.05*RAND()]. The Pearson correlation coefficient for this distribution is R=0.84 and R=0.79 showing that the total rubric score accounts for 70% and 62% of the grader score variance in sections 1 and 2, respectively.

Figure 51: Scatterplot of Rubric Scores vs. Problem Grades for Test 3 Problem 1 (Section 1)

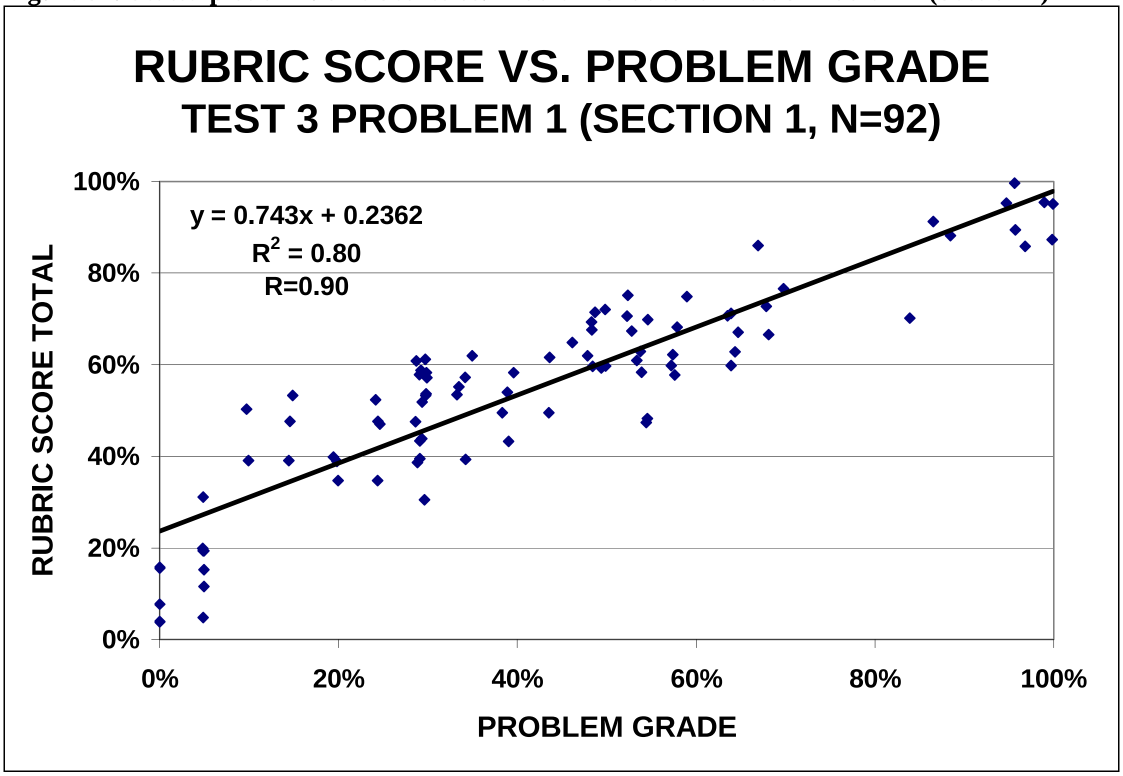
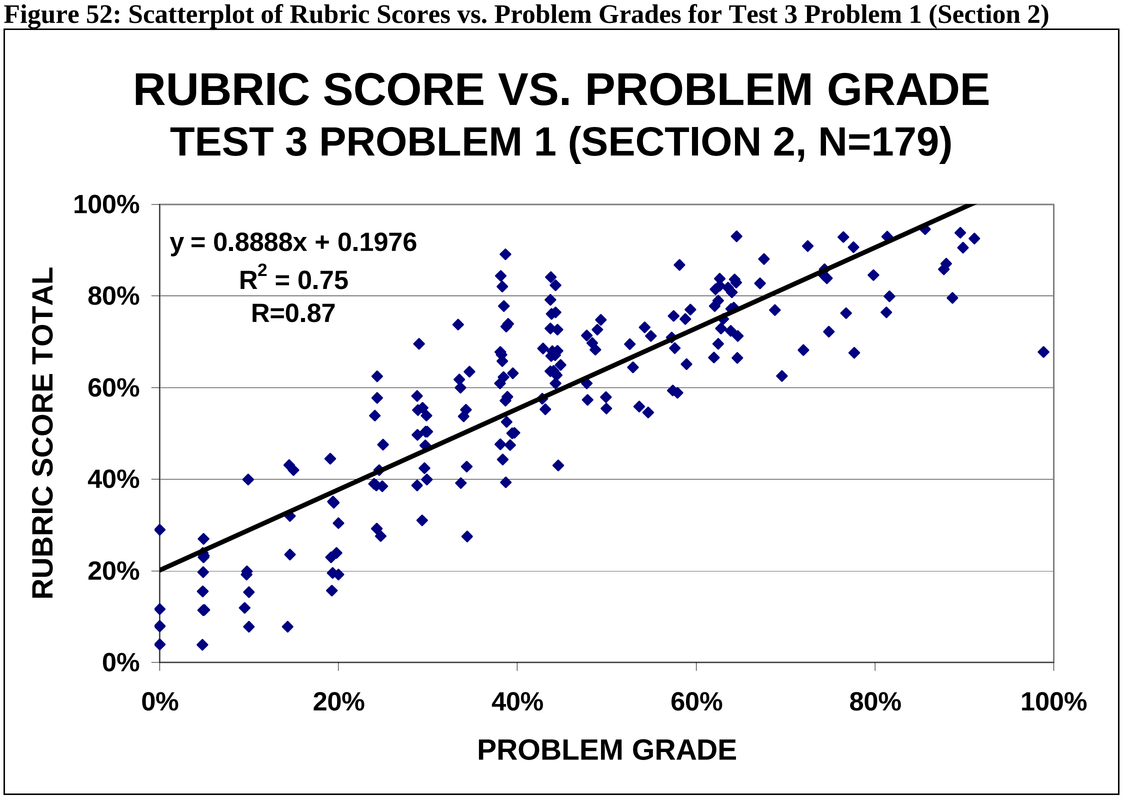


Figure 52: Scatterplot of Rubric Scores vs. Problem Grades for Test 3 Problem 1 (Section 2)



The correlation between the grader's scores and rubric scores for this problem in section 1 are reported in Table 35 and plotted in Figure 53. Section 2 correlations are reported in Table 36 and plotted in Figure 54. Correlations are reported for all students (overall) and separately for the bottom third, middle third, and top third of the distribution of grader's scores.

Table 35: Correlations of Rubric Scores with Grader Scores for Test 3 Problem 1 (Section 1) Excluding multiple choice question C).

	Overall (N=92)	**Bottom Third (0-25%) (N=24)**	**Middle Third (26-50%) (N=35)**	**Top Third (51-100%) (N=33)**
Useful Description	0.64	0.36	0.41	0.54
Physics Approach	0.73	0.42	0.29	0.67
Specific App.	0.84	0.43	0.65	0.86
Math Procedures	0.71	0.91	0.16	0.33
Logical Progression	0.86	0.84	0.30	0.72
All Categories	0.90	0.82	0.54	0.86

Table 36: Correlations of Rubric Scores with Grader Scores for Test 3 Problem 1 (Section 2) Excluding multiple choice question C).

	Overall (N=179)	**Bottom Third (0-30%) (N=59)**	**Middle Third (31-50%) (N=60)**	**Top Third (51-100%) (N=60)**
Useful Description	0.74	0.61	0.17	0.61
Physics Approach	0.78	0.76	0.03	0.32
Specific App.	0.82	0.55	0.28	0.62
Math Procedures	0.72	0.49	0.32	0.18
Logical Progression	0.72	0.63	0.28	0.34
All Categories	0.87	0.79	0.33	0.54

Figure 53: Correlations of Rubric Scores with Grader Scores for Test 3 Problem 1 (Section 1)

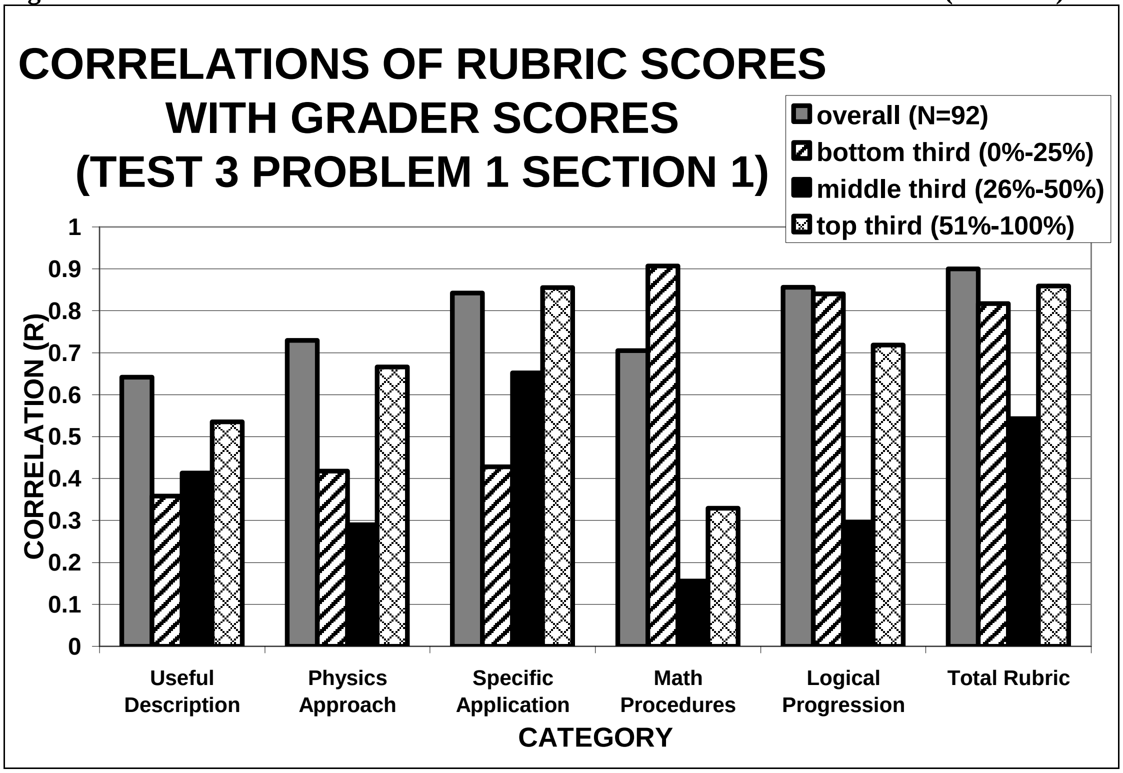
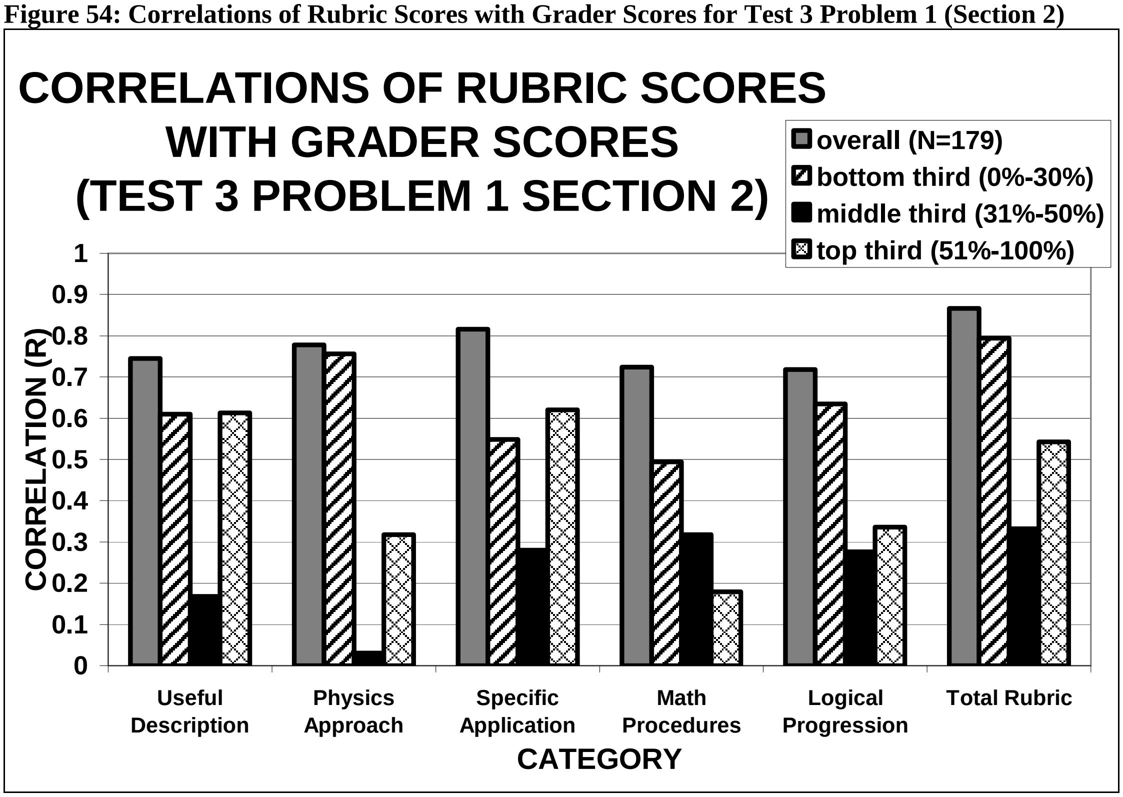


Figure 54: Correlations of Rubric Scores with Grader Scores for Test 3 Problem 1 (Section 2)



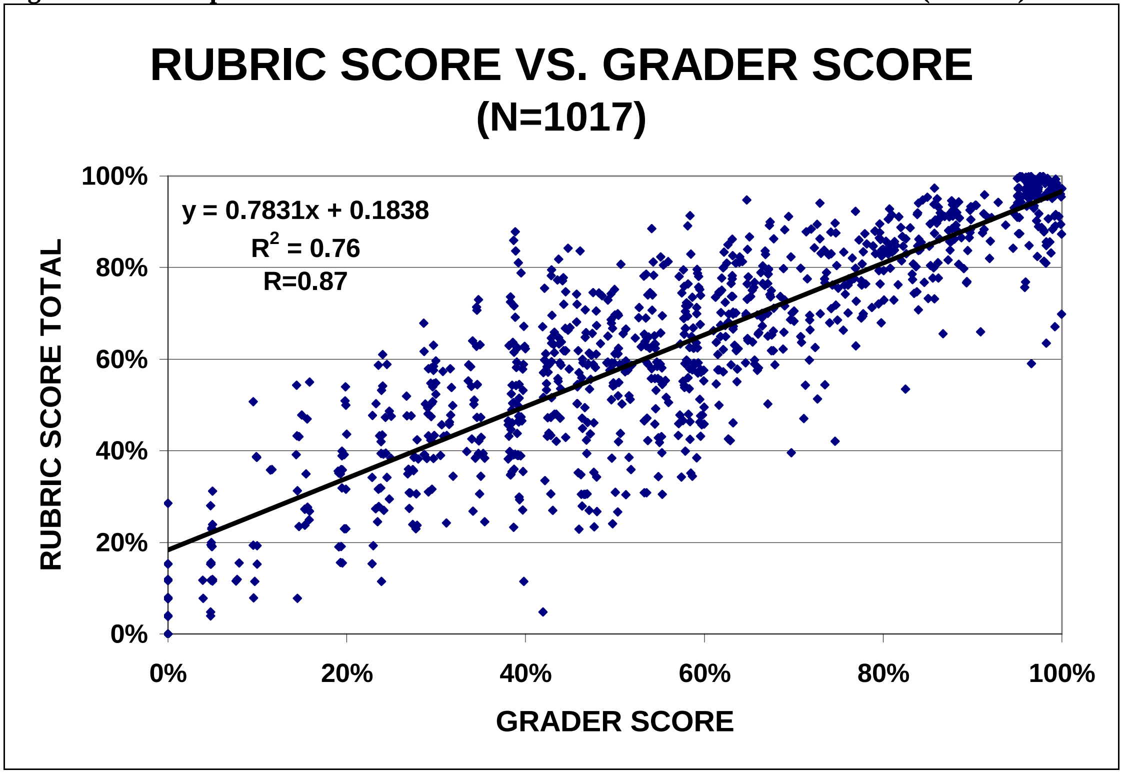
The correlation plots in Figure 53 and Figure 54 indicate differences in the relationships between rubric scores and grader scores for each category, and how these correlations compare for different groups of students (bottom, middle, and top third) and across sections. The “overall” correlations of the rubric scores and grader scores indicate that contrary to previous problems, the relationships to grader scores are similar across categories. In section 1 the description correlation is lower than in section 2, but the logical progression correlation is higher in section 1 than section 2.

In both sections, the middle-scoring group had a lower correlation between the rubric scores and grader scores for most categories than the low and high-scoring groups, with the exception of Useful Description and Specific Application in section 1 and Math Procedures in section 2.

Rubric and Grader Scores on All Problems

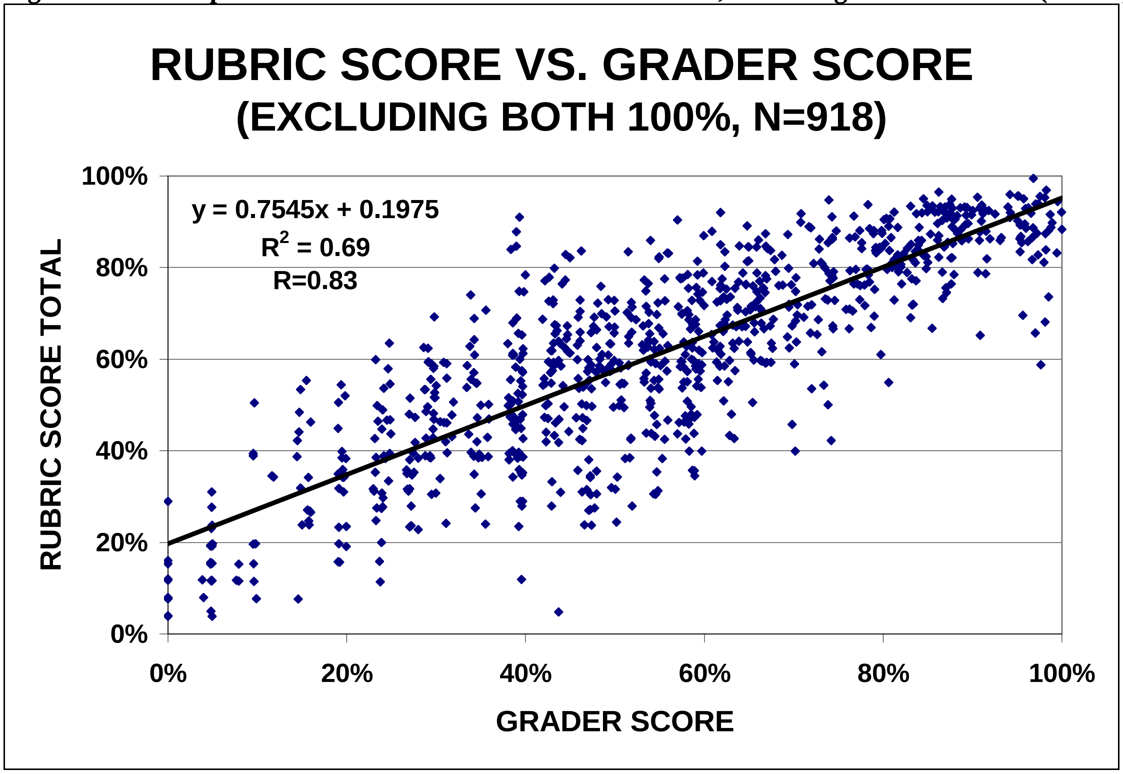
Figure 55 shows a scatterplot of total rubric score versus the grader’s score on all of the papers for all problems. Points are shifted by a small random number so that clusters of scores in the distribution will not be masked by being included as a single point [shifted score=score - score*0.05*RAND()]. The Pearson correlation coefficient for this distribution is R=0.87 showing that the total rubric score accounts for 76% of the grader score variance.

Figure 55: Scatterplot of Total Rubric Score and Grader Score for All Problems (N=1017)



One notable feature of this scatterplot is the cluster of scores on the top right of the graph, indicating several papers that scored near 100% on both the rubric and the graders' criteria. In order to control for the influence of extreme scores on assessing the relationship between the rubric score and grader score, these solutions were removed in subsequent correlation analyses. In the following scatterplot (Figure 56), 99 solutions have been removed because these papers scored 100% on both rubric and by grader, or 0% on both. Most of the 100% scores occurred on the first test; 70 solutions on problem 2.

Figure 56: Scatterplot of Total Rubric Score and Grader Score, Excluding 100% and Zero (N=918)



After excluding the extreme scores, the Pearson correlation coefficient shifted from R=0.87 to R=0.83, showing the total rubric score accounts for 69% of the grader score.

In Table 37 the correlation is calculated separately for each problem and each category of the rubric. The column in the far right of the table (All Problems) summarizes the correlation between category scores and graders' scores. The values in this column indicate Useful Description had the lowest relationship to graders' scores (R=0.55) whereas Specific Application of Physics (R=0.80) was highest, and Physics Approach (R=0.73) and Logical Progression (R=0.74) were higher than Math (R=0.64). The overall (sum) row indicates strong agreement between the rubric and graders' scores for Problem 2 on Test 1 (pigeon hit by a football, R=0.94, 0.92) and lower agreement for the Attwood problem on Test 2 (R=0.72, 0.77).

The overall correlations are repeated in Table 38 and plotted in Figure 57, with values for low, middle, and high-scoring students. In most categories (except math) the relationship of rubric scores to grader scores is lower for the middle scoring group than the low and high-scoring groups.

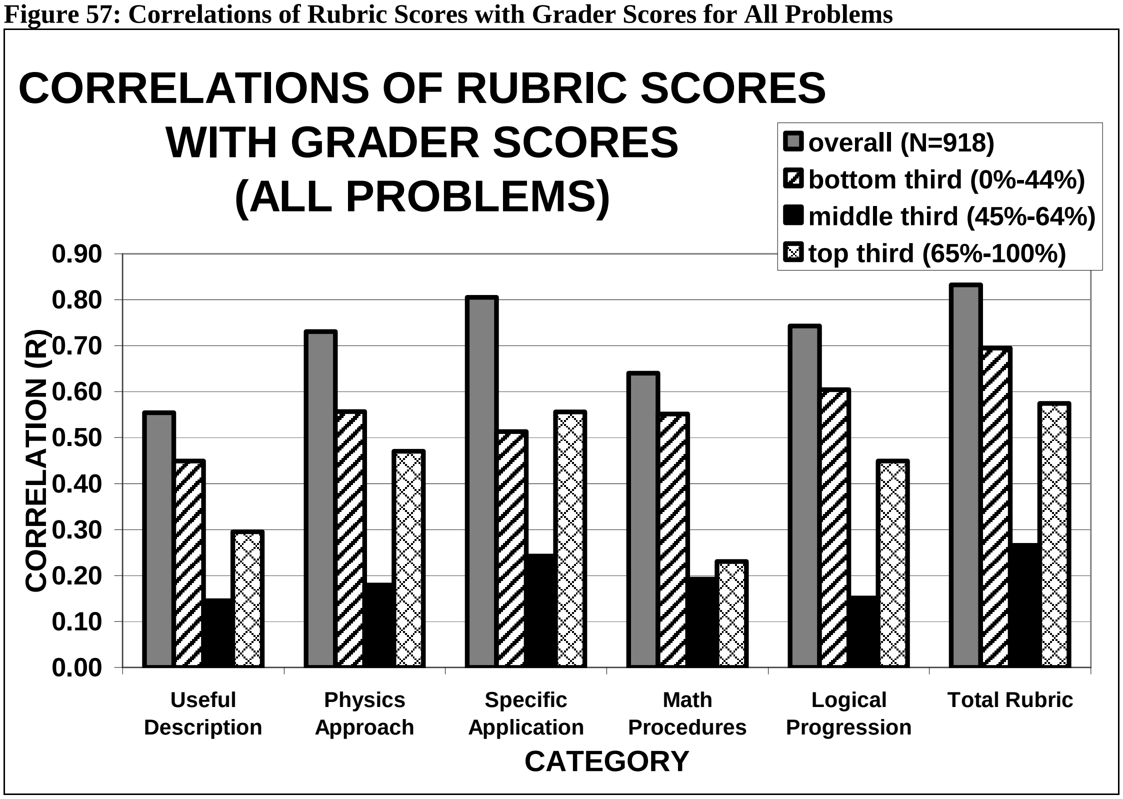
Table 37: Correlations of Rubric Scores with Graders' Scores for Each Problem
(Excluding solutions that scored 100% or 0% for both the grader and rubric)

	Test 1 P2 Section 1 Kine-matics (N=48)	Test 1 P2 Section 2 Kine-matics (N=110)	Test 2 P1 Section 1 Forces (N=92)	Test 2 P1 Section 2 Forces & Circular Motion (N=160)	Test 2 P2 Section 1 Forces & Work (N=81)	Test 2 P2 Section 2 Forces & Work (N=156)	Test 3 P1 Section 1 Energy (N=92)	Test 3 P1 Section 2 Energy (N=179)	All Problems (N=918)
Useful Description	0.51	0.56	0.34	0.45	0.34	0.48	0.64	0.74	0.55
Physics Approach	0.90	0.90	0.67	0.77	0.65	0.69	0.73	0.78	0.73
Specific App.	0.94	0.86	0.90	0.86	0.81	0.74	0.84	0.82	0.80
Math Procedures	0.69	0.71	0.65	0.71	0.47	0.70	0.71	0.72	0.64
Logical Progression	0.82	0.84	0.76	0.80	0.63	0.68	0.86	0.72	0.74
Overall (Sum)	0.94	0.92	0.86	0.89	0.72	0.77	0.90	0.87	0.83

Table 38: Correlations of Rubric Scores with Grader Scores for All Problems

	Overall (N=918)	Bottom Third (0-44%) (N=301)	Middle Third (45-64%) (N=312)	Top Third (65-100%) (N=305)
Useful Description	0.55	0.45	0.15	0.30
Physics Approach	0.73	0.56	0.18	0.47
Specific App.	0.80	0.51	0.24	0.56
Math Procedures	0.64	0.55	0.19	0.23
Logical Progression	0.74	0.60	0.15	0.45
All Categories	0.83	0.69	0.27	0.57

Figure 57: Correlations of Rubric Scores with Grader Scores for All Problems



These overall category correlations suggest that specific application had the strongest relationship to problem grades, and Approach and Logic had a stronger relationship than Description and Math. Another way to look at the relationship of each rubric category score to the problem grade is using a backward multiple regression. The coefficients reported in the following table indicate the contribution of each category score to the problem grade for all sections and problems, by the equation:

Problem grade = B_o + B_1*(Useful Description score) + B_2*(Physics Approach score) + B_3*(Specific Application score) + B_4*(Math Procedures score) + B_5*(Logical Progression score)

where the problem grade is a decimal value (1=100%).

Table 39: Multiple Regression Coefficients for Rubric Category Contributions to Problem Grades
All problems and sections (N=918 solutions)

	Unstandardized Coefficient (B)	Standard Error
B_o (Constant)	.013	.015
B_1 (Description)	.025	.004
B_2 (Physics Approach)	.025	.006
B_3 (Application)	.083	.007
B_4 (Math)	.012	.005
B_5 (Logic)	.041	.008

These coefficients indicate that specific application of physics had the highest contribution to computing a student's problem grade, and logical progression also had a high contribution. Math procedures was lower than Description and Approach. This model accounted for 72±13% of the variance in problem grades (adjusted R Square).

When each section was treated separately in the backward multiple regression, Math and Description were insignificant for section 1 and removed from the model whereas all categories remained in the model for section 2. The model for section 1 accounted for 79±12% of the variance in problem grades and in section 2 accounted for 70±13%. The unstandardized coefficients for the regression model in each section are presented in the following table.

Table 40: Multiple Regression Coefficients for Rubric Category Contributions to Problem Grades by Section

	Section 1 (N=313)	Standard Error	Section 2 (N=605)	Standard Error
B_o (Constant)	-.014	.020	.054	.018
B_1 (Description)			.035	.004
B_2 (Approach)	.029	.009	.028	.008
B_3 (Application)	.111	.010	.062	.008
B_4 (Math)			.017	.007
B_5 (Logic)	.053	.011	.034	.009

In both sections specific application of physics was the highest contribution, and in section 2 the Description and Logic had high contributions whereas in section 1 Logic contributed but Description did not. The Physics Approach contribution was similar for both course sections and Math was low.

These regression results for all problems and sections combined and for each section separately suggest that problem graders give more weight to the specific application of physics in a solution and logical progression and lower weight to mathematical procedures.

Rubric Usefulness for Coaching Students

The rubric can be used to indicate areas of student difficulty for a given problem. For example, rubric scores on this test problem indicated several students in the class received low scores of 1 or 2 for Specific Application of Physics, but received relatively high scores of 4 and 5 for the Physics Approach and Mathematical Procedures. Logical Progression scores were generally in the middle, around a score of 3. For students who appropriately applied a Conservation of Energy approach without an explicit description, the Useful Description was usually scored NA(Solver).

When compared to the standard grading procedure of assigning a single numerical score to a test problem, the rubric provides significantly more information that can be used for coaching students. For example, frequent low scores in a category (such as the low scores in Specific Application) can help focus instruction on modeling this skill and providing guided practice. The rubric only indicates an area of difficulty, however, and a more detailed analysis of the written solution or an interview is required to determine specific difficulties or common responses. This was true for the Test 3 Problem 1 solutions, several of which scored 1 or 2 for specific application. A closer analysis revealed that several students were applying Conservation of Energy to only the middle block, rather than the system of all three blocks.

The rubric also provides instructors information about how the problem statement affects students' problem solving performance, which could be used to modify problems. In the first problem of the third test the problem statement cued on the middle block and student solutions reflected this focus. Additionally, visualization skills were not measured in that problem and the rubric responded with a high frequency of NA(Solver) scores in the description category.

Rubric Usefulness for Problem Selection

The rubric was applied to a range of physics topics tested throughout the semester without difficulty. However, there were some characteristics of problems that did seem to affect the generalizability and meaningfulness of the rubric scores. When processes are not measured for a problem (such as when the description or physics principle is provided), the rubric produces the appropriate Not Applicable scores which shows that this test does not probe that dimension of student learning. If a question is an exercise rather than a problem for the students the rubric produces high scores but makes it difficult to identify student conceptual difficulties for that topic.

The analysis of written work also indicated some characteristics of problems can mask the nature of a student's problem solving processes, such as explicit prompts for procedures or physics cues. For example, a question on the second test (for section 1) explicitly prompted students to draw a free-body diagram in the problem statement and more than half of students did so correctly. In a non-prompted version of the question with a different section of students, fewer included a description of the problem. There was also some indication from the two different versions of the second test problem that a symbolic problem statement makes it more difficult for a student to express their solution in terms of known quantities. In summary, when interpreting rubric scores it is important to consider the structure of the problem and possible bias in problem characteristics.

Degree of Independence of Rubric Categories

One test of the internal structure of the rubric is to assess the degree of independence of the rubric categories, or the extent to which scores on the five categories of the rubric are correlated with each other. Since the rubric scores were shown in the previous section to be correlated with problem grades, the inter-category correlations are computed as a partial correlation that controls for (or "partials out") this relationship to problem grades. The inter-category correlations are computed for each test problem and section separately.

The following tables report correlations among the category scores on Test 1 Problem 2 for each section when controlling for problem grades (kinematics problem in which a pigeon is hit by a football). In the first section, the only statistically significant relationship at the 0.05 level is between Math Procedures and Logical Progression ($R=0.489$, $p=0.001$). In the second section there is a significant relationship between Math Procedures and Logic ($R=0.501$, $p<0.001$), Specific Application and Logic ($R=0.279$, $p=0.004$), Physics Approach and Logic ($R=0.233$, $p=0.019$) and a marginal relationship between Specific Application and Math ($R=0.196$, $p=0.046$). These results indicate that Logic and Math are correlated in both sections, for section two the Logic category is correlated with several other categories, and in both sections the relationships between Useful Description scores and the other categories are not statistically significant.

Table 41: Inter-category Correlations for Test 1 Problem 2 (Section 1, N=48)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	-.328	.245	-.209	-.159
	Sig.	.	.062	.163	.236	.370
Physics Approach	R	-.328	1.000	.203	.097	.260
	Sig.	.062	.	.185	.538	.089
Specific Application	R	.245	.203	1.000	.166	.269
	Sig.	.163	.185	.	.269	.068
Math Procedures	R	-.209	.097	.166	1.000	.489
	Sig.	.236	.538	.269	.	.001
Logical Progression	R	-.159	.260	.269	.489	1.000
	Sig.	.370	.089	.068	.001	.

Table 42: Inter-category Correlations for Test 1 Problem 2 (Section 2, N=110)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.092	-.001	.064	.143
	Sig.	.	.441	.992	.585	.216
Physics Approach	R	.092	1.000	.049	-.008	.233
	Sig.	.441	.	.629	.939	.019
Specific Application	R	-.001	.049	1.000	.196	.279
	Sig.	.992	.629	.	.046	.004
Math Procedures	R	.064	-.008	.196	1.000	.501
	Sig.	.585	.939	.046	.	3.87E-8
Logical Progression	R	.143	.233	.279	.501	1.000
	Sig.	.216	.019	.004	3.87E-8	.

The following table reports partial correlations among the category scores on Test 2 Problem 1 (section 1) when controlling for grades. On this problem, the strongest relationship is between Math and Logic ($R=0.539$, $p<0.001$). Relationships that are statistically significant include Specific Application and Logic ($R=0.357$, $p=0.001$), Physics Approach and Logic ($R=0.271$, $p=0.011$), and Physics Approach and Math ($R=0.255$, $p=0.023$). Once again, the correlations between Useful Description and all other categories are not statistically significant whereas Logic is correlated with several other categories.

Table 43: Inter-category Correlations for Test 2 Problem 1 (Section 1, N=92)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.087	-.038	-.129	.102
	Sig.	.	.442	.735	.263	.358
Physics Approach	R	.087	1.000	.154	.255	.271
	Sig.	.442	.	.157	.023	.011
Specific Application	R	-.038	.154	1.000	.129	.357
	Sig.	.735	.157	.	.246	.001
Math Procedures	R	-.129	.255	.129	1.000	.539
	Sig.	.263	.023	.246	.	1.25E-7
Logical Progression	R	.102	.271	.357	.539	1.000
	Sig.	.358	.011	.001	1.25E-7	.

The following table reports partial correlations among the category scores on Test 2 Problem 1 (Section 2) when controlling for problem grades. On this problem, the strongest relationship is between Specific Application and Physics Approach ($R=0.501$, $p<0.001$). Once again the correlations between Logic and most other categories are statistically significant and Useful Description with other categories are low (near zero) and not statistically significant.

Table 44: Inter-category Correlations for Test 2 Problem 1 (Section 2, N=160)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	-.071	-.002	-.090	-.020
	Sig.	.	.403	.977	.265	.802
Physics Approach	R	-.071	1.000	.501	.165	.433
	Sig.	.403	.	2.24E-10	.050	7.14E-8
Specific Application	R	-.002	.501	1.000	.213	.319
	Sig.	.977	2.24E-10	.	.008	5.01E-5
Math Procedures	R	-.090	.165	.213	1.000	.418
	Sig.	.265	.050	.008	.	5.03E-8
Logical Progression	R	-.020	.433	.319	.418	1.000
	Sig.	.802	7.14E-8	5.01E-5	5.03E-8	.

The following tables report partial correlations among the category scores Test 2 Problem 2 for each section when controlling for problem grades (Attwood problem). On the version of the problem in section 1, the strongest correlations are between Math and Logic ($R=0.558$, $p<0.001$) and Physics Approach and Specific Application ($R=0.470$, $p<0.001$). In section 2 the strongest correlations are also between Math and Logic ($R=0.536$, $p<0.001$) and Physics Approach and Specific Application ($R=0.610$, $p<0.001$), where the approach and application correlation is even higher than in section 1. Contrary to other problems, all of the category scores are significantly correlated to each other, with the exception of Useful Description and Logic in both sections and Useful Description and Math in section 1.

Table 45: Inter-category Correlations for Test 2 Problem 2 (Section 1, N=81)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.431	.238	.149	.155
	Sig.	.	1.58E-04	.034	.202	.172
Physics Approach	R	.431	1.000	.470	.319	.456
	Sig.	1.58E-04	.	3.17E-05	.008	6.33E-05
Specific Application	R	.238	.470	1.000	.329	.402
	Sig.	.034	.000	.	.004	2.67E-04
Math Procedures	R	.149	.319	.329	1.000	.558
	Sig.	.202	.008	.004	.	1.92E-07
Logical Progression	R	.155	.456	.402	.558	1.000
	Sig.	.172	6.33E-05	2.67E-04	1.92E-07	.

Table 46: Inter-category Correlations for Test 2 Problem 2 (Section 2, N=156)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.389	.392	.172	.145
	Sig.	.	2.58E-06	1.38E-06	.048	.086
Physics Approach	R	.389	1.000	.610	.389	.352
	Sig.	2.58E-06	.	3.13E-16	2.86E-06	1.37E-05
Specific Application	R	.392	.610	1.000	.372	.244
	Sig.	1.38E-06	3.13E-16	.	3.99E-06	.002
Math Procedures	R	.172	.389	.372	1.000	.536
	Sig.	.048	2.86E-06	3.99E-06	.	4.62E-12
Logical Progression	R	.145	.352	.244	.536	1.000
	Sig.	.086	1.37E-05	.002	4.62E-12	.

The following tables report partial correlations among the category scores on Test 3 Problem 1 for each section when controlling for grades. In both sections, the strongest relationship is between Math and Logic ($R=0.460$ and $R=0.506$, $p<0.001$). Contrary to other problems, the correlation between Physics Approach and Application is not significant for this problem. There are some differences between the two sections; in section 1 there is a significant correlation between Description and Math ($R=0.391$, $p=0.010$) not observed in section 2, and in section 2 there is a significant correlation between Description and Approach (R=0.269, p=0.005) and between Logic and every category ($p<0.05$) that is not observed in section 1.

Table 47: Inter-category Correlations for Test 3 Problem 1 (Section 1, N=92)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.069	-.052	.391	.245
	Sig.	.	.656	.735	.010	.104
Physics Approach	R	.069	1.000	.191	.075	.259
	Sig.	.656	.	.076	.493	.015
Specific Application	R	-.052	.191	1.000	-.076	.102
	Sig.	.735	.076	.	.477	.337
Math Procedures	R	.391	.075	-.076	1.000	.460
	Sig.	.010	.493	.477	.	5.72E-06
Logical Progression	R	.245	.259	.102	.460	1.000
	Sig.	.104	.015	.337	5.72E-06	.

Table 48: Inter-category Correlations for Test 3 Problem 1 (Section 2, N=179)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.269	.077	.084	.223
	Sig.	.	.005	.421	.382	.018
Physics Approach	R	.269	1.000	.119	.222	.334
	Sig.	.005	.	.129	.005	1.18E-05
Specific Application	R	.077	.119	1.000	.145	.266
	Sig.	.421	.129	.	.061	3.38E-04
Math Procedures	R	.084	.222	.145	1.000	.506
	Sig.	.382	.005	.061	.	2.52E-12
Logical Progression	R	.223	.334	.266	.506	1.000
	Sig.	.018	1.18E-05	3.38E-04	2.52E-12	.

An important goal of this part of the study was to assess the relationships between categories, or the internal structure of the rubric. Table 49 and Table 50 show the overall values for the Pearson correlation coefficient for each rubric category compared to the other categories when controlling for problem grades for section 1 and section 2. The values in the table indicate that the highest correlation in both sections is between Math and Logical Progression (R=0.523 and R=0.568, $p<0.001$). In section 1 the correlation between Physics Approach and Specific Application is lower than in section 2 but both are significant (R=0.302 and R=0.473, $p<0.001$). In sections 1 and 2 the correlation between Useful Description and Math are not significant (R=0.033 and R=0.030, $p>0.05$). Previous research (Foster, 2000) found a significant correlation between Logic and Math regardless of the course structure (R~0.70 without controlling for problem grades). In a reformed course he also found a significant correlation between Specific Application and Physics Approach, and Logic with both the Approach and Application.

Since both the Physics Approach and Specific Application of Physics categories involve physics (and it is difficult to correctly apply an incorrect approach), it is not entirely surprising that the scores are correlated. By construction it is expected that the Description and Math aspects of a solution are relatively independent of the Physics categories, whereas the Logical Progression is an "overall" coherence and consistency of the solution that could be related to each of the other categories.

Table 49: Inter-category Correlations of Rubric Scores for All Problems (Section 1, N=313)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.260	.168	.033	.168
	Sig.	.	5.95E-05	.008	.617	.008
Physics Approach	R	.260	1.000	.302	.161	.343
	Sig.	5.95E-05	.	1.38E-07	.007	1.77E-09
Specific Application	R	.168	.302	1.000	.166	.297
	Sig.	.008	1.38E-07	.	.004	1.05E-07
Math Procedures	R	.033	.161	.166	1.000	.523
	Sig.	.617	.007	.004	.	2.73E-22
Logical Progression	R	.168	.343	.297	.523	1.000
	Sig.	.008	1.77E-09	1.05E-07	2.73E-22	.

Table 50: Inter-category Correlations of Rubric Scores for All Problems (Section 2, N=605)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.098	.109	.030	.007
	Sig.	.	.035	.016	.512	.881
Physics Approach	R	.098	1.000	.473	.322	.469
	Sig.	.035	.	2.37E-32	1.98E-14	1.07E-31
Specific Application	R	.109	.473	1.000	.340	.404
	Sig.	.016	2.37E-32	.	4.61E-17	6.59E-25
Math Procedures	R	.030	.322	.340	1.000	.568
	Sig.	.512	1.98E-14	4.61E-17	.	1.04E-50
Logical Progression	R	.007	.469	.404	.568	1.000
	Sig.	.881	1.07E-31	6.59E-25	1.04E-50	.

Table 51: Inter-category Correlations of Rubric Scores for All Problems and Sections (N=918)

		Useful Description	Physics Approach	Specific Application	Math Procedures	Logical Progression
Useful Description	R	1.000	.156	.118	.049	.054
	Sig.	.	3.46E-05	.001	.195	.141
Physics Approach	R	.156	1.000	.426	.274	.433
	Sig.	3.46E-05	.	7.92E-39	1.40E-15	3.96E-40
Specific Application	R	.118	.426	1.000	.286	.377
	Sig.	.001	7.92E-39	.	7.03E-18	5.57E-32
Math Procedures	R	.049	.274	.286	1.000	.552
	Sig.	.195	1.40E-15	7.03E-18	.	4.41E-71
Logical Progression	R	.054	.433	.377	.552	1.000
	Sig.	.141	3.96E-40	5.57E-32	4.41E-71	.

Summary of Scoring Written Solutions on Exams

The goals for this part of the study included obtaining validity evidence for the response processes of students on written solutions to physics test problems, obtaining validity evidence for the rubric's generalizability (applicability to multiple topics in a semester-long mechanics course), obtaining validity evidence of the external and internal rubric structure using correlational data, and to propose instructor uses of the rubric.

The response process categories of the rubric were observed in the student solutions scored. The external structure of the rubric was assessed by comparing scores on the rubric to course graders' scores for the same solutions. The overall correlation coefficient indicated a strong agreement between a total rubric score and the graders' scores. The correlations for each category suggested that graders gave more weight to the aspects of a solution related to the selection of principles and specific equations used, and lower weight to other aspects of the solution (such as a description and math procedures).

The internal structure of the rubric was measured with correlations between each of the categories. This analysis indicated a strong relationship between Physics Approach and Application of Physics for most problems, consistent with previous results (Foster, 2000) and a strong relationship between Math and Logical Progression. It also indicated that Useful Description scores are relatively independent of Math and Logic, whereas Logical Progression had a strong relationship to all categories with the exception of Description.

The rubric was generalizable (applicable) to all of the physics topics from mechanics scored in this analysis, including motion under constant acceleration in one and two dimensions (kinematics), Newton's laws of motion, work, and conservation of energy. The rubric was applied to problems that differed in complexity and features (such as multi-part problem statements and diagrams provided or missing), but some problem features influenced the interpretation of scores on the rubric. Although the rubric can be used with problems that are "easy" and have explicit prompts, the extent to which scores adequately represent students' problem-solving skills is questionable since in that case, the question is not likely to be a problem for the student.

This study also suggested possible uses of the rubric for instructors and researchers, or its utility. Plots of the frequency of rubric scores for each category of a problem, such as those provided for the problems analyzed in this study, can be used to indicate areas of student difficulty for that particular problem and physics topic and determine features of problems that are or are not useful. This only indicates an area of difficulty, however, and a closer analysis of student solutions is required to further interpret the nature of common student difficulties.

Student Problem-Solving Interviews

Introduction

An important source of evidence for validity based response processes is student problem-solving interviews. In an interview, students are asked to solve physics problems while their actions and voice are recorded. After completing the problem, they are asked to explain their reasoning to an interviewer. The written work is scored using the rubric and then rescored using the interview information. The interview transcripts are also analyzed for evidence of the rubric categories. This gives an indication of whether the processes engaged in by students during problem solving are reflected in the rubric, and the extent to which written solutions are indicative of a student's problem-solving process.

The research questions addressed in this study are listed below, where the number and letter refer to the specific Research Question stated in Chapter 1:

- 1b) To what extent do scores on the rubric reflect the problem-solving processes undertaken by a solver? (*response processes*)
- 1c) To what extent do scores on the rubric support inferences about students' problem-solving skills from other measures of their performance? (*external structure*)
- 3a) To what extent can the rubric distinguish between more- and less-skilled problem solvers?
- 3b) How authentic are the assessment's goals, tasks, and constraints?

In order to obtain validity evidence for response processes and measures of student performance, this interview analysis addresses the following specific questions:

- How does student written work correspond to their self-reported thought processes when interviewed about solving a problem?
- To what extent do the five rubric category processes characterize their self-reported problem-solving processes in the interview?

- What problem-solving processes are self-reported during a problem-solving interview that are not explicitly measured by the rubric?
- How does a written problem solution constructed during the interview correspond to the student's written solutions on similar problems during a test?

Data Collection Procedures

Participants in the problem-solving interviews were students enrolled in an introductory calculus-based mechanics course for scientists and engineers, in a different semester and with a different instructor than the analysis of written exam solutions. Of the 238 students in this course, 13 volunteered to participate in a one-hour problem-solving interview at the end of the term. Ten of these students scheduled a session time. Four interviews took place in the last week of the semester, 2 during finals week, and 2 after finals week. Two students cancelled giving a total of 8 interviews: 7 males and 1 female. In the course approximately one in five students was female, so the sample of women in this study is slightly lower than the class representation.

The final course grades of these eight participants indicate they performed higher than the course average and may not accurately represent the problem solving proficiency of their class. Students 2, 5, 6, and 8 received an A in the course; Students 3 and 4 received an A-; Student 1 earned a B+; and Student 7 a B. The average grade for this class was a B.

During the problem-solving interviews students were asked to work on physics problem(s) while being video and audio taped. They used large sheets of paper and a black marker to record their solution(s) so that the solutions were visible on the video.

Participants were asked to talk out loud while working on a problem if that was comfortable for them, or they could wait and explain their solution at the end. Only one student (Student 1) opted to talk out loud while working. Students were provided a copy of the instructor's equation sheet from the course and their calculator (See the equation sheet in Appendix 7).

Note that the students were enrolled in the same course as the analysis of written exam solutions (introductory calculus-based physics for science and engineering), but in a different term and with a different instructor. The exam problems written by this instructor were not text-book style problems, but were *context-rich* (Heller & Hollabaugh, 1992). They were written to require several decisions from the solver and did not include any illustrations. The problem tasks for the interview were selected to look similar to ones from tests and group problem-solving sessions in their course. The first and most involved task is given in Figure 58. This problem was adapted from previous research (Henderson et al., 2004; Yerushalmi, Henderson, Heller, Heller, & Kuo, 2007). Problem features include: the target of the problem is not explicitly stated, a combination of at least two principles is necessary, and the solver must infer or assume some information. This problem also has the characteristic that it is possible to obtain a correct answer with incorrect or incomplete reasoning (Henderson et al., 2004).

The remaining two problems were designed to be shorter (in anticipation of little available time). One purpose of having additional problems was to make sure that students do not work quickly to finish the interview early, and another purpose was to provide additional opportunities to access the processes of students who finished quickly and may have automated many of their processes. These problems maintained the same context-rich format but only required a single physics principle. Problem two required finding the spring constant of a bungee cord and Problem three involved a car crash at the bottom of a cliff. Students were only given an additional problem if sufficient time (at least twenty minutes) remained after they had explained their reasoning for the previous problem. The interview materials, including the problem statements for all of the problem-solving tasks, are provided in Appendix 7.

Figure 58: First Problem-Solving Interview Task

You are working at a construction site and need to get a 14-N bag of nails to your co-worker standing on the top of the building (9 meters from the ground). You don't want to climb all the way back up and then back down again, so you try to throw the bag of nails up. Unfortunately, you're not strong enough to throw the bag of nails all the way up so you try another method. You tie the bag of nails to the end of a 65-cm string and whirl the string around in a vertical circle. You try this, and after a little while of moving your hand back and forth to get the bag going in a circle you notice that you no longer have to move your hand to keep the bag moving in a circle. You think that if you release the bag of nails when the string is horizontal to the ground that the bag will go up to your co-worker. As you whirl the bag of nails around, however, you begin to worry that the string might break, so you stop and attempt to decide before continuing. According to the string manufacturer, the string is designed to hold up to 500 N. You know from experience that the string is most likely to break when the bag of nails is at its lowest point.

After solving a problem to their satisfaction, each student was asked to go back and explain their solution to the researcher. Questions from the semi-structured interview included the following:

- When you read through the problem, what was the first thing you thought about?
- What did you think about next?
- What was the first thing you wrote down?
- What did you think this question was asking you to find?
- How did you decide to use ___ ? (physics concept or equation)
- If you were solving this problem on an exam, what would you hand in to be graded?
- Have you solved a problem like this before in your physics class? How is that problem similar to or different from this problem?
- While you were working on the problem, was there anything you did in your head that you didn't write down?

The audio files for the eight interviews were transcribed and the written protocols were analyzed using Q.S.R. NVivo® software (http://www.qsrinternational.com/) prescribed code categories or "nodes" that corresponded to the process categories on the rubric and designated responses to specific questions asked during the interview (stated above).

Time Spent on each Problem

One way to characterize the problem-solving behavior of the interview students is to consider the average time they spent working on each problem. The average times students spent working on each problem is listed in Table 52, along with the total number of problems completed during the interview session. The times for the first problem ranged from 6 minutes to 26 minutes, excluding any additional time spent modifying the solution during questioning. For students who solved the problem very quickly (Student 2 and 5), this suggests they had automated some decision processes and perhaps this problem was actually an exercise for them, not a true problem (Schoenfeld, 1985). Students 4 and 7 did not reach a satisfactory answer for the problem and chose to stop at the reported time to explain their thinking. This indicates that most students were unwilling to spend more than 26 minutes working on the problem, and had either reached their tolerance level of frustration at that point or had exhausted their resources for attempting the problem.

Table 52: Time Spent Working on each Problem and Total Number of Problems Completed for Each Interview Student

Interview Student	Time Spent on First Problem	Time Spent on Second Problem	Time Spent on Third Problem	No. Problems Completed
Student 1	14 min 55 sec	11 min 36 sec	--	2
Student 2	6 min 50 sec	6 min 25 sec	10 min 29 sec	3
Student 3	24 min 30 sec	--	--	1
Student 4	26 min 17 sec (+13 min 5 sec)	--	--	0
Student 5	6 min 7 sec (+ 40 sec)	4 min 11 sec	--	2
Student 6	14 min 54 sec	9 min 2 sec	--	2
Student 7	20 min 27 sec (+ 5 min 33 sec)	--	--	0
Student 8	9 min 56 sec	--	15 min 31 sec	2

(Times in parentheses indicate that students changed or added to their solution during the interview questions)

Student 2 was the only student to successfully complete the problem with correct physics reasoning. Student 5 was successful after correcting an error discovered during the interview questioning. Students 6 and 8 obtained the correct answer, but gave incomplete (possibly incorrect) reasoning for parts of the solution. Students 1 and 3 completed the problem using inappropriate physics. Students 4 and 7 did not obtain a final answer in the available time and their approaches included a mixture of confused physics ideas. Two students interpreted the question as finding the height the bag would travel vertically with the maximum string tension value, whereas most students focused on solving for a force. Student four's goal was unclear.

Coding Verbal Statements

All eight students began by drawing a picture of the problem situation and summarizing the information provided. When asked what they thought about first while reading the problem, three students (1, 6, and 8) mentioned this problem description process:

S1: *I was just trying to get an image in my head cuz a lot of times these are written so weird that you have to re-read it three or four times to even figure out what you're trying to see in the picture, so you can start marking down values. But, until you kinda dig through it, you know, even getting a picture in your head, is just kinda confusing.*

S6: *The first thing I thought about was just that it mentioned that the string was most likely to break when the um, bag was at it's lowest point…I can like get a diagram of what that looks like to start. Um. To have some sort of basis to like, base off where I'm gonna go from there. So I kinda just, just kinda to get myself in the mindset of the problem, just kinda drew that even though that didn't prove to be, the most helpful diagram, um. Just something to get started.*

S8: *Well the first thing I uh, thought about was um…I just diagrammed it. I didn't know what to think initially. I just wrote down all the data, diagrammed it.*

Q: *Okay, when you say 'diagrammed' can you tell me more of what you mean by that?*

S8: *Like, I just like, visualized it. Maybe the height had to be from the center of the, center of the thing. I wasn't quite sure exactly what it was but when I drew a picture it made more sense to me.*

Students 3 and 4 mentioned that the problem made them think about circular motion, and Student 7 mentioned parabolic motion because it was like "throwing something". These comments suggest that students were cuing on particular aspects of the problem statement and immediately attempting to categorize it as a familiar type of problem (such as a force, circular motion, or projectile problem).

S3: *That it was like a force problem. And the circular stuff stood out. So you knew that you had to use angular equations. [inaudible] traveling in a circle.*

Q: *So what made you think of forces and circular motion?*

S3: *Umm, well, because of the tension in the string is five hundred Newtons at maximum. And so you're trying to figure out, like, what, if, whether or not the string can hold it at a certain point, so that you needed to use forces for that.*

S4: *Um, that first thing it was gonna be hard. Cuz I'm so bad at the circular, angular momentum. But, at first I wasn't really sure what to do.*

S7: *Um, well first I thought there was gonna be some sort of like parabolic motion or something from here. Um. But then it didn't, it didn't give me enough information to get anything from it…*

Q: *So what made you think of parabolic motion?*

S7: *Well, in the first part they're throwing something up vertically. I guess it's not really parabolic but. Have something go up, and then coming back down.*

Student 5 said they first thought about what the question was asking them to find. The student repeated this response after the second problem, indicating that finding the question is a general procedure they engage in when reading a physics problem statement.

S5: *I want to know and, what's um, I want to find the question. What I want to know in this problem. And like, it's, I find that I want to know how, the height that the bag can reach. The maximum height it can reach and. So, this is the first thing I want to know.*

S5: *Yeah, um. It's the same as that, that question. I want to find what's this question asking me to do. And that, that's it. The first thing I do.*

Student 2 was the only student to explicitly mention physics principles as something they thought about first. It is interesting to note that Student 2 was also the only student to complete the problem both rapidly and correctly on a first attempt.

S2: *That it is not hard…and I should use uh, the equation of the motion and uh, the conservation of energy in this problem.*

The total number of transcript passages assigned to each coding category by researcher 1 is reported in Table 53, and by researcher 2 is reported in Table 54. These tables also report the average number of passages for students 2 through 8 and student 1. Statements that resulted from clarification prompts of a previous statement (can you say more about that or what do you mean by that) were coded as a single statement. As seen from the Average column, most statements pertaining to the rubric categories were coded as evidence of specific application of physics or logical progression. Specific application statements were usually references to particular physics equations and the quantities specific to the problem (such as a velocity, force, or distance). For example, when prompted student 8 stated the velocity of the bag was the same at all points of the swing, which was not obvious from their written solution. Logical progression statements referred to overall steps taken in the solution and explaining reasoning for those steps. The following statement from Student 5 is an example of this category.

S5: *Um, first uh, I find out what I want to know. And I find out what I already know. And I need to build a relationship between them… in this problem I want to know the height so I need to know the velocity. And in order to find the velocity I need to know* *the, use Newton's second law I can find the, the relationship between the force and the velocity. So I build the connection with the known things and the other things.*

In contrast, Student 4 describes their procedure:

S4: *Pretty sure I'm lost.*

Q: *Can you say more about that? What are you, what are you thinking right now?*

S4: *I can't really, I don't really know. I was just trying to put everything I know down, and then seeing what equations eliminate stuff. Um, and what I could plug in. And that didn't get me very far so far.*

As summarized in Table 53 there were 276 total passages coded in the eight interview transcripts that pertained to the rubric categories. On average, students made 32 rubric-related statements. Student one talked out loud while solving the problem and had more: 55 rubric-related statements. The NVivo codings for responses to specific questions were used to summarize verbal statements and the number of these statements that pertained to each question node are not reported here.

Students who spent a lot of time on the problem (3, 4, and 7) had a higher number of statements for some categories, because they attempted several different approaches. Students 3 and 4 made several references to specific physics equations and quantities during these solutions attempts (coded as Specific Application) and performed more mathematical calculations, whereas Student 7 made a higher number of statements pertaining to what they should try next and how that might help them reach the goal (coded as Logical Progression). Student 2 had a high number of Specific Application and Math statements, but this could have resulted from them completing three problems during the interview instead of one.

Table 53: Number of Transcript Passages Assigned to each Coding Node (Researcher 1)
The average represents students #2-8.

Category	S1	S2	S3	S4	S5	S6	S7	S8	Avg
Useful Descrip.	9	5	3	6	6	5	5	4	5
Physics Appr.	10	4	2	3	5	6	5	5	4
Specific Applic.	18	12	12	12	6	6	6	10	9
Math Procedures	7	7	5	6	3	3	4	4	5
Logical Prog.	11	6	9	8	8	9	13	8	9
TOTAL	**55**	**34**	**31**	**35**	**28**	**29**	**33**	**31**	**32**

*Student 1 talked aloud while working on the problems and generally had more statements coded than the other students.

Table 54: Number of Transcript Passages Assigned to each Coding Node (Researcher 2)

Category	S1	S2	S3	S4	S5	S6	S7	S8	Avg
Useful Descrip.	5	1	2	3	3	2	3	1	2
Physics Appr.	5	3	0	1	3	3	7	4	4
Specific Applic.	3	2	0	1	1	3	3	5	3
Math Procedures	1	3	0	1	1	2	1	3	2
Logical Prog.	3	1	2	5	3	1	5	7	3
TOTAL	**17**	**10**	**4**	**11**	**11**	**11**	**19**	**20**	**12**

When asked what they would hand in for a graded exam, all students gave examples of adding more explanation in words to help the grader understand the solution, and most said they would write a picture.

S6: *Um, I would start with two diagrams at the top kind-of. Showing all of this basic information. And then…I would kind-of explain maybe in a phrase or something what each of these different sections were doing, and I'd kind-of put them in a logical order as opposed to here where they're, it's a little bit um, jumping all over the page…just so that it's clear.*

A check of the interview students' final exams revealed that only half of the students actually included words of explanation to help the grader understand their solution. In particular, students 2, 5, 6, and 8 typically included these explanations on their final exam.

When explicitly questioned about what they did in their head and didn't write down, Student 7 described:

S7: *Usually the only thing I write down right away is a picture, so I can see what's going on. Um. But then I'll just have in my head like, if I go from this equation and then I get an answer I can put it into this equation, and then into that equation…Generally I tend to do too much in my head and not write enough stuff down, that's the only, that seems to be where I go wrong.*

Rubric Scores for Written Solutions

The following Figure shows a bar graph of the rubric scores on the first written problem solution from the interview task in each category for the eight interview students. Only the written solution was used for this scoring. Missing scores indicate the category was scored Not Applicable for that solver or NA(Solver). Students 2, 5, and 8 had NA(Solver) scores for their Physics Approach and Students 2 and 8 had NA(Solver) scores for Mathematical Procedures. Students 4 and 7 did not reach a solution in the available time, and their confused problem solving is reflected in their lower rubric scores for most categories and particularly lower scores in their Specific Application of Physics.

These rubric scores only reflect students' written problem solving processes. When their verbal statements were considered in addition to their written work, most students' scores remained the same with a few exceptions. Student 8's Specific Application of Physics score changed from a 5 to a 4 because of incorrect reasoning for the velocity term that was not apparent from writing alone, but was evident from a verbal response indicating the velocity was the same at all points of the swing. Student 2's Logical Progression score increased from a 4 to a 5 because of additional verbal evidence for their reasoning processes. For the students with NA(Solver) scores in Physics Approach and Mathematical Procedures, their scores changed to a 5 as a result of explicit evidence for these categories in verbal statements explaining their procedures. These changes are reflected in Figure 60.

Figure 59: Rubric Scores for Interview Students' First Written Solution

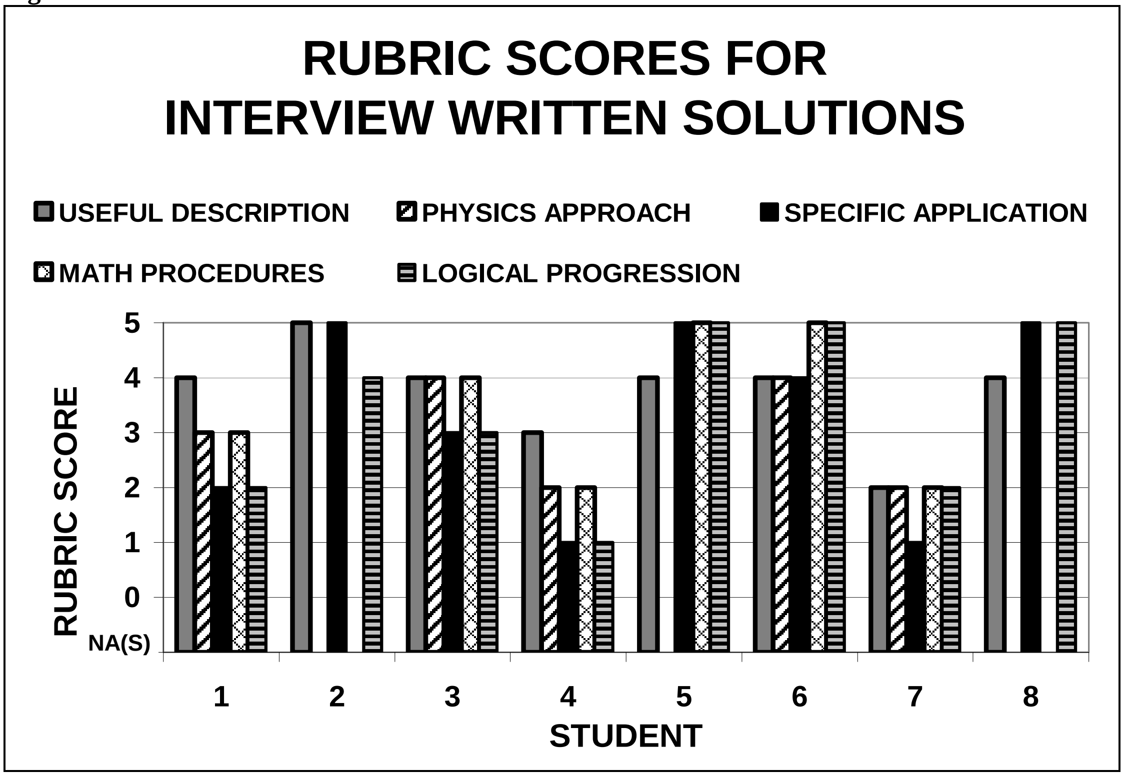
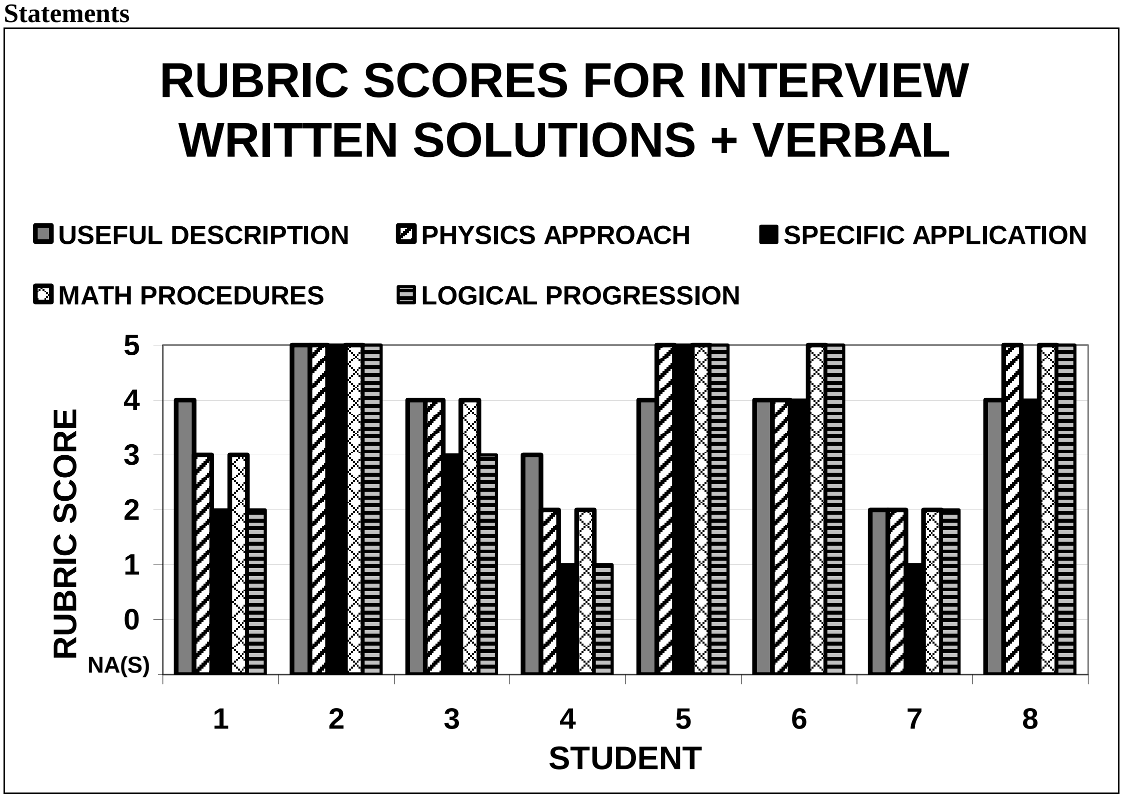


Figure 60: Rubric Scores for Interview Students' First Written Solution, Considering Verbal Statements



Summary of Interviews

In this study eight introductory university physics students each participated in a one-hour interview to determine how well the rubric represents their problem solving processes by comparing their written problem solutions to interview data. For the first problem presented to them, all students wrote down a description, physics equations, and mathematical operations. The scores in the rubric categories were first inferred only from what was written. In the interview the students were explicitly asked about their thought processes (not their rubric categories) and their statements were mapped onto the rubric. The transcripts contained explicit evidence for all five of the rubric categories, with specific application of physics and logical progression having the most coded statements. Since students were prompted to explain their reasoning verbally during the interviews, evidence of logic was much more prominent than is typical of their written work however this was adequately inferred from the written work. Also, there was explicit evidence for the physics approach (selecting appropriate physics principles) that is adequately inferred from the specific equations that are written down. In general, rubric scores of students' written solutions alone were identical with the verbal evidence.

In summary, the process categories of the problem-solving rubric are observed in both written work and verbal interview protocols. There is, of course, much more fine-grained information in the interviews. Also, although the students state that much of what they hand in on a test is a "cleaned up" version of a problem solution that may contain more information than their interview papers, examining their test papers showed no evidence of this. From these interviews, we conclude that rating student written solutions using a problem solving rubric generally gives an accurate, though course-grained, view of their actual problem solving processes.

Summary of Data Collection and Analysis

The goal of this study is to design a simple, fast, and easy to use problem-solving measure for written solutions to physics problems and establish evidence for validity, reliability, and utility. This measure was developed based on the research literature in the form of a rubric, which assigns a separate score for five expert-like problem-solving processes (useful description, physics approach, specific application of physics, mathematical procedures, and logical progression). The studies of the rubric given in this dissertation indicate that this measure is easy-to-use, provides meaningful information, and produces reasonably valid and reliable scores.

The tests with graduate student raters indicated a reasonable level of score agreement with very minimal and non-invasive training. Of course expert teachers with more extensive training gave an almost total level of agreement. The data from rater comparisons also show that the rubric is applicable to a wide variety of topics from mechanics and electricity and magnetism. The analysis of test solutions from a semester-long introductory physics course indicated that the rubric applies to several of the different physics topics in an introductory mechanics course and that the rubric generally agrees with course graders' scores. It also indicated that some problem characteristics mask student problem solving processes, such as overly explicit procedural prompts and physics cues. The rubric provides more fine grained and meaningful information than standard grading by indicating areas of student difficulty that can be used to focus coaching and improve problem writing.

Interviews with introductory physics students indicated that the categories of the problem-solving rubric were observed and their scoring was the same in both written work and verbal interview protocols. In general, rubric scores of students' written solutions were identical with the verbal evidence for those processes. The study concluded that rating student written solutions using a problem solving rubric generally gives an accurate view of their actual problem solving processes which is more course-grained than can be achieved from an interview process but substantially more fine-grained than traditional grading.

CHAPTER 5: DISCUSSION

Introduction

This section summarizes important results from the studies presented in Chapter 4 and how those results address the research questions stated in Chapter 1, fit with the methodology framework in Chapter 3 and when appropriate, how they relate to the literature review provided in Chapter 2.

Evidence for Validity

Chapter 3 defined validity as the "degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests" (AERA et al., 1999, p.9) and outlined several potential sources of evidence for a validity argument. These sources included the content of an assessment, the response processes of examinees, internal and external structural measures, and generalizability.

The literature review presented in Chapter 2 is one source of evidence for the extent to which the rubric categories are consistent with descriptions of physics problem solving processes, or its *content relevance and representativeness.* Studies with experienced and inexperienced problem solvers in physics highlighted that successful solvers use qualitative descriptions or "Useful Descriptions" (Larkin 1979; 1981a; Larkin & Reif, 1979; Reif & Heller, 1982; Simon & Simon, 1978; 1979), use principle-based "Physics Approaches" (Chi et al., 1981, 1982; de Jong & Ferguson-Hessler, 1986; Hardiman et al., 1989), have procedures for the appropriate application of principles or "Specific Application of Physics" (Larkin 1979, 1981a, 1981b; Larkin et al. 1980a; Eylon & Reif, 1984), are skilled at the use of "Mathematical Procedures" to carry out their solution plan (Larkin et al. 1981b; Reif et al., 1976; Van Heuvelen, 1991a), and have strategies for monitoring progress and evaluating results that can contribute to the "Logical Progression" of a solution, or the communication of an organized reasoning pattern (Chi, 2006; Reif & Heller, 1982; Singh, 2002).

A review of existing problem-solving scoring instruments in physics indicated that many of them share similarities with the rubric developed for this study (Blue, 1997; Foster, 2000; Harper, 2001; Huffman, 1997; Murthy, 2007; Ogilvie, 2007) and the categories are relatively consistent with faculty beliefs about problem-solving (Yerushalmi et al., 2007). Comments from raters in the training studies focused on their difficulties interpreting the categories and scores, and did not indicate any strong disagreements with selecting to measure those five processes (categories). Overall, the rubric content was judged to be consistent with descriptions of physics problem-solving by researchers and instructors.

Interviews with students and an examination of written solutions on exams suggested that the rubric adequately reflects students' problem-solving *response processes*. There was evidence of the rubric process categories in the papers scored, and in the verbal statements made by interview students, with high agreement between the two.

A comparison of rubric scores and grader scores on written solutions to exams indicated agreement between the rubric scores and this *external* measure of students' skills. Relationships were strongest for the Physics Approach, Specific Application of Physics, and Logical Progression categories and slightly weaker for the Useful Description and Math Procedures. There was almost always exact agreement between interview rubric scores of written work and verbal statements, with very few instances in which students' reasoning was not apparent from their written work alone.

Correlations between categories of the rubric (its *internal structure*) indicated several overlapping solution aspects that were consistent with expectations. There was a strong relationship between Physics Approach and Specific Application of Physics, whereas Useful Description and Math Procedures were more independent. The Logical Progression category was correlated to several other categories, which is not surprising because this represents an "overall" coherence and consistency to the solution that could have been reflected in other categories.

The analysis of written exams from one semester of a mechanics indicated the rubric is *generalizable* to several physics topics in introductory university physics courses including kinematics, Newton's second law, circular motion, and conservation of energy. The preliminary study also scored problems from conservation of momentum and oscillations, and the studies with training raters included one problem from electricity and magnetism. The rubric was only tested with introductory physics student populations (algebra-based and calculus-based) and instructor solutions, and future work could explore more advanced physics courses and a wider range of topics in electricity and magnetism. The rubric was applicable to all problems tested in these studies, however some problems did not measure a process (such as providing a Description). The rubric responded to this with appropriate Not Applicable scores, but it should be noted that such problems do not provide meaningful information that can be used for research or modifying instruction. Other problem features such as explicit prompts and cues were also found to mask problem-solving processes and threatened the validity of score interpretations made by the rubric.

The second study with raters included revised documentation for the rubric that outlined the purpose of the assessment and distinguished it from grading (*consequences of use*). Throughout the studies reported in Chapter 4, the rubric was never provided to students and it was not tested for this purpose. An extension of this work could include modifications to the rubric language to make it appropriate for students and asking them to use the rubric to self-assess their problem solutions.

Evidence for Reliability

Reliability in this study was interpreted to be the agreement of scores and score interpretations for multiple raters using the rubric. Reliability can be assessed using several statistical measures, including percent perfect agreement, percent agreement within one score, Cohen's kappa, and quadratic weighted kappa, among others. Although the measure kappa is subject to some disagreement among researchers (see Appendix 2) it provided some information regarding different degrees of agreement in various parts of the study.

The preliminary study with two raters indicated that high agreement was obtained after scoring several solutions (N=160) and discussing rubric interpretations periodically. The agreement was judged to be high because kappa indicated substantial agreement (>0.60) in each of the rubric categories and overall (Landis & Koch, 1977) and agreement within one score was over 97±1% in every category. The first study with training experienced graduate students indicated that agreement with a researcher's scores on mechanics and E&M student solutions was fair before the written training (weighted kappa 0.23±0.04 and 0.31±0.05) and increased to moderate agreement (weighted kappa 0.41±0.04 and 0.43±0.04) as a result of the training materials. Although the kappa measures indicated a statistically significant reliability overall ($p<0.001$) on both the mechanics and E&M solution scores, the percent exact agreement was 44±4% and 45±4% after the training which is lower than is desirable for research scoring purposes. The percent agreement within one score was higher (>85%) which suggests the reliability of the rubric scores is sufficient for instructional purposes, such as indicating common areas of student difficulty on a problem or grading.

The second training study included less-experienced raters (first-year graduate students) and involved a shorter training experience with revised materials, producing an overall reliability of weighted kappa 0.32±0.04. Similar to the first training study, the percent exact agreement of scores with the researcher's scores was lower than is desirable for research uses of the rubric (37±4%). Agreement within one score was higher (77±3%) indicating the resolution of the rubric scores is adequate for instructional purposes. Although the second study's brief written training experience (30-35 minutes) produced agreement among raters that was significantly greater than chance, more time and opportunities to discuss interpretations of scores (such as in the preliminary study) resulted in substantially higher levels of agreement. A substantial agreement among raters using the instrument (weighted kappa >0.60) is recommended prior to using the problem-solving rubric for education research studies whereas agreement within one score above 70% or 80% is adequate for some instructional uses. The exact procedures for an expanded written training (somewhere between the two training situations described here) could be the subject of further work.

Evidence for Utility

The rubric's usefulness for research purposes and for instructional purposes were examined separately in the studies. For research on problem-solving processes in physics, raters who are well-trained and initially discuss score interpretations can reach a high level of consistency in their scoring. Scores on the rubric provide information about specific areas of student difficulty for the five rubric processes which can be further interpreted by looking at common errors in the solutions. The rubric was also able to distinguish inexperienced (student) and experienced (instructor) solutions, regardless of the level of detail in the solutions.

For instructional uses, the rubric scores also indicate areas of student difficulty that can be used to focus coaching or modify problem statements to address those difficulties. Agreement within one score that was above 70 or 80%, even for minimal training situations, suggested that the rubric is educationally useful.

In this study the rubric was primarily tested with student solutions from introductory university physics courses, both algebra-based and calculus-based. The physics topics the rubric was applied to were from a standard mechanics course (kinematics, Newton's laws of motion, conservation of energy, conservation of momentum, angular motion, and oscillations) and one problem was from an electricity and magnetism course. Several of the problems had a text-book style (numerical answer and multiple parts) and a some were context-rich. The rubric was found to be applicable to each of these topics and types of problems, with the caveat that some problem features (such as explicit prompts and cues for procedures) make score interpretation difficult. The usefulness of the rubric for other, more abstract topics in physics, for alternate types of problems, or for more advanced courses could be the subject of further work.

Conclusions

In conclusion, an instrument was developed in the form of a rubric for assessing written solutions to physics problems along five aspects: summarizing problem information into a Useful Description, selecting an appropriate Physics Approach based on principles, a Specific Application of Physics to the conditions in the problem, following appropriate Mathematical Procedures, and an organized reasoning pattern or Logical Progression of the solution. Studies of the instrument's behavior in several situations indicated that these problem-solving skill categories were consistent with both the research literature and the processes students engage in while solving problems.

The rubric was applicable to a range of physics topics from introductory university physics courses (mechanics) and a range of problem types corresponding to those found in common textbooks and context rich. Scores on the instrument were highly correlated with other measures such as problem grades, but provide more information than grading by indicating common areas of student difficulty on a problem that can be used to focus coaching and modify problems. The documentation and training materials for the rubric resulted in an overall reliability that was statistically significant in a variety of training situations (weighted kappa $p<0.001$). However, discussion among raters was the only training studied that produced the level of agreement desirable for research purposes, such as comparing instructional pedagogies. Brief written training resulted in a reliability level that is appropriate for some educational purposes (agreement within one score above 70 or 80%).

Overall, this study indicated that it is possible to develop a measure of problem-solving processes on written solutions to physics problems in the format of a scoring rubric. The study provided evidence that the rubric measures meaningful aspects of problem-solving (validity), explored the extent to which its scores are reproducible, and suggested its usefulness for both research and instruction.

REFERENCES

Adams, W.K., & Wieman, C.E. (2006). Problem solving skill evaluation instrument – Validation studies. In L. McCullough, L. Hsu, & P. Heron (Eds.), *AIP Conference Proceedings Vol. 883: 2006 Physics Education Research Conference* (pp. 18-21). Melville, NY: American Institute of Physics.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1985). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

American Psychological Association (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin*, *51*(2, Pt.2), 201-238.

American Psychological Association (1966). *Standards for educational and psychological tests and manuals*. Washington, DC: American Psychological Association.

American Psychological Association, American Educational Research Association, and

National Council on Measurement in Education (1974). *Standards for educational and psychological tests*. Washington, DC: American Psychological Association.

Anderson, J.R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, *94*(2), 192-210.

Anderson, J.R., & Bower, G.H. (1980). *Human associative memory: A brief edition*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Anderson, J.R., Greeno, J.G., Kline, P.J., & Neves, D.M. (1981). Acquisition of problem-solving skill. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 191-230). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Arons, A.B. (1990). *A guide to introductory physics teaching*. New York: Wiley.

Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom*. Thousand Oaks, California: Corwin Press, Inc.

Bagno, E., & Eylon, B. (1997). From problem solving to knowledge structure: An example from the domain of electromagnetism. *American Journal of Physics*, *65*(8), 726-736.

Bédard, Jean, & Chi, M.T.H. (1992). Expertise. *Current Directions in Psychological Science*, *1*(4), 135-139.

Blackburn, S. (1994). *The Oxford dictionary of philosophy.* New York:Oxford University Press.

Blue, J. M. (1997). *Sex differences in physics learning and evaluations in an* *introductory course*. Unpublished doctoral dissertation, University of Minnesota, Twin Cities.

Carney, A. (2006, October). What do we want our students to learn? *Transform*, *1*, 1-6.

Chase, W.G., & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.

Chi, M.T.H. (2006). Two approaches to the study of experts' characteristics. In K.A. Ericsson, N. Charness, R.R. Hoffman, & P.J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21-30). New York, NY: Cambridge University Press.

Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145-182.

Chi, M.T.H., Feltovich, P. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121-152.

Chi, M.T.H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence vol. 1* (pp. 7-75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 10,* 37-46.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213-220.

Cronbach, L.J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational* *measurement* (2nd ed., pp. 443-507). Washington, DC: American Council on Education.

Cronbach, L.J. (1988). Five perspectives on validity argument. In H. Wainer & H.I. Braun (Eds.) *Test validity* (pp. 3-17). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281-302.

de Jong, T., & Ferguson-Hessler, M.G.M. (1986). Cognitive structures of good and poor novice problem solvers in physics. *Journal of Educational Psychology*, *78*(4), 279-288.

de Jong, T., & Ferguson-Hessler, M.G.M. (1991). Knowledge of problem situations in physics: A comparison of good and poor novice problem solvers. *Learning and Instruction*, *1*, 289-302.

de Jong, T., & Ferguson-Hessler, M.G.M. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*(2), 105-113.

Diefes-Dux, H.A., Moore, T., Zawojewski, J., Imbrie, P.K., & Follman, D. (2004). A framework for posing open-ended engineering problems: Model-eliciting activities. *Proceedings of the 34th Annual ASEE / IEEE Frontiers in Education Conference*, Vol. 2 pp. F1A3 – F1A8.

Docktor, J., & Heller, K. (2009). Robust assessment instrument for student problem solving. *Proceedings of the National Association for Research in Science Teaching 82nd International Conference*, Garden Grove, CA.

Dreyfus, H.L., & Dreyfus, S.E. (1986). *Mind over machine*. New York: Free Press.

Dufresne, R.J., Leonard, W.J., & Gerace, W.J. (February, 1995). A qualitative model for the storage of domain-specific knowledge and its implications for problem-solving. Physics Education Research Group. University of Massachusetts-Amherst. Retrieved on March 26, 2007 from http://umperg.physics.umass.edu/topics/model/

Dutson, A.J., Todd, R.H., Magleby, S.P., & Sorensen, C.D. (1997). A review of literature on teaching engineering design through project-oriented capstone courses. *Journal of Engineering Education*, *86*, 17-28.

Elio, R., & Scharf, P.B. (1990). Modeling novice-to-expert shifts in problem-solving strategy and knowledge organization. *Cognitive Science*, *14*(4), 579-639.

Eylon, B., & Reif, F. (1984). Effects of knowledge organization on task performance. *Cognition and Instruction*, *1*(1), 5-44.

Finegold, M., & Mass, R. (1985). Differences in the process of solving physics problems between good problem solvers and poor problem solvers. *Research in Science and Technology Education*, *3*, 59-67.

Flavell, J.H. (1976). Metacognitive aspects of problem solving. In L.B. Resnick (Ed.), *The nature of intelligence* (pp. 231-235). Hillsdale, NJ: Lawrence Erlbaum Associates.

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*(10), 906-911.

Fleiss, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378-382.

Foster, T. (2000). *The development of students' problem-solving skills from instruction emphasizing qualitative problem-solving*. Unpublished doctoral dissertation, University of Minnesota, Twin Cities.

Gick, M.L. (1986). Problem-solving strategies. *Educational Psychologist*, *21*(1 & 2), 99-120.

Halliday, D., Resnick, R., & Walker, J. (1997). *Fundamentals of physics* (5th ed.) New York: John Wiley & Sons, Inc.

Hardiman, P.T., Dufresne, R., & Mestre, J.P. (1989). The relationship between problem categorization and problem solving among experts and novices. *Memory and Cognition*, *17*(5), 627–638.

Harper, K.A. (2001). *Investigating the development of problem solving skills during a freshman physics sequence.* Unpublished doctoral dissertation, The Ohio State University.

Hayes, J.R. (1989). *The complete problem solver* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Heller, J.I., & Reif, F. (1984). Prescribing effective human problem-solving processes: Problem description in physics. *Cognition and Instruction*, *1*(2), 177-216.

Heller, K., & Heller, P. (2000). *Competent Problem Solver – Calculus Version*. New York: McGraw-Hill.

Heller, P. & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics*, *60*(7), 637-644.

Heller, P., Keith, R., & Anderson, S. (1992). Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving. *American Journal of Physics*, *60*(7), 627-636.

Henderson, C., Yerushalmi, E., Kuo, V., Heller, P., Heller, K. (2004). Grading student problem solutions: The challenge of sending a consistent message. American *Journal of Physics*, *72*(2), 164-169.

Howell, D.C. (2002). *Statistical methods for psychology (5th ed.)*. Pacific Grove, CA: Thomson Learning, Inc.

Hsu, L., Brewe, E., Foster, T.M., & Harper, K.A. (2004). Resource letter RPS-1: Research in problem solving. *American Journal of Physics*, *72*(9), 1147-1156.

Huffman, D. (1997). Effect of explicit problem solving instruction on high school students' problem-solving performance and conceptual understanding of physics. *Journal of Research in Science Teaching*, *34*(6), 551-570.

Hull, C.L. (1943). *Principles of behavior*. New York: Appleton-Century.

Ivie, R., & Stowe, K. (2002). The early careers of physics bachelors. *American Institute of Physics Statistical Research Center*, Report R-433. http://www.aip.org/statistics/trends/reports/bachplus5.pdf

Jonassen, D.H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research & Development*, *45*(1), 65-94.

Jonassen, D.H. (2007). *Learning to solve complex scientific problems*. New York: Lawrence Erlbaum Associates Taylor Francis Group, LLC.

Kal, V. (1988). *Philosophia antiqua: On intuition and discursive reasoning in Aristotle.* Leiden, The Netherlands: E.J. Brill.

Kane, M.T. (1992). An argument-based approach to validity. *Psychological Bulletin*, *112*(3), 527-535.

Kane, M.T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, *38*(4), 319-342.

Köhler, W. (1929). *Gestalt psychology*. New York: Liveright Publishing Corporation.

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174.

Larkin, J. (1979). Processing information for effective problem solving. *Engineering Education*, *70*(3), 285-288.

Larkin, J. (1981a). Cognition of learning physics. *American Journal of Physics*, *49*(6), 534-541.

Larkin, J.H. (1981b). Enriching formal knowledge: A model for learning to solve textbook physics problems. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 311-334). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Larkin, J.H., McDermott, J., Simon, D.P., & Simon, H.A. (1980a). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335-1342.

Larkin, J.H., McDermott, J., Simon, D.P., & Simon, H.A. (1980b). Models of competence in solving physics problems. *Cognitive Science*, *4*, 317-345.

Larkin, J., & Reif, F. (1979). Understanding and teaching problem solving in physics.

European Journal of Science Education, *1*(2), 191-203.

Lavoie, B. (2003). *The development and validation of the biotechnology problem-solving skills assessment for community college biotechnology students.* Unpublished doctoral dissertation, University of Minnesota, Twin Cities.

Lear, J. (1980). *Aristotle and logical theory*. London: Cambridge University Press.

Leonard, W.J., Dufresne, R.J., & Mestre, J.P. (1996). Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in solving problems. *American Journal of Physics*, *64*(12), 1495-1503.

Linn, R.L., Baker, E.L., & Dunbar, S.B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, *20*(8), 15-21.

Lissitz, R.W., & Samuelson, K. (2007). A suggested change in terminology and emphasis regarding validity and education. *Educational Researcher*, *36*(8), 437-448.

Maloney, D.P. (1994). Research on problem solving: Physics. In D.L. Gabel (Ed.) *Handbook of research on science teaching and learning* (pp. 327-354). New York: Macmillan.

Marshall, S.P. (1995). *Schemas in problem solving*. New York: Cambridge University Press.

Martinez, M. E. (1998). What is problem solving? *Phi Delta Kappan*, *79*, 605-609.

Mayer, R.E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: W.H. Freeman and Company.

Mayer, R.E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science*, *26*, 49-63.

Mertler, C.A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, *7*(25).

Messick, S. (1989a). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, *18*(2), 5-11.

Messick, S. (1989b). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: American Council on Education and MacMillan Publishing Company.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, *23*(2), 13-23.

Messick, S. (1995a). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, *14*(4), 5-8.

Messick, S. (1995b). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741-749.

Mestre, J. (2002). Probing adults' conceptual understanding and transfer of learning via problem posing. *Applied Developmental Psychology*, *23*, 9-50.

Montgomery, K. (2002). Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. *College Teaching*, *50*(1), 34-39.

Moss, P.A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research*, *62*(3), 229-258.

Moss, P.A. (2007). Reconstructing validity. *Educational Researcher*, *36*(8), 470-476.

Murthy, S. (2007). Peer-assessment of homework using rubrics. In L. Hsu, C. Henderson, & L. McCullough (Eds.), *AIP Conference Proceedings: Vol. 951. 2007 Physics Education Research Conference* (pp. 156-159). Melville, NY: American Institute of Physics.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future.* Washington, DC: The National Academies Press. [http://www.nap.edu]

Newell, A., Shaw, J.D., & Simon, H.A. (1958). Elements of a theory of human problem solving. *Psychological Review*, *65*, 151-166.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Ogilvie, C.A. (2007). Moving students from simple to complex problem solving. In D.H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 159-185). New York: Lawrence Erlbaum Associates Taylor Francis Group, LLC.

O'Kuma, T.L., Maloney, D.P., & Hieggelke, C.J. (2000). *Ranking task exercises in physics*. Upper Saddle River, NJ: Prentice Hall.

Ormrod, J.E. (2004). *Human learning* (4th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Pólya, G. (1945). *How to solve it*. Princeton, NJ: Princeton University Press.

Pólya, G. (1957). *How to solve it* (2nd ed.) Princeton, NJ: Princeton University Press.

Pretz, J. E., Naples, A. J., & Sternberg, R. J. (2003). Recognizing, defining, and representing problems. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 3-30). Cambridge, UK: Cambridge University Press.

Priest, A.G., & Lindsay, R.O. (1992). New light on novice-expert differences in physics problem solving. *British Journal of Psychology*, *83*(3), 389-405.

Reif, F. (1981). Teaching problem solving – A scientific approach. *The Physics Teacher*, *19*(5), 310-316.

Reif, F. (1995). Millikan Lecture 1994: Understanding and teaching important scientific thought processes. *American Journal of Physics*, *63*(1), 17-32.

Reif, F., & Heller, J.I. (1982). Knowledge structure and problem solving in physics. *Educational Psychologist*, *17*(2), 102-127.

Reif, F., Larkin, J.H., & Brackett, G. (1976). Teaching general learning and problem-solving skills. *American Journal of Physics*, *44*(3), 212-217.

Redish, E.F., Scherr, R.E., & Tuminaro, J. (2006). Reverse engineering the solution of a "simple" physics problem: Why learning physics is harder than it looks. *The Physics Teacher*, *44*(5), 293-300.

Schoenfeld, A. (1983). Beyond the purely cognitive: Beliefs, social cognitions and metacognitions as driving forces in intellectual performance. *Cognitive Science*, *7*, 329-363.

Schoenfeld, A.H. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press, Inc.

Schoenfeld, A.H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp. 334-370). New York: MacMillan.

Schoenfeld, A.H., & Hermann, D.J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(5), 484-494.

Schwartz, D. L., Bransford, J. D., & Sears, D. L. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1 - 51). CT: Information Age Publishing.

Secretary's Commission on Achieving Necessary Skills (1991). *What work requires of schools*. Washington, DC: United States Department of Labor. [http://wdr.doleta.gov/SCANS/whatwork/whatwork.pdf]

Shepard, L.A. (1993). Evaluating test validity. *Review of Research in Education*, *19*, 405-450.

Shin, N., Jonassen, J.H., & McGee, S. (2003). Predictors of well-structured and ill-structured problem-solving in an astronomy simulation. *Journal of Research in Science Teaching*, *40*(1), 6-33.

Simon, H.A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, *7*, 268-288.

Simon, H.A. (1976). Identifying basic abilities underlying intelligent performance of complex tasks. In L.B. Resnick (Ed.), *The nature of intelligence* (pp. 65-98). Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, H.A. (1981). *The sciences of the artificial* (2nd ed.). Cambridge, Massachusetts: MIT Press.

Simon, D.P., & Simon, H.A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 325-361). Hillsdale, NJ: Erlbaum.

Simon, D.P., & Simon, H.A. (1979). A tale of two protocols. In J. Lochhead & J.S. Clement (Eds.), *Cognitive process instruction: Research on teaching thinking skills* (pp. 119-132). Philadelphia: Franklin Institute Press.

Singh, C. (2002). When physical intuition fails. *American Journal of Physics*, *70*(11), 1103–1109.

Skinner, B.F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257-285.

Taconis, R., Ferguson-Hessler, M.G.M., & Broekkamp, H. (2001). Teaching science problem-solving. *Journal of Research in Science Teaching*, *38*, 442-468.

Thaden-Koch, T. (personal communication, July 2005).

Thorndike, E.L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monograph Supplements*, *2*(4, whole no. 8).

Thorndike, E.L. (1911). *Animal intelligence*. New York: Macmillan.

Van Heuvelen, A. (1991a). Learning to think like a physicist: A review of research-based instructional strategies. *American Journal of Physics*, *59*(10), 891-897.

Van Heuvelen, A. (1991b). Overview, case study physics. *American Journal of Physics*, *59*(10), 898-907.

Van Heuvelen, A. (1999). Experiment problems for mechanics. *The Physics Teacher*, *33*, 276-280.

Van Heuvelen, A., & Maloney, D. (1999). Playing physics jeopardy. *American Journal of Physics*, *67*, 252-256.

van Someren, M.W., Barnard, Y.F., & Sandberg, J.A.C. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. San Diego, CA: Academic Press Inc.

Walsh, L.N., Howard, R.G., & Bowe, B. (2007). Phenomenographic study of students' problem solving approaches in physics. *Physical Review Special Topics – Physics Education Research*, *3* (020108), 1-12.

Wertheimer, M. (1945). *Productive thinking*. New York: Harper.

Woods, D.R. (2000). An evidence-based strategy for problem solving. *Journal of Engineering Education*, *89*(4), 443-459.

Yerushalmi, E., Henderson, C., Heller, K., Heller, P., & Kuo, V. (2007). Physics faculty beliefs and values about the teaching and learning of problem solving I: Mapping the common core. *Physical Review Special Topics – Physics Education Research*, *3*(020109), 1-31.

Zajchowski, R., & Martin, J. (1993). 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.

APPENDICES

Appendix 1: Early Rubric Drafts

Description of Each Rubric Draft:

Version 1 (Spring 2007)

- Formatted as a table to have same score range for each category

Version 2 (Summer 2007) Used for preliminary study with two raters

- Not Applicable (NA) score split into two scores (Problem and Student)
- Language made more consistent across scores for some categories:
  - 4 = All appropriate and complete
  - 3 = One part missing and/or incorrect
  - 2 = More than one part missing and/or incorrect
  - 1 = All inappropriate or fundamental misunderstanding
  - 0 = All missing
  - NA(Student) = not necessary for this student, as indicated by the overall solution process

Version 3: (Fall 2007) Used for first study with training raters

- Formatted to fit vertically on one page with NA score descriptions in one line at the bottom
- Language changed for scores in some categories:
  - 4 = all appropriate and complete
  - 3 = minor omissions or errors
  - 2 = important part / key feature missing or inappropriate
  - 1 = an attempt is made, but most parts missing/inappropriate/incorrect
  - 0 = the solution does not indicate a [category] or it is all incorrect
- Category Descriptions included with rubric

Version 4: (Fall 2008) Used for second study with training raters, scoring written solutions to exams, and scoring written solutions from interviews

- Score range changed from 0-4 to 0-5
- Returned NA scores to column positions
- Formatted to fit on one landscape page (fewer words) and no asterisks
- Consistent language for a single score across each category:
  - 5 = appropriate and complete
  - 4 = minor omissions or errors
  - 3 = parts missing and/or contain errors
  - 2 = most missing and/or contain errors
  - 1 = all inappropriate
  - 0 = no evidence and necessary

Figure 61: Rubric Version 1

	4	3	2	1	0	NA
Physics Approach	The solution includes an appropriate, complete, and correct statement of basic physics principle(s) in words or as equations.	Most of the physics principles used are appropriate, but one or more principles are missing, inappropriate, or stated incorrectly.	Use of a few appropriate physics principles is evident, but most physics is missing, inappropriate, or stated incorrectly.	All physics principles stated in the approach are inappropriate or indicate a serious physics misunderstanding. A correct solution is not possible.	The solution does not include a statement of basic physics principles, and it is necessary or expected for this problem.	An explicit statement of basic principles is not necessary, or the approach has already been stated in the problem / textbook heading.
Useful Description *	The solution includes an appropriate and useful problem description.	The solution includes a useful description, but one or more elements are missing, inappropriate, or use incorrect physics.	Use of a description is evident, but most of the elements are missing, inappropriate, or use incorrect physics.	An attempt has been made at a description, but it is inappropriate, incomplete, or indicates a serious physics misunderstanding.	The solution does not include a description, and it is necessary or expected for this problem.	A description is not necessary to solve this particular problem, or it has already been given to the solver.
Specific Application of Physics**	The solution indicates an appropriate and complete application of physics that matches the approach and description.	The solution indicates an appropriate application of physics from general approach, but one or more relationships or conditions are missing / incorrectly stated.	Specific app. of physics is evident but does not match the stated approach and description, or an essential physics relationship or condition is missing / incorrectly stated.	The solution includes fundamental physics errors in the specific application [such as treating vectors as scalars.]	The solution does not indicate a specific application of physics and it is necessary or expected for this problem.	Specific application of physics is not necessary to solve the problem; basic principles are sufficient.

	4	3	2	1	0	NA
Mathe-matical Procedures	Appropriate mathematical procedures are used and result in a reasonable answer with numerical values substituted in the last step.	Appropriate mathematics is used but there are minor errors or some instances of early substitution of numerical values.	Use of appropriate mathematical procedures is evident, but most of the math is incorrect, the answer is unreasonable, or there is early substitution of numerical values.	There is evidence of a serious violation of mathematical rules (algebra, trigonometry, calculus, etc.) or the math is significantly easier than that required of a correct solution	There is no evidence of mathematical procedures in the problem solution and it is necessary or expected for this problem.	Mathematical procedures are not necessary to solve this problem, or constitute a very small part of the solution.
Logical Progression	The entire problem solution is focused and organized logically. The steps taken might not be linear, but guide the student toward a solution / converge toward an answer.	The solution is focused and organized, but includes a few logical breaks (inconsistencies) and/or extraneous steps that don't guide the solution. Most steps progress toward an answer.	The solution is somewhat focused and organized, but includes several logical breaks (inconsistencies) and/or extraneous steps that don't guide the solution. Some steps progress toward an answer.	The solution is unorganized. Most steps seem illogical and/or there are several extraneous steps that don't guide the solution. Very few steps progress toward an answer.	Nothing written can be interpreted as logical progression. The entire solution is unorganized, haphazard (random), contains obvious logical breaks, and/or does not progress toward an answer.	Logical progression is not necessary to solve this problem, or constitutes a very small part of the solution (one-step problem).

*****A "problem description" could include: restating knowns and unknowns, defining variables, stating goal or target variable, drawing a picture, stating qualitative expectations, abstracted physics diagram such as a force diagram or motion diagram, a coordinate system; and all variables are defined appropriately [such as m1 and m2 instead of just m]

**A "specific application of physics" includes a statement of definitions, relationships between the defined variables, initial conditions, and assumptions or constraints to the problem [i.e., friction negligible, constant acceleration, static equilibrium, etc.]

Figure 62: Rubric Version 2

	4	3	2	1	0	NA (Prob)	NA (Student)
Physics Approach	The overall solution indicates the solver has an appropriate and complete physics approach.	One basic physics principle of the approach is missing or inappropriate.	More than one basic physics principle of the approach is missing or inappropriate.	All basic physics principles of the approach are inappropriate or indicate a fundamental misunder-standing.	The solution does not indicate a basic physics approach, and it is necessary for this problem / student.	A physics approach is not necessary for this problem. (i.e., has already been stated in the problem or textbook heading)	A physics approach is not necessary for this student, as indicated by the overall solution process.
Useful Description *	The solution includes an appropriate and useful problem description.	One part of the description is missing, inappropriate, or uses incorrect physics.	More than one part of the description is missing, inappropriate, or uses incorrect physics.	The description is inappropriate or indicates a fundamental misunder-standing (i.e., straight-line trajectory for a projectile)	The solution does not include a description, and it is necessary for this problem / student.	A description is not necessary for this problem. (i.e., it has already been given to the solver)	A description is not necessary for this student, as indicated by the overall solution process.
Specific Application of Physics**	The solution indicates an appropriate and complete application of physics to the specific conditions in this problem.	One relationship or condition is missing or uses incorrect physics.	More than one relationship or condition is missing or uses incorrect physics.	The application of physics to this problem is inappropriate or indicates a fundamental misunder-standing (i.e., treating vectors as scalars).	The solution does not indicate a specific application of physics and it is necessary for this problem / student.	Specific application of physics is not necessary for this problem. (i.e., basic principles are sufficient)	Specific application of physics is not necessary for this student, as indicated by the overall solution process.

	4	3	2	1	0	NA (Prob)	NA (Student)
Mathe-matical Procedures	Suitable mathematical procedures are used and result in a reasonable answer with numerical values substituted at appropriate steps.	Suitable mathematical procedures are used with minor error(s) (i.e. sign error, calculation error)	Suitable mathematical procedures are used with error(s), answer is unreasonable and unnoticed, or there is early substitution of numerical values.	Attempted mathematical procedures are inappropriate. (i.e., violate a fundamental rule of arithmetic)	There is no evidence of mathematical procedures in the problem solution and it is necessary for this problem / student.	Mathematical procedures are not necessary for this problem, or constitute a very small part of the solution.	Mathematical procedures are not necessary for this student, as indicated by the overall solution process.
Logical Pro-gression	The entire problem solution is focused and organized logically. The steps taken might not be linear, but guide the student toward an answer.	The solution is focused and organized with a few minor inconsistencies and/or extraneous steps that don't guide the solution.	The solution is focused and organized with multiple inconsistencies and/or extraneous steps that don't guide the solution.	Parts of the solution are focused and organized. There are several inconsistencies and/or extraneous steps that don't guide the solution.	Nothing written can be interpreted as logical progression. The entire solution is unorganized, haphazard (random), and contains obvious logical breaks.	Logical progression is not necessary for this problem or constitutes a very small part of the solution (i.e., one-step problem).	Logical progression is not necessary for this student, as indicated by the overall solution process.

*A "problem description" could include: restating knowns and unknowns, defining variables, stating goal or target variable, drawing a picture, stating qualitative expectations, abstracted physics diagram such as a force diagram or motion diagram, a coordinate system; and all variables are defined appropriately [such as m1 and m2 instead of just m]

**A "specific application of physics" includes a statement of definitions, relationships between the defined variables, initial conditions, and assumptions or constraints to the problem [i.e., friction negligible, constant acceleration, static equilibrium, etc.]

Figure 63: Rubric Version 3

	4	3	2	1	0
Physics Approach	The solver has clearly stated an appropriate and complete physics approach.	The approach is clear but contains minor omissions or errors.	The approach is unclear, or an important physics concept or principle of the approach is missing or inappropriate.	An attempt is made to identify relevant physics concepts or principles, but most of the approach is vague, incomplete, or inappropriate.	The solution does not indicate a basic physics approach, or all of the chosen concepts and principles are inappropriate.
Useful Description *	The solution includes an appropriate and useful problem description.	The description is useful but contains minor omissions or errors.	The description is not useful, or a key feature of the description is missing or incorrect.	An attempt is made, but most of the description is not useful, incomplete, or incorrect.	The solution does not include a description, or all of the description is incorrect.
Specific Application of Physics**	The solution indicates an appropriate and complete application of physics to the specific conditions in this problem.	The specific application of physics to this problem contains minor omissions or errors.	An important specific relationship or condition is missing or applied incorrectly.	An attempt is made, but most of the specific application of physics to this problem is missing or incorrect.	The solution does not indicate a specific application of physics, or all of the application is incorrect.
Mathe-matical Procedures	Suitable mathematical procedures are used during the solution execution.	Suitable mathematical procedures are used with minor omissions or errors.	An important mathematical procedure is missing or is used with errors.	Attempted mathematical procedures are inappropriate, left unfinished, or contain serious errors	There is no evidence of mathematical procedures in the problem solution or all mathematical procedures are inappropriate.
Logical Organi-zation	The entire problem solution is clear, focused, and logically connected.	The solution is clear and focused with minor inconsistencies.	Parts of the solution are unclear, unfocused, and/or inconsistent.	Most of the solution parts are unclear, unfocused, and inconsistent.	The entire solution is unorganized and contains obvious logical breaks.
NA (Prob)	The skill is not necessary for this problem, or constitutes a very small part of the solution.				
NA (Solver)	Explicit statement is not necessary for this solver, as indicated by the overall solution.				

<u>Category Descriptions</u>:

Physics Approach assesses a solver's skill at selecting appropriate physics concepts and principle(s) to use in solving the problem. Here the term *concept* is defined to be a general physics idea, such as the basic concept of "vector" or specific concepts of "momentum" and "average velocity". The term *principle* is defined to be a fundamental physics rule or law used to describe objects and their interactions, such as the law of conservation of energy, Newton's second law, or Ohm's law.

Useful Description assesses a solver's skill at organizing information from the problem statement into an appropriate and useful representation that summarizes essential information symbolically and visually. The description is considered "useful" if it guides further steps in the solution process.
*A *problem description* could include restating known and unknown information, assigning appropriate symbols for variables, defining variables, stating a goal or target, a visualization (sketch or picture), stating qualitative expectations, an abstracted physics diagram (force, energy, motion, momentum, ray, etc.), drawing a graph, stating a coordinate system, and choosing a system.

Specific Application of Physics assesses a solver's skill at applying the physics concepts and principles from their selected approach to the specific conditions in the problem. If necessary, the solver has set up specific equations for the problem that are consistent with the chosen approach.
**A *specific application of physics* could include a statement of definitions, relationships between the defined variables, initial conditions, and assumptions or constraints in the problem (i.e., friction negligible, massless spring, massless pulley, inextensible string, etc.)

Mathematical Procedures assesses a solver's skill at following appropriate and correct mathematical rules and procedures during the solution execution. The term *mathematical procedures* refers to techniques that are employed to solve for target variable(s) from specific equations of physics, such as isolate and reduce strategies from algebra, substitution, use of the quadratic formula, or matrix operations. The term *mathematical rules* refers to conventions from mathematics, such as appropriate use of parentheses, square roots, and trigonometric identities. If the course instructor or researcher using the rubric expects a symbolic answer prior to numerical calculations, this could be considered an appropriate mathematical procedure.

Logical Organization assesses the solver's skills at communicating reasoning, staying focused toward a goal, and evaluating the solution for consistency (implicitly or explicitly). It checks whether the entire problem solution is clear, focused, and organized logically. The term *logical* means that the solution is coherent (the solution order and solver's reasoning can be understood from what is written), internally consistent (parts do not contradict), and externally consistent (agrees with physics expectations).

Figure 64: Rubric Version 4

	5	4	3	2	1	0	NA(Problem)	NA(Solver)
USEFUL DESCRIPTION	**The description is useful, appropriate, and complete.**	**The description is useful but contains minor omissions or errors.**	**Parts of the description are not useful, missing, and/or contain errors.**	**Most of the description is not useful, missing, and/or contains errors.**	**The entire description is not useful and/or contains errors.**	**The solution does not include a description and it is necessary for this problem /solver.**	**A description is not necessary for this <u>problem.</u> (i.e., it is given in the problem statement)**	**A description is not necessary for this <u>solver</u>.**
PHYSICS APPROACH	**The physics approach is appropriate and complete.**	**The physics approach contains minor omissions or errors.**	**Some concepts and principles of the physics approach are missing and/or inappropriate.**	**Most of the physics approach is missing and/or inappropriate.**	**All of the chosen concepts and principles are inappropriate.**	**The solution does not indicate an approach, and it is necessary for this problem/ solver.**	**An explicit physics approach is not necessary for this <u>problem</u>. (i.e., it is given in the problem)**	**An explicit physics approach is not necessary for this <u>solver.</u>**
SPECIFIC APPLICATION OF PHYSICS	**The specific application of physics is appropriate and complete.**	**The specific application of physics contains minor omissions or errors.**	**Parts of the specific application of physics are missing and/or contain errors.**	**Most of the specific application of physics is missing and/or contains errors.**	**The entire specific application is inappropriate and/or contains errors.**	**The solution does not indicate an application of physics and it is necessary.**	**Specific application of physics is not necessary for this <u>problem</u>.**	**Specific application of physics is not necessary for this <u>solver</u>.**
MATHE-MATICAL PROCEDURES	**The mathematical procedures are appropriate and complete.**	**Appropriate mathematical procedures are used with minor omissions or errors.**	**Parts of the mathematical procedures are missing and/or contain errors.**	**Most of the mathematical procedures are missing and/or contain errors.**	**All mathematical procedures are inappropriate and/or contain errors.**	**There is no evidence of mathematical procedures, and they are necessary.**	**Mathematical procedures are not necessary for this <u>problem</u> or are very simple.**	**Mathematical procedures are not necessary for this <u>solver</u>.**
LOGICAL PROGRESSION	**The entire problem solution is clear, focused, and logically connected.**	**The solution is clear and focused with minor inconsistencies**	**Parts of the solution are unclear, unfocused, and/or inconsistent.**	**Most of the solution parts are unclear, unfocused, and/or inconsistent.**	**The entire solution is unclear, unfocused, and/or inconsistent.**	**There is no evidence of logical progression, and it is necessary.**	**Logical progression is not necessary for this <u>problem</u>. (i.e., one-step)**	**Logical progression is not necessary for this <u>solver</u>.**

<u>Category Descriptions</u>:

Useful Description assesses a solver's skill at organizing information from the problem statement into an appropriate and useful representation that summarizes essential information symbolically and visually. The description is considered "useful" if it guides further steps in the solution process. A *problem description* could include restating known and unknown information, assigning appropriate symbols for quantities, stating a goal or target quantity, a visualization (sketch or picture), stating qualitative expectations, an abstracted physics diagram (force, energy, motion, momentum, ray, etc.), drawing a graph, stating a coordinate system, and choosing a system.

Physics Approach assesses a solver's skill at selecting appropriate physics concepts and principle(s) to use in solving the problem. Here the term *concept* is defined to be a general physics idea, such as the basic concept of "vector" or specific concepts of "momentum" and "average velocity". The term *principle* is defined to be a fundamental physics rule or law used to describe objects and their interactions, such as the law of conservation of energy, Newton's second law, or Ohm's law.

Specific Application of Physics assesses a solver's skill at applying the physics concepts and principles from their selected approach to the specific conditions in the problem. If necessary, the solver has set up specific equations for the problem that are consistent with the chosen approach. A *specific application of physics* could include a statement of definitions, relationships between the defined quantities, initial conditions, and assumptions or constraints in the problem (i.e., friction negligible, massless spring, massless pulley, inextensible string, etc.)

Mathematical Procedures assesses a solver's skill at following appropriate and correct mathematical rules and procedures during the solution execution. The term *mathematical procedures* refers to techniques that are employed to solve for target quantities from specific equations of physics, such as isolate and reduce strategies from algebra, substitution, use of the quadratic formula, or matrix operations. The term *mathematical rules* refers to conventions from mathematics, such as appropriate use of parentheses, square roots, and trigonometric identities. If the course instructor or researcher using the rubric expects a symbolic answer prior to numerical calculations, this could be considered an appropriate mathematical procedure.

Logical Progression assesses the solver's skills at communicating reasoning, staying focused toward a goal, and evaluating the solution for consistency (implicitly or explicitly). It checks whether the entire problem solution is clear, focused, and organized logically. The term logical means that the solution is coherent (the solution order and solver's reasoning can be understood from what is written), internally consistent (parts do not contradict), and externally consistent (agrees with physics expectations).

Appendix 2: Reliability Measure Kappa

Cohen's Kappa

There is some disagreement among statisticians surrounding the reliability measure Kappa, for example the argument of pros and cons presented on the following web site: http://www.john-uebersax.com/stat/agree.htm
or the discussion provided on Wikipedia.org for "Cohen's kappa":
http://en.wikipedia.org/wiki/Cohen's_kappa

One reason is that kappa calculates an "expected" level of rater agreement from the observed ratings, and this tends to give an overly conservative measure of agreement. Despite this concern, kappa is still a widely used statistic for measuring the agreement of multiple judges for nominal or rank-ordered categories. An explanation of its calculation is provided below.

(Cohen, 1960; Howell, 2002):

Kappa (κ) is defined as:

$$\kappa = \frac{\sum f_O - \sum f_E}{N - \sum f_E}$$

Where f_O represents the observed frequencies of exact agreement and f_E represents the frequencies expected by chance, and N is the total number of ratings.

The expected frequency comes from the product of the probability for each judge assigning a particular rating. For example, the probability of judge 1 rating something an *i* is the total number of times they rated something as *i*, divided by the total number of items rated. The same is true for judge 2, and the probability of the combination of judge 1 rating an *i* when judge 2 rates a *j* is the product of these two probabilities.

$$probability\ of\ rating\ ij = \left(\frac{\#of\ times\ judge 1\ rated\ i}{N}\right)\left(\frac{\#of\ times\ judge 2\ rated\ j}{N}\right)$$

$$probability\ of\ rating\ ij = \frac{f_{E,ij}}{N}$$

$$f_{E,ij} = \left(probability\ of\ rating\ ij\right)N$$

$$= \frac{\left(\#of\ times\ judge 1\ rated\ i\right)\left(\#of\ times\ judge 2\ rated\ j\right)}{N}$$

Standard error of kappa:

$$\sigma_\kappa = \sqrt{\frac{N\sum f_O - \left(\sum f_O\right)^2}{N\left(N - \sum f_E\right)^2}}$$

Standard error under the null hypothesis that $f_O = f_E$:

$$\sigma_{\kappa o} = \sqrt{\frac{N\sum f_E - \left(\sum f_E\right)^2}{N\left(N - \sum f_E\right)^2}} = \sqrt{\frac{\sum f_E}{N\left(N - \sum f_E\right)}}$$

Significance of kappa:

$$significance = \frac{\kappa}{\sigma_{ko}}$$

Test of the difference between two independent kappas:

$$z = \frac{\kappa_1 - \kappa_2}{\sqrt{\sigma_{\kappa 1}{}^2 - \sigma_{\kappa 2}{}^2}}$$

95% confidence limit: 1.960

99% confidence limit: 2.576

99.9% confidence limit: 3.291

For example, a significance value above 1.960 means the data is significant at the $\alpha=0.05$ level ($p<0.05$), and a value above 3.291 means the data is significant at the $\alpha=0.001$ level ($p<0.001$).

Often agreement measures are calculated from a R x C contingency table. In that case, the expected frequency for a cell in row *i* and column *j* is:

$$f_{E,ij} = \frac{R_i C_i}{N}$$

Where R_i represents the row total and C_j represents a column total, and N is the total number of observations.

Sample Calculation for Cohen's Kappa
(Data from preliminary study with two raters)

Contingency Table for Judge 1 (teacher) and Judge 2 (researcher)

		JUDGE 2: RESEARCHER					
JUDGE 1: TEACHER		**0**	**1**	**2**	**3**	**4**	Row Totals
	0	21	5	4			30
	1	6	39	14	1	1	61
	2	1	32	81	19	1	134
	3	1	2	46	187	57	293
	4	1	1	3	43	180	228
	Column Totals	30	79	148	250	239	746

Score	Observed Agreement f_O (diagonal)	Expected Agreement f_E
0	21	$\frac{30 \times 30}{746} = 1.206$
1	39	$\frac{61 \times 79}{746} = 6.460$
2	81	$\frac{134 \times 148}{746} = 26.584$
3	187	$\frac{293 \times 250}{746} = 98.190$
4	180	$\frac{228 \times 239}{746} = 73.046$
SUM	$\sum f_O = 508$	$\sum f_E = 205.486$

Cohen's kappa:

$$\kappa = \frac{\sum f_O - \sum f_E}{N - \sum f_E} = \frac{508 - 205.486}{746 - 205.486} = 0.5597$$

Standard Error of Kappa:

$$\sigma_\kappa = \sqrt{\frac{N\sum f_O - \left(\sum f_O\right)^2}{N\left(N - \sum f_E\right)^2}} = \sqrt{\frac{746(508) - (508)^2}{746(746 - 205.486)^2}} = 0.0236$$

Standard Error under the Null Hypothesis:

$$\sigma_{\kappa o} = \sqrt{\frac{\sum f_E}{N\left(N - \sum f_E\right)}} = \sqrt{\frac{205}{746(746 - 205)}} = 0.0226$$

Significance:

$$significance = \frac{\kappa}{\sigma_{ko}} = \frac{0.5597}{0.0226} = 24.77$$

So, kappa in this instance is 0.56±0.02 and is significant at the $p<0.001$ level. Note that perfect agreement was 508/746 = 0.68, but accounting for chance the agreement (kappa) was somewhat lower.

Weighted Cohen's Kappa:

Weighted kappa is defined as (Cohen, 1968):

$$\kappa_w = \frac{\sum w_{ij} f_{Oij} - \sum w_{ij} f_{Eij}}{(w_{\max})N - \sum w_{ij} f_{Eij}} = \frac{\sum f_{O(w)} - \sum f_{E(w)}}{(w_{\max})N - \sum f_{E(w)}}$$

Where w_{ij} represents the weight assigned to a particular rating combination ij.

The weights can be chosen as any value representing the weight of the rating. Some examples include linear and quadratic (squared) weights.

Weights for a scale of 0 to 4:

	Rating combination i/j	Linear Weight (w_{max} = 4)	Quadratic Weight (w_{max} = 16)
Perfect agreement	4/4, 3/3, 2/2, 1/1, 0/0	$w_{mzx} - \lvert i - j \rvert = 4$	w_{max}-(i-j)^2 =16
Differ by one	4/3, 3/4, 3/2, 2/3, 2/1, 1/2, 1/0, 0/1	3	9
Differ by two	4/2, 2/4, 3/1, 1/3, 2/0, 0/2	2	4
Differ by three	4/1, 1/4, 3/0, 0/3	1	1
Differ by four	4/0, 0/4	0	0

Weights for a scale of 0 to 5:

	Rating combination i/j	Linear Weight (w_{max} = 5)	Quadratic Weight (w_{max} = 25)
Perfect agreement	5/5, 4/4, 3/3, 2/2, 1/1, 0/0	$w_{mzx} - \lvert i - j \rvert = 5$	w_{max}-(i-j)^2 =25
Differ by one	5/4, 4/5, 4/3, 3/4, 3/2, 2/3, 2/1, 1/2, 1/0, 0/1	4	16
Differ by two	5/3, 3/5, 4/2, 2/4, 3/1, 1/3, 2/0, 0/2	3	9
Differ by three	5/2, 2/5, 4/1, 1/4, 3/0, 0/3	2	4
Differ by four	5/1, 1/5, 4/0, 0/4	1	1
Differ by five	5/0, 0/5	0	0

Standard error of weighted kappa:

$$\sigma_{kw} = \sqrt{\frac{N\sum w_{ij}^{2} f_{Oij} - \left(\sum w_{ij} f_{Oij}\right)^{2}}{N\left(w_{\max} N - \sum w_{ij} f_{Eij}\right)^{2}}}$$

Standard error of weighted kappa under the null hypothesis $f_O = f_E$:

$$\sigma_{kwo} = \sqrt{\frac{N\sum w_{ij}^{2} f_{Eij} - \left(\sum w_{ij} f_{Eij}\right)^{2}}{N\left(w_{\max} N - \sum w_{ij} f_{Eij}\right)^{2}}}$$

Significance:

$$significance = \frac{\kappa_w}{\sigma_{kwo}}$$

Sample Calculation for Quadratic Weighted Kappa

(Same contingency table as before)

		JUDGE 1: RESEARCHER					
JUDGE 2: TEACHER	Score	**0**	**1**	**2**	**3**	**4**	Row Totals
	0	21	5	4	0	0	30
	1	6	39	14	1	1	61
	2	1	32	81	19	1	134
	3	1	2	46	187	57	293
	4	1	1	3	43	180	228
	Column Totals	30	79	148	250	239	746

Score (Row/Column)	Weight	f_O	f_E	$wf_O = f_{O(w)}$	$wf_E = f_{E(w)}$	$w^2 f_O$	$w^2 f_E$
0/0	16	21	1.206	336	19.3	5376	308.84718
1/1	16	39	6.46	624	103.4	9984	1653.7051
2/2	16	81	26.58	1296	425.4	20736	6805.6193
3/3	16	187	98.19	2992	1571	47872	25136.729
4/4	16	180	73.05	2880	1169	46080	18699.668
1/0	9	6	2.453	54	22.08	486	198.69973
2/1	9	32	14.19	288	127.7	2592	1149.4182
3/2	9	46	58.13	414	523.2	3726	4708.4236
4/3	9	43	76.41	387	687.7	3483	6189.008
0/1	9	5	3.177	45	28.59	405	257.33244
1/2	9	14	12.1	126	108.9	1134	980.25201
2/3	9	19	44.91	171	404.2	1539	3637.3995
3/4	9	57	93.87	513	844.8	4617	7603.4678
2/0	4	1	5.389	4	21.55	16	86.219839
3/1	4	2	31.03	8	124.1	32	496.4504
4/2	4	3	45.23	12	180.9	48	723.7319
0/2	4	4	5.952	16	23.81	64	95.227882
1/3	4	1	20.44	4	81.77	16	327.07775
2/4	4	1	42.93	4	171.7	16	686.88472
3/0	1	1	11.78	1	11.78	1	11.782842
4/1	1	1	24.14	1	24.14	1	24.144772
0/3	1	0	10.05	0	10.05	0	10.053619
1/4	1	1	19.54	1	19.54	1	19.542895
4/0	0	1	9.169	0	0	0	0
0/4	0	0	9.611	0	0	0	0
Sum		**746**	**736.4**	**10177**	**6704**	**148225**	**79809.686**

$$\kappa_w = \frac{\sum f_{O(w)} - \sum f_{E(w)}}{(w_{\max})N - \sum f_{E(w)}} = \frac{10177 - 6704}{(16)746 - 6704} = 0.6638$$

$$\sigma_{kw} = \sqrt{\frac{N\sum w_{ij}^{2} f_{Oij} - \left(\sum w_{ij} f_{Oij}\right)^2}{N\left(w_{\max} N - \sum w_{ij} f_{Eij}\right)^2}}$$

$$= \sqrt{\frac{(746)148225 - (10177)^2}{746(16*746 - 6704)^2}} = 0.0185$$

$$\sigma_{kwo} = \sqrt{\frac{N\sum w_{ij}^{2} f_{Eij} - \left(\sum w_{ij} f_{Eij}\right)^2}{N\left(w_{\max} N - \sum w_{ij} f_{Eij}\right)^2}}$$

$$= \sqrt{\frac{(746)79809.686 - (6704)^2}{746(16*746 - 6704)^2}} = 0.0267$$

$$significance = \frac{\kappa_w}{\sigma_{kwo}} = \frac{0.6638}{0.0267} = 24.83$$

So, kappa in this instance is 0.66±0.02 and is significant at the $p<0.001$ level. Note that perfect agreement was 508/746 = 68% and agreement within one was 730/746=98%, but accounting for chance the agreement (kappa) was somewhat lower.

Simulations of Kappa Behavior

One criticism of kappa (and weighted kappa) is that it calculates an expected frequency from the observed ratings of each judge; in other words, it is based on the "tendency" of each judge to assign particular ratings. This has the implication that when one judge always assigns the same score, kappa will be zero because the expected frequency by chance will be the same as the observed frequency.

Simulation One

As an example, consider one judge with a random normal distribution of 100 ratings (score range 0-5) that has a mean of 3 and a standard deviation of 1. This can be simulated in Excel using the function Tools>Data Analysis>Random Number Generation. Those random numbers are then rounded to the nearest whole number to simulate a rating using the function [=ABS(ROUND(cell,0))]. Consider a second judge who always assigns a score of 3. The contingency table for this simulation is below:

		JUDGE 2						
		0	1	2	3	4	5	Row total
JUDGE 1	0				1			1
	1				4			4
	2				28			28
	3				37			37
	4				28			28
	5				2			2
	Column Total	0	0	0	100	0	0	100

Score	Observed Agreement f_O (diagonal)	Expected Agreement f_E
0	0	$\frac{1\times 0}{100}=0$
1	0	$\frac{4\times 0}{100}=0$
2	0	$\frac{28\times 0}{100}=0$
3	37	$\frac{37\times 100}{100}=37$
4	0	$\frac{28\times 0}{100}=0$
5	0	$\frac{2\times 0}{100}=0$
SUM	$\sum f_O = 37$	$\sum f_E = 37$

$$\kappa = \frac{\sum f_O - \sum f_E}{N - \sum f_E} = \frac{37-37}{100-37} = 0$$

Similarly, for weighted kappa the observed and expected frequencies are the same value or are zero.

$$\kappa_w = \frac{\sum f_{O(w)} - \sum f_{E(w)}}{(w_{\max})N - \sum f_{E(w)}} = \frac{1879-1879}{(25)100-1879} = 0$$

In this simulation, the percent perfect agreement is 37/100=37% and agreement within one is (28+37+28)/100=93%, but kappa and weighted kappa are both zero indicating agreement is the same as chance.

This simulation illustrates a situation in which there is high percent agreement within one score, but a low weighted kappa. This low kappa arises either because one of the raters had a tendency to assign the same rating, and/or that there was not much variation in the items being rated.

Simulation Two

Another situation in which weighted kappa produces low values even for high percent agreement is when the distributions of ratings for two raters have the same mean and standard deviation. In this simulation, a list of 100 random numbers with a normal distribution and mean of 3 were produced with various standard deviations. The random numbers were rounded to the nearest whole number to calculate the weighted kappa value for two lists (simulating two raters). In this simulation the distribution of scores for both "raters" had the same mean and standard deviation.

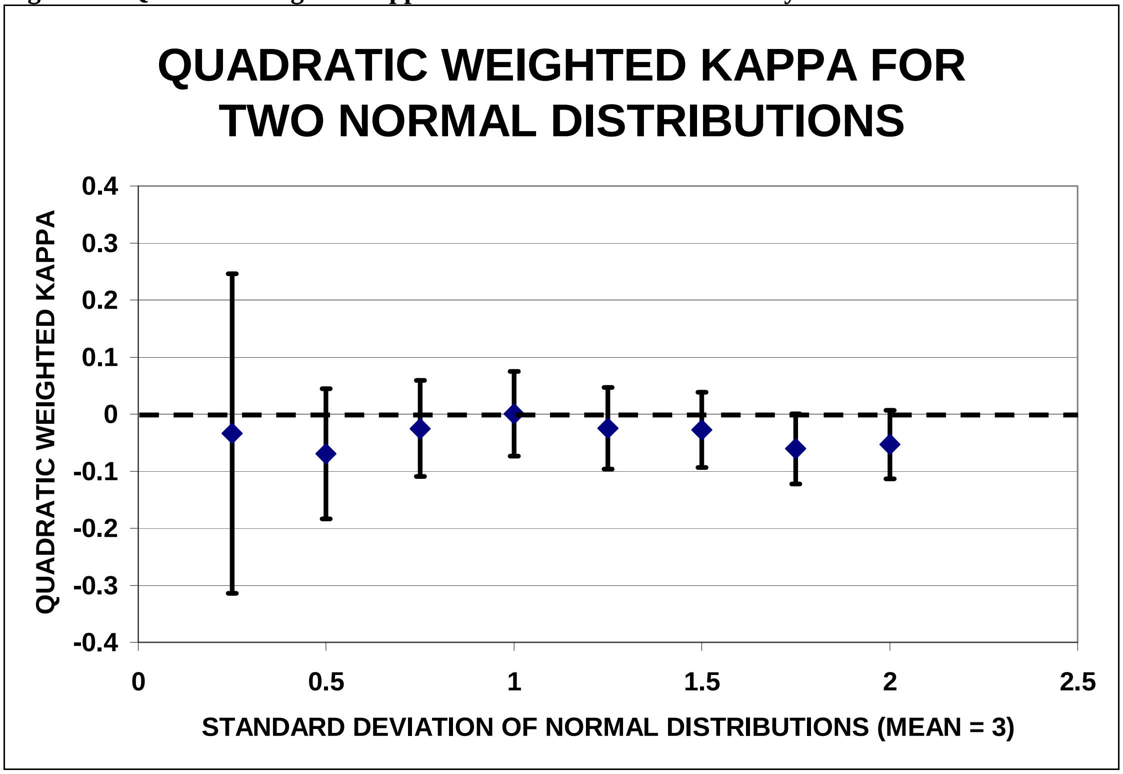
The following table lists weighted kappa values for two normal distributions with the same mean score (mean=3) and as a function of the standard deviations for the two distributions:

Simulated weighted kappa calculations for normal distributions by standard deviation:

Mean	StDev	κ_w	$\sigma_{\kappa w}$	% perfect agreement	% agreement within one
3	0.25	-0.03383	0.279964	88%	100%
3	0.50	-0.06946	0.113957	52%	98%
3	0.75	-0.02525	0.084135	35%	83%
3	1.00	0.000948	0.074243	29%	73%
3	1.25	-0.02497	0.071488	26%	52%
3	1.50	-0.02759	0.066006	15%	51%
3	1.75	-0.06081	0.061664	13%	45%
3	2.00	-0.05312	0.060002	14%	40%

The following figure illustrates that weighted kappa for two random normal distributions of scores with the same mean are at chance level zero (or below chance) regardless of the standard deviation of the distribution.

Figure 65: Quadratic weighted kappa for two normal distributions by standard deviation



Simulation Three

A different method for calculating weighted kappa could be to compute the expected frequency based on scores from ratings of other categories. In this simulation, there are five simulated "categories" (similar to the rubric categories of description, approach, etc.) that each have two lists of 100 randomly generated numbers. Each pair of simulated ratings has the same mean and standard deviation but the categories vary in their mean score. Weighted kappa is calculated in the same way as before, except that the expected frequency is calculated from a set of scores for other categories.

The following table gives values of weighted kappa for the standard calculation (old kappa) and a new value that calculates the expected frequencies from the other four categories (new kappa) for five different categories (CAT1, CAT2, etc.). The new kappa values are higher than the zero (chance) agreement from the standard calculation.

Simulated weighted kappa values with an alternate calculation of expected frequency: (comparing two random rater score distributions with same mean and StDev=1)

	Mean	StDev	Old κ_w	New κ_w	SE	% perfect agreement	% agreement w/in one
CAT1	1	1	-0.04648	0.246904	0.057269	30%	79%
CAT2	2	1	-0.0039	0.25663	0.057361	30%	70%
CAT3	3	1	0.005262	0.260061	0.059461	32%	66%
CAT4	4	1	0.054503	0.252904	0.067291	37%	74%
CAT5	3	1	-0.02954	0.224878	0.059217	28%	68%

This simulation illustrates that one possible alternative to the standard calculation of kappa is to use scores from other categories to calculate the expected frequency of chance agreement, presuming that there is variation in the mean score for those other categories.

When the two category distributions have one rater that is a random normal distribution and one rater that always assigns the same score, the kappa values increase by an even greater amount than when both distributions of rater scores are random.

Simulated weighted kappa values with an alternate calculation of expected frequency: (comparing one random score distribution with flat score at mean; StDev=1)

	Mean	StDev	Old κ_w	New κ_w	SE	% perfect agreement	% agreement w/in one
CAT1	1	1	0	0.465169	0.054639	50%	98%
CAT2	2	1	0	0.427957	0.046439	37%	94%
CAT3	3	1	0	0.434475	0.046938	37%	93%
CAT4	4	1	0	0.321233	0.057169	37%	93%
CAT5	3	1	0	0.467449	0.05195	44%	90%

Simulated weighted kappa values with an alternate calculation of expected frequency: (comparing one random score distribution with flat score at mean; StDev=0.5)

	Mean	StDev	Old κ_w	New κ_w	SE	% perfect agreement	% agreement w/in one
CAT1(2)	1	0.5	0	0.740653	0.04492	75%	100%
CAT2(2)	2	0.5	0	0.788654	0.036606	75%	100%
CAT3(2)	3	0.5	0	0.7951	0.03549	75%	100%
CAT4(2)	4	0.5	0	0.676382	0.142489	67%	100%
CAT5(2)	3	0.5	0	0.729165	0.038591	67%	100%

In this last situation where the distribution of random scores is very narrow (100% agreement within one score) and is compared to a rater that always assigns the same mean score, the weighted kappa value is >0.60 indicating substantial agreement.

Fleiss's Kappa

(Fleiss, 1971)

Fleiss's kappa is a measure of agreement among multiple raters. It's definition is:

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

Its calculation is more involved than Cohen's kappa; the P-bar and Pe-bar equations are derived below.

In the following equations *N* represents the total number of subjects, *n* is the number of ratings per subject (number of raters), and *k* is the number of categories into which assignments are made (in this case a score range of zero to four or five gives a *k* of 5 or 6, depending on the rubric version). The subscript *i* represents the subjects i=1….N and the subscript j represents the categories of the scale j = 1…..k. Then n_{ij} is the number of raters who assigned the *i*th subject to the *j*th category, and p_j is the proportion of all assignments which were to the *j*th category.

$$p_j = \frac{1}{Nn}\sum_{i=1}^{N} n_{ij}$$

The proportion of agreement among the *n* raters on the *i*th subject is:

$$P_i = \frac{1}{n(n-1)}\sum_{i=1}^{N} n_{ij}\left(n_{ij}-1\right) = \frac{1}{n(n-1)}\left(\sum_{i=1}^{N}\left(n_{ij}^{2} - n_{ij}\right)\right) = \frac{1}{n(n-1)}\left(\sum_{i=1}^{N} n_{ij}^{2} - n\right)$$

And the overall extent of agreement is the mean of these P_i's:

$$\bar{P} = \frac{1}{N}\sum_{i=1}^{N} P_i = \frac{1}{Nn(n-1)}\left(\sum_{i=1}^{N}\sum_{j=1}^{k} n_{ij}^{2} - Nn\right)$$

If raters made their assignments at random, the expected proportion of agreement is:

$$\bar{P}_e = \sum_{j=1}^{k} p_j^{\,2}$$

The quantity $1-\bar{P}_e$ measures the degree of agreement attainable over and above what would be predicted by chance. The degree of agreement actually attained in excess of chance is $\bar{P}-\bar{P}_e$, so according to Fliess, kappa represents a normalized measure of overall agreement corrected for the amount expected by chance:

$$\kappa = \frac{\bar{P}-\bar{P}_e}{1-\bar{P}_e}$$

Standard error of Fleiss's kappa:

$$SE(\kappa) = \sqrt{\left(\frac{2}{Nn(n-1)}\right)\left(\frac{\sum_j p_j^{\,2} - (2n-3)\left(\sum_j p_j^{\,2}\right)^2 + 2(n-2)\sum_j p_j^{\,3}}{\left(1-\sum_j p_j^{\,2}\right)^2}\right)}$$

Problem 1 (Calculus-based course):

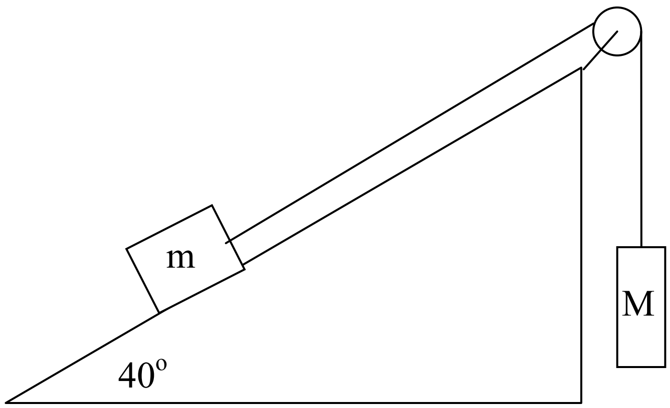
To raise money for a University scholarship fund, the new IT dean has volunteered to bungee jump from a crane if contributions can be found for 10 scholarships. To add some interest, the jump will be made from 42 m above a pool of water. A 30m bungee cord would be attached to the dean. First you must convince the dean that your plan is safe for a person of his mass, 70kg. The dean knows that as the bungee cord begins to stretch, it will exert a force which has the same properties as the force exerted by a spring. Your plan has the dean stepping off a platform and being in free fall for 30 m before the cord begins to stretch.

a) Determine the spring constant of the bungee cord so that it stretches only 12m, which will just keep the dean out of the water. (Assume that the dean is a point-like object).
b) Using the result of a), find the dean's speed 7m above the water.

Problem 2 (Calculus-based course):

The sketch shows a mass, m=3kg, on an inclined plane which is at an angle of 40^o to the horizontal. It is attached to a light string which runs over a frictionless, massless pulley and supports a mass M hanging vertically. The coefficients of friction, static and kinetic, between the mass m and the plane are μ_s=0.4 and μ_k=0.3.

a) If M=5kg, what is the acceleration of the system (magnitude and direction)? (Caution: make sure your frictional force is pointing in the right direction)
b) What is the range of possible values of M such that the system is at rest in equilibrium?



Problem 3 (Calculus-based course):

A 4kg cat (treat it as a point particle) sits on a horizontal floor eyeing a stationary chair of mass 10kg which is a horizontal distance of 1.3m away. The seat of the chair is 0.5m above the floor. The cat jumps up and lands on the seat of the chair just as she reaches the maximum height of her trajectory. She puts out her claws and hangs on. If the chair sits on a part of the floor which has just been waxed, is very slippery and therefore frictionless, what is the momentum of the cat plus chair system just after the cat has landed?

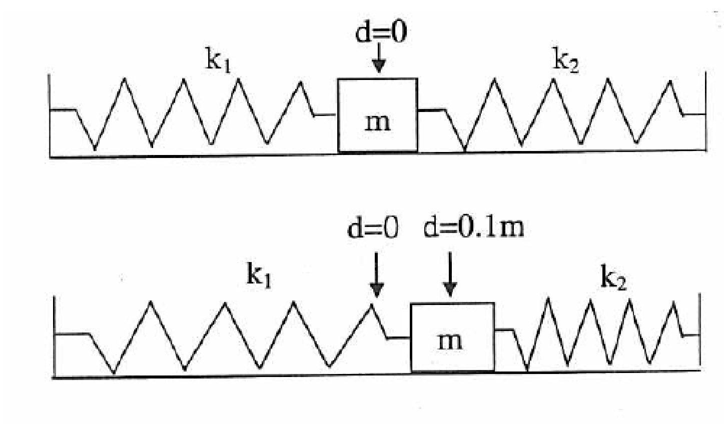
Problem 4 (Calculus-based course):

A uniform rod of mass 3 kg and length 50cm is hinged at one end such that it rotates in a horizontal plane. A putty ball of mass 250gm hits the other end traveling horizontally and perpendicular to the rod with a speed of 1m/s. It sticks to the rod and the rod starts to rotate. What is the kinetic energy of the rod and ball system immediately after the impact? A frictional torque of 0.1 Nm is present in the hinge. Through what angle will the bar rotate before coming to rest? The moment of inertia of a rod of length L and mass M about its center of mass is $ML^2/12$.

Problem 5 (Calculus-based course):

The diagrams show two springs with the same spring constant $k_1=k_2=312$ N/m. One end of each spring is fixed and the other is attached to a mass (m=4kg). The system is free to move in the horizontal direction without friction. The first diagram shows the springs which are neither compressed nor stretched (natural length) with the mass at rest between them at the position d=0. In the second diagram the mass has been moved a distance d=0.1 m, thus stretching the first spring and compressing the second. The mass is released from rest at this position at time t=0.

a) What is the displacement as a function of time: d(t)?
b) What is the velocity at time t = 0.625 s?
c) What is the total energy of the system?

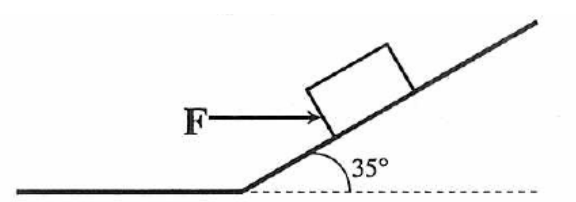


Problem 6 (Algebra-based course):

A batter in a baseball game hits the ball over the center field fence for a home run. The ball is struck 120 cm above the ground with an initial velocity of 40 m/s at an angle of 26^o above the horizontal. A player on the other team makes a great effort to catch the ball, but it flies will above him. At a point just in front of the center field fence, 110 m from where the ball was hit, he leaps straight upward so that his glove reaches a point 3.0 m above the ground. How far above his glove does the ball pass? Neglect the possible effect of air resistance.

Problem 7 (Algebra-based course):

Two movers are unloading a piano of mass 200 kg from a truck using a ramp that is inclined at 35^o above the horizontal. The coefficients of static and kinetic friction between the piano and the ramp are 0.40 and 0.35, respectively. The cell phone of one mover rings and he asks the second mover to hold the piano in place during the call. What magnitude **horizontal** force must the second mover apply to keep the piano stationary? He gets tired after a few minutes, lets go, and jumps out of the way. What is the piano's speed after it slides 3.0 m down the ramp? What fraction of the piano's energy was consumed as work by the friction force?



Problem 8 (Algebra-based course):

An amusement park bumper car of mass *m* (including the occupants) is driven by its exuberant pilot into a head-on elastic collision with a second (target) bumper car that is loaded with junk and is initially at rest. The first bumper car bounces off, with its direction reversed and a speed that is equal to one third of its original speed. Find the total mass of the target bumper car and its contents, expressed as a multiple of *m*. Describe the motion (speed and direction) of the target bumper car after the collision.

Appendix 4: Materials from First Study with Training Raters

Task Instructions

Dear graduate student,

Thank-you for agreeing to help me with my research on physics problem solving assessment! Below you will find instructions for the first part of the task. When you have completed all steps, please return the documents to my mailbox in the envelope provided. I will contact you by e-mail with instructions for Part II.

Jen Docktor
Office 161B, 625-9323
docktor@physics.umn.edu

Instructions Part I:

1) Read the scoring document (rubric) and category descriptions printed on the next page. If there is anything unclear in the wording, make note of it on page 4 of the scoring template Part I.
2) Read the physics problem statement and instructor solution.
3) Look at student solution #1. Use the rubric to assign a separate score of 0, 1, 2, 3, 4, NA(P), or NA(S) for each of the five categories. On the scoring template sheet Part I, record the scores for student #1 and brief notes about your reasoning for each score.
4) Continue the scoring process for student solutions #2-8.
5) Record scoring difficulties on page 4 of the scoring template sheets, and answer the remaining questions.
6) Write your name at the top of each scoring template sheet and the question sheet (four pages). This is only for my reference, and your name will not in any way be associated with results of the study.

Dear graduate student,

Thank-you once again for agreeing to help me with my research on physics problem solving assessment! Below you will find instructions for the *second* part of the task. When you have completed all steps, please return the documents to my mailbox in the envelope provided. I will contact you by e-mail to arrange a brief meeting to discuss your suggestions and comments.

Jen Docktor
Office 161B, 625-9323
docktor@physics.umn.edu

Instructions Part II:

1. If necessary, re-read the scoring document (rubric) and category descriptions printed on the next page.
2. If necessary, read the physics problem statement and instructor solution again.
3. Look at the example scores for student solutions #1-3 and the reasoning for each score on page 5. Compare these scores to your own scores for these solutions from Part I.
4. Look at the student solutions #4-8, which you scored on Part I. Use the rubric to assign a separate score of 0, 1, 2, 3, 4, NA(P), or NA(S) for each of the five categories (you may review your scoring template from Part I as needed). On the blank scoring template sheets for Part II, record your new scores for students #4-8 and brief notes about your reasoning for each score.
5. Continue the scoring process for five new student solutions, #9-13. Record questions and scoring difficulties on page 10 of the scoring template sheets, and answer the questions. Include your name at the top of each scoring sheet. (Note that this is only for my reference, and your name will not in any way be associated with results of the study.)

Sample Scoring Template and Questions:

Student # 1	**Score**	**Notes**
Physics Approach		
Useful Description		
Specific App. of Physics		
Mathematical Procedures		
Logical Organization		

(PART I QUESTIONS)

1. What difficulties did you encounter while using the scoring rubric?
   - Which of the five categories was most difficult to score and why?
   - Which student solutions were the most difficult to score and why?

2. What changes, if any, would you recommend making to the rubric? Why?

3. If you were deciding how to grade these student solutions for an introductory physics course exam, how would you assign points? (out of 20 total points)

(PART II QUESTIONS)

4. What difficulties did you encounter while using the scoring rubric?

5. Were the example scores useful? Why or why not?

6. What further changes, if any, would you recommend making to the rubric?

Problem Solving Rubric Jennifer Docktor [docktor@physics.umn.edu]

	4	3	2	1	0
Physics Approach	The solver has clearly stated an appropriate and complete physics approach.	The approach is clear but contains minor omissions or errors.	The approach is unclear, or an important physics concept or principle of the approach is missing or inappropriate.	An attempt is made to identify relevant physics concepts or principles, but most of the approach is vague, incomplete, or inappropriate.	The solution does not indicate a basic physics approach, or all of the chosen concepts and principles are inappropriate.
Useful Description*	The solution includes an appropriate and useful problem description.	The description is useful but contains minor omissions or errors.	The description is not useful, or a key feature of the description is missing or incorrect.	An attempt is made, but most of the description is not useful, incomplete, or incorrect.	The solution does not include a description, or all of the description is incorrect.
Specific Application of Physics**	The solution indicates an appropriate and complete application of physics to the specific conditions in this problem.	The specific application of physics to this problem contains minor omissions or errors.	An important specific relationship or condition is missing or applied incorrectly.	An attempt is made, but most of the specific application of physics to this problem is missing or incorrect.	The solution does not indicate a specific application of physics, or all of the application is incorrect.
Mathematical Procedures	Suitable mathematical procedures are used during the solution execution.	Suitable mathematical procedures are used with minor omissions or errors.	An important mathematical procedure is missing or is used with errors.	Attempted mathematical procedures are inappropriate, left unfinished, or contain serious errors	There is no evidence of mathematical procedures in the problem solution or all mathematical procedures are inappropriate.
Logical Organization	The entire problem solution is clear, focused, and logically connected.	The solution is clear and focused with minor inconsistencies.	Parts of the solution are unclear, unfocused, and/or inconsistent.	Most of the solution parts are unclear, unfocused, and inconsistent.	The entire solution is unorganized and contains obvious logical breaks.
NA (Problem)	The skill is not necessary for this <u>problem</u>, or constitutes a very small part of the solution.				
NA (Solver)	Explicit statement is not necessary for this <u>solver</u>, as indicated by the overall solution.				

Category Descriptions:

Physics Approach assesses a solver's skill at selecting appropriate physics concepts and principle(s) to use in solving the problem. Here the term *concept* is defined to be a general physics idea, such as the basic concept of "vector" or specific concepts of "momentum" and "average velocity". The term *principle* is defined to be a fundamental physics rule or law used to describe objects and their interactions, such as the law of conservation of energy, Newton's second law, or Ohm's law.

Useful Description assesses a solver's skill at organizing information from the problem statement into an appropriate and useful representation that summarizes essential information symbolically and visually. The description is considered "useful" if it guides further steps in the solution process.
*A *problem description* could include restating known and unknown information, assigning appropriate symbols for variables, defining variables, stating a goal or target, a visualization (sketch or picture), stating qualitative expectations, an abstracted physics diagram (force, energy, motion, momentum, ray, etc.), drawing a graph, stating a coordinate system, and choosing a system.

Specific Application of Physics assesses a solver's skill at applying the physics concepts and principles from their selected approach to the specific conditions in the problem. If necessary, the solver has set up specific equations for the problem that are consistent with the chosen approach.
**A *specific application of physics* could include a statement of definitions, relationships between the defined variables, initial conditions, and assumptions or constraints in the problem (i.e., friction negligible, massless spring, massless pulley, inextensible string, etc.)

Mathematical Procedures assesses a solver's skill at following appropriate and correct mathematical rules and procedures during the solution execution. The term *mathematical procedures* refers to techniques that are employed to solve for target variable(s) from specific equations of physics, such as isolate and reduce strategies from algebra, substitution, use of the quadratic formula, or matrix operations. The term *mathematical rules* refers to conventions from mathematics, such as appropriate use of parentheses, square roots, and trigonometric identities. If the course instructor or researcher using the rubric expects a symbolic answer prior to numerical calculations, this could be considered an appropriate mathematical procedure.

Logical Organization assesses the solver's skills at communicating reasoning, staying focused toward a goal, and evaluating the solution for consistency (implicitly or explicitly). It checks whether the entire problem solution is clear, focused, and organized logically. The term *logical* means that the solution is coherent (the solution order and solver's reasoning can be understood from what is written), internally consistent (parts do not contradict), and externally consistent (agrees with physics expectations).

Mechanics Problem

To raise money for a University scholarship fund, the new IT dean has volunteered to bungee jump from a crane if contributions can be found for 10 scholarships. To add some interest, the jump will be made from 42 m above a pool of water. A 30m bungee cord would be attached to the dean. First you must convince the dean that your plan is safe for a person of his mass, 70kg. The dean knows that as the bungee cord begins to stretch, it will exert a force which has the same properties as the force exerted by a spring. Your plan has the dean stepping off a platform and being in free fall for 30 m before the cord begins to stretch.

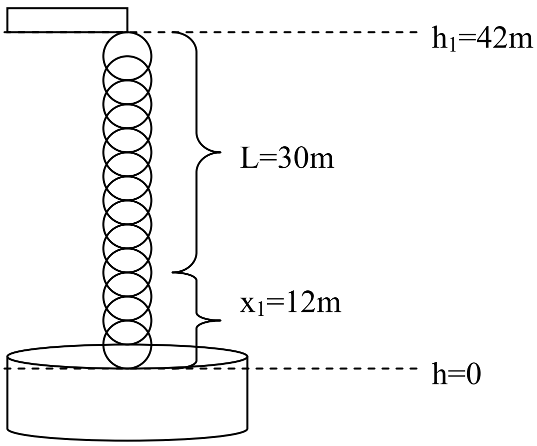
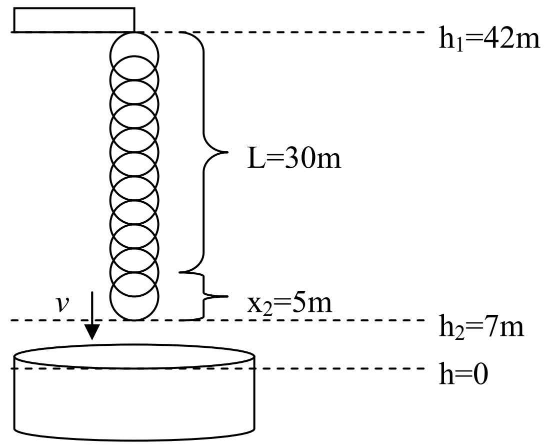
a) Determine the spring constant of the bungee cord so that it stretches only 12m, which will just keep the dean out of the water. (Assume that the dean is a point-like object).
b) Using the result of a), find the dean's speed 7m above the water.

Electricity & Magnetism Problem

You are designing part of a machine to detect carbon monoxide (CO) molecules (28 g/mol) in a sample of air. In this part, ultraviolet light is used to produce singly charged ions (molecules with just one missing electron) from air molecules at one side of a chamber. A uniform electric field then accelerates these ions from rest through a distance of 0.8 m through a hole in the other side of the chamber. Your job is to calculate the direction and magnitude of the electric field needed so that CO^+ ions created at rest at one end will have a speed of 8×10^4 m/s when they exit the other side.

Example Instructor Solutions

(Mechanics Problem) Description

Part a): Find the spring constant, *k*	Part b): Find the velocity of the dean, *v*
	
L=30m; length of the bungee cord h_1=42 m; initial height of the dean x_1=12 m; spring stretch when the dean is at the water surface (h_1-L)	h_1=42 m; initial height of the dean h_2=7 m; final height of dean above the water x_2=5 m; spring stretch when the dean is 7 m above the water (h_1-h_2-L)

Part a): Use conservation of energy: The initial energy is gravitational potential energy at the top of the platform and the final energy is potential energy stored in the stretched spring at h=0.

$$E_{initial} = E_{final} : \qquad mgh_1 = \frac{1}{2}kx_1^2 \qquad \textit{solve for the spring constant } k$$

$$k = \frac{2mgh_1}{x_1^2} = \frac{2(70kg)(9.8m/s^2)(42m)}{(12m)^2} = 400\ kg/s^2 = \boxed{400\ N/m}$$

Part b): Use conservation of energy: The initial energy is gravitational potential energy at the top of the platform and the final energy is kinetic energy, gravitational potential energy at 7m, and potential energy stored in the stretched spring.

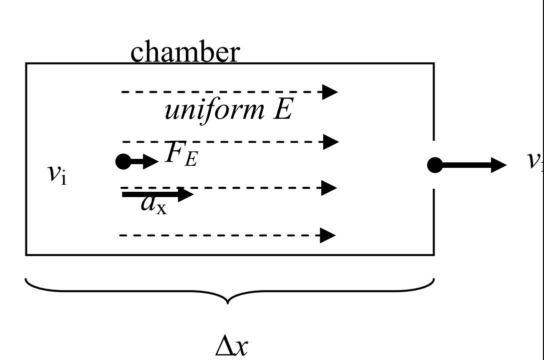
$$E_{initial} = E_{final} : \qquad mgh_1 = mgh_2 + \frac{1}{2}mv^2 + \frac{1}{2}kx_2^2 \qquad \textit{solve for velocity}$$

$$mg(h_1 - h_2) = \frac{1}{2}mv^2 + \frac{1}{2}kx_2^2 \quad \rightarrow \quad 2mg(h_1 - h_2) - kx_2^2 = mv^2$$

$$v = \sqrt{\frac{2mg(h_1 - h_2) - kx_2^2}{m}} = \sqrt{\frac{2(70kg)(9.8m/s^2)(42m - 7m) - (400N/m)(5m)^2}{70kg}} = \boxed{23.3\ m/s}$$

Check: The units are correct for both calculations. The velocity value is reasonable because the dean free-falls for 30m and has velocity $v = \sqrt{2gL} = 24.2\ m/s$ before the bungee spring starts to stretch and slows him down.

(Electricity & Magnetism Problem) **Description**

	v_i= 0; initial velocity of the CO^+ molecule v_f=8x10^{4}m/s; final velocity of the CO^+ molecule E: uniform electric field in the chamber Δx =0.8m; distance to hole in chamber q=1.602x10^{-19}C; charge of a CO^+ molecule m=mass of a CO^+ molecule a_x=acceleration of the CO^+ molecule F_E=force on the CO^+ molecule in the uniform electric field

Target: calculate the electric field, E

Solution Approach 1: Use Newton's Second Law to relate the force on the molecule to its acceleration; use kinematics to write an expression for acceleration in terms of velocity and distance. Assume gravity is negligible. Convert the mass of CO into kilograms per molecule.

$$\sum F_X = ma_X: \quad qE = ma_X \quad \textit{solve for the electric field}: E = \frac{ma_X}{q}$$

$$v_f^2 = v_i^2 + 2a_X\Delta x \quad \textit{solve for acceleration}: a_X = \frac{v_f^2 - v_i^2}{2\Delta x}$$

$$m = \frac{28g}{mol} = \frac{0.028kg}{mol} \bullet \frac{1mol}{6.022\times10^{23} molecules} \approx 4.65\times10^{-26} kg/molecule\, CO^+$$

$$E = \frac{m\left(v_f^2 - v_i^2\right)}{2q\Delta x} = \frac{4.65\times10^{-26}kg\left(\left(8\times10^4 m/s\right)^2 - 0\right)}{2\left(1.602\times10^{-19}C\right)\left(0.8m\right)} = \boxed{1160\, N/C}$$

direction is same as v (to the right.)

Solution Approach 2: Use conservation of energy to relate the electric potential energy transferred to the molecule and its final kinetic energy. Assume gravity is negligible. Convert the mass of CO into kilograms per molecule.

$$E_{final} - E_{initial} = E_{in} - E_{out}: \frac{1}{2}mv_f^2 - \frac{1}{2}mv_i^2 = q\Delta V - 0 \quad OR \quad \frac{1}{2}mv^2 - \frac{1}{2}mv_i^2 = \int \vec{F}\cdot d\vec{s}$$

$$\textit{for uniform electric field}: \Delta V = \int \vec{E}\cdot d\vec{s} = E\Delta x \quad \textit{and} \quad \int \vec{F}\cdot d\vec{s} = F_E\Delta x = qE\Delta x$$

$$\frac{1}{2}mv_f^2 - \frac{1}{2}mv_i^2 = qE\Delta x \quad \textit{solve for electric field}$$

$$m = \frac{28g}{mol} = \frac{0.028kg}{mol} \bullet \frac{1mol}{6.022\times10^{23} molecules} \approx 4.65\times10^{-26} kg/molecule\, CO^+$$

$$E = \frac{m\left(v_f^2 - v_i^2\right)}{2q\Delta x} = \frac{4.65\times10^{-26}kg\left(\left(8\times10^4 m/s\right)^2 - 0\right)}{2\left(1.602\times10^{-19}C\right)\left(0.8m\right)} = \boxed{1160\, N/C}$$

direction is same as v (to the right.)

Written Solutions used for Training

Mechanics Student #1

apply second law:

$$F = ma = kx - mg$$

Because the system has no acceleration at the bottom of the jump:

$$0 = kx - mg$$

$$mg = kx$$

$$k = \frac{mg}{x} = 52.225\ N/m$$

a) 52.225 N/m

$h_1 = 42m$

$h_2 = 7m$

$d = 12m$

$$mgh_1 = \tfrac{1}{2}kd^2 = \tfrac{1}{2}k(d-7)^2 + mgh_2 + \tfrac{1}{2}mv^2$$

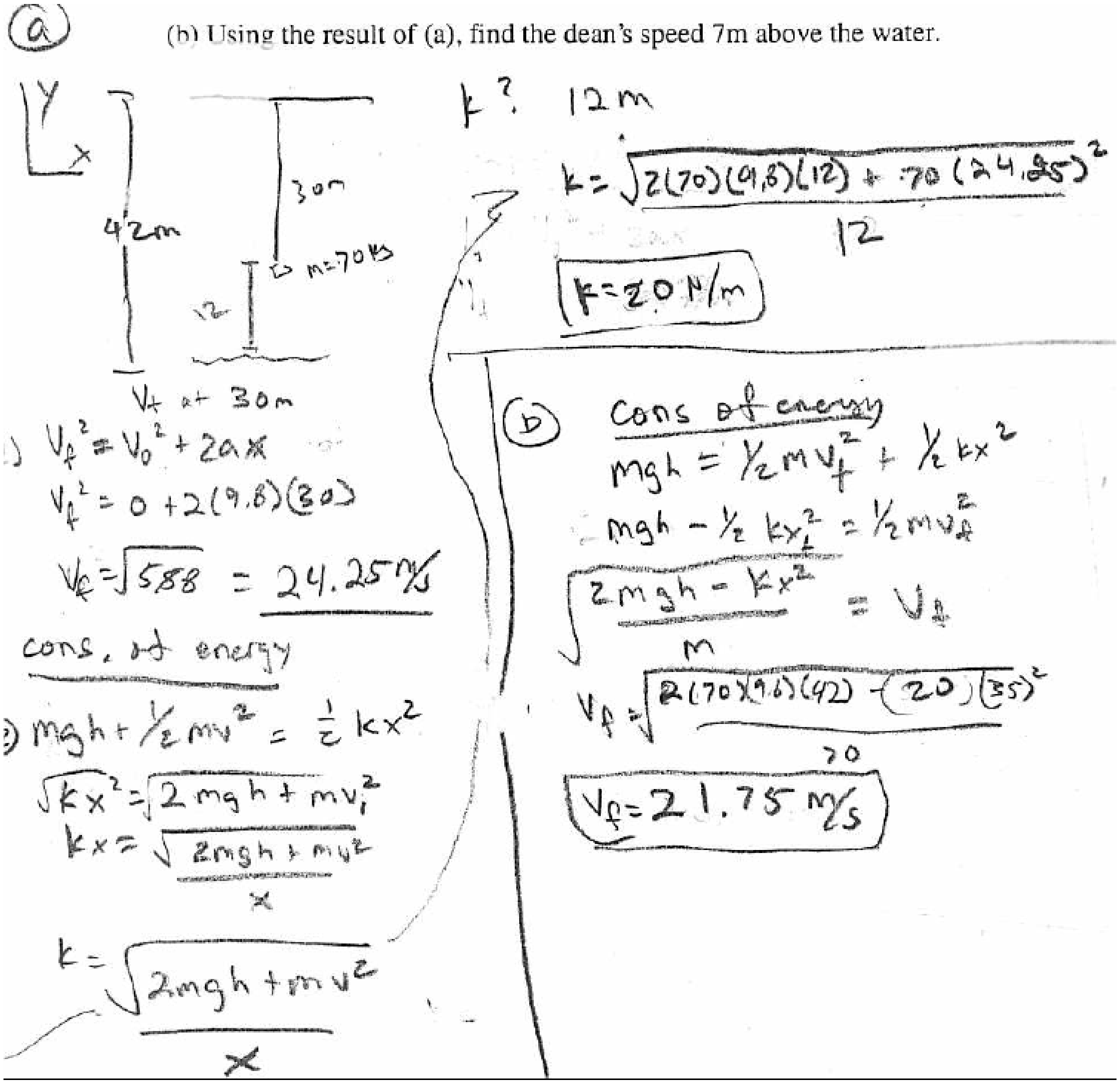
$$mgh_1 = \tfrac{1}{2}k \cdot 25 + mgh_2 + \tfrac{1}{2}mv^2$$

$$2mg(h_1-h_2) = 25k + mv^2$$

$$\frac{2mg(h_1-h_2) - 25k}{m} = v^2 \qquad v = \sqrt{\frac{2mg(h_1-h_2)-25k}{m}}$$

b) v = 25.847 ...

Mechanics Student #2



(a)

(b) Using the result of (a), find the dean's speed 7m above the water.

V_f at 30m

1) $V_f^2 = V_0^2 + 2ax$

$V_f^2 = 0 + 2(9.8)(30)$

$V_f = \sqrt{588} = 24.25 \text{ m/s}$

cons. of energy

2) $mgh + \frac{1}{2}mv^2 = \frac{1}{2}kx^2$

$\sqrt{kx^2} = \sqrt{2mgh + mv_i^2}$

$kx = \sqrt{\frac{2mgh + mv^2}{x}}$

$k = \sqrt{\frac{2mgh + mv^2}{x}}$

k? 12m

$k = \sqrt{\frac{2(70)(9.8)(12) + 70(24.25)^2}{12}}$

$k = 20 \text{ N/m}$

(b) Cons of energy

$mgh = \frac{1}{2}mv_f^2 + \frac{1}{2}kx^2$

$mgh - \frac{1}{2}kx_L^2 = \frac{1}{2}mv_f^2$

$\sqrt{\frac{2mgh - kx^2}{m}} = V_f$

$V_f = \sqrt{\frac{2(70)(9.8)(42) - (20)(35)^2}{70}}$

$V_f = 21.75 \text{ m/s}$

Mechanics Student #3

(b) Using the result of (a), find the dean's speed 7m above the water.

$F_T = kx$ $\quad F_T = F_g$

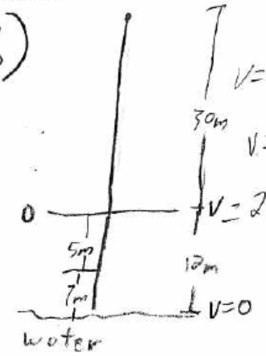
$F_{g\,Dean} = mg$ $\quad kx = mg$

$k\,12 = 70g$

$k = \frac{70g}{12m}$

a) $k = 57.23\ \frac{N}{m}$

b)



$v = \sqrt{2gy}$

$v = \sqrt{60g}$

$v = 24.26\ \frac{m}{s}$

$v_f^2 = v_0^2 + 2ay$

$0 = 588.6 + 24a$

$a = -24.5\ \frac{m}{s^2}$

$v_f^2 = 588.6 - 245.25$

$v_f = 18.53\ \frac{m}{s}$ 7m above the water

Sample Scores used for Training (Mechanics); Rubric version 3

Student # 1	Score	Notes
Physics Approach	**1**	Newton's second law is inappropriate during spring stretch; missing energy conservation for part a); approach in b) is unclear
Useful Description	**NA(S)**	Visualization is unnecessary for this solver; Free-body diagram assumes = forces; defined variables for part b) but not a);
Specific App. of Physics	**2**	Incorrectly assumes acceleration is zero at bottom of jump; does not identify "initial" and "final" energy terms
Mathematical Procedures	**3**	Missing substitution of numerical values during calculations (except d-7); makes a calculation error when finding k in part a)
Logical Organization	**2**	Parts of the solution are unclear due to implicit reasoning; velocity value is greater than free fall after 30 m;

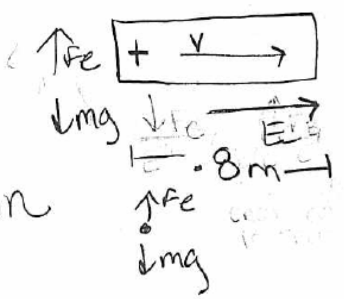
Student # 2	Score	Notes
Physics Approach	**4**	Kinematics is appropriate before spring stretch; conservation of energy approach is explicitly stated
Useful Description	**1**	missing variable definitions; used "h" and "x" w/multiple values; picture is missing variable labels and height/stretch for part b)
Specific App. of Physics	**2**	Does not identify "initial" and "final" energy terms; part b) is missing a mgh term; used incorrect stretch value in part b)
Mathematical Procedures	**2**	Important algebraic mistakes when solving for k (did not need to take square root and incorrectly drops root from k)
Logical Organization	**2**	Should have checked units for k equation in part a) – might have caught inconsistencies;

Student # 3	Score	Notes
Physics Approach	**0**	Equal forces approach is inappropriate; kinematics inappropriate because acceleration not constant; missing energy conservation
Useful Description	**2**	Missing variable definitions such as “y”; picture is missing variable labels; included stretch/height for b)
Specific App. of Physics	**1**	Incorrectly assumes acceleration is constant during spring stretch in part b) and zero in a); doesn’t identify initial and final v’s
Mathematical Procedures	**1**	Missing steps in calculations; early numerical substitutions make procedures difficult to follow (procedures inappropriate);
Logical Organization	**1**	Reasoning for most of the solution is unclear – especially part b); inconsistent assumptions about acceleration

Written Solutions Used for Training
Electricity & Magnetism Student #1

CO: 28 g/mol
UV light
E uniform
CO^+ $v = 8\times10^{14}$ m/s
.8m

Picture:



Question: What is direction and magnitude of E?

approach: use Newton's laws to get force on particle and then use $F_e = qE$ to get E

plan solution

$\Sigma F = ma$

$\Sigma F_y = F_e - mg$

$F_e = qE$

$qE - mg = 0$

$qE = mg$

$E = \frac{mg}{q}$

$q: \; 28\frac{g}{mol} \times 1\,mol = 28\,g \times \frac{1\,kg}{1000\,g} = .028k$

evaluate answer:

reasonable? yes b/c a large electric field is needed to move a molecule of CO .8m

answer Q? yes → direction of E shown in picture

✓ units ✓

execute plan:

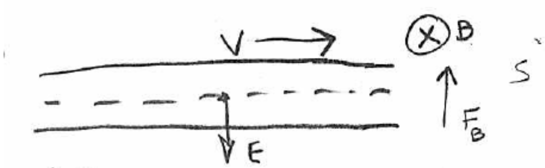
$$E = \frac{(.028\,kg)(9.8\,m/s^2)}{(1.602\times10^{-19}\,C)} = 1.71\times10^{18}\ N/C$$

✓ units:

$kg\,\frac{m}{s^2} \rightarrow N = \frac{N}{C}$ ✓

11

Electricity & Magnetism Student #2



Knowns

CO = 28 grams/mol

d = 0.8 m

$v = 8 \times 10^4 \frac{m}{s}$

$F_{net} = 0$ $F = ma$

$a = \frac{v^2}{r}$

$F_B + F_E = 0$

$F_B = qV \times B$

$F_E = qE$

$F_B - F_E = 0$, V ⊥ E so,

$F_B = F_E$

$F_E = qVB$

$qVB - qE = 0$

$qVB = qE$

solve for B, q cancels and your left w/

$\vec{B} = \frac{\cancel{q} E}{\cancel{q} V} \Rightarrow B = \frac{E}{V} = \frac{E}{8 \times 10^4 \frac{m}{s}}$

Questions: calculate direction and magnitude of electric field needed so CO^+ ions will have speed of $8 \times 10^4 \frac{m}{s}$ when exiting.

Approach: Use Newtons Law to relate the magnetic field + electric magnetic field to find $\vec{B}$ + its direction.

$\frac{\Delta V}{\Delta t}$ $V = \frac{d}{t} = \frac{0.8m}{t}$

$a = \frac{V}{t}$ $V = at$

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

$KE = \frac{1}{2}(28g)(8 \times 10^4 \frac{m}{s})^2$

$KE = 8.96 \times 10^{11}$

$t = \frac{0.8m}{8 \times 10^4 m/s} = 1 \times 10^{-5} \frac{m}{s}$

$a = \frac{v}{t^2}$

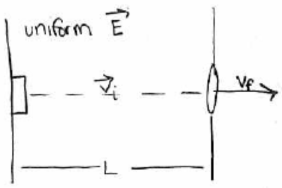
$a = \frac{8 \times 10^4 m/s}{1 \times 10^{-5} s}$

$\Rightarrow 8 \times 10^9 s$

Electricity & Magnetism Student #3

Focus the Problem

Diagram:



$v_f = 8 \times 10^4$ m/s

$L = 0.8$ m

Question:

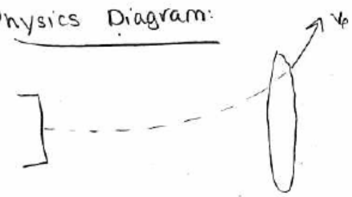
What is the direction & magnitude of the electric field such that the final velocity of the ion is 8×10^4 m/s?

Approach:

Use $\vec{E} = \frac{\vec{F}}{q_0}$, dynamics, Newton's Laws ($F=ma$), & kinematics to determine $\vec{E}$ field, find magnitude 1st then decide direction

Planning the Solution:

Physics Diagram:



$\vec{E} = \frac{\vec{F}}{q_0}$ $F = ma$ $a = \frac{v^2}{R}$

$E_f - E_{in} = 0$

$\Rightarrow E_f = E_i$

Solving the Problem:

1) $a = \frac{v^2}{R}$ } know all but a

2) plug ① into $F=ma$

$\Rightarrow F = \frac{mv^2}{R}$

3) $E = \frac{F}{q}$

$\Rightarrow E = \frac{mv^2}{qR}$

11

Evaluating the Solution

$E = \frac{mv^2}{qR}$ } don't know m

⇒ use CO 28 g/mol:

we will arbitrarily use 1 mol of substance:

$\frac{28\,g}{mol} \cdot \frac{1\,mol}{1} \cdot \frac{1\,kg}{1000\,g} = 0.028\,kg$

⇒ $E = \frac{mv^2}{qR}$

$m = 0.028\,kg$
$v = 8 \times 10^4\,m/s$
$q = -1.602 \times 10^{-19}\,C$
$R = 0.8\,m$

$E = \frac{(0.028\,kg)(8 \times 10^4\,\frac{m}{s})^2}{(-1.602 \times 10^{-19}\,C)(0.8\,m)} = -1.398 \times 10^{27}\,\frac{N}{C}$

↑ implies $\vec{E}$ field is inverted

⇒ $\vec{E}$ direction must be downward.

$\vec{E}$ direction: downward

E magnitude: $1.398 \times 10^{27}\,\frac{N}{C}$

Checking the Answer

units:

$E = \frac{N}{C}$ ⇒ $E = \frac{kg \cdot \frac{m^2}{s^2}}{C \cdot m} = \frac{N \cdot m}{C \cdot m} = \frac{N}{C}$ ✓

Sample Scores used for Training (E&M); Rubric version 3

Student # 1	Score	Notes
Physics Approach	**2**	Use of Newton's second law is appropriate; missing use of kinematics to find acceleration
Useful Description	**2**	Direction of force from horizontal E-field is incorrect; velocity is unclear from picture (constant?) and distance undefined
Specific App. of Physics	**1**	Should neglect the force of gravity; missing analysis of motion in the horizontal direction – missing key specific relationships
Mathematical Procedures	**2**	Missing procedure to convert molar mass to kg
Logical Organization	**2**	Electric field value is unreasonable; direction of force and E-field are inconsistent; answer independent of v and distance

Student # 2	Score	Notes
Physics Approach	**1**	Explicit use of "Newton's law" to equate forces; calculates acceleration and kinetic energy (but approach unclear)
Useful Description	**1**	Velocity is unclear from picture (constant?) and distance not labeled; direction of E-field incorrect; B-field incorrect
Specific App. of Physics	**1**	Incorrectly assumes B-field present; missing analysis of motion in the horizontal direction; missing important relationships
Mathematical Procedures	**1**	Missing procedure to convert molar mass to kg; procedures to solve for target are left unfinished
Logical Organization	**1**	Target variable unclear / inconsistent (B or E?); parts of the solution are inconsistent and unfocused – doesn't reach answer

Student # 3	**Score**	**Notes**
Physics Approach	**3**	Stated approach is appropriate but actual approach is missing kinematics or energy
Useful Description	**2**	Inconsistent use of variables L and R; should indicate direction of E-field and force on the picture; should label R
Specific App. of Physics	**1**	Incorrectly assumes circular motion; assumption to use 1 mol unjustified; assumes neg. charge; missing important relationships
Mathematical Procedures	**2**	Missing an appropriate procedure to convert molar mass to kg
Logical Organization	**3**	Answer unreasonable and unnoticed; stated approach does not match actual approach

Appendix 5: Materials from Second Study with Training Raters

Task Instructions

Introduction:

In this task you will be asked to assess the quality of two student solutions to a physics exam problem using a prescribed scoring technique. Your scores and comments will help me improve the assessment instrument for my thesis research and will also (hopefully) help you reflect on your own teaching practices. Complete the following instructions by yourself - after the written task we will have a class discussion. Return the scoring sheet to me (Jen) before you leave.

Instructions for the scoring task:

1. Read the scoring document (rubric) and category descriptions printed on the next page. If there is anything you find unclear in the wording, write down your comments on page 2 of the scoring sheet (last page of the packet).
2. Read the physics problem statement and think about how solve it. Briefly write down your thoughts in the white space beneath the problem.
3. Check your approach using the instructor solution (other side). Note that there are two possible solutions, and the problem requires a unit conversion.
4. Read each of the scored example solutions **A-E** with scores at the top and score comments in the boxes. Important features are also listed below:

a) Logical progression is good (the solution process is clear) but the application of physics is incorrect

b) Physics approach and math calculations are unnecessary for this solver (NA – Solver)

c) The solution is unfocused and does not progress to an answer

d) Example of a score "1" in physics approach

e) A description is unnecessary for this solver (NA-Solver)

5. Look at student solution **F**. Use the rubric to assign a separate score of **0, 1, 2, 3, 4, 5, NA(Solver), or NA(Problem)** for each of the five categories. On the scoring sheet, record the scores for student solution **F** and any relevant notes. Refer back to the example scores **A-E** as necessary.
6. Continue the scoring process for student solution **G**.
7. Answer the question on page 1 of the scoring sheet. Record comments and scoring difficulties on page 2 of the scoring sheet.
8. Await further instructions for a class discussion. Before you leave, return the scoring sheet to Jen. Results of the task will be reported to you by e-mail within a week.

Sample Scoring Template and Questions:

Student F	Score	Notes
Useful Description		
Physics Approach		
Specific App. of Physics		
Mathematical Procedures		
Logical Progression		

Questions:

1. What features do you usually look for when scoring a student exam paper?

Comments about the rubric scoring activity:

2. What difficulties did you encounter during this activity?

   i) Difficulties understanding the scoring task

   b. Difficulties using the scoring rubric

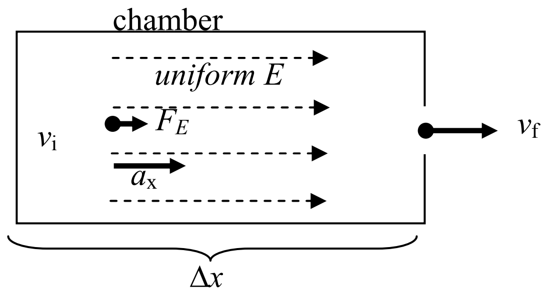
Additional comments:

Problem

You are designing part of a machine to detect carbon monoxide (CO) molecules (28 g/mol) in a sample of air. In this part, ultraviolet light is used to produce singly charged ions (molecules with just one missing electron) from air molecules at one side of a chamber. A uniform electric field then accelerates these ions from rest through a distance of 0.8 m through a hole in the other side of the chamber. Your job is to calculate the direction and magnitude of the electric field needed so that CO^+ ions created at rest at one end will have a speed of 8×10^4 m/s when they exit the other side.

Example Instructor Solution

Description

	v_i= 0; initial velocity of the CO^+ molecule v_f=8x10^{4}m/s; final velocity of the CO^+ molecule E: uniform electric field in the chamber Δx =0.8m; distance to hole in chamber q=1.602x10^{-19}C; charge of a CO^+ molecule m=mass of a CO^+ molecule a_x=acceleration of the CO^+ molecule F_E = force on the CO^+ molecule in the uniform electric field

Target: calculate the electric field, *E*

Solution Approach 1: Use Newton's Second Law to relate the force on the molecule to its acceleration; use kinematics to write an expression for acceleration in terms of velocity and distance. Assume gravity is negligible. Convert the mass of CO into kilograms per molecule.

$$\sum F_X = ma_X : \quad qE = ma_X \quad \textit{solve for the electric field}: E = \frac{ma_X}{q}$$

$$v_f^2 = v_i^2 + 2a_X \Delta x \quad \textit{solve for acceleration}: a_X = \frac{v_f^2 - v_i^2}{2\Delta x}$$

$$m = \frac{28g}{mol} = \frac{0.028kg}{mol} \bullet \frac{1mol}{6.022\times10^{23}\,molecules} \approx 4.65\times10^{-26}\,kg/molecule\,CO^+$$

$$E = \frac{m(v_f^2 - v_i^2)}{2q\Delta x} = \frac{4.65\times10^{-26}kg\left((8\times10^4 m/s)^2 - 0\right)}{2(1.602\times10^{-19}C)(0.8m)} = \boxed{1160\,N/C}$$

$$\textit{direction is same as } v\,(\textit{to the right}.)$$

Solution Approach 2: Use conservation of energy to relate the electric potential energy transferred to the molecule and its final kinetic energy. Assume gravity is negligible. Convert the mass of CO into kilograms per molecule.

$$E_{final} - E_{initial} = E_{in} - E_{out} :$$

$$\frac{1}{2}mv_f{}^2 - \frac{1}{2}mv_i{}^2 = q\Delta V - 0 \quad OR \quad \frac{1}{2}mv^2 - \frac{1}{2}mv_i{}^2 = \int \vec{F}\cdot d\vec{s} - 0$$

$$\textit{for uniform electric field}: \Delta V = \int \vec{E}\cdot d\vec{s} = E\Delta x \quad and \quad \int \vec{F}\cdot d\vec{s} = F_E \Delta x = qE\Delta x$$

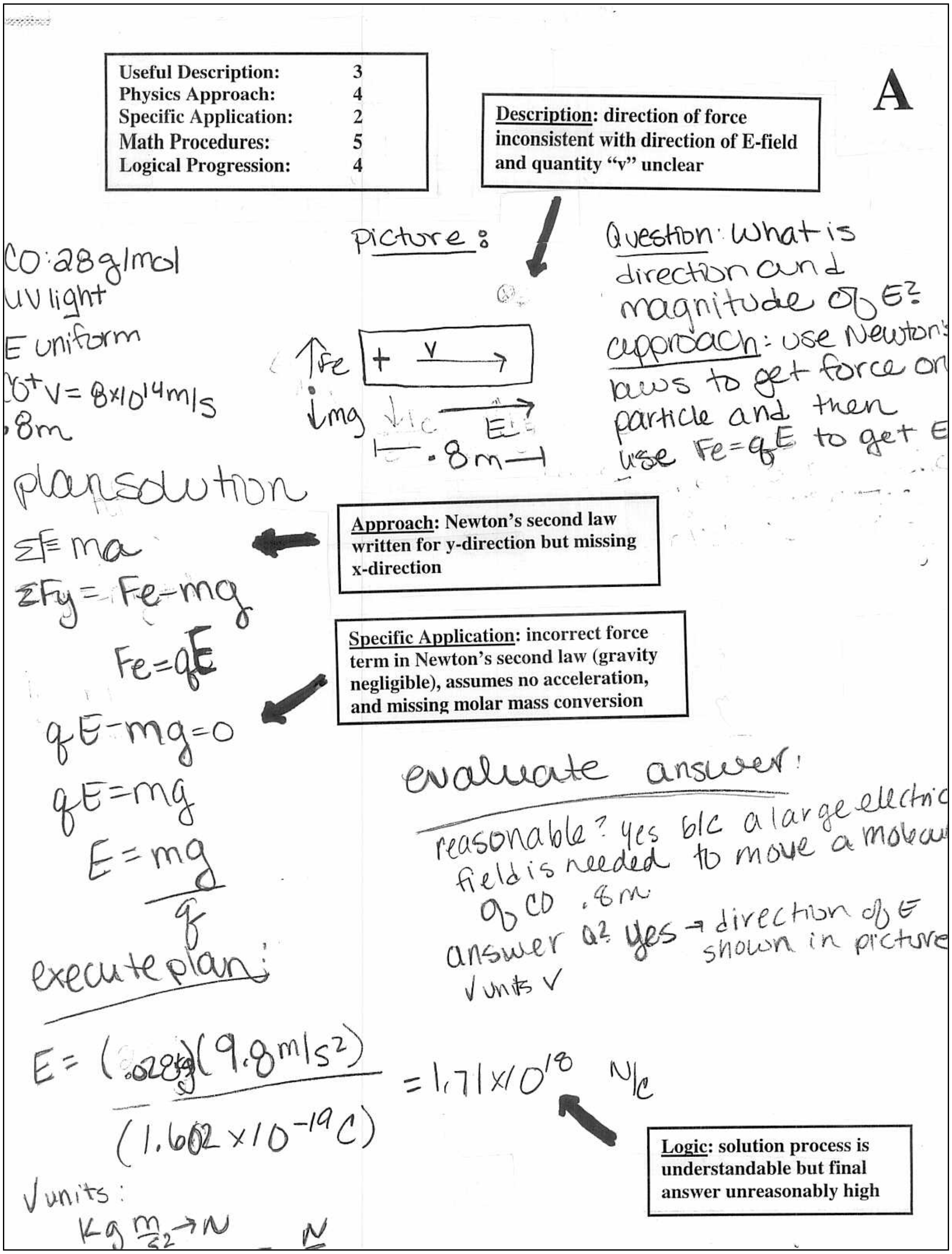
$$\frac{1}{2}mv_f{}^2 - \frac{1}{2}mv_i{}^2 = qE\Delta x \quad \textit{solve for electric field}$$

$$m = \frac{28g}{mol} = \frac{0.028kg}{mol} \bullet \frac{1mol}{6.022\times10^{23}\,molecules} \approx 4.65\times10^{-26}\,kg/molecule\,CO^+$$

$$E = \frac{m(v_f^2 - v_i^2)}{2q\Delta x} = \frac{4.65\times10^{-26}kg\left((8\times10^4 m/s)^2 - 0\right)}{2(1.602\times10^{-19}C)(0.8m)} = \boxed{1160\,N/C}$$

$$\textit{direction is same as } v\,(\textit{to the right}.)$$

Written Solutions used for Training



Useful Description:	**3**
Physics Approach:	**4**
Specific Application:	**2**
Math Procedures:	**5**
Logical Progression:	**4**

A

Description: direction of force inconsistent with direction of E-field and quantity "v" unclear

CO: 28 g/mol
UV light
E uniform
CO⁺ v = 8x10¹⁴ m/s
.8m

Picture:

Fe
mg
+ v
E
.8m

Question: What is direction and magnitude of E?

approach: use Newton's laws to get force on particle and then use Fe=qE to get E

plan solution

ΣF = ma

ΣFy = Fe − mg

Fe = qE

Approach: Newton's second law written for y-direction but missing x-direction

Specific Application: incorrect force term in Newton's second law (gravity negligible), assumes no acceleration, and missing molar mass conversion

qE − mg = 0

qE = mg

E = mg / q

execute plan:

$$E = \frac{(.028 kg)(9.8 m/s^2)}{(1.602 \times 10^{-19} C)} = 1.71 \times 10^{18} \; N/C$$

√units:
Kg m/s² → N   N

evaluate answer:

reasonable? yes b/c a large electric field is needed to move a molecule of CO .8m.

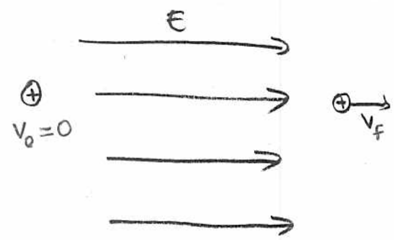
answer Q? yes → direction of E shown in picture

√units √

Logic: solution process is understandable but final answer unreasonably high

Useful Description:	5
Physics Approach:	NA(S)
Specific Application:	5
Math Procedures:	NA(S)
Logical Progression:	5

B



$\Delta x = 0.8 m$

$V_0 = 0$

$V_f = 8 \times 10^4 m/s$

Description: assigns appropriate symbols for quantities

$m = 28 g/mol$

$\frac{28 g}{1 mol} \cdot \frac{1 mol}{6.022 \times 10^{23} atoms} \cdot \frac{1 kg}{1000 g} =$

$m = 4.65 \times 10^{-26} kg$

Question: Find E field so that $V_f = 8 \times 10^4 m/s$.

$\vec{E}$ field points in the same direction as desired velocity (can ignore gravitational force)

F_e: $qE = F_e = ma$

Approach: general equations unnecessary for this solver (NA)

$E = \frac{ma}{q}$

$V_f^2 = V_0^2 + 2a\Delta x$ ($V_0 = 0$)

$a = \frac{V_f^2}{2\Delta x}$

$E = \frac{m V_f^2}{q \, 2\Delta x} = \frac{[kg]\cdot[m]^2}{[s]^2 [C][m]} = \frac{N}{C}$

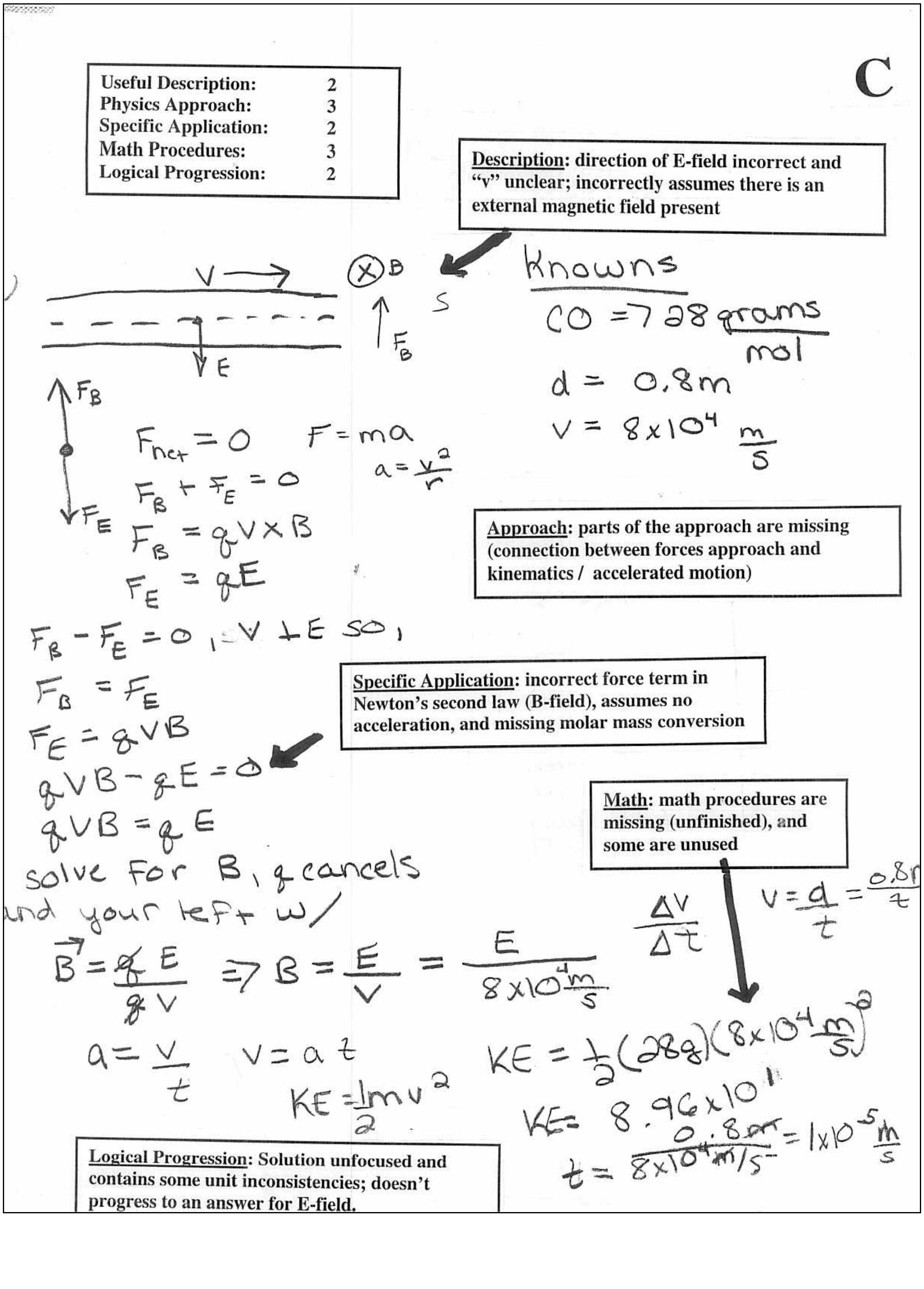
$\frac{V}{m} = \frac{N}{C}$

$N = \frac{kg\,m}{s^2}$

units ok ✓

$E = 1162 \frac{N}{C}$

Math: numerical calculations are not shown but the answer is correct (unnecessary step for this solver)



C

Useful Description:	2
Physics Approach:	3
Specific Application:	2
Math Procedures:	3
Logical Progression:	2

Description: direction of E-field incorrect and "v" unclear; incorrectly assumes there is an external magnetic field present

Knowns

CO = 28 grams/mol

d = 0.8m

$v = 8\times10^4 \frac{m}{s}$

$F_{net} = 0$ $F = ma$ $a = \frac{v^2}{r}$

$F_B + F_E = 0$

$F_B = qV\times B$

$F_E = qE$

$F_B - F_E = 0$, $V \perp E$ so,

$F_B = F_E$

$F_E = qVB$

$qVB - qE = 0$

$qVB = qE$

solve for B, q cancels and your left w/

$\vec{B} = \frac{qE}{qV} \Rightarrow B = \frac{E}{V} = \frac{E}{8\times10^4 \frac{m}{s}}$

$a = \frac{v}{t}$ $v = at$ $KE = \frac{1}{2}mv^2$

$\frac{\Delta v}{\Delta t}$ $v = \frac{d}{t} = \frac{0.8m}{t}$

$KE = \frac{1}{2}(28g)(8\times10^4 \frac{m}{s})^2$

$KE = 8.96\times10^{1}$

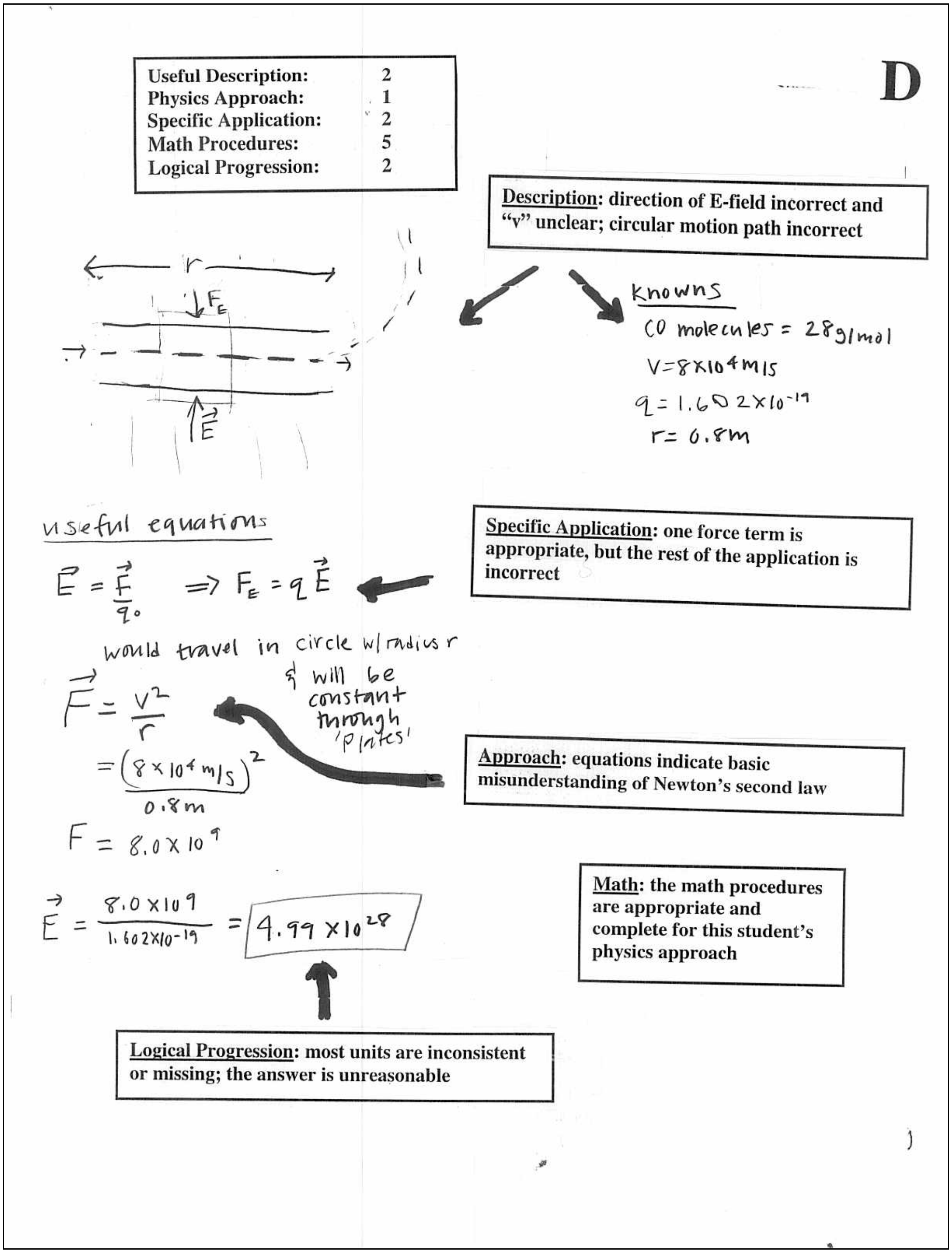
$t = \frac{0.8m}{8\times10^4 m/s} = 1\times10^{-5} \frac{m}{s}$

Approach: parts of the approach are missing (connection between forces approach and kinematics / accelerated motion)

Specific Application: incorrect force term in Newton's second law (B-field), assumes no acceleration, and missing molar mass conversion

Math: math procedures are missing (unfinished), and some are unused

Logical Progression: Solution unfocused and contains some unit inconsistencies; doesn't progress to an answer for E-field.



Useful Description:	2
Physics Approach:	1
Specific Application:	2
Math Procedures:	5
Logical Progression:	2

D

<u>Description</u>: direction of E-field incorrect and "v" unclear; circular motion path incorrect

Knowns

CO molecules = 28 g/mol

V = 8x10^4 m/s

q = 1.602x10^-19

r = 0.8m

useful equations

$\vec{E} = \frac{\vec{F}}{q_0} \Rightarrow F_E = q\vec{E}$

<u>Specific Application</u>: one force term is appropriate, but the rest of the application is incorrect

would travel in circle w/ radius r

& will be constant through 'plates'

$\vec{F} = \frac{v^2}{r}$

$= \frac{(8\times10^4\ m/s)^2}{0.8m}$

$F = 8.0\times10^9$

<u>Approach</u>: equations indicate basic misunderstanding of Newton's second law

$\vec{E} = \frac{8.0\times10^9}{1.602\times10^{-19}} = \boxed{4.99\times10^{28}}$

<u>Math</u>: the math procedures are appropriate and complete for this student's physics approach

<u>Logical Progression</u>: most units are inconsistent or missing; the answer is unreasonable

E

Useful Description:	**NA(S)**
Physics Approach:	**5**
Specific Application:	**5**
Math Procedures:	**4**
Logical Progression:	**4**

Question: what is the Direction and magnitude of an Electric Field is required to move a positively charged CO^{+} ion from rest to a velocity of 8×10^4 m/s in a Distance of .8 m.

Description: **description is missing but solution process correct, so unnecessary (NA-Solver)**

Direction of E = → in the Direction of motion. (the Direction of the Force on the positively charged ion)

Approach: Use conservation of Energy (system defined as everything in the box)

$E_F - E_i = E_{in} - E_{out}$

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

E_i = Electric potential energy = $\Delta V \cdot q$

$\Delta V = -\int E \cdot ds$ (assume we are applying a constant E)

$= E d$

$E_{in} = 0$

$E_{out} = 0$

Logical Progression: **reasoning for unit conversion value is missing**

$E_F = E_i$

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

$\frac{2mv^2}{dq} = E$

$$\frac{2(4.65 \times 10^{-26} \text{ kg})(8 \times 10^4 \text{ m/s})^2}{(.8\text{m})(1.6 \cdot 10^{-19} \text{C})} = \boxed{4650 \text{ N/C}}$$

unit check $\frac{kg\, m^2/s^2}{m \cdot C}$ = N/C ✓checks out

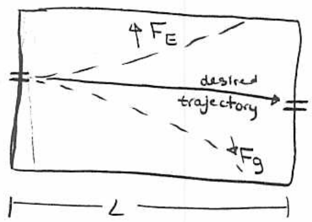
units of E = N/C

big but reasonable considering the high velocity it must obtain

Math: **minor algebra mistake (factor of 2) when solving for E**

F

picture:



$q = e$ $(= +1.602 \times 10^{-19}\,C)$
$v_i = 0\ m/s$
$v_f = 8 \times 10^4\ m/s$
$m = 28\,g/mol$
$L = 0.8\,m$

Question: What magnitude and direction of an electric field should be used for charged particles to reach a velocity of 8×10^4 m/s and experience no net force to make it through the hole on the other side?

Approach: Use Newton's Laws to find the value of $\vec{F}_E$ and then use that information to solve for $\vec{E}$.

$F_{total} = F_E + F_g = 0 \rightarrow F_g = |-F_E| \longrightarrow mg = qE$

and $\vec{E} = \frac{\vec{F}_E}{q_0}$ $\vec{F}_E = q\vec{E}$ $\boxed{E = \frac{mg}{q}}$

$F_g = mg \rightarrow m: \frac{28\,g}{mol\ CO} \cdot \frac{1\,mol}{6.02 \times 10^{23}\,molecule} \cdot \frac{1\,kg}{1000\,g} = 4.65 \times 10^{-26}\,kg/molecule$

$F_g = (4.65 \times 10^{-26}\,kg)(9.8\,m/s^2) = 4.56 \times 10^{-25}\,N$

$4.56 \times 10^{-25}\,N = qE \rightarrow E = \frac{4.56 \times 10^{-25}\,N}{q} = \frac{4.56 \times 10^{-25}\,N}{1.602 \times 10^{-19}\,C}$

11

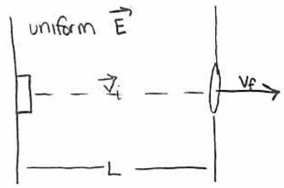
$\boxed{E = 2.85 \times 10^{-6}\,\frac{N}{C} \text{ straight upward}}$

check units: $\frac{N}{C}$ is correct for

G

Focus the Problem

Diagram:



$v_f = 8 \times 10^4$ m/s

$L = 0.8$ m

Question:

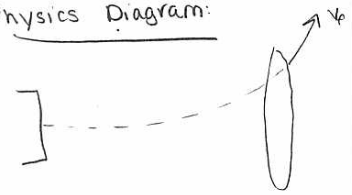
What is the direction & magnitude of the electric field such that the final velocity of the ion is 8×10^4 m/s?

Approach:

use $\vec{E} = \frac{\vec{F}}{q_0}$, dynamics, Newton's Laws $(F = ma)$, & kinematics to determine $\vec{E}$ field; find magnitude 1st then decide direction

Planning the Solution:

Physics Diagram:



$\vec{E} = \frac{\vec{F}}{q_0}$ $F = ma$ $a = \frac{v^2}{R}$

$E_f - E_{in} = 0$

$\Rightarrow E_f = E_i$

Solving the Problem:

① $a = \frac{v^2}{R}$ } know all but a

② plug ① into $F = ma$

$\Rightarrow F = \frac{mv^2}{R}$

③ $E = \frac{F}{q}$

$\Rightarrow E = \frac{mv^2}{qR}$

1

Evaluating the Solution

$E = \frac{mv^2}{qR}$ } don't know m

$\Rightarrow$ use CO 28 g/mol:

we will arbitrarily use 1 mol of substance:

$$\frac{28\,g}{mol} \cdot \frac{1\,mol}{1} \cdot \frac{1\,kg}{1000\,g} = 0.028\,kg$$

$\Rightarrow E = \frac{mv^2}{qR}$

$m = 0.028\,kg$
$v = 8 \times 10^4\,m/s$
$q = -1.602 \times 10^{-19}\,C$
$R = 0.8\,m$

$$E = \frac{(0.028\,kg)(8 \times 10^4\,\frac{m}{s})^2}{(-1.602 \times 10^{-19}\,C)(0.8\,m)} = -1.398 \times 10^{27}\,\frac{N}{C}$$

↑ implies $\vec{E}$ field is inverted

$\Rightarrow \vec{E}$ direction must be downward

$\vec{E}$ direction: downward

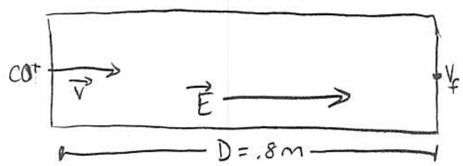
E magnitude: $1.398 \times 10^{27}\,\frac{N}{C}$

Checking the Answer:

units:

$E = \frac{N}{C} \Rightarrow E = \frac{kg \cdot \frac{m^2}{s^2}}{C \cdot m} = \frac{N \cdot \cancel{m}}{C \cdot \cancel{m}} = \frac{N}{C}$ ✓

H



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

$V_i = 0$

$V_f = 8E4\ m/s$

$\Sigma F = F_E = ma$

$F_E = Eq$

Q: Calculate direction & magnitude of Electric field.

Approach/Solve:

Because CO^+ is a positive charge, the electric field will be moving in the direction of CO^+.

To find magnitude of E:

$\Sigma F = F_e = ma$

$m_{CO^+} = 28g = .028\ kg$

$V_f^2 = \cancel{V_i^2}^{\,0} + 2a\Delta x$

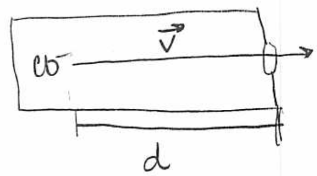
$V_f^2 = 2aD$

$a = \frac{V_f^2}{2D} = \frac{(8E4\ \frac{m}{s})^2}{2(.8m)} = 50{,}000\ \frac{m}{s^2}$

$F_e = ma = .028\ kg \times 50{,}000\ \frac{m}{s^2} = 1400 = Eq$

$E = \frac{1400}{q}$

I



Question: Find the direction and magnitude of the electric field needed to move the CO molecules through the hole at 8×10^4 m/s

$d = .8$ m

$\vec{v} = 8 \times 10^4$ m/s

Approach: use conservation of energy assume gravity to be negligible in comparison to electric force

System: CO^+, box, Earth

t_i = when CO^+ molecule is at rest

t_f = when CO^+ molecule is leaving the box

$E_i = -\int \vec{F} \cdot d\vec{s}$

$E_f = \frac{1}{2} m v^2$

$E_{input} = q_0 \vec{E}$

$E_{output} = 0$

Quantitative Relationships

$\vec{E} = \frac{\vec{E}}{q_0}$ $\quad PE = -\int \vec{F} \cdot d\vec{s}$ $\quad E_f - E_i = E_{in} - E_{out}$

$KE = 1/2 m v^2$

$\sum F = 0$

15

We want the velocity and field parallel to make this motion.

$-\int \vec{F} \cdot d\vec{s} = -q\vec{E}d\cos 0° = -qEd$

Now all that's left is

$\frac{1}{2}mv^2 + q\vec{E}d = q_0\vec{E}$

$\frac{1}{2}mv^2 = q_0\vec{E}(1-d)$

$\vec{E} = \frac{\frac{1}{2}mv^2}{q(1-d)}$

$m_{CO^+} = 28\frac{g}{mol} \cdot \frac{mol}{6.022 \times 10^{23} ions}$

$m_{CO^+} = 4.65 \times 10^{-23}$ g/ion $= 4.65 \times 10^{-}$ Kg

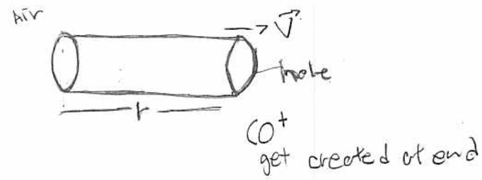
$q_0 = 1.602 \times 10^{-19} C$ because it has a +1 charge

$\vec{E} = \frac{\frac{1}{2}(4.65 \times 10^{-20} kg)(8 \times 10^4 m/s)^2}{1.602 \times 10^{-19} C\ (1 - .8m)}$ units $\frac{N}{C}$ ✓

$\vec{E} = 4.64 \times 10^9$ N/C !

That is ridiculously high for the field but then again making a particle move from rest to 8,000 m/s in less than a meter is ridiculous too.

J



$V = 8\times10^4$ m/s
$r = 0.8$ m
CO = 28g/mol.

Question.

- Calculate the direction and magnitude of the electric field needed so that ions created at rest at one end will have a speed of 8×10^4 m/s when the exit the othe side.

Approach

Use coulom's Law. to find out the magnitube of the electric field

Direction should be. right side

Solution.

$F = K_e \dfrac{q_1 q_2}{r^2}$ $\quad K_e = 9.00\times10^9 \dfrac{N\cdot m^2}{C^2}$

$\vec{F} = q\,\vec{v}\times\vec{B}$

F=

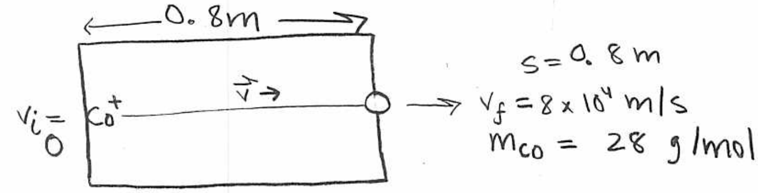
$B = \dfrac{F}{q_1 v} = \dfrac{K_e q_1 q_2}{r^2}\cdot\dfrac{1}{q_1 v} = \dfrac{K_e q_2}{r^2\cdot v} = \dfrac{9.00\times10^9 \frac{N\cdot m^2}{C^2} \times 28\,g/mol}{0.64\,m^2\cdot 8\times10^4\,m/s} = 4.92\times10^6 \dfrac{N\cdot g}{C^2\cdot m}$

Unit ✓ kind of weird unit

answer ✓

reasonable?. ✓ yes

K



Question: Calculate the direction and magnitude of the electric field needed so that Co^+ ions created at rest at one end will have a speed of 8×10^4 m/s when they exit the other side

Approach: Use conservation of energy

system: Co^+ particle

intial time: right as Co^+ enters the box

final time: right as Co^+ leaves the box

$E_i = 0 \quad v_i = 0$

$E_f = \frac{1}{2}mv^2 \quad E_{in}$ = electric potential energy

$E_{out} = 0$

Electric potential energy $= \Delta V q = -q\int \vec{E} \cdot ds$

because the electric field is constant

$PE_e = -qE\int ds$

$PE_e = -qES$ ← just want magnitude so can leave negative sign off

$E_f - E_i = E_{in} - E_{out}$

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

$E = \frac{\frac{1}{2}mv^2}{qs}$

check units

$\frac{1}{2}mv^2 \rightarrow$ energy units

$E \cdot s, q$
$V \cdot C$ = energy units

units ok ✓

11

$$|E| = \frac{\frac{1}{2}mv^2}{qs} = \qquad J = n \cdot m$$

$CO - e^- = CO^+$

$.028\ Kg - 9.11 \times 10^{-31}\ Kg = 0.028\ Kg$

$$|E| = \frac{\frac{1}{2}(0.028\ Kg)(8 \times 10^4\ m/s)^2}{(1.602 \times 10^{-19}\ C)(0.8\ m)}$$

$$\boxed{|E| = 6.99 \times 10^{26}\ N/C}$$

units ✓

$$\frac{Kg \cdot m^2/s^2}{C \cdot m} = N/C$$

evaluate

✓ units OK

✓ seems like a rather large magnitude for an electric field but it would make sense because it takes a lot to accelerate a CO^+ particle from rest to 8×10^4 m/s in such a short distance

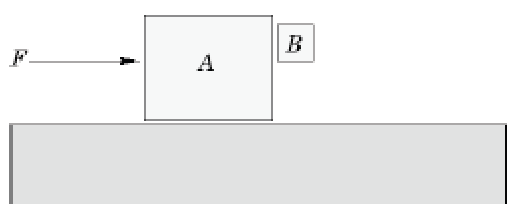
The Electric field should be pointing right because it is the only force on the CO^+ particle and since the particle is being accelerated to the right, the force should be to the right.

Test 1 Problem 1
A block of mass m = 2.5 kg starts from rest and slides down a frictionless ramp that makes an angle of $\theta = 25^o$ with respect to the horizontal floor. The block slides a distance *d* down the ramp to reach the bottom. At the bottom of the ramp, the speed of the block is measured to be $v = 12$ m/s.
(a) Draw a diagram, labeling θ and d. [5 points]
(b) What is the acceleration of the block, in terms of g? [5 points]
(c) What is the distance *d*, in meters? [15 points]

Test 1 Problem 2
A punter kicks a football during a critical football game. The ball leaves his foot at ground level with velocity 20.0 m/s at an angle 40^o to the horizontal. At the very top of its flight, the ball hits a pigeon. The ball and the pigeon each stop immediately and both fall vertically straight to the ground from the point of collision.
(a) With what speed is the ball moving when it hits the pigeon? [10 points]
(b) How high was the ball when it hit the pigeon? [10 points]
(c) What is the speed of the ball when it hits the ground? [5 points]

Test 2 Problem 1 (version 1)

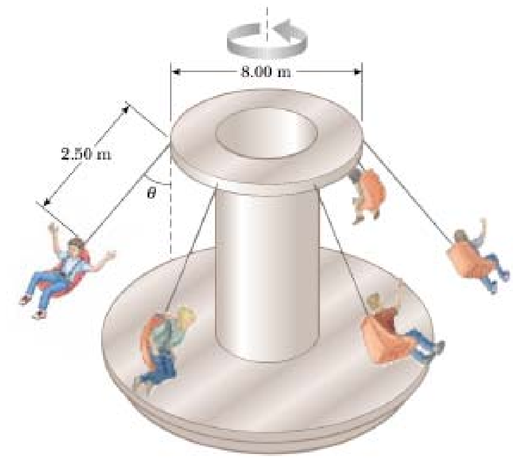
The mass of block *A* is 75kg and the mass of block *B* is 15kg. The coefficient of static friction between the two blocks is $\mu = 0.45$. The horizontal surface is frictionless. What minimum force *F* must be exerted on block *A* in order to prevent block *B* from falling?



Test 2 Problem 1 (version 2)

An amusement park ride consists of a rotating circular platform 8.00 m in diameter from which 10.0-kg seats are suspended at the end of 2.50-m massless chains (see figure). When the system rotates, the chains make an angle $\theta=28.0^{o}$ with the vertical.

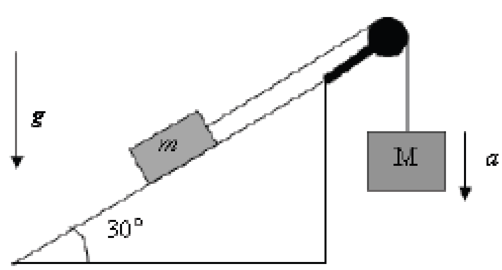
a) What is the speed of each seat? [5 pts]
b) Draw a free-body diagram of a 40.0-kg child riding in a seat [5 pts]
c) Find the tension in the chain [15 pts]



Test 2 Problem 2 (version 1)

A block of mass m = 3 kg and a block of unknown mass M are connected by a massless rope over a frictionless pulley, as shown below. The kinetic frictional coefficient between the block m and the inclined plane is $\mu_k = 0.17$. The plane makes an angle 30^{o} with horizontal. The acceleration, a, of the block M is
1 m/s^2 downward.

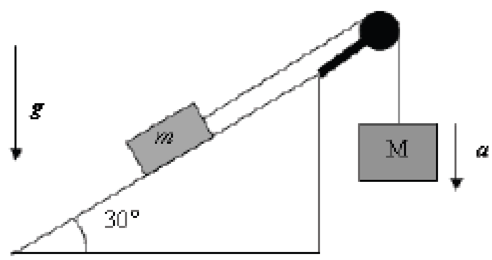
(B) Draw free-body diagrams for both masses. [5 points]
(C) Find the tension in the rope. [5 points]
(D) If the block M drops by 0.5 m, how much work, W,is done on the block m by the tension in the rope? [15 points]



Test 2 Problem 2 (version 2)

A block of known mass *m* and a block of unknown mass *M* are connected by a massless rope over a frictionless pulley, as shown. The kinetic frictional coefficient between the block *m* and the inclined plane is μ_k. The acceleration, *a*, of the block *M* points downward.

a) If the block *M* drops by a distance *h*, how much work, *W*, is done on the block *m* by the tension in the rope? Answer in terms of known quantities [15 pts]
b) Now let the mass m=3kg, the coefficient of kinetic friction between the block *m* and the inclined plane be $\mu_k = 0.17$, and the acceleration *a*, of the block *M* be 1 m/s^2 downward. How much work, *W*, is done on the block *m* by the tension in the rope if the block *M* drops by 0.5m? [5 pts]
c) If the inclined plane were frictionless, would the total work done on both blocks by the tension in the rope increase, decrease, or stay the same?



Test 3 Problem 1:

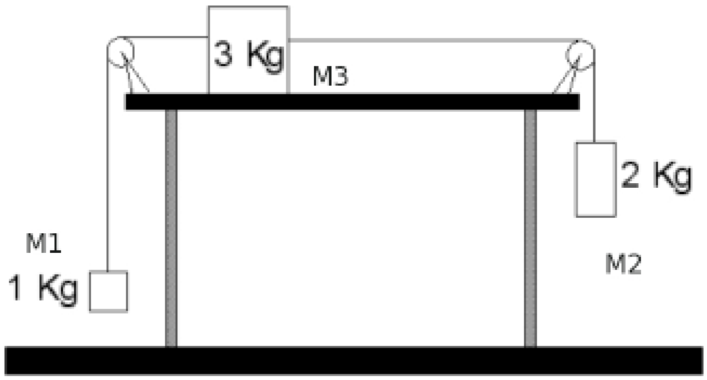
The system of three blocks shown is released from rest. The connecting strings are massless, the pulleys ideal and massless, and there is no friction between the 3kg block and the table.

(A) At the instant M_3 is moving at speed v, how far *d* has it moved from the point where it was released from rest? (answer in terms of M_1, M_2, M_3, *g* and *v*.)

(B) At the instant the 3 kg block is moving with a speed of 0.8 m/s, how far, d, has it moved from the point where it was released from rest? [5 pts]

(C) From the instant when the system was released from rest, to the instant when the 1 kg block has risen a height h, which statement (1, 2 or 3) is true for the three-block system? (1) The total mechanical energy of the system increases. (2) The total potential energy of the system increases. (3) The net work done on the system by the tension forces is 0. [5pts]

(D) Now suppose the table is rough and has a coefficient of kinetic friction $\mu k = 0.1$. What is the speed, v, of the 3 kg block after the 2 kg block drops by 0.5 m? (Assume again that the system is released from rest.) [5pts]



,

Test 3 Problem 2

A Radium-226 atom at rest emits an alpha particle (a 4He nucleus). The energy released during the process is $7.8 \times 10-13$ Joules (roughly 5 MeV).

(A) What is the speed of the alpha particle? [10 points]

(B) What is the speed of the the remaining Radon-222 nucleus? [10 points]

(C) What is the kinetic energy of the Radon-222 nucleus? (The masses of the nuclei may be taken to be $226m_o$,$4.0m_o$, and $222m_o$, respectively, where $m_o = 1.66 \times 10^{-27}$kg). [5 points]

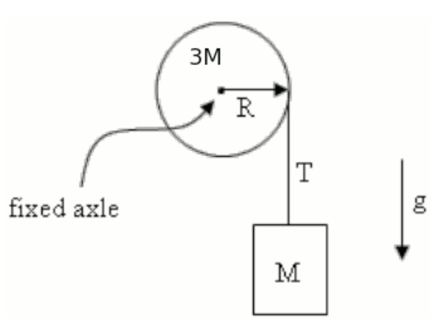
Test 4 Problem 1 (version 1)

A hollow cylinder of mass 3M and radius R rotates on a horizontal frictionless axle through its center. A weight of mass M hangs vertically from a light string wrapped around the cylinder. When the system is released, the falling weight causes the cylinder to turn as the string unwraps. The moment of inertia for a hollow cylinder is mr^2.

(A) Assuming the string does not slip on the cylinder, what is the tension in the string? (Answer in terms of given quantities)[10 pts]

(B) What is the acceleration of the mass? [5 pts]

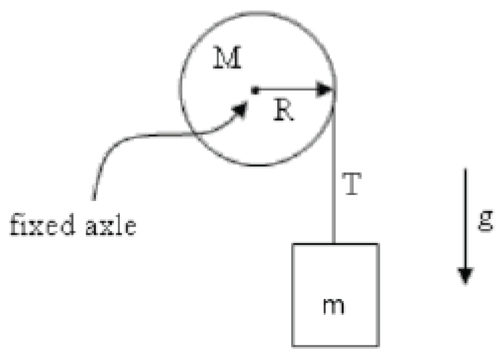
(C) After the mass has fallen a distance *h*, how fast is it moving? (Do not use kinematics.) [10 pts]



Test 4 Problem 1 (version 2)

A massless rope is wrapped around a hollow cylinder ($I = MR^2$) of radius R whose central axis is fixed in a horizontal position. A mass m hangs from the rope and, starting from rest, moved a distance d in time Δt.

(A) Draw a FDB for the hanging mass m. [5 pts]
(B) What is the mass M of the cylinder (in terms of known quantities)? [10 pts]
(C) Now let the cylinder have radius R = 12 cm, the small mass m = 4kg, the distance d = 180cm, and let the time be Δt = 2 seconds. What is the mass M, in kilograms? [5 pts]
(D) Through what angle θ does the cylinder rotate during the 2 seconds? [5 pts]



Test 4 Problem 2 (version 1)

A cockroach of mass m sits on the rim of a uniform disk of mass $4m$ that can rotate freely about its center – like a merry-go-round for roaches. Initially the cockroach and disk rotate together with an angular velocity ω. Then the bug walks halfway to the center of the disk.

(A) What is the new angular velocity of the roach-disk system? [10pts]
(B) What is the ratio K/K_o of the new kinetic energy of the system to its initial kinetic energy? [10 pts]
(C) What accounts for the change in kinetic energy? [5pts]

Test 4 Problem 2 (version 2)

A Physics 1301 student (mass m) runs (speed v) along a line tangent to the edge of a motionless merry-go- round and jumps on at the very outside. The merry-go-round has the shape of a uniform disk, with $I = 12\ MR^2$. Pictured is the top view.

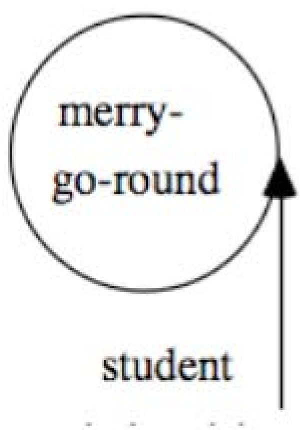
(A) Is angular momentum conserved in this interaction? Why or why not? [5 pts]

(B) Is kinetic energy conserved? If it is, say why. If it's not, where did the energy go? [5 pts]

(C) Find these two ratios:

1. final angular momentum to initial angular momentum *Lf/Li*;
2. final KE to initial KE, *Kf/Ki* . [5 pts]

(D) Find the final angular velocity of the merry-go-round + student system. [10pts]



Test 3 problem 1: Examples of student solutions

The most common application error (25% of students) was to only consider the kinetic energy of block 3, rather than the kinetic energy of all three blocks. An example of this application error is shown below.

(A) $\Delta PE_{M_2} = KE_{M_3} + \Delta PE_{M_1}$

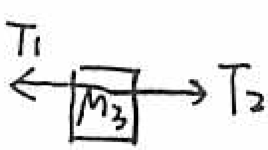
$$M_2 g(\Delta h) = \frac{1}{2} M_3 v^2 + M_1 g(\Delta h)$$

$$M_2 g(\Delta h) - M_1 g(\Delta h) = \frac{1}{2} M_3 v^2$$

$$g\Delta h(M_2 - M_1) = \frac{1}{2} M_3 v^2$$

$\Delta h = d$ so $\boxed{d = \dfrac{\frac{1}{2} M_3 v^2}{g(M_2 - M_1)}}$

Another common error was to apply Newton's Second Law with incorrect reasoning that the tension in each string was equal to the weight of the hanging masses. At least 15% of students misapplied Newton's Second Law with this reasoning. An example is provided below.



(A) $M_2 g - M_1 g = M_3 a \Rightarrow a = \dfrac{M_2 g - M_1 g}{M_3}$

$$S = \frac{V_t^2 - V_0^2}{2a} = \frac{V_t^2}{2a} = \frac{V_t^2}{\frac{2(M_2 g - M_1 g)}{M_3}} = \frac{M_3 V_t^2}{2(M_2 - M_1)g}$$

For some student solutions, the final answer is correct but the reasoning is unclear. An example is shown below. For this student, it is possible that the answer was obtained using correct reasoning (F represents net external forces) but it is also possible that the student used false reasoning, such as the T=Mg error.

(A) $$a = \frac{F}{m} = \frac{M_2 g - M_1 g}{M_1 + M_2 + M_3}$$

$$2ad = \Delta V^2$$

$$d = \frac{\Delta V^2}{2a} = \frac{(M_1 + M_2 + M_3)V^2}{2(M_2 g - M_1 g)}$$

Appendix 7: Interview Materials

Recruitment Flier

Volunteers Needed

to participate in a University of Minnesota
research study on physics problem solving

- We are looking for volunteers to participate in a one-hour problem solving interview.

- You will be asked to solve physics problems similar to ones in your physics class and explain your reasoning to a researcher. During the interview, you will be video and audio taped.

- You will receive $25 upon completion of the interview.

- For more information, contact Jennifer Docktor (docktor@physics.umn.edu or 612-625-9323)

Consent Form

CONSENT FORM
Using Computers as Personal Problem Solving Coaches

You are invited to be in a research study of physics problem solving. You were selected as a possible participant because you are enrolled in an introductory physics course at the University of Minnesota and you volunteered. We ask that you read this form and ask any questions you may have before agreeing to be in the study. This study is being conducted by: Leon Hsu, Department of Postsecondary Teaching and Learning; Ken Heller, Department of Physics; and Jennifer Docktor, Department of Physics.

Background Information
The purpose of this study is to investigate the problem solving processes used by students in introductory physics courses. This information will be used to design and modify a problem solving assessment rubric.

Procedures:
If you agree to be in this study, we would ask you to do the following things:

1. Attempt to solve one or two physics problems printed on a worksheet for approximately 30 minutes. During this time your actions will be videotaped and your voice will be recorded.
2. Participate in a 20-30 minute interview in which you explain your written work and reasoning processes to an investigator. You will continue to be video and audio taped during this interview.
3. Allow researchers to access your physics course grade at the end of the semester. This information will only be used to compare your problem-solving performance on the interview tasks with your overall performance in the course. All academic records will be kept private.

Risks and Benefits of being in the Study
The study has no appreciable risks. We hope that you will acquire additional practice solving physics problems similar to those in your physics course.

Compensation:
If you complete the procedures listed above, you will receive payment of $25 upon completion of the problem-solving interview.

Confidentiality:
The records of this study will be kept private. In any sort of report we might publish or presentation we might make, we will not include any information that will make it possible to identify a subject. Research records will be stored securely and only researchers will have access to the records. Video and audio tapes will only be accessible to the researchers and will be destroyed three years after the completion of the study.

Voluntary Nature of the Study:
Participation in this study is voluntary. Your decision whether or not to participate will not affect your current or future relations with the University of Minnesota or the Department of Physics. If you decide to participate, you are free to not answer any question or withdraw at any time without affecting those relationships.

Contacts and Questions:
The researchers conducting this study are: Leon Hsu, Ken Heller, and Jennifer Docktor. You may ask any questions you have now. If you have questions later, **you are encouraged** to contact Leon Hsu at 378 Appleby Hall, 612-625-3472, lhsu@umn.edu.

If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher(s), **you are encouraged** to contact the Research Subjects' Advocate Line, D528 Mayo, 420 Delaware St. Southeast, Minneapolis, Minnesota 55455; (612) 625-1650.

You will be given a copy of this information to keep for your records.

Statement of Consent:
I have read the above information. I have asked questions and have received answers. I consent to participate in the study.

Signature:_________________________________ Date: __________________

Signature of Investigator:____________________ Date: __________________

IRB Code #0903S60722
Version Date: March 30, 2009

Physics 1301 Equation Sheet

- If $\vec{r}(t)$ is the position of the object as a function of time then velocity is $\vec{v}(t)=\frac{d\vec{r}}{dt}$ and acceleration is $\vec{a}(t)=\frac{d^2\vec{r}}{dt^2}$.
- When the acceleration is a constant $\vec{a}$ then $\vec{r}(t)=\vec{r}_o+\vec{v}_o t+\frac{1}{2}\vec{a}t^2$.
- For motion in a circle of radius *R*, $v=R\omega$, $s=R\phi$, and the centripetal acceleration is $a_c=\omega^2 R$.
- Newton's Laws: $\vec{F}=m\vec{a}$ and $\vec{F}_{12}=-\vec{F}_{21}$
- Common forces include static friction ($F\le\mu_s F_N$), kinetic friction ($F=\mu_k F_N$), gravitational force ($F=mg$), drag ($F=\frac{1}{2}\rho A C_D v^2$) and the spring force ($F=-kx$)
- Kinetic energy is $\frac{1}{2}mv^2$, work is $W=\int\vec{F}\cdot d\vec{x}$, gravitational potential energy is $U_g=mgh$, and the spring potential energy is $U_s=\frac{1}{2}kx^2$.
- Rotational physics: $K=\frac{1}{2}I\omega^2$, $\tau=I\alpha=Fr\sin\theta_{rF}$, $\theta=\theta_o+\omega_o t+\frac{1}{2}\alpha t^2$, $L=I\omega$.
- Moments of inertia: $I=\sum_i m_i R_i^2$
    - For objects on axes through the center of mass: MR^2 - hollow cylinder, $\frac{1}{2}MR^2$ - solid cylinder, $\frac{2}{5}MR^2$ - solid sphere, $\frac{2}{3}MR^2$ - hollow sphere, $\frac{1}{12}ML^2$ - thin rod.
    - For parallel axes: $I=I_c m+Md^2$
- The acceleration due to gravity on Earth is 9.8 m/s^2 or 32.2 ft/s^2.
- Some conversions: 1 meter = 3.281 feet, 1 in = 2.540 cm, 1 lb = 4.448 N
- The solutions to the quadratic equation $0=ax^2+bx+c$ are given by $x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$.
- The following derivatives and integrals may be useful:
    - If $y(t)=Ax^m$, where *A* and *m* are constants, then $\frac{dy}{dx}=Amx^{m-1}$; $\int y\,dx=\frac{A}{m+1}x^{m+1}$, for $m\neq-1$

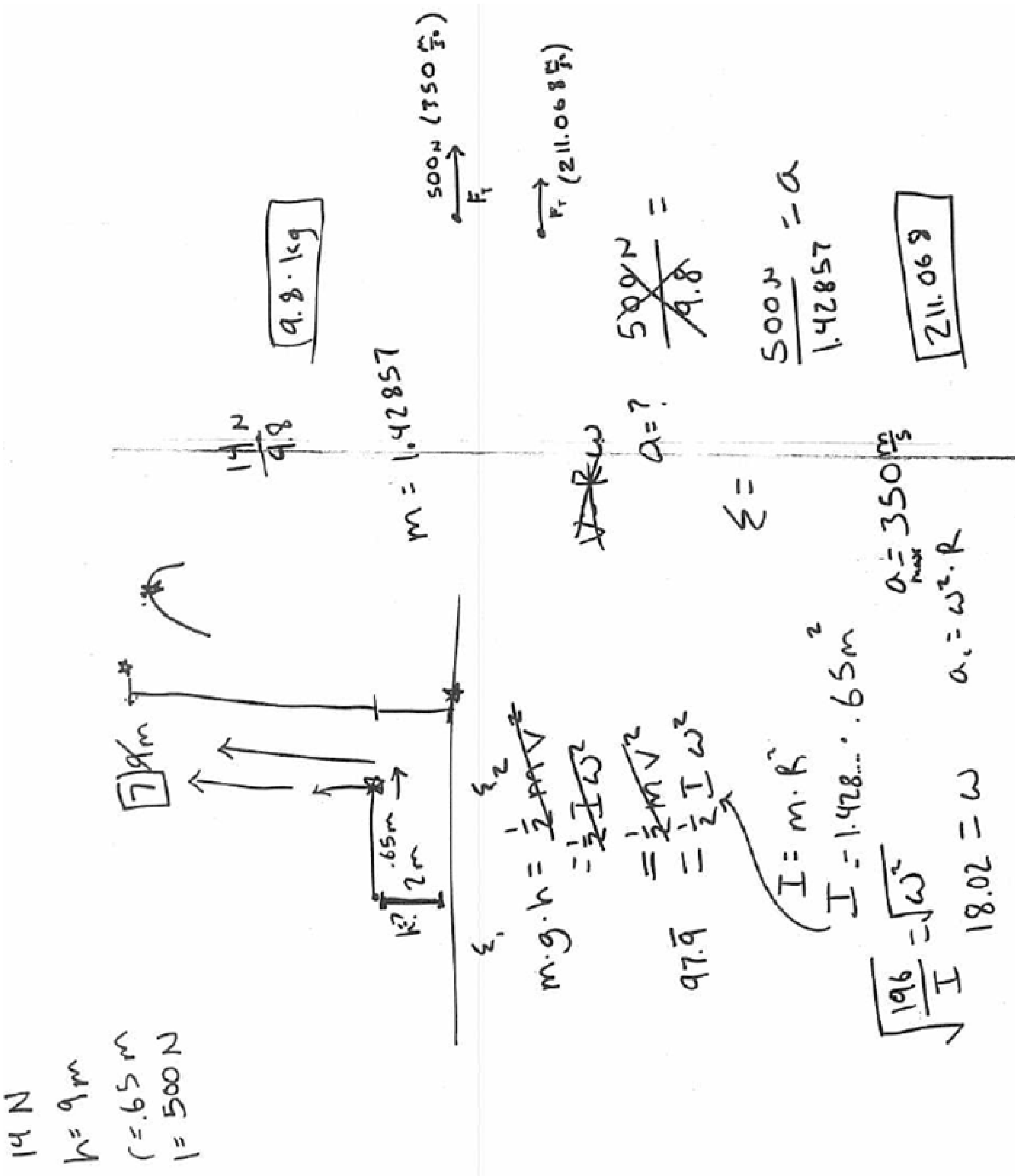
Problem-Solving Tasks

Problem 1:
You are working at a construction site and need to get a 14-N bag of nails to your co-worker standing on the top of the building (9 meters from the ground). You don't want to climb all the way back up and then back down again, so you try to throw the bag of nails up. Unfortunately, you're not strong enough to throw the bag of nails all the way up so you try another method. You tie the bag of nails to the end of a 65-cm string and whirl the string around in a vertical circle. You try this, and after a little while of moving your hand back and forth to get the bag going in a circle you notice that you no longer have to move your hand to keep the bag moving in a circle. You think that if you release the bag of nails when the string is horizontal to the ground that the bag will go up to your co-worker. As you whirl the bag of nails around, however, you begin to worry that the string might break, so you stop and attempt to decide before continuing. According to the string manufacturer, the string is designed to hold up to 500 N. You know from experience that the string is most likely to break when the bag of nails is at its lowest point.

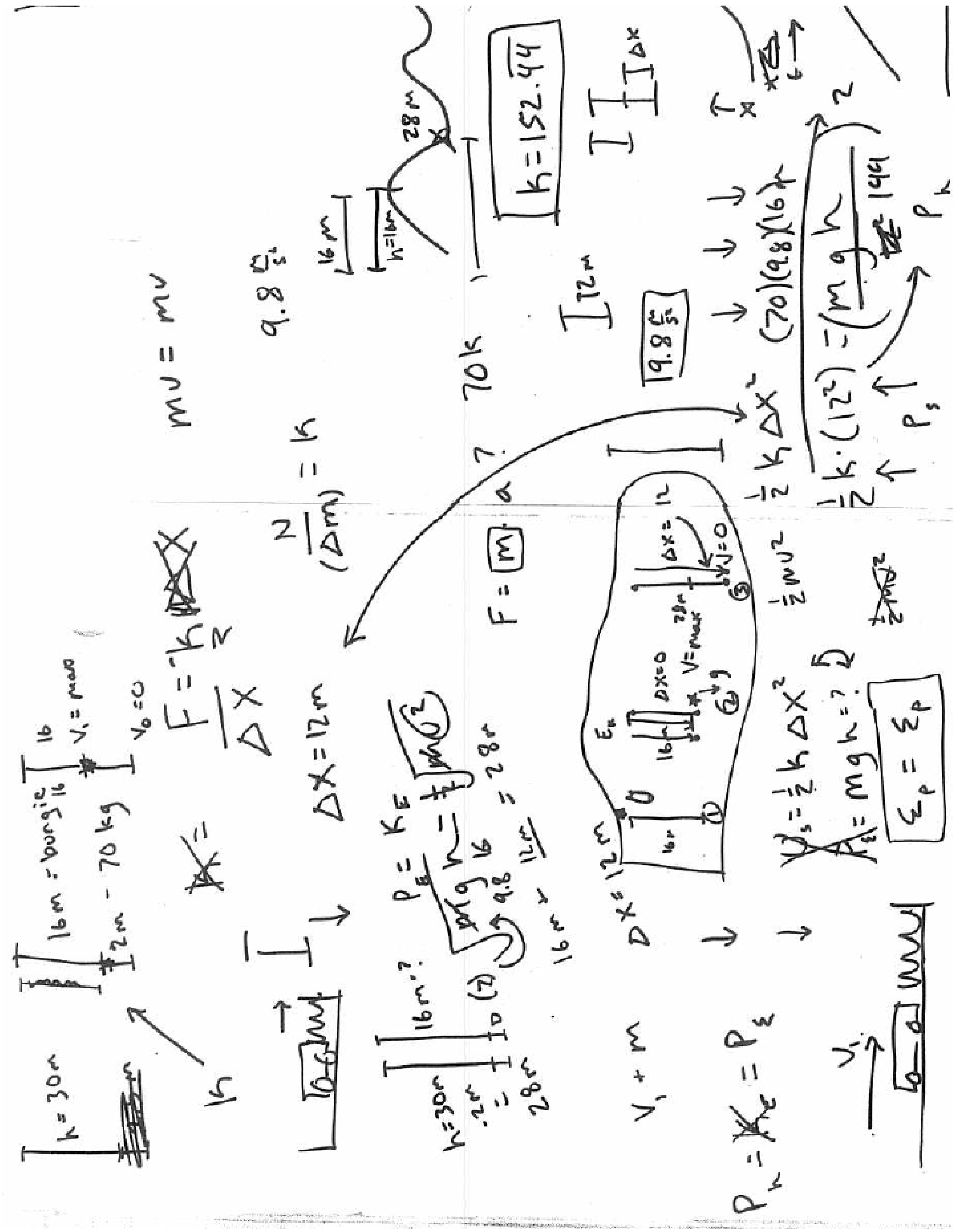
Problem 2:
To raise money for a University scholarship fund, you want to have the IT dean bungee jump from a crane if contributions can be found for 10 scholarships. To add some interest, the jump will be made from 30 m above a 2.5 m deep pool of Jello. A 16-m long bungee cord would be attached to the dean's ankle. First you must convince the dean that your plan is safe for a person of his mass, 70 kg. The dean knows that as the bungee cord begins to stretch, it will exert a force which has the same properties as the force exerted by a spring. Your plan has the dean stepping off a platform and being in free fall for the 16 m before the cord begins to stretch. You must determine the elastic constant of the bungee cord so that it stretches just enough to keep the dean's head out of the Jello. The dean is approximately 2 meters tall.

Problem 3:
You have a summer job with an insurance company and have been asked to help with the investigation of a tragic "accident." When you visit the scene, you see a road running straight down a hill which has a slope of 10 degrees to the horizontal. At the bottom of the hill, the road goes horizontally for a very short distance becoming a parking lot overlooking a cliff. The cliff has a vertical drop of 400 feet to the horizontal ground below where a car is wrecked 30 feet from the base of the cliff. Was it possible that the driver fell asleep at the wheel and simply drove over the cliff? After looking pensive, your boss tells you to calculate the speed of the car as it left the top of the cliff. She reminds you to be careful to write down all of your assumptions so she can evaluate the applicability of the calculation to this situation. Obviously, she suspects foul play.

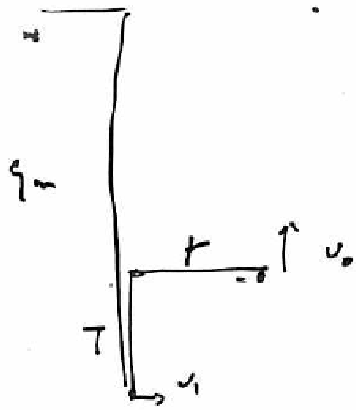
Student 1 Problem 1



Student 1 Problem 2



Student 2 Problem 1



$v_0^2 - 0^2 = 2gS$ $S = 9 - 0.65 = 8.35m$

$\therefore v_0^2 = 2gS$

$\frac{1}{2}mv_0^2 + mgr = \frac{1}{2}mv_1^2$

$v_0^2 + 2gr = v_1^2$

$m\frac{v_0^2}{R} = T - mg.$

$m(\frac{v_1^2}{R} + g) \le T = 500N$

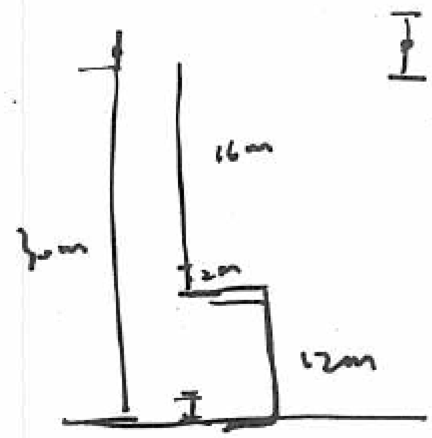
$m(\frac{v_0^2+2gr}{R} + g) \le T$

$m(\frac{2g(S+R)}{R} + g) \le T$

$mg(\frac{2(S+R)}{R} + 1) \le T$

401.7N

Student 2 Problem 2

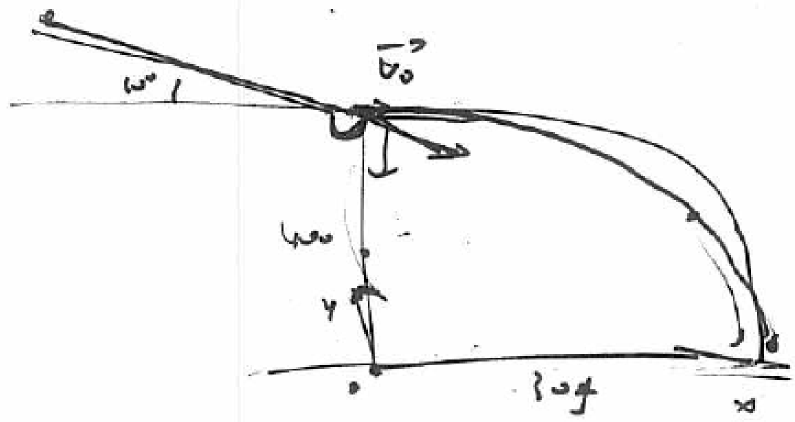


$x = 12\,m$ $h = 30\,m$

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

$$K = \frac{2mgh}{x^2} = 285.83\,N/m$$

Student 2 Problem 3



$h = 400 ft$

$s = 30 ft$

$30/3.281 =$

$\vec{v_x} = \vec{v_0}$

$\vec{v_y} = 0$

$$\begin{cases} \frac{1}{2} g t^2 = h \\ v_0 t = s \end{cases} \quad t = \frac{s}{v_0}$$

$\vec{v_x} = \vec{v_0}$

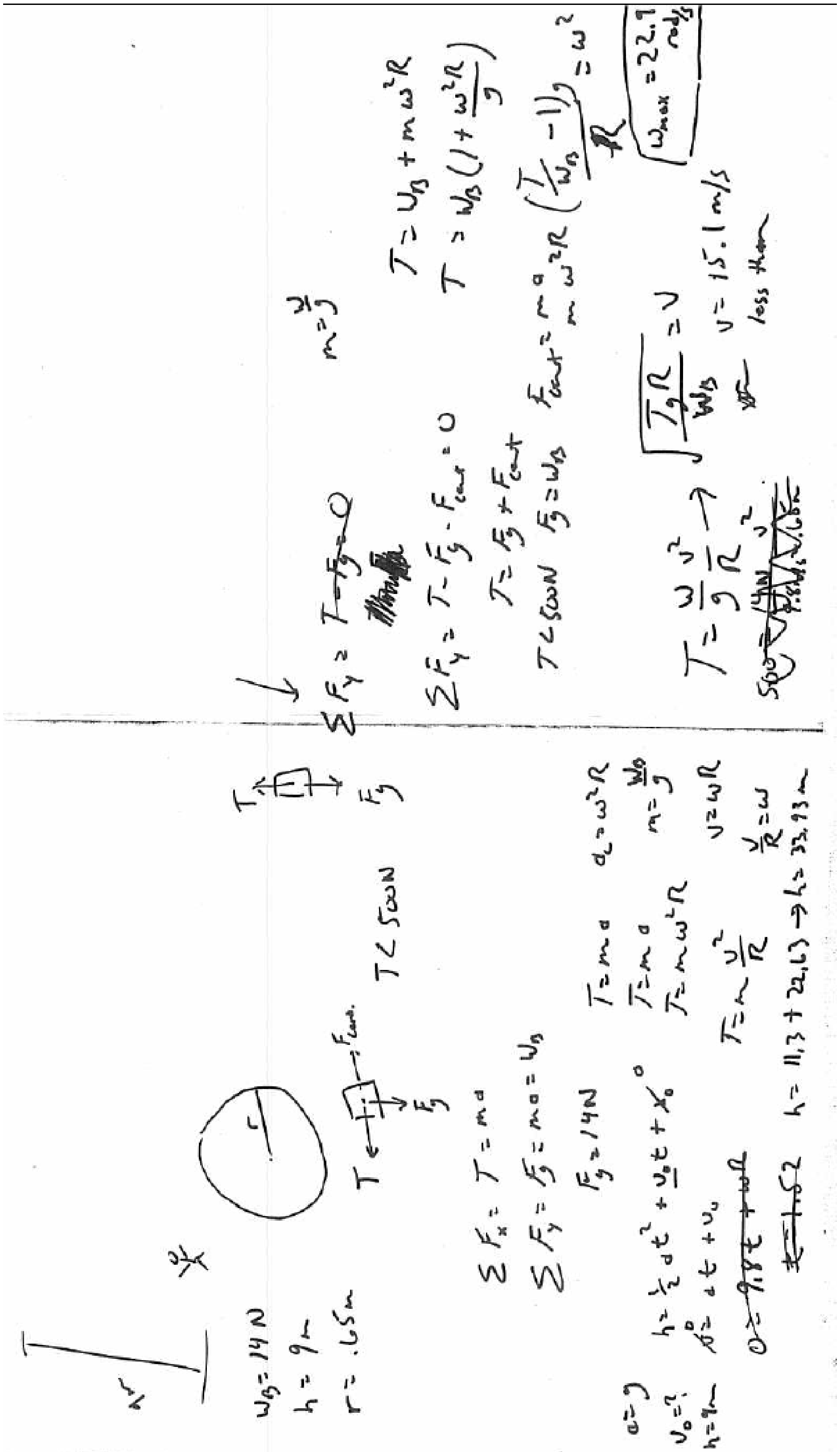
$\vec{v_y} = \frac{1}{2} g t_2^2$

$$\frac{gs^2}{2v_0^2} = h$$

$$\sqrt{\frac{gs^2}{2h}} = v_0 = 1.8 m/s$$

$$= 6.019 ft/s$$

Student 3 Problem 1



$W_B = 14\,N$
$h = 9\,m$
$r = .65\,m$

$T < 500N$

$\Sigma F_x = T = ma$
$\Sigma F_y = F_g = ma = W_B$
$F_g = 14N$

$T = ma$ $a_c = \omega^2 R$
$T = ma$ $m = \frac{W_B}{g}$
$T = m\omega^2 R$ $v = \omega R$
$T = m\frac{v^2}{R}$ $\frac{v}{R} = \omega$

$a = g$
$v_0 = ?$
$h = 9m$

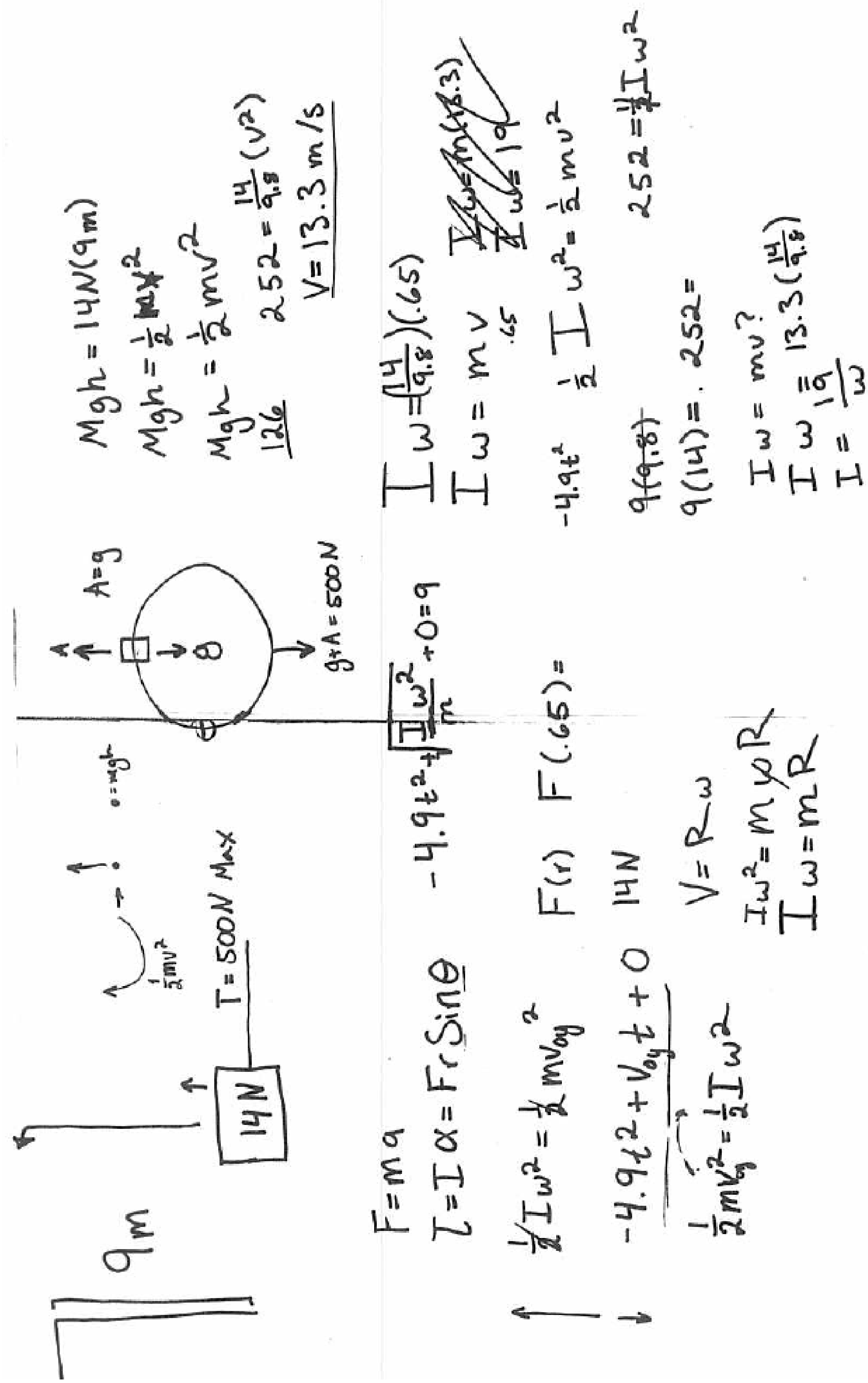
$h = \frac{1}{2}at^2 + v_0 t + x_0$
$v = at + v_0$
~~$0 = 9.8t + \omega R$~~
~~$t = 1.52$~~
$h = 11.3 + 22.63 \rightarrow h = 33.93\,m$

~~$\Sigma F_y = T - F_g = 0$~~
$\Sigma F_y = T - F_g - F_{cent} = 0$
$T = F_g + F_{cent}$
$T < 500N$ $F_g = W_B$ $F_{cent} = ma = m\omega^2 R$
$m = \frac{W}{g}$

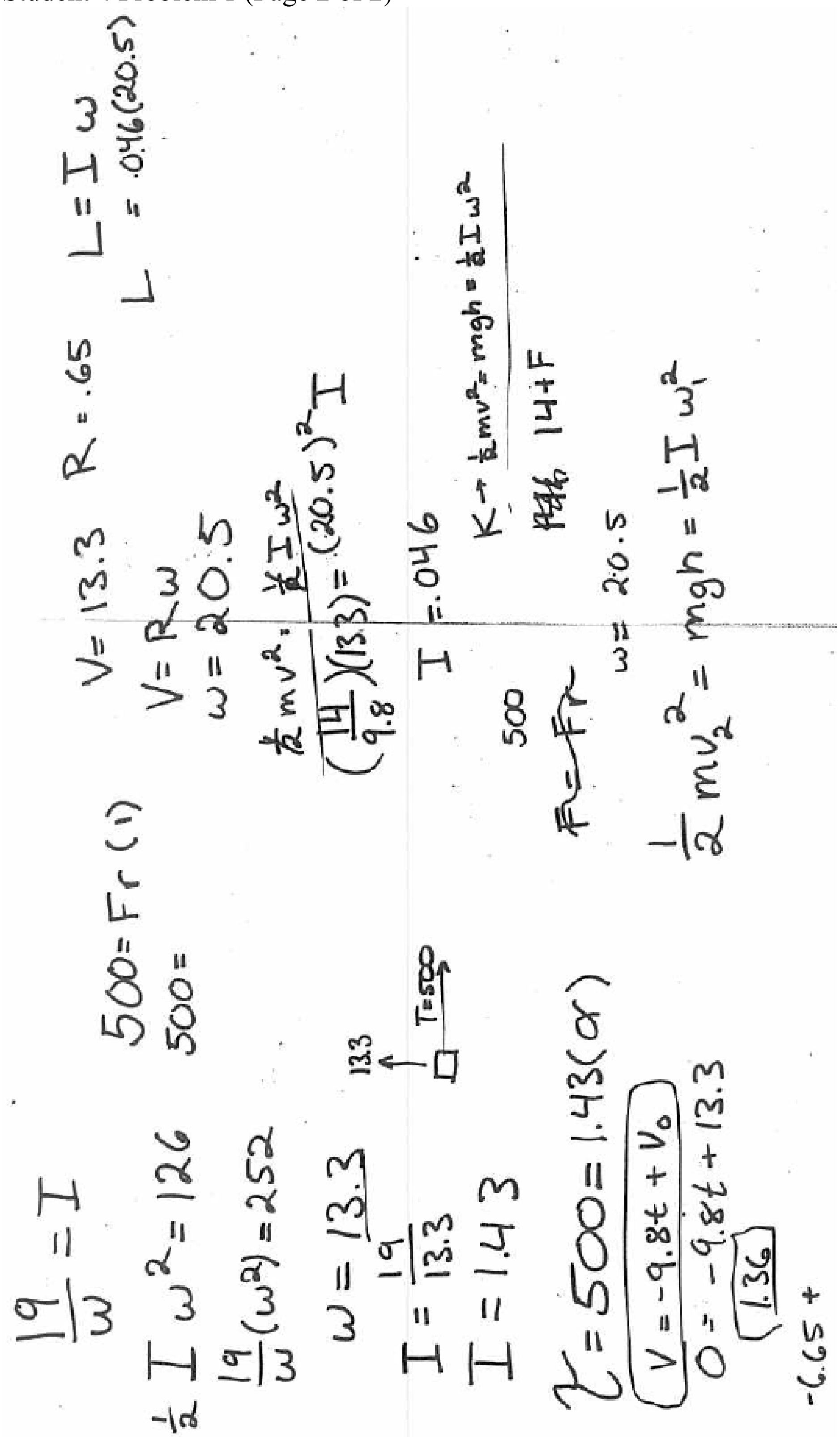
$T = \frac{W}{g}\frac{v^2}{R} \rightarrow \sqrt{\frac{TgR}{W_B}} = v$
$v = 15.1\,m/s$
less than

$T = W_B + m\omega^2 R$
$T = W_B\left(1 + \frac{\omega^2 R}{g}\right)$
$\frac{\left(\frac{T}{W_B} - 1\right)g}{R} = \omega^2$
$\omega_{max} = 22.1\,rad/s$

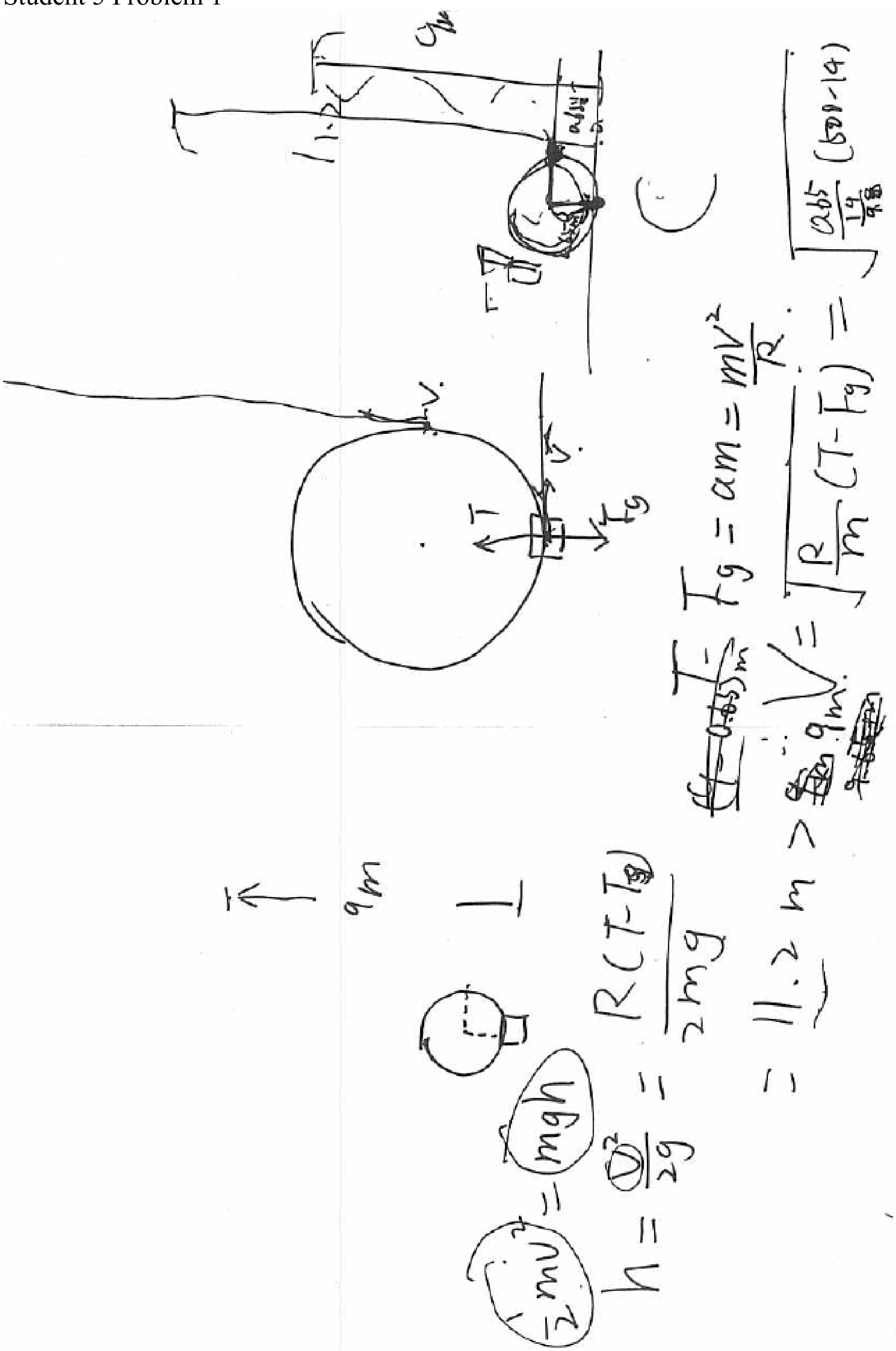
Student 4 Problem 1 (Page 1 of 2)



Student 4 Problem 1 (Page 2 of 2)



Student 5 Problem 1



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

$k = \frac{2mgh}{x^2}$

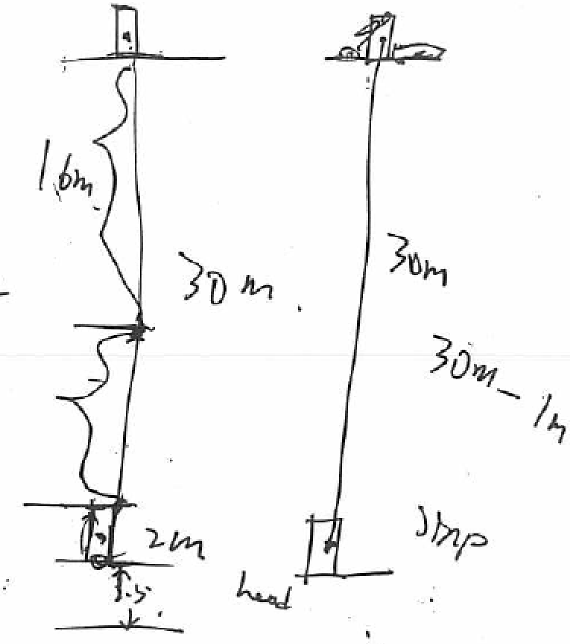
$\Delta h.$

$= \frac{2 \times 70 \times 9.8 \times 30}{12^2}$

$= 285.8 \ N/m.$

$X = 30 - 16 - 2$

$= 12.$



30 m - 16 m. - 2m

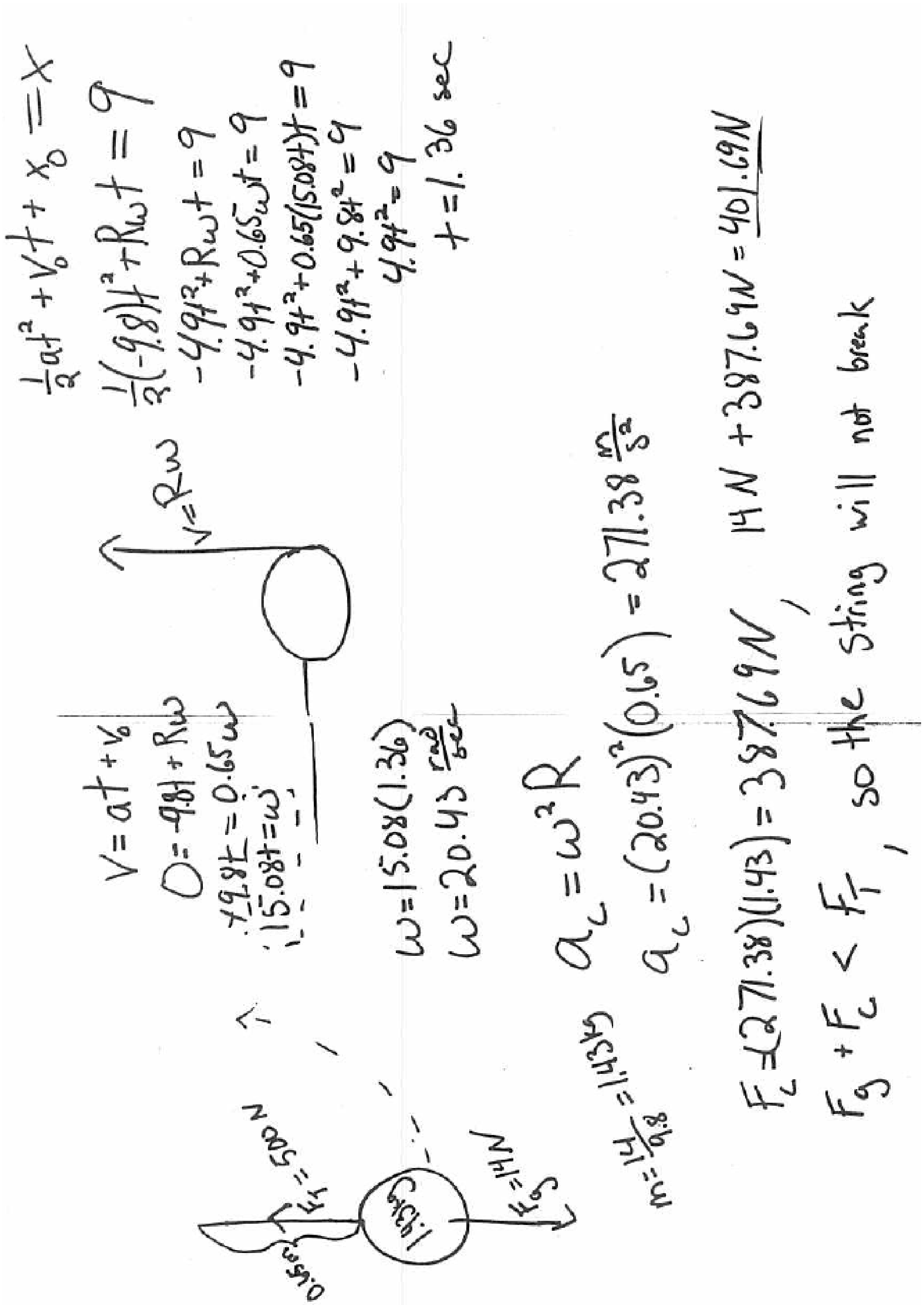
12 m

N/m

~~k~~ $F = -kx.$

$k = \frac{F}{x} = \frac{N}{m}$

Student 6 Problem 1



$\frac{1}{2}at^2 + V_0t + x_0 = X$

$\frac{1}{2}(-9.8)t^2 + R\omega t = 9$

$-4.9t^2 + R\omega t = 9$

$-4.9t^2 + 0.65\omega t = 9$

$-4.9t^2 + 0.65(15.08t)t = 9$

$-4.9t^2 + 9.8t^2 = 9$

$4.9t^2 = 9$

$t = 1.36$ sec

$V = at + V_0$

$0 = -9.8t + R\omega$

$9.8t = 0.65\omega$

$15.08t = \omega$

$\omega = 15.08(1.36)$

$\omega = 20.43 \frac{rad}{sec}$

$m = \frac{14}{9.8} = 1.43kg$

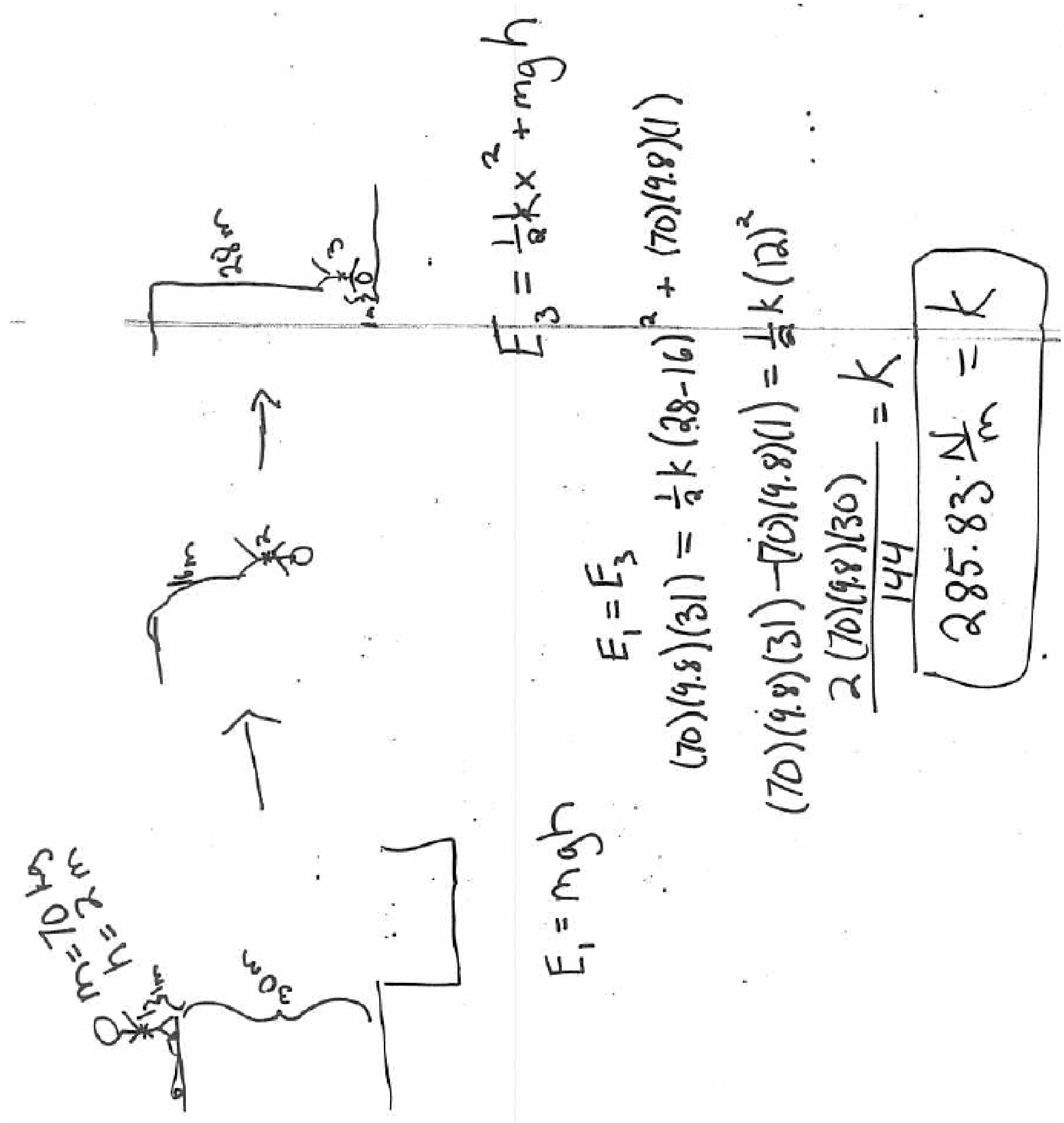
$a_c = \omega^2 R$

$a_c = (20.43)^2(0.65) = 271.38 \frac{m}{s^2}$

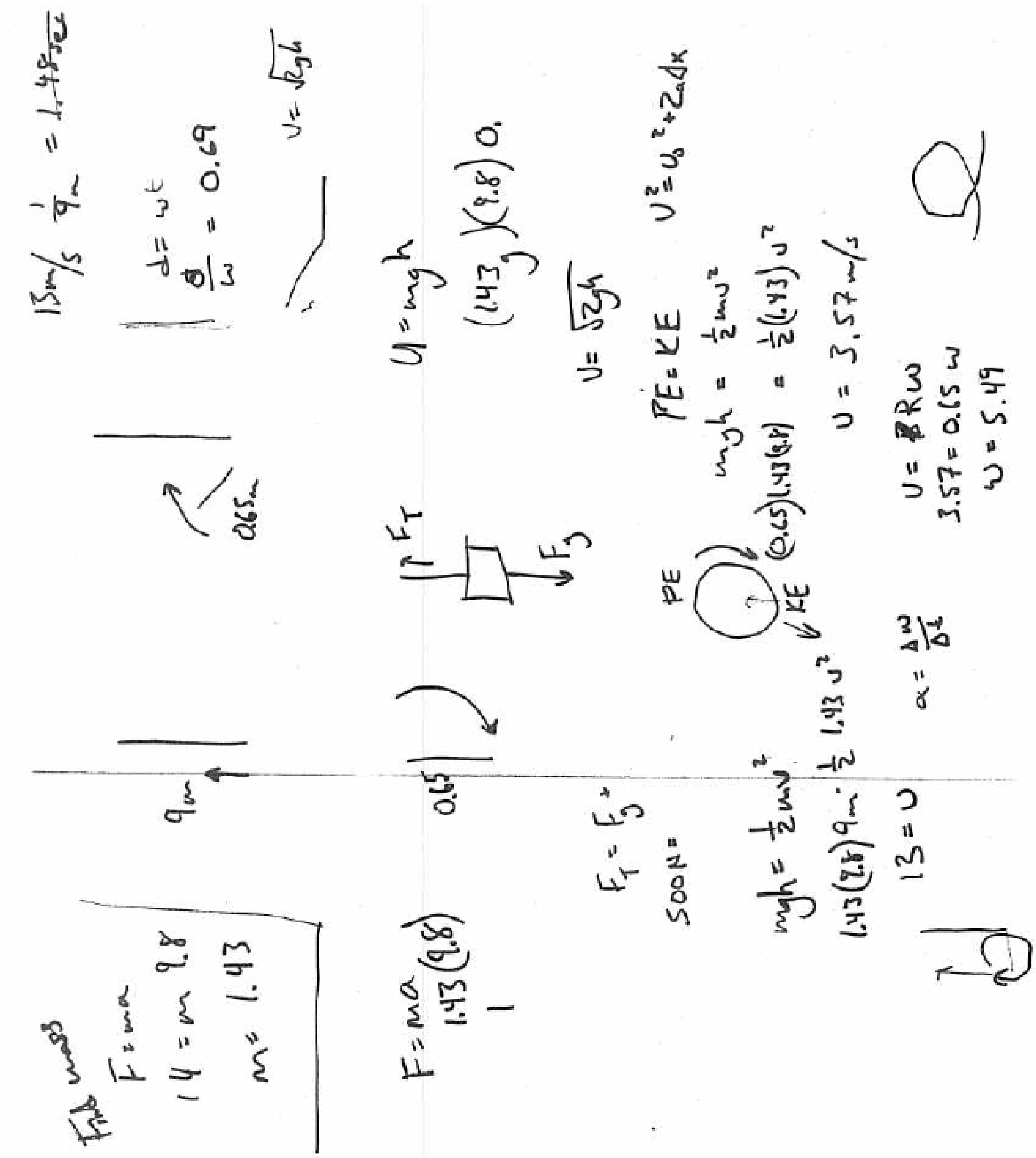
$F_c = (271.38)(1.43) = 387.69N$, $14N + 387.69N = \underline{401.69N}$

$F_g + F_c < F_T$, so the string will not break

Student 6 Problem 2



Student 7 Problem 1



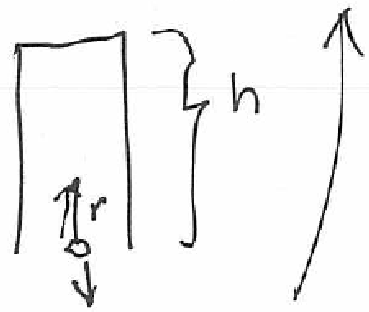
Student 8 Problem 1

$h = 9 \text{ m}$

$W_n = 14 \text{ N}$

$r = 65 \text{ cm} = 0.65 \text{ m}$

$T_m = 500 \text{ N}$



$m = \frac{W_n}{g}$

$\frac{1}{2} m v^2 = mgh$

$v = \sqrt{2gh}$

$T - mg = \frac{mv^2}{r}$

$T = \frac{2ghm}{r} + W_N$

$T = \frac{2hW_m}{r} + W_N$

$= W_N \left(2\frac{h}{r} + 1\right)$

$= 402 \text{ N}$

Doesn't Snap.

Appendix 9: Interview Transcripts

Problem-solving interview #1

Wednesday May 6, 2009 2:00-3:00 p.m.
Appleby Hall conference room 351

Note: Student #1 talked out loud while working on the problems whereas Students #2-8 typically worked in silence.

Summary of audio file: 01:03:15 receives the first problem (nails), 01:18:10 done, 01:24:16 receives second problem (bungee), 01:35:52 finishes second problem, 01:55:15 end of interview

01:01:52
I: 01:02:04 OK, so I'm going to give you a problem
S: Mmm kay
I: …to work on, and it should look kind-of similar to ones that you've done in your, um 1301 physics class, and I want you to just um solve it like you would an exam problem
S: Mmm kay
I: 1:02:20 and um, then when you think you're, you know, done uh, you can let me know and then I'll ask you, go back and ask you some questions about it
S: OK
I: 1:02:30 If you wanna talk aloud during it, if you're comfortable doing that you can,
S: OK
I: otherwise you can just kinda work it out like you usually would
S: 01:02:34 well what would work best for you? Like, do you want, do you want, 'cuz I don't, I don't mind. I mean, do you want me to sort-of explain it orally as I'm going through the steps? Or,
I: If that's comfortable for you
S: Yeah
I: you can do that
S: I don't mind at all, whatever
I: [laughs] OK then, and then I can ask you some questions
S: yeah, sure
I: when we get done, so don't, you know um, if you make a mistake just kinda cross it out,
S: Yeah
I: don't crumple anything up [laughs]
S: 01:02:55 New page, yeah, I gotcha
I: use as much paper as you need
S: OK
I: Um, so
S: Eh, do you want it a certain size, like so it's readable, or
I: I think with this [video camera, marker], with this it should be
S: should be OK?

I: it should be OK, so.
S: Ok, alright.
I: So. [mumbles, ok…press start here] [beep - turns on video camera] OK, well, here's, here's the problem
S: Mmm kay. 01:03:15
pause, reading problem 15 seconds [1:03:15-1:03:30], starts writing with marker & silence [1:03:31-1:05:02]
S: [quietly] 01:05:02 Mmm kay [silence, tapping marker on paper 1:05:03-1:05:21] Alright [sniff]. Okay. So this problem is asking if we spin a certain mass around on a string that's sixty-five, uh, rather point six five meters, sixty-five centimeters, how fast do we need to spin it so when the string is horizontal to the ground it will go up nine meters [writing].
S: 01:05:45 The problem's a little ambiguous 'cuz it doesn't say where you're spinning it, it just says you're holding it in your hand, so, for the sake of the problem I'm just gonna assume that this point is, boy I don't know, we can just treat this as the ground, so, somehow it's going through the ground but [sniff] we'll say that, cuz it doesn't say what the height of the actual person is. Well, or we could just estimate it. A person's what, two meters tall? Okay. So we'll say this is two meters, right here. So when it's horizontal, we need to go up another seven meters in this direction [sniff].
S: 01:06:26 So to figure this out, we first need to convert fourteen Newtons, which is a weight, into mass. Fourteen Newtons, and a Newton is…um…nine point eight times a mass. [sniff] I think. So, yeah. We want to figure out the mass, that should just be…fourteen divided by nine point eight. Which is [using calculator] one point four three? [quietly…eight five seven] So that's the mass of the object that we're spinning, right there [sniff].
S: 01:07:25 So, what we need to do is figure out how high we need to get it and I think we're gonna need to use conservation of energy there, so…[sniff] at this point right here it's gonna be entirely kinetic, then at this top point, at seven meters from when it's released 'cuz now…uh, this is the actual height that we're working with, starts here and ends here, seven meters, at this point right here it's gonna be entirely potential energy, which is mass times gravity times height. So that's our second energy and now our first energy is gonna be one half mass, velocity squared. S: 01:08:12 However that's a…energy for something moving at a linear…motion, I think. Angular. Here we go, rotational physics. So in this case, we're not gonna use that. K for that is one half "I" double-u squared? No, that's not right. Because we don't have an… I. We don't have a moment of inertia here cuz it's not a solid spinning I think it's just [sniff] something spinning on a string and the mass of the string is negligible. So, maybe we…can work with this. Okay. 01:08:55 [mumbles, that would be squared?] Um, [sniff] how do we convert that? [Pause 5 sec] to angular velocity 01:09:13
S: 01:09:24 Angular acceleration is the second derivative of position. [pause 10 sec, sniff]. Oh no, I think we do use moment of inertia. Cuz moment of inertia is I equals the sum of mass times the sum of radius squared, and in this case it's just one object so, here's our mass times our radius squared, and we know both of these things. Mass is one point four two eight, blah blah blah-blah, times radius which in this case is point six-five meters. So that should, that's our moment of inertia [sniff] and then from that I

think we can go straight to kinetic energy, we don't have to worry about velocity, or anything like that.
S: 01:10:14 We could deal with kinematics, we could figure out what the initial velocity would have to be and then…use kinematics to figure out when it would peak right here, at the vertex of that parabola, but, I think it's gonna be easier to use conservation of energy here. So we have E-one, E-two. [sniff] Mmm kay.
01:10:36 So now that we have our I value, we'll go back to this one-half I omega squared [sniff] and the uh…okay so the definition of omega..oh no, here we go. Maybe we can use kinematics. Vee equals a radius times our angular velocity. So that might actually be easier. 01:10:59 [pause mumbles: I think, for velocity then (?)] That's unknown. [pause]
01:11:10 uh, well we know our final velocity, cuz we can work backwards then. Um [pause]. Huh. [sniff]. Do we need to? Well, that'd be using kinematics. We can still go with conservation of energy here. Cuz now that we know I all we need to find is omega and we know all three pieces of our potential energy at the top.
01:11:36 So we can calculate [sniff] the sum of energy in the system, assuming it's conserved. Right now just knowing this mass and how high it has to be…which would be, one point four two eight five seven kilograms times nine point eight meters per second squared and then our height which is, we're gonna say seven, just assuming it's a two-meter tall person [sniff], which gives us…nine seven point nine repeating.
S: 01:12:10 All right. So now we just need to plug in I there and then set the equation equal to omega. So, [sniff] multiply both sides by two. Then you get hundred ninety six, divided by I equals omega squared, under a root on each side [sniff] gotta solve for omega. So one ninety six divided by our moment of inertia, which we figured was one point four two eight five seven times point six five. Uh, squared? Is that squared? It is squared. I almost missed that. [sniff] [quietly, five, squared [typing]] Mmm kay. And then take the square root of this, we should be on the right track. 01:13:23 [typing in calculator] Mmm kay.
01:13:34 So our angular velocity, as far as I understand, is reading at eighteen point zero two. That's angular velocity.
01:13:46 Now. We know how fast it has to be spinning, but the question is…whether or not at that speed, can a mass of one point four two…hold…on a string that can only hold five hundred Newtons of force. And that's spinning...pretty fast, although it's not much of a mass, I can't really estimate, um, off the top of my head. But, we should be able to figure it out. [sniff]. But how? Five hundred Newtons, that's an acceleration…times a mass. And we know the mass, so we can, we can break it down and figure out [sniff] the acceleration just like we did for the fourteen Newton weight. So we need five hundred Newtons divided by nine point eight, which is our g. Equals, oh wait…No, I did that wrong. Our *a* is unknown, for this…five hundred Newtons divided by…our mass, one point four two eight seven, equals…our a.
01:15:00 So [typing in calculator…five seven] Okay. Then it comes out to be exactly two hundred fifty. That's reassuring. Okay. Or is it? Oh, dear lord. Three hundred fifty meters per second. That doesn't look right. Mmm kay. Well, [laughs] um, so [sniff] this is the maximum, oh no, five hundred Newtons is a pretty big amount, and that's a pretty small mass, so that might make sense. Does that make sense? Oh boy. We'll see.

01:15:45 Okay, so this is the a-max, that's the maximum acceleration that the string can hold given this mass. And mass is a constant, so now I just need to figure out what the acceleration is given this omega. Which, I'm sure there's an equation for. Dum, dum, dum. [pause]
01:16:10 Okay, here we go. Centri-peetal acceleration equals omega squared times r. Fabulous. [sniff] Mmm kay. Well we could be pretty close. So. Um. We know omega and we know r. Eighteen point zero two [typing in calculator] squared, and then…times r, which in this case is point six five. Mmm kay, well I think I might have done it right then.
01:16:43 Because…the acceleration up of the, the centripeetal acceleration in this direction which is pulling on the string, so if we were to draw a force diagram, this is our force of tension right here, this right here has a max of five hundred Newtons, right here, which…in the case of the mass is gonna be three hundred and fifty meters per second squared. And the actual acceleration, given this mass, came out to be two hundred and eleven point zero six eight meters per second squared. Which is fabulous. Which means…that this…should work just fine.
01:17:28 Given we only need to get it seven meters high. Now, the problem could have ended up totally [laughs] different, cuz this arbitrary two-meter high mark, which is assuming that the person's spinning it at two meters off of the ground. I just put that in there, so I would probably wager a guess that since they didn't give us a value they meant maybe it's just spinning from the zero point, or the point you're spinning it from is nine meters below this point…but uh, they said nine meters from the ground. So…assuming you have another two meters I think it's really likely that uh, you'd have to be spinning it so fast that it wouldn't hold the tension. So. Pretty much it though. 01:18:10
I: Mmm kay, so
S: [mumbles…for that problem?]
I: 01:18:12 you're satisfied with that answer?
S: 01:18:13 Yeah, I'm, I'm 'satisfied' with it [laughs]
I: [laughs]
S: we'll go with that, that's a good word
I: [laughs] 01:18:17 okay. Um, now I'm just gonna ask you a few questions about what you were thinking while you were, I mean, you did pr – you did a really good job of explaining it
S: Oh, okay.
I: while you were going along [overlapping with next line]
S: 01:18:27 I hope I didn't get too talkative, my mind's all over the place with these.
I: 01:18:30 So, I just wanted to know, like, after you read the problem.
S: Mmm hmm.
I: 01:18:34 What was the first thing that you were thinking about?
S: 01:18:36 Um, I, I'm, I was just trying to get an image in my head cuz a lot of times these are..written so weird that you have to re-read it three or four times [interviewer laughs] to even figure out what you're trying to see in the picture, so you can start marking down values. But, until you kinda dig through it, you know, even getting a picture in your head, is just kinda confusing.

I: 01:18: 56 Okay, so, so you're saying first you kinda wrote down some values, and
S: Mmm hmm.
I: and tried to get a picture a picture of the – of it in your head?
S: 01:19:04 Yeah, just figure out what the situation is that I'm trying to emulate, just based on the metrics they've given me.
I: Okay. Um, is this something that you usually do when you're solving problems…
S: 01:19:15 Oh yeah.
I: in physics?
S: 01:19:17 yeah, I'd say so, it's always good to get the raw numbers down and then figure out, you know, how they sort of fit in to the overall picture.
I: 01:19:25 Mmm kay. And then after you did that, what was the next thing you started thinking about?
S: 01:19:29 Um, well once I know what they're trying to…figure out, and once I think that I, I know…ah, what I'm trying to figure out given what they've told me, I, I try to sort of get a road map at least, and, at least with classical kinematics, Newtonian kinematics, there are usually a couple of approaches, conservation of momentum, conservation of energy, or just, raw kinematics. And in this case, I could have done two things…I ended up doing um, kinematics in this case, figuring out velocities or accelerations at each point, but I also could have done something with conservation of energy, which is essentially finding out the total amount of energy in the system and then setting that equal at different points. So. 01:20:15
I: So how did you make that decision?
S: 01:20:19 Um, [laughs] it's kinda messy. I mean, I was sort-of digging through, trying to figure out what would be the, the easiest thing for me to do. Cuz…technically speaking you, you can always use all of these options in physics. Everything can be broken down into forces if you really want to. But, a lot of times there is…one way that's gonna be a lot easier, just given your variables. Cuz, and at the end of the day whatever you use these variables are gonna cancel out, but they might not cancel out until you have a huge equation that takes up the whole page, in some cases, so. I just started writing down, uh, different raw steps before I plugged in any of the numbers said, okay once I find this number I can fill in this, and once I find that, then I can set it equal to this, or can find these two equations and cancel out variables, and. In this case, it wasn't too messy or convoluted, but, I still, you can see, I went through that process where, uh, do I need this, do I not need this, and. You sort-of feel your way through until you know what's gonna work. 01:21:14
I: 01:21:16 Mmm kay. And then…what did you do after, after you've written down a few things?
S: 01:21:22 Well, once I had a rough road map, I, I just started plugging in numbers. And, a lot of times they give you sort-of goofy little things, like normally, they'd give you a mass, in this case you got fourteen Newtons, but you know you're on Earth, so…okay, we know the acceleration and, we can figure out the mass based on fourteen Newtons so you do that. You start breaking down little things into the pieces you need, and then you start plugging those little pieces into the bigger picture. And then start following that road map that you had planned.

S: 01:21:53 And hopefully all goes well, you know, and you end up with the right answer. It's always good to check, you know, once and a while, you know, once I start getting these raw numbers it's like, oh god, does three hundred fifty really make sense in this context. So, I guess though, seeing there's such a small mass you'd have a pretty high acceleration. 01:22:09
I: 01:22:10 Okay, so when you say "check it" you mean, like, the value?
S: 01:22:14 Yeah, yeah, just just the raw numbers, you know. If you're getting acceleration, like three hundred thousand, you know, you've probably done something wrong, you know. Just, just little things like that. You're not, you know, doing super specific checks here and there, otherwise it's gonna take forever but, if you end up with values that look a little off-the-wall, chances are you missed a square root, you know, you missed an exponential, or, something little like that that you just, you know, copied down wrong. 01:22:38
I: 01:22:40 Mmm kay. Um, so if this were on an exam, how do you think it would be graded?
S: 01:22:47 Um, what, what do you mean by that? How would it be graded?
I: 01:22:50 Well, well, um. What do you think the TA would look for on your paper?
S: 01:22:54 Um, I think, boy, if I were the TA I'd want to know roughly what my game plan was, uh, what this person's game plan was, uh in terms of problem solving and then go through each step to see if the actual abstract of the problem had the right idea and then, once I s-, you know, figured out okay, he knew what he was doing, then I'd start looking at the raw numbers, but. Especially in the case of physics, it's so easy to know exactly what you're doing and just miss a square root, to make your, data totally worthless you know, so if, if it were up to me and of course I'm not a physics TA but, if, if I were grading it I'd focus more on the, you know, broader abstract than each little paint stroke, you know. 01:23:35
I: 01:23:38 Mmm kay. Um. Are you up for a second problem?
S: 01:23:44 Uh, well sure. How much time do we have? I don't mind at all.
I: [laughs] That took just about twenty minutes.
S: Oh, yeah. Sure. That's just fine.
I: Um.
S: 01:23:52 And if you want to stop me, like, halfway through [interviewer laughs] to go to the interview thing, that's cool too.
I: 01:23:56 No, this, this is good.
S: Okay.
I: This is really, you're, this is really useful.
S: Oh good, good.
I: So. Get a fresh sheet of paper. Same kind of thing as before.
S: Okay.
I: Um, here's a new problem. 01:24:10
S: Okay. Well, if only we had these on tests, you know [referring to large paper] It'd be a blast. [sniff]. [1:24:16]
S: 01:24:23 You wanna have the new IT dean bungee jump from a crane. [laughs] That's fabulous. Okay. Thirty meters above a two point five meter high pool of [laughs] Jello. [inaudible, I'll buy that] Okay. Sixteen meter long bungee. Okay [writing].

S: 01:24:49 [inaudible? So I have y, which is] thirty meters. Then our pool of Jello is another two point five meters deep [sniff]. And we have a…sixteen meter bungee. Okay. 01:25:12
S: 01:25:16 Oh yeah [sniff]. And the dean is two meters tall, and that's hanging from his ankle. So…we have elasticity in this, but hopefully our dean doesn't stretch. He's another two meters. Those are attached right there, so that's static 01:25:37
S: 01:25:40 And he is also seventy kilograms [sniff]. Mmm kay. Alright. So it's gonna act like an ideal spring so we can use Hooke's law…Um. Okay.
S: 01:25:57 So we're trying to figure out the k value or the spring constant of this bungee cord right here, that's how much it stretches, so according to Hooke's law for ideal springs, um, let me rewrite it first [sniff]. Your total force equals negative k, that's the spring constant, times your change in x, the change in distance in this case. 01:26:21
S: 01:26:24 Oh, okay. And we don't want him to fall into the Jello. So…This value right here doesn't…mean anything for the problem. Cuz we wanna keep him out of the Jello. I'm wondering what the purpose of the Jello pool is. Just a flashy advertisement. Okay.
S: 01:26:49 So…we should be able to figure out the force here, and we should be able to figure out the change in x, or our maximum change in x, so we can rewrite it as this. F over change in x equals k, our spring constant. Now. 01:27:06 [writing].
S: 01:27:10 We have a total of thirty meters that we can fall [sniff] minus two meters for the dean's height. So that's gonna give us twenty eight meters of total bungee stretch, if you like. There's our two meters right here, here's our dean, here's the bungee which was originally sixteen meters times what, what's the stretch in that. So we have a sixteen meter rope that's stretched twenty eight meters, and I'm clearly a little behind on my fourth grade math…okay, twelve [laughs]. Sorry. So that's gonna be…change in x is a stretch of twelve meters.
01:28:04 So, we know that change in x equals twelve meters, and we should be able to figure out our force because we know his mass, and we know the height he's falling from, and all the forces. Is the mass times an acceleration. 01:28:26 [pause, inaudible? Then why would it matter] 01:28:42 Because acceleration doesn't change, velocity changes. [sniff]. [Pause]. Hmmm. 01:28:55 So maybe we need to do this in two parts, before it stretches and after it stretches. Two different time frames. So we have a twenty eight meter stretch. Total. So [?] the first frame, change in x equals zero. [length is] Sixteen meters. [? Here's the] Second frame. Change in x equals twelve. Total length is twenty eight meters. 01:29:30
01:29:35 We need to figure out the force, and we should know the mass. The acceleration is a little confusing….because throughout this whole thing it's just gonna be g which is 9.8, but it would clearly matter how high he jumps cuz you would think it would stretch more. Handy dandy equation chart. 01:30:04 [pause, x seconds?, tapping marker]. Oooh. Mmm kay.
S: 01:30:25 We are gonna use conservation of energy here, because…we know our total potential energy…at this point. Before he falls at all, so let's just, okay. Bungee cord's around there. Now his total height from when it stretches is sixteen meters even though his height from the ground is twenty eight. So….as far as I understand, since we know what spring potential energy equals, U-s, which is one half k, spring constant, and then

change in x squared, [sniff] and then we know…? What potential energy is. So…which is gonna be mass times gravity times height [sniff].
01:31:29 So if we set them equal to each other, assuming energy is conserved, can we solve that. And would the height be sixteen. I think it would. No, it wouldn't though, because [sniff] right here when it's stretching you still have a downward acceleration due to gravity it's just being counteracted by the spring ? Huh. So…I think we have to do this in two parts. 01:32:05 [pause, tapping]
S: 01:32:12 Maybe I'm overthinking this. Maybe I was right with my first idea. Just using kinematics. That's a force divided by the change in x and we know the change in x. What is the force? Um. [sniff]. M g. Okay. We know the mass, we know the acceleration. But there's more than that. Is it…force? Yeah…..Hmm. 01:32:55 [sniff].
S: 01:33:04 Well I'm not confident about either of these, but… at least I know I have all the pieces to solve this and I think this makes sense. The only thing I'm a little bit worried about is what h should equal. If h would be the total length until he stops moving, or the height until the spring starts pulling back on him. Because in this frame, it should be entirely kinetic, right here. So E-k plus no potential energy, but there is potential energy because he hasn't fallen all the way yet, he hasn't stopped. [sniff]. Yet…He's gonna be slowing down. [pause 4 sec] Huh.
01:34:04 Well. Let's give this a shot, and set 'em equal to each other. Just to see what we get, see if it makes any sense…So. One half k change in x squared. Change in x is twelve meters. So we have twelve squared times k times one half equals m g h. We know all three of these variables. H is gonna be sixteen, g is nine point eight, and our mass is seventy kilograms. So once we figure out that, we can divide this side…multiply by two, and then divide it by twelve squared, which is one forty-four. So…my final guess is gonna be this 01:35:06 [typing into calculator..sixteen…times two…divided by one forty-four] 01:35:18 That could be right. It's a high spring constant. We have a k-value of one fifty two point four repeating. But…we have a relatively large mass, seventy kilograms. That's pretty big. And it fell from sixteen meters. So we've got a lot of energy here. And if we wanna slow that down from the distance of another twelve meters, you're gonna need a pretty thick, pretty tense spring. So. That's my final answer. 01:35:52
I: 01:35:58 Okay. So, now we're gonna kind of do the same thing as before [laughs]
S: Mmm kay, okay.
I: 01:36:06 And it's okay if you repeat what you've said before, too.
S: Okay.
I: 01:36:08 Um, so when you, when you read through the problem
S: Mmm hmm
I: What was the first thing you were thinking about?
S: 01:36:14 Um, like last time, I really wanted to…draw the situation first, and this time it was a lot more straight-forward, I guess. It was, the way it was described, it gave you sort-of the pieces in the order that you would think about 'em being there, versus the last one, it was like, here's this and this and this and this and you kinda have to put 'em together. But here it's like, he's falling from this height, he's this heavy, the bungee cord is this long, he's gonna be falling, you know, this distance until…he reaches this length, so it's nice in the sense that we're only working in one dimension here, so…as

you can see the only drawings, if you can even call them that, are just line drawings. Just have to figure out the distance between certain points and different heights. 01:36:55

I: 01:36:58 Okay, so, so again you were drawing, drawing the situation and labeling it.

S: Yep,

I: with some things

S: 01:37:06 Yep. Yeah, just filling in, um spaces. Just sort of abstracting numbers so visually you can sort of see okay, he's falling this far, and then this event happens, and then he's this distance so we have to, you know, just visually subtract these pieces out, and then you can sort of compare situations er, like, I did here, which is really helpful, um, you can sort-of visualize different frames, you know, in time, you can say, okay, in the first frame it's like this, in the second frame it's stretched to this point, and all the energy is in this state, and then all the energy is in this state.

I: 01:37:39 Okay, so can you say again what was the first picture there?

S: 01:37:42 Okay well, this, I tried to draw, and I-I don't even know if these are correct in this context but, I was thinking okay, before he even falls, he's sixteen meters above when the bungee cord would start acting on him. Because it's sixteen meters tall, so. It's just sort-of looped up, and this is his potential energy. He's not moving anywhere, it's just sitting at this height. 01:38:05 And in the second frame right here, um, he's fallen completely and the bungee cord hasn't started acting, but it's completely tense. So there's no distance left for it to stretch but it hasn't actually stretched yet, so there's no change in x right here. And he's sixteen meters below where he began. And now right here…the third frame, he's completely stopped so he should have a velocity of zero and it's stretched completely. Which means our total change in x right here is twelve plus the original sixteen meters. 01:38:34

I: 01:38:35 Okay, so you've kind-of labeled the different, the different times

S: 01:38:39 Yeah, yeah. Just to sort-of visualize how it changes through time.

I: okay, and then once you had done that, what did you consider next?

S: 01:38:49 Ah, yeah well at this point I-I couldn't figure out whether or not I should…um…go with conservation of energy or kinematics again, cuz I started off knowing since they say it's an ideal spring you can use Hooke's law to abstract that which is k times change in x equals your force. You can rearrange that force divided by change in x equals your k-value, spring constant. 01:39:12 And, I pulled that apart and I'm like, okay force equals mass times acceleration, of course we know the mass, but I'm really confused about the acceleration part because the only acceleration we have is nine point eight meters per second squared, due to gravity. And that's not gonna change. What changes is your velocity. So you're gonna have more energy in the system but the actual acceleration shouldn't change, so…I was kind-of at a dead end. 01:39:36 And if someone explained this to me it'd probably, you know, be like oh, of course, but. I can't quite figure it out, so. I went in to conservation of energy then, knowing that the potential energy of a spring also has the spring constant as a variable and then we can figure out potential energy just based on how high this mass is, and…how great of a mass it is and then the acceleration due to gravity so. Assuming energy is conserved my best guess was that you would have full kinetic energy right here in your second frame, before the spring starts pulling. When he's at his maximum

velocity, right there. So I set those two equal to each other at that second frame. 01:40:18
I: 01:40:21 So, what did you set equal to each other?
S: 01:40:23 Um, sorry, yeah. I have um, well I can sort-of abstract it like this, your potential energy equals your potential energy here. And this potential energy is how high he is, so, you know, if you lift something up and drop it, that's, you know, how much, uh, energy is in the system. And if you pull a spring back that's also a way to store energy. So assuming no energy is lost we can set these two equal to each other right here. 01:40:48 Now what I wasn't sure about is if I had to deal with this velocity factor right here, which would be one half m v-squared. And part of me feels like I do. Or I did, [laughs] in a sense. But I couldn't quite figure out how to use that, so. 01:41:04 I-I looked at it in the second frame. And the only reason the frame is important is because that's figuring out your total height, your initial potential energy. So I, I wagered that sixteen meters was your total height, but it could have been twenty eight meters because he doesn't stop moving until he's twenty eight meters below where he started. Sixteen meters below where he started he's at his maximum velocity. So. I'm just not sure based on the definition of potential energy, which would have been the right one. But I think if I had picked the uh, right height, this would be an effective way to solve the problem, with conservation of energy. I was just, a little ambiguous as to which height I should use. 01:41:47
I: 01:41:49 Okay, and so then…um, what did you do after you set the energies equal to each other?
S: 01:41:57 yeah, well, once I kinda gave up on that guessing match with myself…uh, and I figured I should probably finish it at some point, I could have been here for thirty more minutes trying to figure out which one of these I should have picked. But, I kinda arbitrarily picked the second frame. I said, okay, sixteen meters, that'll be our initial height. So we'll, we'll base our potential energy on that, so. I just plugged in the numbers right here, right here's the potential energy for height, and here's your potential energy for the spring, so, you know, there's my mass, there's the acceleration due to gravity, and there's my arbitrary height. Now right here there's my spring constant, and there's the change in x squared which is twelve so you would just divide that side by one forty-four and then multiply each side by two, which gives us k. 01:42:40 [pause]

I: 01:42:44 Okay, so when you said you just arbitrarily went with
S: Yeah
I: With that value, now if you were at an exam,
S: Mmm hmm
I: Yeah, is that something that you would do, or
S: Yeah, well okay I guess
I: or how would you decide
S: [I guess] 01:42:55 arbitrary isn't the right, it wasn't completely arbitrary, it's, uh, obviously not. I , you see I gave five minutes of thought trying to figure out what it was, it clearly wasn't arbitrary. But, I-I, I did feel a little hopeless at the end, just cuz I couldn't quite figure it out. Although…I must have picked it for a reason, I mean, it's

not like I just randomly pick things, I shouldn't have said that. I definitely picked it for a reason, now what reason I picked it for, I mean, whether it's just this gut feeling or it makes the most sense to me, I-I really couldn't tell ya why I picked this instead of twenty eight, although I think if I would have picked twenty eight, clearly he would have had a heck of a lot more energy and we would have had a much higher spring constant. 01:43:38 So…I-I, again, I can't really estimate what the spring constant should be, you know, there's some quantities where you can just guess, like, you can visualize what five kilograms is and you can visualize what sixty miles an hour is, but you can't really visualize what a hundred fifty-two point four means for a spring constant. You can't say, oh, that's a really tense spring, you know? But, I,I think if it was twenty eight that would have been a huge number. That would have been really large, so. I, I guess that's probably why I picked it, just cuz it was the lesser of two evils, maybe. I don't know. You, you don't happen to have the answers do you? 01:44:16
I: Yep, yep. I'll
S: Oh, you do?
I: give them to you at the end [laughs]
S: 01:44:18 Oh! [laughs] Oh [inaudible]
I: 01:44:21 For now, right now I'm just, I'm just trying to get an idea for how you make these decisions
S: Yeah, sure, sure.
I: when you're working on a problem
S: Sure
I: 01:44:28 Um, so. Um, so, you kind-of had a, you made a decision
S: Mmm hmm
I: Based on which one you thought might give you a more reasonable answer?
S: 01:44:36 Yeah, that and I…boy, I'm just, I'm really not sure…what I would do, because if I did set it equal to twenty eight meters, my thinking is that, it would imply that, assuming full conservation of energy, it would be free falling for twenty eight meters, which would mean it's having a constant acceleration of nine point eight meters per second applied to it, which would mean that at the bottom then it would be one half m v-squared based on that, but it's not a free fall, it's a free fall until it hits this sixteen meter point, and at that point the spring starts acting on it, because the spring is completely tense and once it starts stretching out then it's no longer free fall and that acceleration isn't being applied anymore. 01:45:29 Rather, it is being applied, but kinematically speaking…if you were to graph it the derivative would not look like it was being applied. It wouldn't be this anymore, and if we were, so, let's say this is x, rather, this is x, and this is t right here, so if you're taking the first derivative of that, you know, okay we have velocity right here, and since we know it's a constant slope that's um, a constant acceleration if you take the second derivative, so here's our acceleration right here, but when the spring starts acting, what we should see instead is something, we have the velocity here and then the oooop, it starts leveling off. So you have the acceleration right here and then, booop, the acceleration goes down. 01:46:13 Right? Doing this, and then it should slow down and reach a stopping point right there.

So my guess is right at this point, would be, where you measure the potential energy. At that t, so. 01:46:26
I: 01:46:28 Okay, so you're saying the acceleration is going down? It's not,
S: 01:46:32 Y-Well, well the effect of acceleration. Once the spring starts acting it's not gonna be falling at the same rate any more, so even though it's being accelerated by gravity at that rate, uh, in terms of motion it's not going to be falling as fast anymore because the spring's going to start slowing it down and eventually it's gonna stop which means that as this gets lower and lower, your velocity should actually equal zero at some point. 01:46:55 So it should go up up up, then it hits the spring and oooop, goes back down to zero, right…and then up, goes back up again and you know, assuming it bounces it would start doing that, but. This is the point right here, this sixteen meter point where the spring starts acting. 01:47:14 And right here, is the uh, the total length, the twenty eight meter point when it should be completely stopped, right there. So…that, that was my best wager as to what to pick for the height. We would set the height to sixteen rather than twenty eight. 01:47:29 [pause 5 sec] That's a really round-about way of reasoning, but…I-I couldn't think of anything better. 01:47:39
I: 01:47:43 Okay. Um, do you know what the units are for k?
S: Um,
I: Or did you think about that at all?
S: 01:47:48 I think k is just a, I think it's a, I don't think it has a unit actually. I think, cuz it's just a coefficient, just like the, the coefficient of friction doesn't really have a unit, I think…because it's a coefficient it doesn't have a unit…I-I don't think so. So, in terms of yeah, in terms of solving a problem and seeing if I end up with the right units, I don't think that would work with k. I could be wrong. But I think it's a dimensionless quantity. Just a…01:48:16
I: Is there a way you could check?
S: Just a scalar. 01:48:19 Well, if we looked at k, we would have, in terms of units we would have a force, mm kay, and a force is Newton's right? Over a change in x, so, a change in meters. So, Newton over change in meters, you know? Uh, it doesn't sound familiar to me, I've never seen a unit like this before. So if there is one, I don't know. I-I know there are certain like, tensile strengths and like, um, certain constants with materials, like in mechanical engineering they do a lot of that stuff and like, Young's modulus and stuff and those all have units, but those are, those are based on SI units, you know like how many meters can you stretch it with a certain force before it breaks, but I think…spring constants k are dimensionless. Pretty sure. And if they did have a unit it would be Newtons over change in meters, so. 01:49:10
I: Okay.
S: Yeah.
I: 01:49:13 Have you done a problem like this one in your class before?
S: 01:49:15 Yeah, yeah. A while back. I think it was the…second quiz, when we would have done spring constants and stuff. [sniff] And I can't remember if we did conservation of energy back then or not. If we had approached that topic yet. So, I sort of had another, you know, tool in my toolbox in the sense that I can figure out the total amount of potential energy stored in the spring um, as a function of k, its, its spring constant, whereas before I might have just had to use kinematics, saying okay it

stretched so far so it had a certain force applied to it, and, this force equals you know, this velocity at a certain point. 01:49:53 So. I think I could have solved the problem in another way, just figuring out what his velocity would be after a sixteen meter free fall, and then once we have that velocity, figure out, maybe convert that to momentum? Maybe that's what I should have done. Is done that, and then a conservation of momentum once we have this velocity right here, and then knowing velocity null equals zero, uh, okay, now that I'm rethinking that might have made more sense, might have been easier. I'm not gonna redo the problem, but [laughs]. 01:50:22
I: Well, what made you think of momentum in that...
S: Um
I: problem?
S: 01:50:26 Well, sine we know, it's gonna be really easy to figure out this right here, cuz right here, I'm sure we can use potential equals kinetic energy [sniff] So once he falls sixteen meters and assumed complete free-fall, it should be really easy to figure out his velocity right here. [writing] Because with mgh, the masses cancel out right here, and then we know that's sixteen and this is nine point eight…so, we can solve for this really easy. We can just, honestly square that, and then, times it ooops, put that under the radical, and then multiply that by two [sniff] and then right there we have the velocity, and then, once we have the velocity at this time frame right here, once he's fallen sixteen meters, then maybe we could convert that to momentum? 01:51:14 We would have m v equals m v. Although he's not changing mass, so. Why would w-we wouldn't need to convert it to momentum. Maybe we could just stay in kinematics then. And just have that velocity but, once we know this velocity now, would we multiply that by it's …mass to get the energy? See that's what I was confused about then, because once we have that, then wouldn't you say, okay, our kinetic energy, once it's all motion, equals our potential energy of the spring, and in this sense, then this is just the potential energy of the height, so. 01:51:57 I, I thought about this in the middle, actually and I-I figured why would I convert it to that if ultimately it's just equal to that anyways…So I kinda just ended up skipping that step cuz I couldn't figure out, okay, now that we know the velocity…and the mass, how we figure out what the spring constant's gonna be, based on that. 01:52:21 And I think it would have been easier if the problem was on a horizontal plane, cuz then we could have ignored gravity. I think then we could have solved it differently. Say, there's a spring right here and, it's going, you know, it's going to be one right here but we don't have worry about gravity still acting on it with the spring whereas when you're falling down this way, even as the spring is stretching, gravity is still pulling on you, right now. So, in a sense even if he jumped off and it was completely tense, of course the spring's still gonna stretch because he's still putting a load on that spring. 01:52:54 Whereas if you were in this situation and you had a car, you know, set right next to a spring, well the spring's not gonna stretch, it's just gonna sit right there. [sniff] So I think in that aspect, if it would have been, you know, a horizontal plane, I would have been really confident about my answer. But since it's vertical, and while the spring's squishing you still have an acceleration in that direction, that's what made me a little bit unsure about my answer. 01:53:18

I: 01:53:20 Now when you say, kinematics, can you tell me what you mean…what you, what you mean by kinematics?
S: 01:53:26 Um, I-I think, roughly, in a physics sense anyways, the definition of kinematics is just, analyzing things based on motion. So, you're just looking at uh, how fast something's spinning, how fast something is moving. How fast something is accelerating, versus sort-of these ethereal quantities, like how much work is being applied, how much energy is in the system, you know. 01:53:48 I don't think that you would say that potential energy is um, pure kinematics because, you know, just having a picture hanging on the wall there's nothing really in motion, even though there can be physics applied to that, you know, and you can say okay there's that much energy stored in that painting if it were to fall, that's not really kinematics. 01:54:07 Whereas if that painting were falling then we could say okay, it's accelerating at this speed, it's going to reach this point at this speed, it can travel this distance based on this time frame, so I think it's using those, you know variables there, versus these ethereal energy and momentum quantities. 01:54:22
I: 01:54:23 So is the force something that, um, is kinematics, or not?
S: 01:54:27 Um, boy, is force part of kinematics. I'm not sure. Would I say it's in-part of kinematics? I don't think I would.
I: Okay.
S: 01:54:37 I don't think I would define it as pure kinematics. 01:54:40
S: 01:54:42 Or maybe I would. Mass times acceleration. Boy, now I'm second-guessing. Yeah, maybe it is. 01:54:47
I: 01:54:48 Okay, is there anything um, else that you wanna add, anything that you
S: Um..
I: were thinking about while you were solving this that you didn't write down?
S: 01:54:56 I wanna apologize for the mess. [interviewer laughs] I mean, I guess in one sense it really represents how I'm thinking, for better or for worse. You have all these great, like IT-honors kids who have everything sorted on graph paper and I'm just this, ugly wrong mess, everywhere. Oh, poor you. [laughs] No, I-I think I'm good. 01:55:15
I: 01:55:16 Okay. Well, thank-you for volunteering your time.
S: Sure.
I: [beep, turns off camera] and I have [murmuring, okay that stopped]…Um, so I have one more thing for you to sign,
S: Sure
I: verifying payment and stuff.
S: 01:55:34 Ooh. The fun part.
I: The fun part. [laughs]

Problem-solving interview #2

Thursday May 7, 2009 9:00-10:00 a.m.
Appleby Hall conference room 351

Summary of audio file: 00:01:14 receives first problem (nails), 08:04 finishes first prob, 13:35 receives second problem (bungee), 20:00 finishes, 27:15 receives third problem (car/cliff), 30:00 and 32:45 asks for clarification, 37:44 starts explaining, 51:48 done talking

S: 00:00:01 can I ask you to explain this question for me if some words I cannot understand
I: 00:00:06 Yep, yep. You're free to ask questions about it if there's something, something that you don't understand, um…And then, um, I'm, I'm gonna ask you to solve it just however you would solve, like, an exam problem
S: Mmm hmm
I: 00:00:22 Um, and then…uh, afterwards we can go back and I'll ask you some questions kind of, like, what you're, uh, what you were thinking when you were going through it.
S: Mmm hmm
I: I mean, if you wanna say some stuff while you're working on it you can too, but it's whatever you feel comfortable with.
S: [quietly] 00:00:36 Okay
I: 00:00:38 So if you don't feel comfortable talking while you're, while you're working on it that's okay too. Um. So I'm gonna make sure that this is set up and then I'll give you the problem. 00:00:47 [clicks, beep, turn on camera].
I: 00:01:08 Okay, so. Here's the copy of the, of the problem. 00:01:14 [shuts door ?]
Quiet / working on the problem 00:01:14 to 00:08:04
00:03:48 starts writing
00:06:50 hear typing in calculator
S: 00:08:04 I have solved it
I: Okay, so. [*moving to sit at table with student*]
I: 00:08:12 Okay. So now I'm just gonna, um…ask you a few questions.
S: Mmm hmm.
I: 00:08:12 Okay, so. After you read the problem, what was the first thing that you were thinking about.
S: 00:08:30 Uh…that it is not hard
I: [laughs]
S: 00:08:38 And, I should use uh, the equation of the motion and uh, the conservation of energy in this problem
I: 00:08:50 Okay, so you were thinking about conservation of energy, and when you said 'equation of motion' which one, which equation are you talking about?
S: 00:08:59 Uh, I used this one to sol- to find the velocity at this point. 00:09:07
I: 00:09:09 Okay. Um so before you even did that, um, I noticed that you wrote this down.

S: Mmm hmm.
I: 00:09:17 Okay. Um, why did you decide to write that down.
S: 00:09:21 Mm, this picture, or [inaudible, this graph?]?
I: Yeah.
S: 00:09:24 I just want to make the question clear in the graph so that is easy to solve the problem
I: 00:09:31 Okay. So is that something that you usually do..
S: Yeah
I: when you're solving a problem ?
S: Yeah
I: 00:09:38 Mmm kay. Um. So tell me, um, was this the first thing that you wrote down?
S: Yeah.
I: 00:09:48 Okay. And what did you do here?
S: 00:09:50 So, at this point, this is the building, it is nine meters, and this is the radial of the string and I thought at this point the horizontal is, eh, the velocity is v, v-knot.
S: 00:10:09 And according to, as opposed to at the, this point, and the zer- uh, the velocity is zero.
S: 00:10:16 So, according to this equation I find v-not squared should equal to two g times s which s is [inaudible, refers to?] the distance between these two, two part.
S: 00:10:30 And if the string should not be broken at the lowest point, and I find the equation so that this, and this is v-one at this point, the lowest point. And they are, should satisfy this equation if it is not uh, broken.
S: 00:10:51 And use the conservation of energy I find at this point, the energy at this point should be equal to at this point. And I set up these two equation and find v-one squared should be equal to this thing and I plug in, and then I find, v-knot squared gives you this, this part, I plug in, and then I take the g out and finally the equation should be like this.
S: 00:11:22 And I find in the right – ah the left part, uh, the force is four hundred and one point seven Newton and which is smaller than five hundred Newtons so, I think it is possible to do that. 00:11:38
I: It's possible?
S: 00:11:39 Yeah, even [?] at the highest point the velocity is zero.
I: Okay. 00:11:47 Um, so, so when you um, solve problems in class, is this how you usually solve them?
S: 00:11:56 Yeah.
I: Mmm kay. Um, do you usually - I noticed that you um, plugged in numbers at the very end.
S: 00:12:03 Mmm hmm.
I: Is that, is that something that you usually do…too?
S: 00:12:06 Yeah, cuz maybe in the progress I can cancel some [mass? math?] terms [inaudible]. [pause 10 seconds]
I: 00:12:21 Okay. And when you were doing this was there, was there anything that was going on in your head, or that you did in your head that you didn't write down?

S: 00:12:29 No, I think I have show [showed?] all I thought on the paper. [pause 15 seconds]
I: 00:12:52 Um, do you want to try another problem [laughs]? 00:12:54
S: 00:12:56 Yeah, but it is right or not?
I: Oh, I'll show you the solution at the end. 00:13:00
S: Okay
I: So [laughs] 00:13:02 I'll make you wait just a little bit longer, but, um. I'll give you one more problem and then um we can…[pause, putting fresh paper on table] talk about it afterwards again. 00:13:12 [pause, ~20 seconds to find new problem]
I: [gives student problem] Okay [time 00:13:34]
[student reading problem 00:13:35 – 00:15:25 starts writing]
00:19:10 hear typing in calculator
00:20:00 [inaudible noise; I am done?]
I: Okay, are you…
S: [inaudible? Telling finished?]
I: 00:20:08 Okay….So now on this one…after you read the problem, what was the first thing you did, or you thought about.
S: 00:20:23 Uh, this problem should be just something about conservation of energy and how the spring, like the spring and the uh, potential energy of the gravity
I: 00:20:42 So how did you decide that…what, what about it told you, made you think about energy?
S: 00:20:48 So uh, the total energy from the top to the bottom should be equal to the, uh potential energy of the spring at last. If the person at this point to, the velocity should be zero 00:21:04
I: Okay, 00:21:07 um, so that was the first thing that you were thinking when you read the problem?
S: 00:21:12 Yeah.
I: 00:21:14 Okay, and then so what did you do, what did you write down?
S: 00:21:17 So I draw the picture and the height is thirty meters, and I found if the person at, this is the person, and this is the pool and the height is, the person's height is two meters tall and I find that actually that the bungee should be stretched to twe-twelve meters 00:21:39
I: 00:21:41 Okay, so how did you get that twelve meters?
S: 00:21:44 Like I used thirty meters two, ah minus two meters. And this is this length, that is twenty eight meters. And originally I thought that, oh let me see, that the bungee is sixteen meters so I minus that and find that it should be twelve meters. 00:22:06
I: 00:22:08 Okay.
S: 00:22:09 And, actually the person fall down to thirty meters. And this, this energy should be equal to the uh, the potential energy of the spring, and, which is equal to one-half k x-squared. And [inaudible g?] is equal to thirty meters, and k is equal to [??that thing] and plug in the value and find this one. 00:22:34
I: 00:22:35 Okay. So, so you're saying when you plugged in the numbers here you used h is thirty?
S: Yeah.
I: 00:22:44 Okay, and then x you used the twelve?

S: Mmm hmm.
I: Is that right? Okay um….
I: 00:22:51 and then, you have the units Newtons per meter?
S: 00:22:56 Uh, yeah.
I: 00:22:57 Okay. How did you know to use that…
S: 00:23:01 You mean
I: units?
S: 00:23:05 How did…?
I: Yeah, how did you know the units? Is it something you remembered about…k?
S: 00:23:10 Yeah.
I: Okay
S: 00:23:11 Also you can calculate from, like it's just m, m g that is Newtons, and this is meters, and this is meters squared, just like this, cancel [? inaudible]
I: 00:23:24 And get Newtons per meter?
S: Yeah.
I: Okay. [pause 6 seconds] 00:23:30 So um, if you're solving something on an exam, like, that's gonna be graded,
S: Mmm hmm
I: um do you usually write more than this, or is this about the amount that you usually write?
S: 00:23:43 Uh, more than this. I would maybe write some, some words to explain why I write like that.
I: 00:23:51 Okay. Like what kind of words would you write?
S: 00:23:55 Yeah. This, this part is my center of what I'm going to write and I explain that.
I: 00:24:00 Okay, so if you were trying to explain more in words, um, what would you explain to the TA who's grading it? 00:24:11
S: 00:24:14 Just write how I find this twelve meters and [?height?] of thirty meters, and like this part I use, I write conservation of energy.
I: Okay.
S: 00:24:25 Of this whole system. And then [?like this?] I'm talking [plug in?] value, and maybe I write uh, if the person, if I want the person should be safe and the constant should be like this.
I: 00:24:43 Okay. Um, and do you ever, um try and check your answer when you're solving problems?
S: 00:24:51 Yeah, this one I solve my answer, like you see at first I think I get wrong answer cuz I make the height to be eighteen meters, I forget to uh calculate this part, and after check I changed the h to be thirty meters. 00:25:11
I: 00:25:13 Okay so, so was, was there something about this number that made you decide to go back and check? Or did you just think about it? [laughs]
S: 00:25:23 I just go over the problem
I: Okay
S: 00:25:25 and see whether somewhere I was wrong and I found that.
I: [pause, 13 seconds] 00:25:41 Mmm kay. Uh, was there anything else that you were thinking about that you didn't write down? Thinking about in your head?

S: 00:25:53 Uh, yep, yes. You're going to consider the height of this, this person, that I just write here each side should consider this person to be a ball, so I consider this person's center of mass is here in the middle of his height is one meter and draw it like this and this and this to the lines is also thirty meters. Just didn't write it on the, on this paper. 00:26:20
I: 00:26:23 Okay. So you were thinking something about the center of, center of mass being at
S: 00:26:30 Yeah, the mo-, I thought about the motion of the center of mass. And that is thirty meters. 00:26:36
I: 00:26:39 Okay. [pause]
I: 00:26:46 Do you wanna try one more? [laughs]
S: Uh, yep.
I: Okay, you're pretty quick, so we'll do [laughs] we'll try one more
00:26:55 [hear papers shuffling]
I: 00:27:13 Okay, and this will be the last one. 00:27:15
00:29:27 starts writing
S: 00:30:00 So…I'm confused about like, uh this, the cliff is four hundred feet tall. Like this one is four hundred, and I don't understand below where car is wrecked thirty feet from the base of the cliff [inaudible]
I: 00:30:20 So the base means the bottom
S: Here?
I: Mmm hmm.
S: 00:30:24 And this thirty meters, uh thirty feet….So, what does wrecked mean?
I: 00:30:39 Wrecked means it's, that's where it's like crashed, the car is um, like damaged.
S: 00:30:47 Okay. So this picture would look like this, right?
I: 00:30:55 Yep, those distances are right, yeah.
00:30:54 to 00:32:45 silence
S: 00:32:45 I still do not understand the question. Um, they say ten degrees here, and then a short distance that is parking lot, and then overlooking a cliff, is the cliff this side or this side?
I: 00:33:06 Um, this is the cliff, right here,
S: Yeah
I: so it just means that it's a high distance. 00:33:13
S: 00:33:14 So the car goes this down and like this? Goes this way?
I: 00:33:23 Yeah, they're saying that the car is found thirty feet away from this base here.
S: 00:33:29 Yeah, and I'm going to try to find the velocity here?
I: 00:33:35 Um, it says to…'calculate the speed of the car as it left the top of the cliff'.
S: 00:33:46 So if a car fells down [?] to here, and then like, goes this way, and just try to find the velocity at this point, right?
I: 00:33:59 That's the top of the cliff, yeah.
S: 00:34:02 Okay.
00:34:03 to 00:34:29 silence
00:34:30 starts writing

00:35:11 typing in calculator
00:36:58 [inaudible student murmur]
I: 00:37:09 Are you satisfied with that answer, or do you have a question? 00:37:11 [pause 17 seconds]
S: 00:37:28 I don't know whether my understanding of this question is right. 00:37:32
I: 00:37:40 Well, you can tell me what you're thinking [laughs]
S: 00:37:44 [?] Uh, just this is four hundred feet high and this thirty feet, the distance, and the car, the [?] is like this, and these, these two equation. I set this, the velocity horizontally is v-knot, and v-knot times t should be equal to this distance, and one half in this equation should be the height of the cliff, and I cancel t and no, plug in t should be equal to s over v-knot, find in this, in this equation. And found this one, and v-knot should be equal to this. 00:38:30
S: 00:38:32 And I change uh, the thirty feet to meters in order to calculate, cuz the g, then you need this g is uh, m-squared, m s squared over s-squared. So I find v-knot is one point eight meters per second. I don't know whether v is right in this understanding of this question. 00:38:58
I: 00:39:01 So when you were converting, did you use something on the equation sheet, or..? How did you convert the feet to meters? 00:39:11
S: 00:39:13 You mean why I ?
I: 00:39:14 No, no, how, like, what, what numbers did you use?
S: 00:39:20 Uh, you mean, ch- convert feet to meters?
I: 00:39:24 When you calculated this number, I guess, what, what were the numbers that you, that you used.
S: 00:39:28 uh, g is equal to nine point eight, like here.
I: Okay
S: 00:39:32 Uh, I didn't see g is equal to also to this one. I didn't see that, so I changed these two heights to be meters.
I: 00:39:40 Okay.
S: 00:39:41 Cuz the units of the g is this one, not the feet. But if I see something, g like that, I use this one to calculate. 00:39:50
I: 00:39:52 Okay, so what did you get for the, for the new s?
S: 00:39:57 Uh, new s, that is thirty divided by three point two eight one. That is [murmurs, to this one?]
I: Okay, so you got something in your calculator…
S: Yeah
I: …for that?
I: 00:40:12 Okay, and is that the same thing you did for h?
S: 00:40:14 Yeah, change the same unit
I: 00:40:18 Okay, so you divided it by three point two eight
S: 00:40:21 One, yeah
I: okay…
I: 00:40:27 So how um, how about these equations here. Is that something that you remembered, or did you
S: Yeah
I: look on the equation sheet somehow?

S: 00:40:38 Actually something I remembered, but this equation is from this one, cuz the, uh, initial position and velocity on hori- oh, vertical line is zero, so just divide, uh just cancel these two part, and then it's like this one.
I: 00:40:57 Okay, so you remembered the one half g t-squared?
S: Yeah.
I: [clears throat] 00:41:04 Okay, and is it…the same thing for this one? You, you remembered…that one?
S: 00:41:12 Uh, yep, cuz there's no acceleration in the horizontal line, can cancel the third part and first part. 00:41:19
I: [pause] 00:41:28 Okay, and when you had these two equations then, how did you decide to solve one for t?
S: 00:41:36 Cuz, I'm going to try to find the velocity of v-knot. And…and there are both t in these two equations, and I try to make a connection with the t, so like t equal to s over v-knot and plug in to this one. As a bridge, like. 00:41:58
I: 00:42:02 Okay, um…So you could have also solved for this t, so is there a reason why you solved for this t and not this t?
S: 00:42:13 Cuz if I solve this t, I should have a square root and I don't like to use square root during the calculation. So I solve, found the second equation, solve t.
I: 00:42:27 Okay, so you thought that that one might be easier…?
S: 00:42:35 Yeah.
I: [pause 11 seconds] 00:42:46 Kay, is there anything that you're still thinking about in this problem?
S: Um
I: [laughs]
S: 00:42:52 I'm thinking that maybe I should use g equal to thirty-two point two feet and try to find whether they are equal.
I: 00:43:02 You can go ahead and try it if you, if you want to try and check it that way.
S: [typing in calculator, murmurs while writing; v equal] 00:43:25 They are the same thing.
I: 00:43:31 Kay, so then you calculated this was the velocity in feet per second
S: 00:43:36 yeah, yeah, and they are equal
I: 00:43:38 Okay, so how did you know they were equal?
S: 00:43:40 Just divided by this three point two eight
I: 00:43:46 Okay, and then you got this, this number again?
S: 00:43:49 Mmm hmm.
I: 00:43:50 The one point eight? [pause]
I: 00:43:58 When you do problems in class, are they usually in meters, or do you sometimes have feet?
S: 00:44:04 Yeah, sometimes have feet.
I: 00:44:06 Sometimes have feet? Okay.
I: 00:44:12 Now do you think this looked familiar, have you solved a problem like this before? [quietly…in class?]
S: 00:44:17 Yeah, in class.
I: Okay. What kind of problem was it?
S: 00:44:25 You mean, like this one?

I: 00:44:27 Yeah, do you remember one that was like this one?
S: 00:44:33 Mmm, almost the same thing, like, from a cliff or something really high and have a, have no initial velocity in the v-vertical line but have a velocity in the horizontal line, and try to find the distance maybe. Or, maybe the height. Just these three, these three terms and they give you two and like you to find the other one. Questions like that. 00:45:01
I: 00:45:07 So when you read this problem did you think about a problem that you might have solved before?
S: 00:45:16 Um [pause] 00:45:22 No.
I: 00:45:23 No, you just, you just knew that this might be helpful
S: Yeah
I: 00:45:27 these equations might be helpful
S: Mm hmm.
I: in this situation? 00:45:29
[pause, 11 seconds]
I: 00:45:41 Okay, so [laughs] did this make you more confident in your answer…checking it a second way?
S: 00:45:48 Mmm [laughs] If the understanding is right, I think this should be right. 00:45:54
[pause, 18 seconds]
I: 00:46:11 Kay, was there anything more you were thinking about?
S: 00:46:15 Uh, I didn't use the ten degrees to the horizontal, that here…Maybe just some descrip-, description of the question, maybe not useful in problem. 00:46:30
[pause 11 sec]
I: 00:46:41 Kay, so you're not really sure what the ten degrees
S: 00:46:46 Yeah
I: …about that?
S: 00:46:49 I think that is just the description of the problem and not useful. 00:46:54
I: 00:47:05 Kay, is there anything else you were thinking about?
S: 00:47:09 And I think that the sentence like, have a short um [inaudible, where is that?], had a very short distance become a parking lot, which is horizontally, this sentence make this question is easier, cuz if there is not this part, this short distance, velocity initial will be this direction and I have to divide it into direction to solve the equation. But, if there is this distance, the initial velocity, that this, ah, the direction should be the horizontal, so make the question more easier. 00:47:46
I: [pause] 00:47:57 Kay, was there anything else here [laughs] that you, you were thinking about?
S: 00:48:03 Uh…like, maybe like our professor may-, will ask us to maybe set a, uh, like how do you say that..? [quietly] Like this is x, this is y, and this is z, this is zero?
I: 00:48:23 Like a coordinate axis?
S: 00:48:24 Yeah, yeah. x-y coordinate, to solve the equation. But I think, maybe it's not…there's no need to do that. 00:48:32
I: 00:48:37 So when your professor solves problems, does he usually do that?
S: 00:48:42 Yeah, and he will like, uh make the velocity into two components. One is in the x axis and one is in the y axis. 00:48:53

I: 00:48:55 So in a problem like this one, he would, would he have the x-direction on this velocity?
S: 00:49:02 Yeah, maybe, I think maybe he will v-x should equal v-knot and v-y equal to, at this point is equal to zero, and maybe at this point is something like this point, v-x still equal to v-knot and v-y equal to like, [writing] g t-squared. Maybe this one, t, t-one, t-two, t-two squared, like that. Maybe he will write like that. 00:49:34
I: 00:49:36 Okay, so your professor might have some more labels…
S: 00:49:39 Yeah
I: like subscripts?
S: 00:49:42 And also in vectors.
I: And write it as vectors?
S: Yeah…00:49:49 It is helpful to solve more complex uh, questions, like that. But, this one…[pause] 00:49:55
I: 00:50:12 So do you sometimes use these too? Like, the vectors?
S: Uh, yeah. 00:50:17
[pause, 24 seconds]
I: 00:50:42 Is there anything else that your professor usually does…different than, than what you might have done?
S: 00:50:49 Mmm, I think just this point. He will make two components and write the vectors. 00:50:56
I: [pause 20 seconds] 00:51:16 Kay, was there anything else that you were thinking about? 00:51:21
S: 00:51:30 Um, no.
I: No? [laughs]
S: 00:51:36 Okay, was there anything on the, on the other problems, that you wanna add? Or that, um, you were thinking about?
S: 00:51:48 No, I'm just curious whether the answers were right or not.
I: Okay….00:51:58 Kay, well we can, we can look at 'em. [murmurs, I'll shut this, turn this off, camera beeps]
I: 00:52:10 I'm not sure that I have the answer to the car one, but I think I have the other ones. [goes to get problem solutions].

Problem-solving interview #3

Thursday May 7, 2009 11:30 a.m. -12:30 p.m.
Appleby Hall conference room 351

Summary of audio file:
01:30 receives problem, 26:00 done, 45:30 discuss payment / solution

I: 00:00:03 Okay so, um…you got that all signed [*referring to consent form*]
S: Mmm hmm
I: 00:00:08 Okay, um. And what I'm gonna do is I'm gonna, um, give you a problem to solve and this is, I think this should look like the same equation sheet that you use in, in thirteen oh one. And um, I'm gonna ask you to write it on this paper with the marker, just like you would an exam problem, like how you would solve it out for an exam problem. And if you want to you can talk aloud while you're working on it, but if you're not comfortable doing that you can just, solve it and then, you know I'll ask you some questions afterward about what you were thinking when you were,
S: Okay
I: when you were working on the problem
S: 00:00:48 Do you want it somewhat big for the…
I: Yeah, yeah. I mean, I might get up and zoom in a few times [laughs] if I need to.
S: Okay
I: But, yeah. I mean, just, um, probably a little bigger than you would normally write… on a piece of paper.
S: Okay
I: 00:01:02 Um, so do you have – you know, if you make a mistake, kind-of cross it out cuz we don't have a way to erase, um [inaudible on there? Or anything like that?]. Do you have any questions about that?
S: No.
I: No? You ready to start? [laughs] Okay.
I: 00:01:15 So just let me know, you know, when you're…kind-of satisfied with your answer and, and we'll go back and talk about it. So, here's, here's the, here's the problem. [beeps, camera turned on] 00:01:30
00:01:30 Reading the problem
00:02:31 starts writing with marker
00:26:00 [inaudible noise from the student]
I: 00:26:03 Are you, have you solved it to your satisfaction [laughs]?
S: 00:26:05 I think so.
I: 00:26:06 You think so? Okay [*moving to sit at the table with the student*]
I: 00:26:13 So now we're just gonna go back and I'll ask you a few questions about what you were thinking when you were working on the problem
S: 00:26:20 Mmm kay.
I: 00:26:25 So, so when you read the problem, what was the first thing you thought about?
S: 00:26:31 Problem, that it was like a force problem

I: 00:26:36 Okay
S: [?inaudible] 00:26:38 And the circular stuff…stood out. So you knew that you had to use angular equations. 00:26:48
I: 00:26:50 Okay
S: [?] traveling in a circle
I: 00:26:53 So what made you think of forces and circular motion?
S: 00:27:00 Umm…00:27:05 well, because of the tension in the string is five hundred Newtons at maximum. And…so you're trying to figure out, like, what, if, whether or not the string can hold it at a certain point, so that you needed to use forces for that 00:27:24
I: 00:27:25 Mmm kay. [pause]
I: 00:27:31 So, so what was the first thing that you, you did?
S: 00:27:36 Um, the first thing I did was I drew kind-of like the basic thing of what it means to be down in the problem, that's the easiest way to conceptualize it I think. And then writing down some of the basic, um…measurements or whatever, that you need to have in the problem. It helps you to figure out what you need to know, or what you already know. 00:28:04
S: Umm. 00:28:08 And then drawing a free-body diagram I think is the easiest way to figure out what forces you need to solve for, and what you already have. And like how they interact with each other. 00:28:21
I: 00:28:23 Okay. So it looks like you have a couple.
S: 00:28:25 This one was a wrong one, cuz I didn't read that the lowest point, so I had to, had a horizontal instead of, of where the tension should be perfectly vertical. 00:28:36
I: 00:28:37 Mmm kay. So at first you wrote it,
S: In that
I: you drew the forces
S: Yeah
I: for the string horizontal?
S: 00:28:45 Mmm hmm.
I: Okay. And…um, 00:28:50 after you had done that, were those equations for that picture?
S: Yeah.
I: [?] Okay.
S: 00:28:58 And then I also made a mistake there. It was accelerating, which really, it's, if you don't want the string to break it should be static, so it shouldn't be moving. It shouldn't, yeah. 00:29:13
I: 00:29:15 Okay, well can you say more about what you were thinking when you were writing down some of these things?
S: 00:29:21 Um, well, I split 'em up into the x and y components. And so, and then I went about and did that. Um, since in this one the only one in the y-direction is force of gravity so then that's equal to your weight. But, that was wrong anyways. And then I knew that the acceleration, er, the centrifical acceleration should be in circular motion so you had to change it from linear to uh, uh, circular motion. 00:30:03
I: 00:30:05 Okay, so when you say change it, what do you, what do you mean?

S: 00:30:08 Uh, just [?] the relationships between angular and linear motion, and you can substitute it in there.
I: 00:30:18 Okay, so you were using an equation here for cen, centripetal acceleration
S: 00:30:25 Yeah, and then the relationship between linear velocity and angular velocity
I: 00:30:32 Okay. And then you solved for…this one here? Is that?
S: 00:30:40 Yeah, so then you get the angular velocity, omega, from the linear velocity divided by the radius.
I: 00:30:51 Okay and then what did you do with that?
S: Umm…[tapping noise?]
I: [? Laughs]
S: 00:31:02 I substituted because originally I had that the tension was equal to the mass times the acceleration, which the acceleration would bee…omega squared r. And if you substitute the omega into that, then the velocity's in the radius axis. 00:31:26
I: Okay. 00:31:34 And then after that, did you go…here, or here?
S: [?] 00:31:39 But then this is…this is the one that should be more right than this, so.
I: 00:31:48 Okay, so what made you change your mind?
S: 00:31:50 When I was reading through the problem like and I saw 'the lowest point'.
I: Okay.
S: 00:31:54 I didn't read through the problem carefully.
I: 00:31:59 Okay so, can you tell me about when you, when you changed your mind and, um wrote a new, new picture?
S: 00:32:08 Well, cuz it's at the lowest point everything should be in the vertical direction. So it makes it a lot easier. Umm. And so then you just have your force into the y-direction, that'll make your tension positive with the force of gravity negative and the centrifical force negative also, and since it's, you're trying to figure out what the maximum is without breaking, you're gonna put that equal to zero then, and then maximum tension is equal, is less than five hundred, or, equal to five hundred, so. 00:32:49
I: 00:32:50 So you had initially written this and then crossed it out. Um, so what, what made you change your mind?
S: 00:32:57 Um, the centrifical, probably.
I: 00:33:01 Kay, so you just thought, you were missing something?
S: 00:33:04 Yeah. [quietly] I might be wrong.
I: 00:33:08 Okay, so then, then, then you, in this equation you added the force here. 00:33:13 [pause] 00:33:17 Or I mean, subtracted that,
S: Yeah
I: 00:33:20 that new force? Okay, and then once you had that equation…you solved for tension there?
S: 00:33:29 Yeah.
I: Is that right?
S: 00:33:30 Cuz the force is equal to the weight, and then just, the centrifical is maximum acceleration, and that's the same stuff that I did right here. Just, use this mass is equal to the weight that the gravity, you can pull out the weight. And then you get this. This is the square root of this. Equal to the maximum angular speed that you have. 00:33:58

I: Okay…00:34:03 so you have a, an answer here, of the maximum angular speed?
S: 00:34:08 Mmm hmm.
I: 00:34:13 Okay. And you have some units here. How did you know to use those units?
S: 00:34:21 Because angular speed's in radians per seconds and since everything in the problem was in metric units, you don't have to do any type of conversions. Everything was in Newtons and meters, so. You convert the centimeters to meters, and everything cancels out. Should.
I: 00:34:45 Okay, so one thing you did right away was convert the centimeters
S: Yeah
I: to meters when you wrote it down. 00:34:51
I: 00:34:57 Okay. So is this uh some, something that you usually do when you're solving a problem?
S: 00:35:06 converting things?
I: 00:35:07 Well I mean, is this kind-of a typical
S: Yeah
I: 00:35:09 process? [pause]
I: 00:35:18 So, um, if you were solving this for an exam,
S: Mmm hmm
I: like that's being graded, would you solve it the same way that you just did now, or would there be something different?
S: 00:35:31 Um, I'd probably do the same thing just…I'd cross out this stuff because it wasn't at its lowest point so it wouldn't be right. So I would have just done this stuff. 00:35:42 [pause] 00:35:48 And put some of this stuff over here, the basic stuff.
I: 00:35:54 Okay, so you would still show some of this stuff?
S: 00:35:56 Yeah, I would show the pictures and everything. 00:35:58
I: 00:36:06 Okay, so when you were, I saw you looking at the equation sheet a few times, what were you looking for on that?
S: 00:36:14 Umm. [pause] Nothing in particular…oh I looked at the centripetal acceleration equation. Um. That was about it, cuz the rest of the stuff we didn't really need. So…00:36:38
I: 00:36:40 Have you seen a problem like this before in your class?
S: 00:36:46 Umm. Not exactly. But things similar. [quietly, I don't think] - We haven't had anything like this exactly.
I: 00:36:59 Okay. Did parts of it look familiar?
S: 00:37:04 Yeah. The swinging, it looks familiar. Like a long time ago though. And 00:37:10 [pause] 00:37:22 Um…kind-of the, the mom- the maximum force thing was familiar.
I: 00:37:36 Okay, so the way that it's stated?
S: 00:37:37 Yeah.
I: 00:37:45 So while you were working on this problem was there anything else that you were kind-of thinking about that you didn't write down?
S: 00:37:54 Um…probably.
I: [laughs]
S: 00:37:58 I don't show my steps very well sometimes 00:38:01

S: 00:38:09 I don't really know how exactly but…I think that I didn't solve for, like whether or not it would get that high yet. 00:38:22
I: 00:38:25 Okay, when you wrote these equations in the corner.
S: Mmm hmm.
I: Um, what were you…
S: 00:38:33 I was thinking about the, like whether or not the velocity would get up to the highest height. Like the, maximum velocity, angular velocity would translate into a linear velocity and whether that would get to the nine meters. Which I didn't do for this thing. So. 00:38:49
I: 00:38:54 Okay so you were writing down some equations…um, in terms of the height?
S: 00:39:00 Mmm hmm.
I: 00:39:14 So is there anything, um, you would change, still, on this solution? 00:39:18
S: 00:39:21 Yeah, I'd figure out whether or not this could get up to nine meters, if it was in linear terms.
I: 00:39:30 Okay, so how would you go about doing that?
S: Um…
I: 00:39:38 We have time if you wanna [laughs] if you wanna try it out.
S: 00:39:42 Oh. Just say the 00:39:45 [pause] 00:39:56 since you know that the velocity at the top can be zero, to get right there, you have the acceleration and you know the velocity, which is just um, omega r, so then you just times by this by the r so it would be [writing] equal to nine point eight two plus [pause] omega r. So. 00:40:24 [pause, using calculator and writing]
S: 00:40:55 So we know it takes one point five two seconds from here. [the height, inaudible] We need a height. 00:41:07 [?]
S: 00:41:14 Your height is nine meters and your acceleration is nine point eight, since you're starting at x-knot it's zero, so 00:41:23 [pause] 00:41:35 then you can just solve for that. [?Can you?] 00:41:40 [tapping]
I: 00:41:59 So you said you were thinking about solving for the height here?
S: 00:42:02 Uh, I think so. This, the question's kinda confusing cuz it doesn't actually ask for a specific thing that you're trying to find. 00:42:13
I: 00:42:16 Okay, so, so what did you find?
S: 00:42:20 The like, maximum velocity that the rope can hold, is what I found. So. 00:42:29
I: 00:42:40 So do you know if it'll still
S: [inaudible]
I: …oh, go ahead [laughs]
S: 00:42:45 Well, I guess with time you can just put it into the thing with acceleration and the velocity and see what height it got.
I: 00:42:54 Do you wanna try that?
S: Sure 00:42:56
I: [laughs]
S: [inaudible? Pause, typing in calculator]
S: 00:43:43 So according to this it can get high enough
I: 00:43:46 Okay, so what did you get for the final…?

S: 00:43:48 thirty three point nine meters
I: 00:43:51 Mmm kay.
I: 00:43:57 So you just took the one half a t-squared plus…
S: 00:44:02 Yeah, the velocity, the initial velocity times time
I: 00:44:07 Kay. [pause]
I: 00:44:13 So are you satisfied with this solution, or is there anything you're still thinking about?
S: 00:44:19 For the most part. The centrifical force sometimes is really confusing to me. So. Might be that, that's where I think I made the biggest mistake
I: 00:44:33 Kay so, what, uh can you say more about
S: well like,
I: what's confusing?
S: 00:44:39 They say it's like a fake force or whatever, so. Just kind-of a weird, it's hard to visualize sometimes I guess you could say. And then put it into more mathematical terms. 00:44:54
I: 00:44:57 Kay, is there anything else that you're still thinking about?
S: 00:45:01 Mmm, Not really. The relationships are correct. Just…you know. For the most part. I think. 00:45:12 [pause 12 seconds]
I: 00:45:24 Okay. Well, unless there's anything else [laughs] that you think um, is important.
S: 00:45:30 Not really.
I: 00:45:33 Kay. Well I think we're…I don't think we have time for a second problem so, [beep - turns camera off].
S: 00:45:39 Okay.
I: Okay. Um, and I have a, a solution I can give to you [laughs, inaudible].
S: 00:45:53 Is it somewhat correct?
I: 00:45:54 Somewhat, yeah

Problem-solving interview #4

Friday May 8, 2009 10:00-11:00 a.m. (arrived a little late)
Appleby Hall conference room 351

Summary of audio file: 1:30 receives problem, 2:40 student comments about problem; 13:52 minutes says lost, 21:00 interviewer says can use more paper if need to; 25:47 student says he's done; 33:38 interviewer says have time if want to try calculation; 46:43 resumes discussion; 56:48 end interview / discuss solution & payment

I: 00:00:02 Alright…and that - there's a copy of that for you to keep, um, afterwards [*referring to consent form*]. And this should look
S: Yeah
I: like the same equation sheet, um, from, from class. And so, um, [quietly] get this set up [*camera beeps*]
I: 00:00:27 So I'm gonna give you a problem
S: Mmm hmm
I: And it should look kinda similar to [closing squeaky door] problems that you've had like, on exams,
S: Okay
I: and stuff. Um, and I just want you to try and solve it like you would in an exam, um, and use the paper and the marker. You might need to write a little bit bigger than usual so I can see it
S: Okay
I: 00:00:47 on the camera. Um. And…you know, just, if you're comfortable you can talk out loud while you're working on it,
S: Sure
I: If you're not comfortable doing that, um, you can just wait and we'll talk about it
S: Okay
I: 00:00:59 at the end, um, so. When you're satisfied with, with your solution, or, [laughs]
S: Uh-huh
I: you wanna, you wanna be done with it
S: Yeah
I: 00:01:09 Just let me know and we'll, we'll talk about it a little bit more. I'll ask you a few questions.
S: Okay
I: Um, so. Do you have any questions
S: Um
I: before I give you the problem?
S: 00:01:19 I don't think so.
I: [laughs] Okay. [pause] Here's the problem.
S: Mmm kay. 00:01:30
S: 00:02:38 [sniff] Mmm kay, I was bad at um, circular motion, so. [hear writing, sniffs]

S: 00:13:52 [sniff] Pretty sure I'm lost
I: [laughs] Pretty sure you're lost?
S: Yeah
I: Can you say more about that? What are you, what are you thinking right now?
S: 00:13:59 Um…okay I'm just, I'm pretty bad at the rotational energy stuff. So…I got, I'm not sure if I should even be using, um, a t-squared plus v-t plus x, but I think…I'm just having trouble tying in the fourteen Newtons, the tension, and how that's um, converts into the y-axis. [sniff] 00:14:40
I: 00:14:47 So what have you tried so far?
S: 00:14:50 I can't really,
I: [laughs]
S: 00:14:52 I don't really know. I was just trying to put everything I know down, and then seeing what equations el-eliminate stuff. Um, and what I could plug in [*sniff*]. And that didn't get me very far so far. 00:15:07
[00:15:18 hear writing]
[00:16:08 typing in calculator]
I: 00:21:02 You can feel free to use another paper if you wanna start [*laughs*]
S: 00:21:05 Oh, okay.
I: 00:21:06 …writing on another sheet.
[00:21:48 hear paper shuffling - started writing on new paper?]
S: 00:25:47 Um, I'm done
I: You're done? [laughs]
S: 00:25:50 Yeah
I: Okay, so we'll take some time now and talk about
S: Mmm kay.
I: What you're thinking about this problem
S: Mmm kay.
I: [laughs] 00:26:00 So when you, when you read it, when you read the problem
S: Mmm hmm.
I: …what, what did you think about?
S: 00:26:06 Um…that first thing it was gonna be hard. Cuz I'm so bad, at the circular, angular momentum but, at first I wasn't really sure what to do and then, um, I remembered…this is where I started having a little more confidence was…um, kinetic energy and momentum, cuz I figured the kinetic energy of it going around would be equal to the kinetic energy as, as soon as you let go [*sniff*] and then, um, I figured at that point that kinetic energy should be equal to the potential up here. And then, um, I just couldn't figure out how to…find out the…connection between the kinetic energy here and then the m-g-h there. Or actually the, uh, angular kinetic energy and the energy there. 00:27:19
S: 00:27:22 And for some reason some of this stuff was, I couldn't figure out where to put it in, um. Gravity right here, I don't think I figured in to my equations anywhere. I'm not sure where to put that one in. [pause 4 sec] Um. 00:27:42
I: 00:27:45 So even before you got to this point
S: Uh-huh

I: Um, can you say more about what you were thinking about when you wrote some of, some of this stuff down?
S: 00:27:52 Yeah, I just wanted, I knew force equals m-a is in like every physics problem, so I put that one down. And then, I knew there was angular, um, momentum and torque. So I knew I would have to be using one of these, either this or the, the force times radius. Um. But then I didn't, I wasn't sure, uh, the alpha. You couldn't, you don't know, I figured it out over here and that was like, ridiculous, like three hundred meters per second squared. So I knew that was wrong. Um. 00:28:34
S: 00:28:35 And then I was trying to…to get the, I forget what you call it but, um…projectile motion. Cuz that I, for some reason I always go back to that and it never seems to help me when, but. And that didn't really help. Cuz that's solving for t and t wasn't even in here anywhere. But once I figured t wasn't in there, I figured kinetic energy would be, um…and then I was trying to think what the force would be. And I couldn't find the force unless I knew *I* and alpha. So. 00:29:25
I: So, when you wrote this down,
S: Uh-huh
I: what, what force are you talking about there?
S: 00:29:31 Um, the torque. And actually, the torque would be…would be the force this way, right? I'm not quite sure about that. [inaudible, relate this thing here so] 00:29:56
S: 00:30:00 Yeah, that's really what got me was the torque, and the alpha, and the double-u, the omega.
I: 00:30:21 Can you say more about what, what this picture is?
S: 00:30:24 Yeah, it said, um…after a little while of moving your hands back and forth you notice that you no longer have to move your hand to keep the bag moving in a circle which means the force going up and it's, that's the, where gravity opposes the force the strongest. Or um, yeah…[quietly] I think. But
I: 00:30:59 So you're, so you're saying this A is the force?
S: 00:31:01 Um, yeah. That has to be somehow to counteract gravity. So. These two have to be equal or A has to be greater than, and then down here it said the rope would break at five, five hundred Newtons so I knew that gravity plus whatever this is, must equal five hundred Newtons otherwise it'll break, or it has to be less than that. Um. And I'm not sure if that is torque or if that is…something else. 00:31:43 [pause 23 seconds]
S: 00:32:06 Not sure.
I: 00:32:14 So then you have some other things written down here…
S: 00:32:18 Mmm hmm. Um…v, velocity equals radius times omega. Um. I think I used, no I didn't but…I probably could have, I forgot about that one.
I: 00:32:42 If, if you were gonna use it, what would you do with it?
S: 00:32:47 I think um, I wanna find double-u, cuz I keep getting…I was getting a different answer for double-u over here I think than back here, um. So I'd plug in the radius, point six five times um, I believe this is pretty close, and this is correct. So I'd plug in um, v and r and get double-u. And then I can use that either in kinetic energy or…mmmomentum. Angular momentum. And I'm not sure if angular momentum helps me here. I think it's more of kinetic energy. 00:33:31
I: 00:33:35 We have a few minutes if you wanna try

S: 00:33:38 Okay.
I: …try that. 00:33:38
S: [quietly while writing] 00:33:43 so velocity…r… [typing in calculator] 00:34:08 Kay I got double-u over here is thirteen point three and then I got, um, over here was twenty point five. So, obviously one of them's wrong. It's probably that one. So now I have, I have omega. And [pause 8 seconds] and we can't use this cuz this is probably wrong to. This, um, I-value…So I don't know how I would find the I-value. Or wait, now I have, okay. 00:35:03
S: 00:35:14 Kay the velocity of the angular motion should equal the velocity of the…the linear up and down right at that point. So. And so should the energy. 00:35:37 [*writing*]
S: 00:35:50 I'm really, hoping this is right. This equation right there.
S: 00:35:54 [*writing, typing in calculator*]
S: 00:36:39 Now I have *I* equals point zero four six. Which…yeah, probably not right [*typing in calculator*].
S: [murmuring, ?] 00:37:08 equals that, I think so.
I: 00:37:13 Why do you say it might not be right?
S: 00:37:16 Um, cuz I'm looking at…these other uh, moment of inertia values like, two-fifths m r-squared. And if you just plug in like, [*typing in calculator*] two-fifths times, I don't know, one kilogram [*typing in calculator*] divided by, er times...point five meters, actually that won't work, you get point two 00:38:03 [pause 9 seconds]
S: 00:38:12 That's off by a factor of ten. I'm just trying to relate that to some object I can picture and then, but I don't know if I should be doing that or not. But, okay, if it was right…then I have *I*. 00:38:33
S: [pause 11 seconds] 00:38:44 Mmm, I don't know. I'm lost. 00:38:46
I: 00:38:53 Did you have something in mind before, solving for omega and I?
S: 00:39:00 Um. Well now I think I have both omega and I here and 00:39:09 [pause 9 sec]
S: 00:39:18 Okay. [pause shuffling paper]
[00:39:34 hear writing]
S: 00:39:42 Well if I put five hundred Newtons into the torque, in that equation t equals I-alpha, then I get some ridiculously huge number for alpha. 00:39:54
S: [pause, typing in calculator] 00:40:07 So maybe I don't, hmmm. [pause 28 seconds] [00:40:37 writing, 00:41:00 typing in calculator, 00:42:05 writing again, 00:45:51 typing in calculator]
S: 00:46:53 [*sighs*] Kay now I know, I know my velocity's off, I think. Cuz I plugged, um, v equals [*writing*] use that, and then to figure out when v equals zero, um, I plugged in the thirteen point three into v-knot and zero over there, so my t was one point three six. 00:47:35
S: [*typing in calculator*] 00:47:40 And, and then I plugged one point three six into…the…that equation up there. And I got…um, height of eleven point three six which is over nine, which…that should have, with the numbers I used I should have got exactly nine. So now I know my velocity is wrong. So, I would just have given up on this problem.
I: [laughs]

S: At this point 00:48:09
I: What would you hand in, to be graded?
S: 00:48:14 Um, I'd probably just write this right here. [*writing*] V, we'll just call it v two squared, equals, m-g-h. Um, I'd do all the diagrams and then…um. I don't even know. 00:48:57 [pause 8 seconds] 00:49:05 I'd probably figure out that, write that and try and um, just put in what my calculations were. And, you know, make it clear that I, I wasn't getting [inaudible it, anywhere?].
I: Which calculations would you include?
S: 00:49:22 Um, well I'd, I'd include that just to say, you know, m-g-h equals that, since I already wrote that. But that's turning out wrong, so I don't know. 00:49:37
I: 00:49:40 So the m-g-h and the one-half m v-squared?
S: 00:49:42 Yeah. [*typing in calculator*]
S: 00:50:03 I still got v equals thirteen point three. Maybe that's wrong. 00:50:17
I: 00:50:21 How do you, how do you decide when you want to…quit working on a problem…
S: Um
I: or keep going with it?
S: 00:50:28 Well since, uh we ha-, we have timed tests, so usually what I do is, I'll just write down this, and then I'll write down um, the diagram. And then I'll go to the next one and I'll do the same and then, whichever one seems like I know what I'm doing more I'll finish that one. And then I'll go back to the other one. And then, um, usually the multiple choice I either get it or I don't, and I'm not gonna work too hard on those. So I'll just save like ten minutes of time for those. So I'll finish one problem, work on the next one until I have ten minutes left and then, do the multiple choice. 00:51:15
I: 00:51:16 So you usually have two problems on an exam?
S: 00:51:18 Yeah, there's two problems and then five multiple choice. And I think that's the way with all the IT physics. Yeah, this problem I'd probably get like, five points out of twenty five. Or maybe ten points or something. 00:51:40
I: 00:51:43 Do you get points for the picture? Or is that just something that you put down?
S: 00:51:47 Um, I th - it depends. Some, I think different people grade 'em. But sometimes they'll give you a ton of points if you…put down the picture and the equations, and then other times they just give you like, three points or five points for that. So it's really just, you might as well just put those down anyways. And I think some, if I don't do that I'll get mixed up somewhere, and it just, it's a lot of, it's a lot more help. 00:52:24
I: 00:52:25 Kay. So you said it helps you? So how does, how does it help you?
S: 00:52:28 Well if I'm, if I'm looking for energy and stuff, and then I can kind-of picture like, okay, it's going around it has…kinetic energy. [?] It let goes at some point. And I can look at that, and then it's going up, um, I look over here, I drew something there for…um. So then I know the kinetic energy there, is, probably if my hypothesis is not, it's equal there. And then, I know that there, it goes up nine meters, um. And it'll stop at the top so, potential energy equals kinetic energy equals…kinetic angular…energy. And it's just easier if I'm looking at something and I get an answer, and then I can look at that and then I'll say okay, that's what that means, or, that's

where that fits in, or if it, if, you know, I get a negative number for velocity then I know obviously I'm, I'm wrong. I have to look back. 00:53:46
I: 00:53:50 So you don't just use it at the beginning? Sometimes you go back..
S: Yeah
I: to the picture?
S: 00:53:54 It's usually the beginning, and then I'll go through some calculations, and once I get an answer I'll look back at it, and then, see if that makes any sense from the diagram. [pause 5 seconds]
S: 00:54:11 And then, at the, the very end I'll just look at it, if everything makes sense. Cuz usually I can picture if it's, um, like, like I said the, I was looking at the inertia, and if I just plugged in some object that, you know, had some inertia and then if I plug in the values for this and it's like, uh, you know, two orders of magnitude off then I know I'm probably not right, otherwise they're being mean 00:54:47
I: They're being mean? [laughs]
I: 00:54:50 Okay, so you try to compare the value
S: Yeah
I: …that you get
S: 00:54:54 To some kind-of, normal object. Like one of the tests we were getting like, it was a cannon, and then we were trying to find the velocity after it shot something off, we were getting something like six hundred meters per second and you know, if I would have just wrote that down and said okay that makes sense that'd have been really bad, so. Just like double checking that um, in your head and stuff, or using the diagram. 00:55:26
I: 00:55:30 Was that on a, on a group problem, or on an individual?
S: 00:55:35 That was…that was group. Yeah. And I don't think anyone in our, in our discussion got it. It was really difficult…But it seemed like we were doing it right. I, I had no idea where we were going wrong. And then we kept getting answers that were like double what it should have been. 00:56:01
I: 00:56:04 Okay is there…we're kinda running out of time. Is there anything else that you wanted to say about this problem?
S: Um
I: That you were thinking about
S: 00:56:12 I should probably work on angular momentum stuff. But otherwise, no…I think I have to figure out what direction torque is in. I can't remember that….or what, what it acts on. It's pulling in, or if it's…[? inaudible, should be determined?] I don't know. But I think that's it. 00:56:48
I: Mmm kay.
S: 00:56:52 Alright
I: [quietly] I'll just shut this off and then I have um, a copy of a solution that I'll give to you [laughs]
S: Oh,
I: So that you're not like… agonizing over this problem for the next couple of days
S: [laughs]
I: 00:57:07 And I do need one more signature from you about the payment, um. So.
S: Okay.

Problem-solving interview #5

Tuesday May 12, 2009, 2:00 – 3:00 p.m.
Tate Lab of Physics room 160

Summary of audio file: 00:01:35 receives first problem, 00:03:00 starts writing, 00:06:13 asks question, 00:07:42 says done, 00:13:44 changes mind about distance (initially added to 9+.65 m instead of subtracting), 00:14:36 realizes velocities different, makes changes to solution until 00:16:04; 24:09 starts second (bungee) problem, 25:52 asks what Jello is, 28:20 finishes second problem, 37:38 discuss what hand in to be graded, 44:14 discuss solutions/answers

I: That's for you to keep
S: Okay.
I: if you wanna keep a copy of it [consent form]
I: Okay, so. I have a, a problem I'm gonna give you [laughs]
S: Yep
I: And, um you can solve it just like you would a problem in your physics class
S: Okay
I: Like, for an exam
S: Okay
I: Um, and you can write with this marker on here and, yeah. And then um, you might wanna write a little bit bigger than, than what you usually do so
S: Okay
I: I can see it
S: Okay
I: on the camera. Um, and this should look like the same equations that, from class.
S: Okay
I: And if you have a calculator you can use yours, otherwise there are a couple…
S: I, I have [inaudible, getting calculator from bag]
I: Okay. And then uh, if you're comfortable talking while you work on the problem you can do that [laughs]. If you're not comfortable doing that, um, you can let me know when you're finished and then I'll ask you some questions.
S: Okay.
I: afterwards about what you were thinking. But you, do whatever's comfortable for you.
S: Okay.
I: Um, so. Are you ready?
S: Uh yeah I'm ready.
I: Do you have any questions?
S: Nope. 00:01:13
I: And I'll be back here just um, focusing the camera and stuff, so. Try to ignore me [laughs]. Here's the problem.
S: Only one problem?
I: Yeah, if we have time for another one…

S: If I have something I don't understand can I ask you?
I: Yep, yep. You can ask me. That's fine. 00:01:35 [camera beeps]

[00:03:00 starts writing]
S: 00:06:13 I have a question.
I: Okay.
S: Is this problem to find whether I can throw the bag to the top of the building?
I: Um, well it says um…
S: That, I didn't find the question
I: Okay.
S: But I think it is asking whether I can throw the bag to my co-worker at top of the building.
I: Okay.
S: So I solved the maximum height it can reach is larger than 9 so, it can reach.
I: Okay. Are you satisfied with that answer,
S: Uh
I: or do you want to take any more time?
S: Uh, think I will...
I: [laughs] Okay.
S: read the question again. [00:06:55]

S: 00:07:42 Yep, I am done.
I: Okay, now I'm gonna go back and ask you a couple questions about what you were thinking
S: Okay
I: So. When you first read this problem, what was the first thing that you thought about?
S: 00:07:57 Mmm. I want to know and, what's um, I want to find the question. What I want to know in this problem. And like, it's, I find that I want to know how, the height that the bag can reach. The maximum height it can reach and. So, this is the first thing I want to know.
I: Okay, so you first read it and,
S: Yeah
I: looking for the question?
S: Yeah
I: Okay, um so then what did you start thinking about after, after that?
S: 00:08:33 Um, first uh, I find out what I want to know. And I find out what I already know. And I need to build a relationship between them. So I can find answer. So and, like in this question. Uh, I have something know, known here. Like the weight of the bag and the, um, the radius of the circle. And also the, here's the maximum um, [murmurs, say that?] the string only can hold up maximum, the maximum force the string can hold is five hundred Newtons. And so.
S: 00:09:23 And, and in this problem I find that the, I want to know the height so I need to know the velocity. And in order to find the velocity I need to know the, the…use Newton's second law I can find the, the relationship between the force and the velocity. So. And I build the connection with the known things and the [?] other things.

I: Okay, so how did you decide to use Newton's second law?
S: 00:09:56 Cuz um, like it's in a circle motion and um. Um, like I said I want to find the velocity. But I don't know, I only, the thing I only know in this question is the force like, the tension. The maximum tension is five hundred Newtons. And um, so I think Newton's second law is the best way to connect the force and the velocity. And um, like a m is equal to net force. And uh, so T minus m g is equal to a m. And a I know is, a is equal to v-squared over r and then I build this, then I got v. If I got v, use the energy conservation and I got h.
I: Okay. Did you consider any other physics um, before you chose Newton's second law and energy?
S: 00:10:58 Um, actually, mm I didn't. I just…
I: [laughs]
S: Come up with these two and then solve.
I: Okay….And so then, um, what did you do once you had that? Um, those relationships?
S: 00:11:21 Um, I. If I had got these relationships and plug the numbers T and v and r and I got um, eleven point two meters. And then I found that it's higher than the building. Building is nine, nine meters from the ground. So, I think I can, it is larger than that I can throw the bag the top of the building to my co-worker.
I: Okay, and then I see you crossed something out here. The nine meters. Can you explain…
S: Yeah.
I: …that to me?
S: 00:11:57 Yep, that's a little problem, cuz. A man is, a man stand at the ground, and it said that it's a circle and uh…the, the time I set up, I set off the bag is when the string is, string is horizontal so, I think the initial height is the radius of the circle. So this is the building, nine meters. So the bag is not, not set off from the ground, it is set off from the, here. And I considered this is the radius of the circle. But maybe not, cuz there is person, so we don't consider the height of person so we have to consider the circle.
I: Okay. So how did you consider that?
S: 00:12:58 Mmm. Like, [? How] Uh, well, we're, the bag like this. On the circle. And then we don't know the height of the person. So, I think…it should be like this. Circle like this. And then the bag is set off from the, when the line is horizontal. So this height I consider is to be the [? point six five] meters.
I: Okay, and so then when you crossed this out what, what does that mean?
S: Um. [pause] 00:13:44 Actually, cuz, actually I compared this, this one, eleven point two is the height from here to here. [writing] So I choose, actually not this should use this [re-writing] Nine minus zero point six five.
I: Okay, so the first time you added it?
S: Yeah.
I: And now, now you [laughs] subtracted it?
S: 00:14:17 Yeah, cuz, Yeah. Cuz we need to compare this height with this height [inaudible].
I: Okay. So when you found um, this velocity. What, where is that velocity? What is that velocity?

S: 00:14:36 Uh, this velocity is, initial velocity. Like, um, from this point. Cuz, this is height when bag set up, um get off my hand, so this velocity – Oh! Oh sorry.
I: Well, no, I mean, I'm asking what [laughs] what you think.
S: The velocity is this at this point. 00:15:05 [pause, murmurs at this point? Inaudible, kinetic energy]
I: 00:15:28 We have time if you wanna make some changes
S: Okay
I: to your solution [laughs]
S: Okay
I: Or if you wanna use another piece of paper that's okay too. 00:15:35
S: [murmuring and from here…this point] 00:16:04 Okay. [writing] Nine. 00:16:10
I: 00:16:15 [laughs] Okay
S: I change this to nine.
I: Okay, so can you explain why you changed that to nine?
S: 00:16:18 Yeah. The problem is I forget the velocity is from here, and it is not from this point. So, um, so I should…uh, I should actually I should know the velocity from here. But, and from this point to this point [?and then it goes down?] Only force that does work on the bag is the…gravity. Do work on the bag.
I: Okay
S: 00:16:54 And the, and the, so. I-I can use the energy conservation like, yeah. The kinetic energy all transferred to the potential energy. So the h is, so I can get the h from the highest point to the lowest point. So it's eleven point two and just compare it with nine.
I: Okay. Now is there anything else that you were thinking about uh, that you didn't write down when you were working on this problem?
S: 00:17:35 Mmm. [pause] Anything else. 00:17:46 I didn't write the 'according to the Newton's second law' and conservation of energy. And that's all.
I: Okay, so when you do this, these kinds of problems on exams, do you usually write out you know, 'according to Newton's second law…'
S: 00:18:03 Yeah. I have to write that cuz maybe lose points for that.
I: You can lose points for that?
S: 00:18:19 Yeah, if I did that. Because maybe the grader maybe didn't know what did I mean. How to say, clearer.
I: Okay. What else do you usually write down for a graded problem?
S: 00:18:25 Mmm. I write down the conclusion, like, you ask me whether you can reach the building. So at last I say that um, after um, that, so, so it can reach the top of building.
S: 00:18:46 And um, and I will say every um, every…Like velocity. I will use something else to explain at which point this velocity is and the height. And uh, something like that. And at the beginning of the problem I will state what, what's this mean, what's that mean.
I: Okay. Um. I saw you draw a picture, too.
S: Yeah.
I: Is that something that you usually do?
S: Yeah. Yeah.

I: Okay.
S: 00:19:20 You have to do that cuz it will make the, it will simplify the problem.
I: Okay, and do you always draw a picture?
S: Ah yeah.
I: Okay.
S: It's a habit, I think.
I: A habit? [laughs] Okay. Um. Do you ever get points for that, on tests?
S: 00:19:41 Uh, graph.
I: Yeah, or a picture?
S: 00:19:43 Picture. I have no idea about that but I think, I've done pretty good on the exams so far. Mmm, I think picture is important because, and graphs, cuz in the discussion, in the discussion group the teacher told us that we should draw, draw pictures before we solve the problem.
I: Okay. Um, and then when you get an answer uh, do you ever think about if it might be right or not, or?
S: 00:20:28 Yeah. I will double check it. Like at the end of the exam I will come back to beginning to find whether, which progress goes wrong, or which number I write wrong.
I: Okay, so how do you check? What do you check?
S: 00:20:46 Uh, this problem. I will use the number I get and uh, and uh, plug in to the problem and the, to solve the, the things I already know. To solve it again and check if it's right with, with the question.
I: Okay.
S: 00:21:14 It's one way and the…Can't think of. Mmm.
S: 00:21:26 Sometimes I just do the question again. [laughs, inaudible]
I: Do you have time to do that, on a test?
S: 00:21:33 Uh, usually do. But the…on a quiz it's difficult, maybe I don't have enough time. Usually I have time.
I: Usually?
S: On a quiz.
I: Okay.
S: Do this here.
I: Have you seen a problem like this before?
S: 00:21:52 Uh, this kind of problem, no. But I…I know this kind of, it's different from this, this problem but it has something in common.
I: Okay, so what do you think is in common?
S: 00:22:10 Like, Newton's second law. Just finding the velocity of the motion, velocity and the force. So, the first time I see the question I was thinking about this and the, the other question I have done before there is also, also use the Newton's second law and uh, actually they are different. Maybe they are different from the question but uh, actually they are the same in the way of doing them.
I: Okay, so how was this problem different from one you've done before?
S: 00:22:50 Um, like, combine the circular motion and the, also the, the linear motion. And this is the one difference. And then [pause 5 seconds] It's very long. [laughs]
I: It's very long? [laughs]

I: Does it look similar uh, to what you usually get on your exams?
S: Mmm.
I: Or, you said this was longer?
S: 00:23:18 Longer than that but, it's, the, the way is similar. Cuz you are asked to do something. You are asked to design someth-something. And always, like that. And um. And I think it's…that's all.
I: Okay. Do you wanna try a second problem?
S: Okay.
I: [laughs] Okay. 00:23:44 [murmurs, get you some fresh paper]
S: [inaudible]
I: 00:23:52 And then it's kind-of the same as before, um. If you're comfortable talking aloud while you work on it, that's okay. If not, I'll ask you questions
S: Okay
I: afterwards again. 00:24:09

S: 00:25:01 [sighs] I think I have done this question before.
S: 00:25:52 I have a question. What's the meaning of this two point five meter pool of…?
I: Jello is like a, um, it's a food that is…
S: Food?
I: Yeah, it's a food [laughs] that is like, um, a gelatin or like a gel.
S: 00:26:11 Um, a gem?
I: Gel, so it's kind of thick, like a thick liquid.
S: A liquid?
I: Kind-of like a liquid,
S: Okay
I: but a little thicker.
S: Okay.
I: [laughs] 00:26:21
[00:26:21 writing]
S: 00:26:38 Oh.

S: 00:28:20 I am finished.
I: Okay [moves to sit at table] So when you read this problem, what was the first thing that you thought about?
S: 00:28:32 Yeah, um. It's the same as that, that question. I want to find what's this question asking me to do. And that, that's it. The first thing I do.
I: Okay, so what did you think this question was asking you to do?
S: 00:28:49 Um, asking me to find the, the requirement for the spring constant so that the people who try bungee jump can, won't, won't die.
I: Won't die? [laughs]
S: [inaudible?] 00:29:11 His height reach the top of the Je-Je-Jello. [pause] Yeah. This line. [referring to problem statement]
I: 'You must determine the elastic constant?'
S: This, this

I: That sentence?
S: 00:29:33 Yeah. This is the question.
I: That's the question? Okay. So after you found the question, then what did you think about?
S: 00:29:40 Um, I want to think about the like, this is what we want to know, the con - spring constant. So I-I [?] by writing [?out] the, the the motion of the people and the, uh, the things, the height of the, the heights. All kinds of heights. Uh, including the people's height. And the, the mass of people. So I need to c-, uh, uh combine these things together and the, I-I-I have to analyze the, what's happen in this motion.
S: 00:30:18 So, uh, I think uh, it's uh, energy conservation. So the people from this, the highest point to lowest point and the initial velocity is zero and the final velocity is also zero. So, there is only the potential energy, um, [inaudible let's see] that the potential energy transferred to the spring, I forget the name of the- [pause]
I: The energy in… the spring energy?
S: 00:30:57 Yeah, it's not that name, but you, you can understand that. Transfer to the stretch of the spring. That kind of energy. And the, this is the spring potential energy and this is the gravity potential energy, and that equal so, I can solve that.
I: Okay. So when you wrote down this, these numbers for x what, what were you doing there?
S: 00:31:27 Um, I want to find the stretch of the, uh, spring. Uh, this is the, not the spring but is similar to a spring. Uh, mmm, cuz I need to know, this is the person. And the…person is two meters height. And this height, this height [?] reach the lowest point. Mmm, and uh, I want to find the stretch of the spring so the spring is sixteen, sixteen meters long so before it gets sixteen it is not stretched. So after sixteen so this length is, this stretched length, and uh so, I used the whole thirty meters minus the initial length of the, the spring, the rope. Then I minus the height of the person and I found this, this part. This twelve meters. 00:32:35
I: Twelve meters?
S: Yep. And the, this is stretch of the rope.
I: Okay, so first you…solved for k
S: 00:32:50 This is the k.
I: Yeah
S: 00:32:52 And I want, I want to find the k.
I: Okay. And then you…you put the numbers in.
S: [inaudible]
I: Okay. And what did you get for k?
S: 00:33:09 Uh, Two hundred and eighty five point eight.
I: Okay. And um, how did you find the unit for that?
S: Oh
I: Did you…
S: 00:33:21 I remember that.
I: You remembered that?
S: 00:33:23 Also, you can think of a way like, [writing] k is equal to um the, in a spring the force is equal to minus k x. So k is equal to F over x, so, so the force, the unit of force is N. The unit of x is m. So it's [inaudible, this units N/m].

I: Mmm kay…So was there anything else you were thinking about um, when you were working on this problem?
S: 00:34:00 Oh, the problem is the h, what's the h of the, in this question. Actually, mmm I think if the person is standing at this point, so, actually the h is the third height. So the change of the h, the height, and the um, if the person is standing from here, and then the third[?] h should be [writing] thirty meters. But if the person is lying here, [laughs] so, it shouldn't be thirty meters. It should be thirty meters minus one meters.
I: Okay so when you draw the, when you drew this, um, what is this height?
S: 00:34:54 It's the the, the center of mass of the person
I: Okay
S: 00:34:58 I regard this as the rectangle. The person. So it's the center of mass of the person, I regard is just in the middle the person. The person is two meters height and this is one meter.
I: Okay. So then, in this one it matters if they're standing up
S: Yes
I: And you said, or laying, lying down
S: 00:35:24 So in this question I think he he must standing at there
I: Okay
S: 00:35:29 But I am wondering in this problem, this question. But this little problem.
I: Kay was there anything else you thought about?
S: Mmm. 00:35:45 At first I think the two point five meters of the Jello and I think maybe that it has some connection with the two point five and then maybe the person can get into the Jello but after I read this question here, then I know that there's no relationship with the two point five meters. In this question.
I: Okay. Anything else that you thought about in this question?
S: 00:36:14 Mmmm. [pause] No. 00:36:21
I: [laughs] Okay. If you were…
S: 00:36:23 I think uh, I think that I have done a similar question before.
I: A similar question before?
S: 00:36:29 Harder than this.
I: Harder than this?
S: [laughs] Yeah.
I: How - , What was different about that question?
S: 00:36:33 This one is, I remember seeing in one three zero one the example in the, the lecture, and ah, he asked us a similar question. But difference is the person has a little initial velocity and the most, many of us don't know why it has that initial velocity, and actually he, he jumped off of here and then at this point [?inaudible] it has an initial velocity. So it's a little harder than this.
I: So, so how would you do that one?
S: 00:37:12 Actually, the same but, I just need to add [write?] [writing] kinetic energy here. Cuz there is both kinetic energy and the gravity potential energy. 00:37:28 [pause]
I: 00:37:38 Okay, and then if you were going to hand this in to be graded, what would you hand in?
S: 00:37:43 Oh, not this.

I: Not this? [laughs]
S: [laughs] No. 00:37:47 Zero points. Uh, first I would draw graph and the clearer, make every quantity clear. And then I always said according to the energy conservation and then said write this, and then this, this, and this. And then just, some more sentence make these things [sense?] And very, very careful about this question. And that's that's in an exam.
I: Okay. So do you usually solve it first, and then write it again for the grader…?
S: Um,
I: Or what do you usually do?
S: 00:38:33 I, in a quiz I will always come up with ideas in my mind. And I don't solve the right number, but, after reading this question I don't think, I won't write, I just draw a picture on the other paper, and then I will think how can, can solve the question. And the, and then I will come up with an idea in my mind and I will write it on the paper. I won't write it directly on the paper.
I: 00:39:05 Okay, so you keep it in your head first.
S: 00:39:08 Yeah. I must know if I can, whether I can solve it, or not. Before I write.
I: Okay, so how do you know if you can solve it?
S: 00:39:18 Like in this question, actually I read this problem, I know I can solve it.
I: [laughs] Okay, how did you know?
S: 00:39:27 In one, uh, one situation is oh, I've done this before so I can definitely do it. I can solve it. The other one, after I read it, like I said before, I find what I want to know and I find what I already know and I think oh whether is there any relationships with the, and I think the theorem like, the Newton's law, the energy, the momentum. Like that. And then – oh. I got, I got oh I can find the relationship. Then I know I can solve it. Usually it works, but in some hard problems it will take some time. Maybe I will write some things. [pause] That's it. 00:40:26
I: That's it? [laughs] Okay. Um. I think that's all we're going to do today. So did you have anything else that you wanted to say about, about when you're solving problems, what you think about, how you usually do them?
S: 00:40:43 Mmm, I think um, an important thing is you must understand what you are asked to do. Cuz sometimes we're always making mistakes when I, when we solve the question like, we just finished reading the question and oh, I think he is asking us to do, to find a v, the velocity. But anyway, actually it's wrong so we need to read the question carefully to find what are we asked to do.
I: Mmm hmm.
S: [inaudible]
I: 00:41:32 Did you take a physics class before one three zero one?
S: No, I, I think I…in my high school but not in college.
I: Okay, is this class different than the one you took in high school.
S: 00:41:48 High school? Uh, definitely.
I: Definitely?
S: 00:41:52 Very different. I'm from another country and a, different way of thinking. I think it's very interesting here.
I: Okay.
S: 00:41:59 Do lots of experiments, and that's very good.

I: So what do you think is interesting about here, or different?
S: 00:42:08 Oh well, actually, make some result of problem into the experiment. In lecture the, our professor will do three or four experiments. Like interesting experiment. And it will make, make it easier to understand. And it will, seems more interesting. Like in my high school I, we do the, we solve questions most of the time. 00:42:46
I: Most of the time?
S: Yeah. 00:42:50 So that's why I am good at solving questions.
I: So you've solved a lot of physics questions?
S: Yeah, before in my high school. [?inaudible]
I: Okay. Is there anything else you can think of?
S: 00:43:07 About the question?
I: Well, just about um, solving problems and what you do when you solve problems.
S: 00:43:15 Yeah, I must remember if you want to solve the question you must have something in your mind. Like you have some theorem, theorem in the book, you must understand all of them and you must remember um, understand, not simply understand, like you must know the Newton's second law. It's a combination of the force and the motion. And it's very important. I think that. Some question just, give you a force and ask you a velocity and some students just don't know what, what to do. So, you must know every theorem in the book. What's their physics meaning. 00:44:00
I: The physics meaning of the theorems?
S: Yeah.
I: Okay.
S: We learn that a lot. 00:44:07
I: Okay. Well, I don't have any more questions for you [laughs]
S: 00:44:14 Okay.
I: So, I'll shut this off. [camera beeps] And I have one more thing um, for you to sign so I can give you the the money. And I do have um, a copy of, of an example solution..
S: 00:44:34 Oh.
I: to give you too.

Problem-solving interview #6

Wednesday May 13, 2009, 9:30 – 10:30 a.m.
Tate Lab of Physics room 160

Summary of audio file: 00:01:20 receives first problem (nails), 00:02:20 starts writing, 00:16:14 completes first problem, 00:24:34 receives second problem (bungee), 00:25:42 starts writing, 00:33:36 finishes second problem, 00:44:11 done with questioning

I: 00:00:06 Okay. Did you have any questions
S: Nope
I: …about that [*consent form*]? Okay. So I'm gonna give you a problem to solve and then you can use the paper, use as much paper as you need
S: Okay
I: and the marker. And you might need to write a little bit bigger than you usually do so I can see it on the camera. Um. But just solve it like you would an exam problem.
S: Okay.
I: And it's okay if you don't [laughs] know what to do. I'm more interested in you know, what, what you're thinking about while you're solving the problem. So if you're comfortable talking out loud while you do that, you can do that. If you're not comfortable doing that I'll just ask you some questions at the end.
S: Okay
I: to explain your solution to me.
S: Okay
I: Um, so do you have any questions about that?
S: Nope. And this is the equation sheet?
I: Yeah, that should look like the one…
S: Yeah
I: …from your class.
S: Yeah
I: And if something's not on there it probably means you don't need it [laughs] Um. Kay. I'll give you…the problem. [pause 5 sec] And I'll, you can just try to ignore me, I'll just be zooming in and out with the camera back [?] here. Just let me know when you, when you've solved it, to your satisfaction.
S: Mmm kay. 00:01:20
[00:02:20 starts writing]
[00:09:50 typing in calculator]
S: 00:16:14 [?] Mm kay.
I: Are you satisfied with your solution?
S: As satisfied as I'm going to get
I: Okay. Now I'm just gonna go back and ask you some questions about what you were thinking.
S: Alright.
I: when you were solving the problem. [pause, moves to seat at table] So when you first read the problem, what was the first thing that you were thinking about?

S: 00:16:36 The first thing I thought about was just that it mentioned that the string was most likely to break when the um, bag was at it's lowest point. All I can like [?] get a diagram of what that looks like to start. Um. To have some sort of basis to like, base off where I'm gonna go from there. So I kinda just, just kinda the get myself in the mindset of the problem, just kinda drew that even though that didn't prove to be, the most helpful diagram, um. Just something to get started.
I: 00:17:03 So this is what you drew, first?
S: Yes.
I: Okay. And then what did you think about next?
S: 00:17:09 Well then, I was thinking about um, equations for circular motion. And I knew that when, if you release the bag of nails at this point the velocity would be going this way, and that is equal to um, the radius times the angular velocity. And so then I knew that the velocity will need to be large enough so that it would go up the nine meters to reach the friend at the top of the building. So I just used some, uh, a position equation to set up, um, that. And then I used the equation for velocity 'cuz at the top it would not have to have any velocity. The minimum um, velocity would just have, it would, at the highest point it would have a velocity of zero. [?] Set those equal to, solve for um, omega and plug that in. And solve for time, plug in and solve for omega.
I: Okay. And then once you had omega, what did you…
S: Well then
I: do next.
S: 00:18:06 I found the acceleration. Um. Using the radius and omega. And then I um, found the uh, centripetal force.
I: Okay, so how did you find centripetal force?
S: 00:18:23 Um, well I knew that the force of gravity on the bag of nails is fourteen Newtons, so if you divide by the acceleration of gravity you get a mass for that. And then I did mass times the acceleration right here. Solve for that. [pause]
I: Okay, and then what did you do?
S: 00:18:44 Then I just compared the, those two forces to the force of tension which was given as five hundred Newtons. As long as they weren't greater than that the string wouldn't break.
I: Okay, so your final answer…
S: 00:18:58 Is the string will be fine.
I: It'll be fine?
S: Yeah.
I: Okay. So, were there other things that you were thinking about, you know, that you didn't write down?
S: 00:19:08 Um….aside from just some like, trying to sort which equ, or do [?] the equations and [?] not really. I pretty much knew I had to work with circular motion and then some position. Equations.
I: Okay. So, how did you decide on, to use the position equations?
S: 00:19:29 Um, well this one was just standard so I figured [laughs] it was kind-of just the standard position equation, so I just started with that to see where I ended up. And then once I got here and had two variables, I knew I had to use another equation that had, um those variables in them so that I could solve for one of them and then get down to just one variable.
I: Okay. Was there something in the problem that made you think of the position?
S: 00:19:54 Um, well, I don't know. When you're thinking of like, it's almost like projectile motion. Throwing the, er swinging the bag of strings around and then trying to get it up to a

certain height. So, for all our projectile motion problems we've always used position equations so just, it seemed like it'd be something I'd have to work in there.
I: So what made you think of forces?
S: 00:20:17 Um, well, when it gives you a, a tension, which is a force, I figured you had to um, find the opposing forces to find just whether or not that tension would, would hold. Y-you know in fact you know the tension is given in Newtons, just clues you in right away that, to use forces. 00:20:36
I: Okay, and then as you were solving it, was there, um, anything you thought about to check your answer?
S: 00:20:53 Um…to me it just kind of seemed reasonable that if the string is able to hold five hundred Newtons, and it's not that heavy of an object you're swinging around, so the force of gravity isn't going to be that great in comparison, it, you'd have to be swinging it extremely fast for the, to exceed the force of tension. So, it seemed reasonable to me that it wouldn't break, so. Just kinda, it's not that I had a specific way to go back and check, but it just didn't feel unreasonable. 00:21:26
I: Mmm kay. Um, now if you were uh, handing this in for an exam, what would you hand in to be graded?
S: 00:21:38 Um, I would start with two diagrams at the top kind-of. Showing all of this basic information. And then I would probably um, list like, you know, uh, you know, trying to solve, I would kind-of explain maybe in a phrase or something what each of these different sections were doing, and I'd kind-of have [?put] them in a logical order as opposed to here where they're, it's a little bit um, jumping all over the page but, just explain you know, here I'm trying to find the um, time it takes for the bag to reach the, uh, top of the building. And then here I'm uh, using…my equation for velocity to find the angular velocity. And using angular velocity to find the acceleration and I'm using that to find force.
I: Okay, so you would write more words?
S: 00:22:29 I would just, I like to put like phrases at the top of each section as to just basically what that section is solving for. Just so that it's clear.
I: Mmm kay. And do you, do you usually include a picture?
S: 00:22:45 Ah, yeah. I always include a picture.
I: Alright. Um, and how, how do you decide what to include on a picture?
S: 00:22:54 I pretty much include all the information that's given in the problem. Cuz I start with the picture and so I haven't always figured out exactly everything that's going to be important. So I figure it doesn't hurt to have a piece of information that I don't need. But, I'd rather have it there when I'm going back and need it later than not have it.
I: Mmm kay. Did this problem look familiar at all?
S: 00:23:18 Um, it looked like maybe like, parts of two problems kind-of combined together. But, uh, not, I haven't seen a problem exactly of this type before. It's like you know, I've seen projectile motion or throwing things. But, I haven't necessarily seen the, when it's being released from, um, circular motion. 00:23:40
I: Okay. [pause] Let's see. We, we probably have time to do another problem.
S: Okay.
I: 00:23:50 If you're up to doing another problem. [laughs] Unless there's anything else you want to say about, you know, what you were thinking…
S: No, I think [laughs] think I'm good. 00:23:59
[00:24:17 inaudible? Give you a clean sheet of paper? Use this Inaudible]
I: 00:24:22 Alright. Some kind-of thing. If you wanna say anything while you're working on the problem, uh, if you're comfortable, that's fine. Otherwise I'll ask you questions again maybe [?], afterwards.
S: Mmm kay. 00:24:34

[00:25:42 hear writing]
[00:32:04 typing in calculator]
S: 00:33:36 Alright.
I: Okay. Are you satisfied with your…
S: Yeah.
I: your answer. Alright. I think the same thing as before, I'll just kind-of ask you um, what you were thinking, and it's okay if you repeat information. [laughs]
S: Okay.
I: 00:33:54 So when you read the problem what was the first thing that you thought about?
S: 00:33:57 Well, this was like a problem we did in class.
I: Okay.
S: 00:34:01 So, that was the first thing I thought about, I'm like how did we set this up in class. And so I tried to remember exactly how we set it up in class, 'cuz I figured it would work again. The same set-up, so I knew that um, we'd have to use conservation of energy here. And so I drew the energy at three points and knew I'd only need to use two of them to set it up. But, once I got to the equation. But I just, to get all my information out there I needed to have three points. So I just have before he jumped, before the cord starts to stretch, and then when he comes to a rest just before hitting the pool of Jello. 00:34:38
I: Okay. And then [pause, closing door?] what did you think about next?
S: 00:34:50 Well then, I knew that I would need the energy at this point because that's when I'm going to get the spring constant [?jump] of potential energy of this spring of the bungee cord here. So I knew that I had to use this point when I set, went into conservation of energy. And then [that?] this point was just easier to pick because it was only potential energy and here we have kinetic energy and potential energy and, it just wouldn't have been as convenient as that. So, I decided that I would look, compare the energy here and here. So, I wrote out what the energies would be in those situations. And then, proceeded to…[?]
I: 00:35:32 Okay, so when you have m-g-h here, what did you use for h?
S: 00:35:38 I used thirty-one because I took his, I took, I measured at the, from the middle of his body every time because [once?] body flips at some point during the jump it'd be, that's just convenient so you don't have to worry about when it flips. And um, so, it's the thirty meters up for the uh, platform and then one meter up to go halfway up his body. So thirty-one…is the height and then his mass is seventy [?]
I: Okay, and then in the, in the other energy part, what did, what did you use?
S: 00:36:14 Well, I, you're solving for k so I didn't know that, but we knew that the uh, displacement before, it was, it ah stretched out to twenty eight meters because that makes the thirty plus the two meters of his height so he doesn't hit the pool. And it was a sixteen meter um, bungee cord so we knew that it's, that had to be it's x-value there. And then I used one for the height here because again, I was measuring from the middle of his body still. So he did have the potential energy for that one meter that it came up, so. Had that there. And essentially what happens is when you subtract that out it's like you just used the thirty for the height because it subtracts that on the other side. To be, essentially seventy times nine point eight times thirty one minus one, so.
I: 00:37:09 Okay, and then you solved for k?
S: Yes.
I: Okay, and how did you decide, um, the units for k?
S: 00:37:20 I went through and figured out what the units were on each of these items. Um. I probably, it would have been better if I had written those down all along, but. I got to this stuff and just went through and figured out what the units were at each step of the way. 'Cuz here you have Newton-meters over meters squared. So. The meters cancelled out. 00:37:43
I: Okay. So was this problem different from the one you did before?

S: 00:37:54 Um, it was a very similar set-up to the one we did in class. And like, uh, that's how, that's, I distinctly remember measuring from the middle of the body. That's like, the main thing I remember from when we set up the problem like this in class. But it was something to the effect of like, something, the bungee cord, someone jumping off of something or falling off of something. So. Um. I remembered that would be the, that was like the main [?thing I took away from the] problem in class. It was very similar to what we did in class.
I: 00:38:27 Mmm kay. Um, so if you were handing this in to be graded for an exam, um, would you hand in this, or something different?
S: 00:38:40 I would basically hand in this. I would probably include units at every step of the way maybe, along here. Just so that I would be, you know, make sure that that's that and that it's clear where, what every variable is. And maybe just, you know, um…maybe even include like a sentence of explanation or something. With some of this where it might not be clear as to like where I pulled some of these numbers from. But. Basically…
I: Can you say more about what you would..
S: Um
I: what you would say?
S: 00:39:12 Um, I might write a sentence er like, just [?asterisk] or something like, you know. X is equal to the maximum, um, stretch of the bungee cord minus it's you know, initial length. And that um, all energies are being taken uh, from the middle of the dean's body. Just so that's clear. I think that's probably what I would add.
I: Mmm kay. And, um. Did you use the equation sheet at all?
S: 00:39:51 Uh, to get the…potential energy for a spring. I thought it was that, but I just thought I would confirm. But I knew the regular potential energy.
I: Okay. For the, for the previous problem, did you use the equation sheet?
S: 00:40:09 Ah yes, to get, to confirm um, velocity and uh, acceleration. The, how they related to angular…um velocity.
I: Mmm kay. Um, so was this um, typical of what you usually do?
S: 00:40:29 Yeah, I would say it was.
I: Mmm kay. Um, and then is there anything else that you thought about…
S: 00:40:39 Um, I briefly considered using forces and then immediately decided against it.
I: So why did you decide against it?
S: 00:40:46 Because energy seemed so much easier, to use conservation of energy than trying to figure out um, the force, um, of the bungee cord and gravity and…um, just conservation of energy seemed like it would be a much cleaner way to go [?].
I: Okay, is there something that made you think it would be easier, or cleaner?
S: 00:41:09 Um, probably just the fact that when we did a problem like this in class it was when we were in the conservation of energy unit, so it's like, what I associate it with [laughs]. I really, I mean, I, if I had sat down to really think about it I'm sure I could have found a way that would have worked with forces, too. But. This just, popped into my head as, seemed like it would be more straightforward to me, so. 00:41:30
I: 00:41:36 Alright. Is there anything else that you usually do when you're solving problems? Or you usually think about?
S: 00:41:44 Nah, it's pretty much, start with the diagrams and then just kinda, see where they take me. [laughs]
I: Okay so, um. You know, after you get the diagrams, that next step of deciding what to do next um, how do you, how do you usually decide what to do next?
S: 00:42:02 Um, well usually when you look at the problem you have a feel for, it's either you know, there's either one or two concepts that you can kind-of sense that they're trying to get at. So then here I could tell you know, they were you-, well here I looked at, I knew you had to find something that had a spring constant. So it's either you're gonna use the force of the spring or

the energy of the spring. So I just, and I, when I decided that energy I knew I would have to use conservation of energy. Um, usually though I kind-of just look at the problem and see what kind of key concepts they seem to be getting at, and I think about what equations would go along well with that.
I: So what do you think of as key concepts?
S: 00:42:38 Um, well, you'd be looking at you know, energy and energy conservation. Some have that[?], have momentum in there, um. Forces, um. Equations of, of motion. Uh, circular motion or springs. Just, kind-of the general like, almost units that we covered in class. 00:42:58
I: Okay. And then once you've decided on something, then you start looking at the equations?
S: 00:43:09 Yeah. And usually I'll start…I'll, in my head I'll try to like, think about, about, to go about halfway through it but still looking like this is going to work. Not necessarily solving it but just seeing okay, do I have enough information to make this work. And if I find myself being able to get that far, I usually start writing the stuff down, and going with it 'cuz I figure that I'll be able to solve it then at that point.
I: So you do some things in your head first?
S: 00:43:33 I like to just, think to make sure, do I have enough information to use these equations, or am I picking equations that I don't have enough information to use, and should look at other equations maybe.
I: Mmm kay. So how do you decide if you have enough information or not?
S: 00:43:48 Well I just look at the, basically I look at the variables in the equations and see if I can pull them off my diagrams, or, or manipulate the information in the diagrams to get the information.
I: Mmm kay. Is there anything else that you want to say about how you usually…solve problems in physics?
S: 00:44:11 I think I've basically covered it.
I: Okay. Thank you for [laughs] volunteering. I have one more thing for you to sign to verify the payment. [pause] And I do have a copy of, um, a solution to each problem that I can show you, too.

Problem-solving interview #7

Friday May 15, 2009, 9:30 – 10:30 a.m.
Tate Lab of Physics room 160

Summary of audio file: 01:53 starts first problem, 09:26 says stuck, 21:35 says stumped again, 23:20 talk about process (what thought about first), 38:32-44:05 working on revising calculations (find v to go up 9 meters), 53:02 if solving for an exam, what would hand in; 59:33 have you seen a problem like this before, 01:05:09 interviewer says out of time

I: 00:00:06 Kay. Did you have any questions about the consent form?
S: Nope [laughs]. Pretty straightforward.
I: Okay. Um, and you brought a calculator with you?
S: I did.
I: [laughs] it looks pretty similar to the one…
S: [inaudible]
I: Okay.
S: Exactly, yeah. Yep…Tell you what, these calculators are kinda backwards, but I'm so used to using it now.
[camera beeps] 00:00:27 It was the only thing I could use in high school.
I: Oh.
S: So.
I: [inaudible, non-graphing, pretty um, pretty basic]
S: Yeah
I: Kay well, um, I have a problem for you to solve.
S: Alright.
I: I guess. And you can use the marker and the paper there, and you might need to write a little bit bigger than usual
S: okay
I: so I can see it on the uh, video camera. Just solve it like you would an exam problem, or you know, show as much work as you would if it was [inaudible] kind-of like an exam. And you know, feel free to take as much time as you want, use as much paper as you want, and then, if you're comfortable talking out loud while you're working on it you can do that, otherwise we'll go back at the end and I can ask you some questions about what you were thinking
S: Okay
I: so just do whatever's most comfortable for you, um. And then…yeah and then it doesn't matter, you know, if you don't know how to solve it. That's okay too that gives me information
S: Alright
I: about what, what you're thinking about in the problem, so. Um, you can either let me know when you're, you're satisfied with your solution or when you just
S: Completely stumped?
I: wanna be done with it [laughs]

S: Alright.
I: Yeah, so, here you go. 00:01:44
I: And just try to ignore me. I'll be back here zooming in and out with the camera. 00:01:53

[when start writing? 00:05:38 or earlier?]
S: 00:09:26 So, I'll just explain where I am right now, what I'm getting stuck on. So I know that it has to, it's rotating around a radius of uh, point six five meters. And um, that the maximum tension it can have is five hundred Newtons. What I'm a little confused on is, um, there's no rotational speed given. And it says that there's no um, that you're not adding any force so it seems like gravity's the only thing but that doesn't quite make sense [laughs]. So that's where I am. 00:10:11
I: 00:10:31 Have you decided what the question, or what the problem's asking?
S: It seems to be asking what the tension is….yeah. And whether that's greater than the design tension. 00:10:54
[hear writing]
S: 00:21:35 I'm kind-of stumped [laughs] It's the exact same problem I had trouble on on the final.
I: Oh, you had a problem like this on your final?
S: Pretty similar. Not exactly the same, but the same sort of idea.
I: Okay, can you say more about what you're stumped on?
S: 00:21:53 Uh…well pretty much the same thing as before. There's no, there's no time and you don't have a speed. So I'm kinda confused on um, well I'm also confused on what forces are acting exactly. Since there isn't…there isn't any added energy from the person spinning it. But since I don't know, well I know, I *think* I know how fast it's going, um, but I don't know any sort of time or, yeah. So that's where I'm confused.
I: If you did have the time, what would you use that for?
S: 00:22:45 Um, let's see…If I knew a time I could find a distance and then I could find acceleration. And then I could find um, a force using Newton's law. 00:23:04 That- That's the approach I was taking and then it just kind-of ended [laughs]. Um. Yeah. 00:23:19
I: Well if you want we can talk about it a little bit and then if you think of something, you know, you can
S: Sure
I: go with it. Um. So when you first read through this problem, what did you think about first?
S: 00:23:36 Um, well first I thought there was gonna be some sort of like parabolic motion or something from here. Um. But then it didn't, it didn't give me enough information to get anything from it. And so finally I got down to this string part. Which I'm still, it's a little odd that, they give you nine meters and then say you can't throw it up that high, it's kind-of like great[?], Not quite sure if I should be getting something from that but I don't think so. Um. Yeah, right away it's circular motion. Um. Yeah.
I: So what made you think of parabolic motion [?]?
S: 00:24:32 Well, in the first part they're throwing something up vertically. I guess it's not really parabolic but. Have something go up, and then coming back down. So it's

not, I thought it would be something like that where you have to calculate gravity and, if it actually makes it or not. Um, but that didn't play in at all. [laughs] [pause 10 seconds]
I: So that's what you were thinking about, what, what did you write down first when you were reading through the problem?
S: 00:25:20 Um. I wrote that down first. Cuz I knew for sure you'd have to find a mass from this Newtons. Um. And then, that second because I thought you'd have to, there'd be something with the building but it didn't actually have anything to do with it.
I: So how did you calculate the mass?
S: 00:25:49 Um, Newton's law. Force equals mass times acceleration. So. Pretty sure this has the acceleration already in it. And so then you just divide out gravity [murmurs, inaudible] Yeah, I think.
I: Okay. So you took the fourteen Newtons and divided it by that
S: Yeah
I: nine point eight?
S: Yep….00:26:15 So then I had a mass, um….Yeah, then I got stuck on the circular motion. [laughs] Um. Yeah, so then I just tried to look for equations that would help me solve the problem.
I: Mmm kay.
S: 00:26:43 First like, trying to balance the forces. So you know it has to be less than the force of tension. And there's a force of gravity, and there's some sort of a…maybe angular momentum, or angular….[?well you do need a] angular acceleration so you can find force. But I didn't know what to put there.
I: So you think that there's something more in this force equation?
S: 00:27:23 Yeah. Um…Well cuz it seems like when you swing something, it, there's, it has more energy than if you just like dropped it. But, yeah.
I: So what makes you think that it has more energy? 00:27:53
S: 00:28:02 Um, I don't know that's just a guess I made at this point. Now that I think about it, um. 00:28:06
S: 00:28:17 I always like to think of amusement park rides. [laughs] Like if you're on one of those big ship [?rides, ones?] and you're on the very bottom there's a lot of force. But I guess if you actually just fell and stopped and like…[?] be more or equal.
I: So when you say at the bottom there's a lot of force, what, what is that or what are you thinking about?
S: Um, [?pushed] Um, but in the actual component forces, [?] yeah. That-that's the part I got stuck on. Cuz I wasn't sure if all that force is from gravity and it's just like, being directed or whatever. Yeah, I don't know.
I: So how did you decide um, these equations that you wrote down? 00:29:24
S: Um. [pause] 00:29:36 Well I wanted to find accelerat-, angular acceleration eventually so I could find the force that is um, that would be directed inwards, but, cuz that's related to the tension. Um, then I only knew the mass and the radius so I had to find equations that didn't have time. [murmurs, inaudible] And I knew [?N]…And I tried to do something with potential energy and kinetic energy to see, but then, I don't know.
I: Okay so you…you have uh, PE=KE. So can you explain what that is?

S: 00:30:51 Well, at the top you have a lot of gravitational potential energy and at the bottom you have a lot of kinetic energy from the gravitational energy. So I guess in this instance I was thinking all the force just coming from gravity, um. And so.
I: When you say at the top, where are you talking about?
S: 00:31:16 At the top of the swing, with the rope.
I: Mmm hmm.
S: 00:31:20 Um. And so I think from that I can get the velocity since the potential energy is pretty straightforward and I knew the mass. So then I was able to find velocity, and I could find um,…what's it called. Angular velocity from the linear velocity. Um, that's where I pretty much got stuck. 00:31:50
I: Okay. Did you have a reason why you were calculating angular velocity?
S: 00:32:01 Um, hoping it would lead me to something. [laughs] And maybe be able to find angular acceleration. But that didn't really pan out.
I: Okay, so if you wanted to find angular acceleration, what, what would you need?
S: 00:32:26 Um, let's see. Well, this equation, you need a distance. Um. Or if I just used this one here. But just…[?]
I: Which one?
S: This one. [? Inaudible, don't have ___ or any time] 00:32:53
[pause]
S: 00:33:49 I dunno. Circular motion always confuses me. [laughs] It seems like it should be really straightforward and then I get lost.
I: So what confuses you about it?
S: 00:34:08 The, like, I don't know. It just, it seems like it should be like normal motion like, linear motion. And then there's, you have different symbols and everything's a little bit different. And especially um, like where different vectors are pointing really confuses me.
I: Vectors for what?
S: 00:34:35 So like, acceleration and um, velocity makes sense but acceleration is really, like how it points inwards. 00:34:45
S: 00:35:24 If I assume…[writing] only, I'm just curious to see what this number is. [laughs] If you assume gravity is the only thing acting on it. [typing in calculator] Yeah. [inaudible, got fourteen Newtons again] Yeah it's, it's too small. 00:36:09
I: You're saying it's too small…?
S: 00:36:23 Yeah, like, in relationship tension seems, like usually they like make, in the problems they make it pretty borderline, so.
I: So you're saying the fourteen Newtons is too small, compared to the five hundred…?
S: To the five hundred, yeah. 00:36:44
S: 00:37:47 Well and here's the other thing too, is it's, you think that if you release the bag of nails when the string is horizontal to the ground that it will reach a co-worker. Um. If you knew that, then you can calculate how, exactly how fast it's going.
I: If you knew…what?
S: 00:38:07 If you knew that if you release it when it was horizontal that it would go nine meters vertically, then you could calculate. Um. How fast it would have to go to get that high.

S: 00:38:27 [laughs] The problem's a little, I don't know, it…that you're supposed to put that in or not. 00:38:32 So.

[pause, writing]

S: 00:44:05 [laughs] I'm still stuck on how to get the, how [?] to use for the velocity.

I: What did you calculate here?

S: 00:44:15 So here I assumed that you could in fact, um, get it going spinning fast enough that it would go up vertically nine meters. Um. And so…I think this is the equation that you'd use to calculate that. Um. So then I got another velocity which is different than this velocity. Um, that…

I: Mmm kay, so you said, you were talking about how fast you need to spin it?

S: Mmm hmm.

I: Okay, so is that what you mean by this velocity? What is this velocity?

S: 00:45:05 Yeah, well this is like, you release it and then it's going up, so. Like that…[writing, murmuring inaudibly] and then it goes up vertically nine meters. How fast it would have to go to go that far.

I: Okay, so you're talking about velocity…going upward?

S: Yeah. 00:44:35 [pause]

S: 00:46:02 [?have to be] some of the equations are related. This angular velocity to the force, but I don't, I can't find the connection. 00:46:14

I: 00:47:24 What, what did you write at the top of the paper?

S: 00:47:27 That was just how long it would take to travel nine meters.

I: Mmm kay, how did you calculate that?

S: 00:47:36 Um, I just…multiplied to cancel meters. Well actually, divided. Um. [?see is that, inaudible] 00:48:16 But um, that's still like, that's how long it takes to get nine meters but I don't really know how far it's traveling here[?], I-I'm confused. I don't know how to relate that back to what we're trying to find.

I: Kay, so what are you trying to find?

S: 00:48:40 Um, [laughs] tension eventually. But I don't…yeah. I don't know how to get tension from the…I was just hoping to like, stumble into it [laughs] and it hasn't worked.

I: Did you have any equations that have tension in them?

S: 00:49:12 Well, like this original had tension in it. But then I got stuck. And I know that um, let's see. 00:49:21

S: 00:49:32 What I don't, I just need to know that the force of tension, I know the force of tension is five hundred maximum. Um. So I just need to know that the force um, outwards is either less than or greater than that. So I guess that's what I was trying to find.

I: Kay, when you say the force outward, what does that mean?

S: 00:50:07 Um, I'm not completely sure [laughs] that's why I'm having trouble finding it. I guess [?inaudible] Um. Yeah.

I: So is this a force different than gravity? 00:50:21

S: 00:50:26 I think so. Um, 'cuz….well it's some of the gravity redirected, I-I don't know. [pause 17 seconds]

S: 00:50:55 But yeah, that's the part that confuses me. If it's already, you added some motion to get it spinning vertically, so it seemed like there's some, some energy in there

that's not just gravity. Um. Although it says you're not adding any energy right now. But it seemed like there would be some sort of like, um, angular momentum or something that would be in addition to the force provided by gravity…But the fact that it's not mentioned in the problem makes me think I'm probably wrong [laughs] 00:51:44
I: 00:52:09 So when you drew this picture of the forces…
S: Mmm hmm.
I: what is this arrow there?
S: 00:52:18 Um, that's the direction I think force of tension is working so it's, resisting the, it's keeping the object from flying off in that direction.
I: Okay. So this is the direction of the, this force?
S: Yeah.
I: Okay. Is this the string?
S: Yeah
I: This is the string and this is…okay. 00:52:40
I: 00:53:02 So if you were solving this on an exam, um, what would you hand in?
S: 00:53:10 Um, well usually it ends up looking something like this.[laughs] Um, let's see. I'd. Yeah. I'd, um. [pause 17 seconds] Probably stop here and just, um. Yeah I don't know how to get any closer to the answer I don't…I might try to make it easier for people to understand my thinking, but that's probably it.
I: Okay, how would you make it easier for the grader to understand the solution?
S: 00:54:08 Um…like, just find the mass from Newtons here and then kind-of box it off. [writing] Mass and then, um. [laughs, pause]
S: 00:54:41 Like, yeah, basically just explain by writing it out, kind-of like, separating stuff so that they can see that one thing isn't…yeah. [laughs] Pretty much.
I: Do you usually write the words down while you're working on the solution, or do you wait until the end?
S: 00:55:04 I wait until the end.
I: You wait until the end?
S: [laughs] 00:55:07 Which leads to a lot of mess, but. Um, yeah. Cuz a lot of times I won't know…basically I'll either know right away how I'm gonna get the answer or I won't really know and I'll just kind-of see what I can find, and if that gets me close to the answer. Um. So I don't write stuff down cuz I don't always know where I'm going. [laughs]
I: So you said you can tell right away?
S: Yeah.
I: So how, how do you tell whether or not you're gonna be able to do it?
S: 00:55:45 Well, I'll be…usually within the first two or three minutes I can see if I, I'll know if I have a logical set of steps I can go through that I know right now. And it's just working my way through. Or, if I don't, if there's some-like here, I didn't know how to get from, I knew, I didn't know quite a few things. And since I didn't know I couldn't draw any conclusions. So, I can see I don't know enough information and I'm missing something. So then I can't, I don't have like a logical pathway to follow. 00:56:22
I: 00:56:30 So when you said you try and see if you some steps you can
S: Mmm hmm

I: …go through, do you usually keep that in your head, or do you write some things down?
S: 00:56:40 In my head.
I: In your head?
S: 00:56:43 Yeah, it's…usually the only thing I write down right away is a picture, so I can see what's going on. Um. But then I'll just have in my head like, if I go from this equation and then I get an answer I can put it into this equation, and then into that equation. And then, yeah.
I: So how do you come up with those steps? 00:57:07
S: Um. [pause 9 sec] 00:57:24 I guess, like if you see a linkage between the, the physical things. So like…if it's something falling you know the potential energy is gonna change into kinetic energy so then you have, you know that there's a, that the answer you get from once you solve for kinetic energy can be used to find potential energy. Something like that. 00:57:47
I: So after you, you said you usually draw a picture?
S: Mmm hmm.
I: at first. So what do you use that for?
S: 00:58:06 Um, primarily just to make sure that I understand what's going on and, it also allows me to put everything that I know from the problem out. So like in this instance, when I drew this and then like, I read a little further and then I saw that it was spinning, that's not happening [?where] this is happening. And so then I can, I have a clear indication of what the problem's saying. And also the person who's grading it knows what I'm thinking. [laughs] Um. But yeah, primarily it lets me put all the information down where I can see it all at the same time. And where I can see if there's anything [?standing out]
I: Okay, so after you have that picture, is that when you start thinking about the links?
S: 00:59:02 Um, yeah. If I haven't seen anything yet, like haven't right away seen how [I'm gonna have to go from here to here?] then I'll look at the picture and try to figure it out.
I: Okay, so sometimes you think about that right away?
S: 00:59:15 Yeah. Or while I'm drawing the picture, at the same time. 00:59:18
I: 00:59:33 So have you seen a problem like this before in your class?
S: 00:59:36 Yeah. [laughs] No like I said it was on the final. Um, and there were problems like it in the book. And I hated them. [laughs] So I got this problem and I'm like, Oh no. Um, yeah. I don't, obviously don't know how to solve it.
I: Was there things about this problem that look different?
S: 01:00:02 Um…not really, it's pretty much, I mean there's certain things that are different like, the having it in Newtons was different. Um. But it's just a conversion thing. Um. Yeah. Pretty much.
I: Okay, so what about it is, have you seen before, or what looks similar?
S: 01:00:30 Well, basically you're trying to find the force to keep it in a circle. And so there's…I mean there's different like, ways the problem is applied but like, the classic one that, the one I remember from the book, but I don't think I did very many of these which I should have but um, it was like you have a motorcycle doing a loop. And so it's basically the same thing. You have to have a certain amount of force to keep it on the

circle. And this is just, you have to have a certain, you know you have a certain amount of force and you can't exceed that. So that's not quite the same thing but it's a similar type of problem.
I: Mm Kay so, um. Did having a previous problem like that help you at, at all when you were working on this problem?
S: 01:01:28 Not in this instance. Cuz I didn't know how to solve this problem, so it didn't help me here. Um, but a lot of times yeah it will.
I: Okay, so how does the previous problem help you?
S: 01:01:42 Well you just um, you know the steps to go through. Um. And you know what things you need to actually solve for, whereas like here I was confused. I didn't really know what I needed to get to the answer, so. I just solved for a bunch of stuff [laughs] Um, where, if you've seen the problem before you know, oh I need these three things and then I can find this and calculate the answer.
I: Okay, so when you said you need these three things, or, what are you talking about?
S: 01:02:18 Um, let's see….Um, well like I guess if you have…if you have like any sort of problem with a ramp and then like, need to know the speed at the bottom you know it's related to the height so you need to find the height. And you know you can, use certain like, this equation to get it. The potential height.
I: The m-g-h?
S: 01:02:56 Yeah, like that. And so you know whenever there's a ramp and you have something going down it, this is gonna be somehow related. So I guess that's how like, that would be one of the three things you know you have to find. To find, whatever. It depends on the problem, too.
I: Okay, so when you see a ramp in a problem, you immediately think about the height,
S: Mmm hmm.
I: and this velocity?
S: Yeah.
I: Okay, so when you saw this circular, um, swinging problem, what did it make you think of?
S: 01:03:41 Um. Oh shoot I didn't know how to solve this on the test! Um, yeah, and then I had the, well I knew from the test I had an equation. On the test it was a multiple choice answer. And I had an equation that I used and I got a answer that wasn't one of the possible answers so that confused me. Um. And I saw that the equation wasn't on this sheet. So that's kind-of like, it kind-of broke down there. So that, the connection didn't work. 01:04:12
I: 01:04:24 When you said something about that motorcycle problem
S: Mmm hmm.
I: …and, like you might know the steps from the previous problem. What do you mean by the steps?
S: 01:04:36 Um, well like in this instance, there's the speed you have to go in to make it around so it's kind-of the exact opposite of this problem. So if I actually knew how to solve this one too, then you just kind-of like work backwards, but. I got stuck. I don't know what to do. Didn't know how to do this one so don't know how to do this one either.
I: Okay well I think that we're about out of time. 01:05:09

S: [laughs]
I: Is there anything more you wanna say about um, how you usually solve problems or what's going on in your head while you're working on problems?
S: 01:05:20 Um…nothing I can think of. Generally I tend to do too much in my head and not write enough stuff down, that's the only, that seems to be where I go wrong.
I: What makes you say that?
S: 01:05:38 Um, well just like working with other people and, especially in physics we have a group and you solve group problems. And they'll write everything down and I'll try to figure it out in my head [camera beeps] and I won't get there. They seem to be a lot more successful just writing stuff down.
I: Mmm kay.
I: Mmm kay. Well, I think we'll call it a day. [laughs]
S: [inaudible] think the camera's off.
I: I have one more thing for you to sign.
S: Okay.
I: For the payment. And then I do have a solution I can, I can give you too.

Problem-solving interview #8

Tuesday May 19, 2009, 3:00-4:00 p.m.
Tate Lab of Physics room 160

Summary of audio file: 2:55 receives first problem (nails), 12:51 says done, 23:03 receives second problem (car cliff), 38:34 says done, 53:04 sign payment form & discuss solution

00:00:20 (hear beep, camera turning on?)
S: 00:00:44 Uh, what's the date?
I: Oh, I think it's…let's see, the nineteenth today? [laughs] It's going by pretty quick. [laughs]
S: Yeah. 00:00:57
I: 00:01:04: All right. Uh, do you have any questions about this? [*referring to consent form*]
S: Uh, no. So I just have to uh, solve the problems, or?
I: Yep, I'll give you a problem, um, and then you can just solve it on this sheet of paper using a marker. And then if, if you're comfortable talking about it while you're working on it you can do that. If not, we can just wait until, until you've um, finished and then I can ask, I'll ask you some questions about what you were thinking about while you were working on the problem.
S: Um,
I: And it's okay if you don't know how to solve it. Because that also helps me understand what you were thinking about [laughs]. Um, so. You might wanna write, you know, a little bit bigger so I can see it on the camera.
S: 00:01:55 I have pretty bad handwriting.
I: [laughs] That's okay. [laughs]
S: All right.
I: Um, so are you, are you ready…
S: Yeah
I: to start? Okay. And I have a couple calculators here. Um. If you brought one that you prefer to use you're welcome to use that too.
S: Oh, yeah.
I: [murmuring, you brought one? Okay.]
S: [?unclear? Once you use the…hard to use anything else]
I: [laughs] Yeah.
I: Mmm kay. And I'll, I'll be back here, you know, using the camera so you can just try and ignore me [laughs]
S: Okay.
I: [hear camera beep] Okay, are you ready to get started?
S: Yeah
I: Okay, here's a copy of the problem.
S: Alright. [00:02:55]

[00:04:27 hear writing]
S: 00:12:51 Done.
I: Done? [laughs]
S: Yeah
I: Okay. Are you satisfied with that answer, or do you wanna take any more time?
S: Um, I'm satisfied.
I: You're satisfied?
S: Yeah.
I: Okay. Well now, I'm just gonna go back, and have you um, tell me some more about what you were thinking while you were working on this.
S: 00:13:15 Well, basically um, I was thinking about uh, conservation of energy. Like, it says like, the guy predicts he will just have enough energy in the system to launch it at that height. So basically I want the gravitational potential energy to equal the initial kinetic energy. And then if I can solve for the velocity I know that at the lowest point, that the tension and the weight has to move in a centripetal…acceleration. So um, once I solve for that particular v I just plug it in to this and solve for the T that's required for that. If it's less than the maximum it um, it should be good.
I: Okay, so when you first read through the problem, um, what was the first thing you thought about?
S: 00:14:23 Well, the first thing I uh, thought about was um, was the uh, I-I just diagrammed it. I didn't know what to think initially. I just wrote down all the data, diagrammed it.
I: Okay, when you say 'diagrammed' can you tell me more of what you mean by that?
S: 00:14:48 Like, I just like, visualized it. Maybe the height had to be from the center of the, center of the thing. I wasn't quite so sure exactly what it was but when I drew a picture it made more sense to me.
I: Okay. So you wrote down uh, some, some of the values given in the problem?
S: 00:15:16 Yeah and I converted them to SI so calculations would be easier once I solve.
I: Okay.
S: 00:15:25 And basically I just solved it symbolically because um, if I plug in numbers I might get confused.
I: Okay, so you waited until…
S: 00:15:37 Until it was in its simplest form.
I: Okay.
S: 00:15:41 Because I could cancel out g and m.
I: Okay so what did you use for h? You used…
S: 00:15:51 height of the building from the ground.
I: From the ground? Okay.
S: Yep.
I: So how did you know to use conservation of energy?
S: 00:16:03 Uh, basically um. The only, the only force I know that's um, acting externally to the system would, would be gravity. And that's a conservative force. And um, basically there's no time dependence here, so. So then I knew that it had to have something to do with energy.

I: 00:16:40 [laughs] Okay. So you knew right away that's what you were, that's what you were thinking?
S: 00:16:48 Yeah, because it didn't have much of a time dependency and it, and it says that the guy thinks he can reach it. That means it should have enough energy to go [?up] there.
I: Okay….So when you have this, this velocity here, and here, um. What is that velocity?
S: 00:17:15 Uh, that's the um, the total velocity in the whirling bag of nails.
I: Okay, so the bag of nails…
S: 00:17:25 Is like going in a circle so um, there's a tangential velocity to its trajectory.
I: Okay. [pause 10 sec] So if you were solving this problem on an exam. What would you hand in to be graded?
S: 00:17:53 Oh, Basically I would give, give like a bigger explanation of things. Like, in exams I think I sometimes write cliff notes.
I: Cliff notes? Okay. What do you mean by cliff notes?
S: 00:18:07 Like I say okay, by conservation of energy so and so. Once we know that velocity then we know that by Newton's, Newton's second law that the sum of all forces is the mass times acceleration. And then, then I say plugging in the velocity and simplifying and all that.
I: Okay. Do you also include a picture…
S: 00:18:37 Yeah.
I: Like you've drawn here?
S: Yeah.
I: So, do you usually do that um, because it helps you, or is it something that's graded?
S: 00:18:53 It helps me because then um, I don't have to keep re-reading the problem to get the data. Just right there.
I: Okay
S: 00:19:01 And it's also all in SI units so, then at the end I don't make conversion errors.
I: Okay, you do the conversions right away?
S: Yeah.
I: [pause] Mmm kay. Um. Have you ever solved a problem like this one before?
S: 00:19:22 Um. Not this one in particular.
I: Kay, were there some things that were similar about this problem, to other problems?
S: 00:19:31 Uh, we did a lot of uh, problems like this in the group sessions.
I: Okay…How was it like this, I guess?
S: 00:19:45 More harder I think
I: Harder?
S: Yep
I: This one was harder, or the group problems were harder?
S: 00:19:50 The group problems.
I: The group problems?
S: Yep.
I: [pause] Did you use the equation sheet at all?
S: No.

I: No? Okay. Um. So is this, are these equations things that you had remembered?
S: 00:20:16 Yeah…I also knew my answer was right because the dimensions were correct. Like, the ratio of two lengths cancel out so then I know the answer is Newtons.
I: Okay.
S: 00:20:39 If that wouldn't happened I would have been very concerned. [laughs]
I: Okay, is that something you usually check?
S: Yeah.
I: Mmm kay. So you check it in your head?
S: Yeah.
I: Kay, is there anything else that you were thinking about while you were working on this problem that you didn't write down?
S: 00:21:04 Um…no, not really.
I: [pause 15 sec] Mmm kay, so you didn't write anything, any picture down about these forces. So how did you decide that this one was negative?
S: 00:21:34 Uh, because I just, in my head I just established a coordinate system pointing up here.
I: Okay, so you decided that um, in your coordinate system the T, the tension was positive?
S: 00:21:52 Uh, because it's at the lowest point it needs tension upwards to move it at a circle, but at the same time there is gravity pulling it down.
I: Okay, and then you said that was equal to the mass times v-squared over r?
S: 00:22:13 Yeah. Centripetal acceleration.
I: Okay…Okay, well I think um, unless there's anything else you wanna tell me about what you were thinking uh, we could do another problem.
S: Okay.
I: If you're interested.
S: [?]
I: Alright. I'll challenge you once more. [shuffling paper] 00:22:39
I: Mmm kay. Here's another question. 00:23:03
[00:26:02 hear writing]
[time spent on problem: 15:31]
S: 00:38:34 Okay
I: Okay.
S: Yep.
I: You're satisfied with your solution?
S: Uh, Yeah.
I: Yeah? Mmm kay. So now like, like last time I'll just ask you to go back and…tell me what you were thinking while you were working on this problem. So, so when you read it, what was the first thing that you thought about?
S: 00:38:59 Drawing the diagram.
I: Drawing the diagram?
S: 00:39:02 Because this looked, it sounded very confusing.
I: Okay, so that's the first thing that you did?
S: 00:39:09 Yeah. I just like, okay um, representated by ten degrees, but in the end I didn't even end up using it but, that's just good to have in there. Like, and then as I

drew it I summarized the information I drew, I [???point] And then um, I explain my ration-rationale right here. Assuming no frictional effects energy is conserved. But the thing is um, they never give us this h in the problem. So basically I-I had to assume, assume like um, a reasonable h for it.
I: Okay
S: 00:39:56 In order, in order for the crash to be plausible. That it's a accident. [laughs]
I: [laughs] Okay. So did you assign a, a value to this h then?
S: 00:40:11 No, just a symbol here that represents h. And then I solve for it
I: Okay
S: 00:40:17 The h that will be required, that if it goes with a velocity from the energy stored here, it will crash thirty feet under. Assuming no air resistance or friction.
I: Okay, after you had your picture drawn, what was the next thing that you did?
S: 00:40:38 Well the next thing I thought really really hard about it because [laughs] they never gave us this h. Or the length here, so I can't even use trigonometry to get it.
I: Okay, so when you were thinking about it…
S: 00:40:54 I'm like, maybe they want me to state my assumption on h itself. Because they never gave it. So I'm like, my assumption is h is bigger than that.
I: H is bigger than
S: At-
I: …point five six?
S: 00:41:12 It has to be bigger for the crash to be, for the crash to work. Cuz then all the energy will be stored here and it will fall.
I: Okay, so how, how did you do that? I guess, tell me what you did.
S: 00:41:30 So um, I applied conservation of energy between this point and that point. One and two. Then at two it goes with a velocity then, we know the velocity and there is no air resistance it becomes a kinematics problem.
I: Okay.
S: 00:41:51 So um, I know the kinematics for constant acceleration, um, then I just need have it fall, fall h prime amount or four hundred feet um, in-in a certain amount of time. And during that time it also has to go thirty feet. Or L-feet, which is L over here. So basically I used that information to solve for time. And once I know the time I know in that time it has to go that distance. So I just plugged in time, plugged in the v I got for here. And then once I had that I just rearrange and solve for h. And then I said, as long as h is big-greater than or equal to that, it's plausible. 00:42:50
I: Okay. So where did you get um, the kinematics equations from?
S: Uh
I: Like this one, and the y, and the x?
S: 00:43:03 Off the top of my head [laughs]
I: [laughs] Okay. Um.
S: 00:43:11 Because I know there's no acceleration in the horizontal, just the vertical. And that's the acceleration due to gravity.
I: Okay, so in the y-direction you have this acceleration from gravity,
S: Yeah
I: but in the x-direction you don't have…

S: 00:43:27 You just have a velocity. There's no y-velocity, it's purely horizontal. Because he um, it says so right here.
I: Okay. Um, so…how did you interpret the question, or what did you think you were, you needed to find in this question?
S: 00:43:52 In this question initially I thought that I would find the v that would fulfills that condition and they would give an h or an L with the theta and then I could solve it. But then I, they never gave an h or an L. So, and this is a real life situation so, um, they probably wanted me to find a reasonable answer. So I said it's reasonable as long as that much energy is in the system. 00:44:25
I: Mmm kay. [pause]
I: 00:44:38 So…um, have you seen a problem like this before?
S: 00:44:41 Not like this, where they, they just assume h like this.
I: Mmm kay. Um, and what, what did you use for g in your equations?
S: 00:44:58 Just g. It cancels out in the end.
I: Oh, okay. So in the end you canceled it out?
S: 00:45:04 That's why I use symbols because it lets me cancel out numbers..And then it's also simple because L-sq, distance squared over distance is just distance so the units make sense too. 00:45:22
I: [pause] 00:45:34 Mmm kay. Was there um, anything else you were thinking about while you were working on this that you didn't write down?
S: 00:45:43 Um, not really. I wrote it down.
I: You wrote it down?
S: 00:45:49 Yeah. Except my rationale for interpreting the problem as, assume a good h.
I: Okay, so if you were handing this in to be graded on an exam, is that, what would you hand in? [quietly: I guess]
S: 00:46:10 Um, then I would write um, we, then we can interpret the problem as um, finding a po-finding a realistic h for it. And then, and then since it's probably, the height is more than one feet it's, it's probable in reality. I would say it's possible because that reason. 00:46:40
I: 00:46:47 Okay. So I'm just curious, more about um, after you've drawn a picture when you're solving a problem, you know, how do you make a decision about what physics to use?
S: 00:47:05 Um, basically my first decision was um, was over here there's no friction. It's a car that's rolling down um. If we neglect the friction between the axles and stuff. Then um, we can simplify the problem [?] And that would be pretty realistic I guess. Yeah. 00:47:37
I: So when you were thinking about physics you thought first about some of the friction?
S: Yeah
I: Ignoring friction
S: 00:47:43 Friction and air resistance over here. Then I'm like, um, probably the air resistance is minimal for a car. They're designed that way.
I: They're designed that way? [laughs]…Okay. And then, you said assume no frictional effects, energy is conserved. So how did you- what made you think of energy?

S: 00:48:16 Um, well basically I knew initially there was no velocity but there was a height. Then it rolls down and converts into kinetic energy here. And then if we take a reference, um, reference frame starting here. Then uh, h from here to here is the energy um, potential energy. Then it just converts into kine-kinetic energy. And then once I know this velocity it becomes just a kinematics problem.
I: Okay, how did you know that it became a kinematics problem?
S: 00:49:00 Um. How'd I know?
I: Yeah [laughs]
S: 00:49:09 Because uh, you can't really store energy in a horizontal distance.
I: Mmm kay.
S: 00:49:17 Like, you have gravitational potential energy but it's not like there's a spring or anything there. So then I know if it falls here in a certain time if I can get that time I can get the time it takes to go over here. So it's just some kinematics thing.
I: Okay, so you decided not to use energy…here because you said something about it can't be stored in a horizontal…
S: Yeah
I: distance? Okay.
S: 00:49:51 And so then L would have [?no meaning] and we're given that information.
I: So you knew you needed to use the L?
S: For sure.
I: For sure? Okay. And so, having that, that horizontal distance and the vertical distance, that made you think of kinematics?
S: 00:50:14 Yeah. Because then I figured if I can get the time it takes to fall, I can get also the time it takes to go that way. And then if I plug in the time I know how far it should go. 00:50:27
I: 00:50:39 Kay. Is there…anything else that you were thinking about?
S: 00:50:46 Not really.
I: Okay, was this similar to any problems that you've done in class?
S: 00:50:54 No, because of that.
I: Because of that height?
S: 00:50:58 Because like, assuming stuff. Normally we're given values, but. We're not given a value so I'm like okay, then must be an assumption problem.
I: Okay, how about this part of it. The second part of the problem. Did that look similar to any other problems you've done?
S: 00:51:21 Yeah, we've done a lot of problems where a car drives over a cliff.
I: Oh really?
S: Yeah.
I: Okay. So those other problems that you've solved, did those help you in any way to solve a new problem?
S: 00:51:36 I think so. Because then I after I did it I realized as long as I know what variables, I know then I can cancel them out somehow. Like I knew I could cancel out time. It never gives time.
I: It never gives time?
S: 00:51:56 It doesn't give like, the time
I: Mmm hmm

S: 00:51:58 and time is not really the answer, I don't think. Because you're solving for h. So um, if I can connect the time between these two components…like, after we did a lot of problems like I figured I can get rid of time that way. 00:52:16
I: Okay. [pause] 00:52:26 So these two problems that we did, um, do you think that they're, they're consistent with what you usually do when you're working on problems?
S: Yeah. 00:52:35
I: 00:52:54 Mmm kay. Unless you have anything else that you want to, want to tell me about what you think about while you're solving a problem…
S: 00:53:04 Nope.
I: Okay. So, then I have one more thing for you to sign um, turn that off [camera beeps]. …to verify that I've given you money [laughs]
S: Okay