

TOWARDS AN UNDERSTANDING OF HOW STUDENTS USE  
REPRESENTATIONS IN  
PHYSICS PROBLEM SOLVING

by

PATRICK BRIAN KOHL

B.S, Western Washington University, 2001

A thesis submitted to the  
Faculty of the Graduate School of the  
University of Colorado in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
Department of Physics  
2007

This thesis entitled  
Towards an Understanding of How Students Use Representations in  
Physics Problem Solving  
written by Patrick Brian Kohl  
has been approved by the Department of Physics

---

(Committee Chair)

---

(Committee Member)

Date \_\_\_\_\_

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

HRC Protocol # 0205.21

Kohl, Patrick Brian (Ph.D., Department of Physics)

Towards an Understanding of How Students Use Representations in Physics Problem  
Solving

Thesis directed by Assistant Professor Noah D. Finkelstein

Skill with different representations and multiple representations is highly valued in physics, and prior work has shown that novice physics students can struggle with the representations typically used in solving physics problems. There exists work in PER examining student use of representations and multiple representations, but there have been no comprehensive attempts to understand what factors influence how introductory students succeed or fail in using representations in physics. This thesis is such an attempt, and is organized around four main goals and results. First, we establish that representation is a major factor in student performance, and uncover some of the mechanisms by which representation can affect performance, including representation-dependent cueing. Second, we study the effect of different instructional environments on student learning of multiple representation use during problem solving, and find that courses that are rich in representations can have significant impacts on student skills. Third, we evaluate the role of meta-representational skills in solving physics problems at the introductory level, and find that the meta-representational abilities that we test for in our studies are poorly developed in introductory students. Fourth, we characterize the differences in representation use between expert and novice physics problem solvers, and note that

the major differences appear not to lie in whether representations are used, but in how they are used.

With these results in hand, we introduce a model of student use of representations during physics problem solving. This model consists of a set of practical heuristics plus an analysis framework adapted from cultural-constructivist theory. We demonstrate that this model can be useful in understanding and synthesizing our results, and we discuss the instructional implications of our findings.

## Acknowledgements

I gratefully acknowledge all of the contributors to this work, including (but not limited to) the entire PER group at the University of Colorado at Boulder and David Rosengrant, Eugenia Etkina, and Alan Van Heuvelen at Rutgers, the State University of New Jersey.

Furthermore, I wish to thank the thesis committee members for their considerable time investment. Their input and assistance has been much appreciated.

We also acknowledge funding through an NSF graduate fellowship, an NSF CAREER grant, and from the APS, the AIP, the AAPT, and PhysTEC.

## Table of Contents

Chapter 1: Introduction	1
SECTION I: LITERATURE REVIEW AND METHODS	
Chapter 2: Review of prior work on representations	7
Chapter 3: Review and evaluation of available theoretical frameworks	21
Chapter 4: Project goals and methods	32
SECTION II: INITIAL STUDIES	
Chapter 5: Establishing representational effects on problem-solving	35
Chapter 6: Effect of different representational environments on representational skills	68
Chapter 7: Fine-grained examination of representational and meta-representational skills	88
Chapter 8: Section II summary	123
SECTION III: A DEEPER LOOK	
Chapter 9: Representation, concept, and context: FMCE and BEMA data	129
Chapter 10: Representation-dependent cueing and the role of prior knowledge	149
Chapter 11: Multiple representation use in different physics course environments	167

Chapter 12: Task-level analysis of multiple representation use by experts and novices	202
SECTION IV: SYNTHESIS AND IMPLICATIONS	
Chapter 13: Model and heuristics	236
Chapter 14: Review and conclusions	253
References	264
APPENDICES	
Appendix A: Homework and quiz problems, chapters 5-7	271
Appendix B: Student reasons for choosing representations	282
Appendix C: BEMA and FMCE, standard format	284
Appendix D: Revised FMCE	304
Appendix E: Chapter 10 recitation problems	312
Appendix F: Multiple representations usage survey	321
Appendix G: Car motion representations	323

## List of Tables

Table I: Student performance on representationally-varied homework problems	42
Table II: Student performance on quizzes given with and without a choice of problem representation	43
Table III: Statistical significances of quiz performance differences	48
Table IV: Reasons for choosing a quiz representation (part 1)	53
Table V: Reasons for choosing a quiz representation (part 2)	54
Table VI: Student homework performance, reform section	71
Table VII: Student quiz performance, reform section	72
Table VIII: Statistical significances of quiz differences, reform section	73
Table IX: Student success broken down by consistency or inconsistency of their strategies across representation	103
Table X: Thesis goals and coverage by chapter in Section 2	125
Table XI: FMCE scores, categorized by whether the problem involved representational or conceptual translations	137
Table XII: Pre and post test scores on original and revised FMCEs	154
Table XIII: Student performance on trial 1 problems, chapter 10	161
Table XIV: Student performance on study problems 1-4 for Rutgers and CU, chapter 11	184
Table XV: Student use of pictures and FBDs on problems 1-4	185
Table XVI: Student performance on challenge problem, broken down by number of forces identified and types of pictures used	186



Table XVII: Further analysis of performance on the challenge problem	189
Table XVIII: Challenge problem performance vs. force and picture use combined	192
Table XIX: Expert and novice performance on electrostatics problems	215
Table XX: Time for experts and novices to complete the study problems	217
Table XXI: Solution complexity, experts vs. novices, car problems	218
Table XXII: Solution complexity, electrostatics problems	220
Table XXIII: Thesis goals and chapter coverage thereof	254

## List of figures

Figure 1: Four kinds of representation	1
Figure 2: Representation use in dual-coding theory	22
Figure 3: The mediated action triangle	29
Figure 4: Frames of context	30
Figure 5: Isomorphic homework problems	39
Figure 6: Spectroscopy quiz examples in two formats	47
Figure 7: Example multiple-representations quiz solution	58
Figure 8: Example multiple-representations exam problem	78
Figure 9: Representational content of the lectures studied	80
Figure 10: Representational content of the exams studied	82
Figure 11: Representational content of the long-format homeworks	83
Figure 12: Graphical and pictorial versions of the spectroscopy quiz	96
Figure 13: Verbal and pictorial versions of the pendulum quiz	98
Figure 14: Pictorial and mathematical versions of the spring quiz	100
Figure 15: First two problems from the FMCE	132
Figure 16: Problems 22 and 24 from the FMCE	134
Figure 17: First three problems from the BEMA	135
Figure 18: Problems 28 and 29 from the BEMA	141
Figure 19: Original and revised versions of FMCE problem 1	152
Figure 20: Original and revised versions of FMCE problem 39	153
Figure 21: Original and revised versions of FMCE problem 44	156
Figure 22: Graphical and vector representations of a chapter 10 quiz	160

Figure 23: Chapter 11 multiple representations electrostatics problems	173
Figure 24: Electrostatics challenge problem	174
Figure 25: A multiple-representations exam problem from CU	178
Figure 26: Sample questions from the multiple representations survey	180
Figure 27: Representational content of CU and Rutgers exams	182
Figure 28: Problem success vs. number of forces identified	188
Figure 29: Example Schoenfeld plot describing a physics novice	211
Figure 30: Example representation use vs. time graph for a novice	212
Figure 31: Order of representation use in novice student solutions	213
Figure 32: Representations of an unsuccessful novice's problem solution	224
Figure 33: Representations of a successful novice's problem solution	226
Figure 34: Representations of a successful expert's problem solution	229
Figure 35: Contextual constructivist triangle	239
Figure 36: Contextual constructivism adapted to problem solving	240
Figure 37: Three categories of representation use in problem solving	244
Figure 38: Student solution to a multiple representations problem	246
Figure 39: Triangle representation of the solution in Figure 38	247
Figure 40: Expert representation use as a function of time	248
Figure 41: Scan of successful expert solution of the pulley problem	249
Figure 42: Triangle representation of an expert's solution	250

## Chapter 1: Introduction

Skill with different representations and multiple representations is highly valued in physics, both as a tool for understanding basic concepts[1, 2] and as a means to solve difficult physics problems.[3-8] By ‘representations’, we refer to the many ways in which one can communicate physical concepts and situations. For instance, in kinematics one often uses the example of a car accelerating constantly from rest. One can express this motion with graphs of position, velocity, or acceleration versus time, with a written description of the motion, with equations appropriate to such motion, or with a series of snapshots depicting the motion (Figure 1).

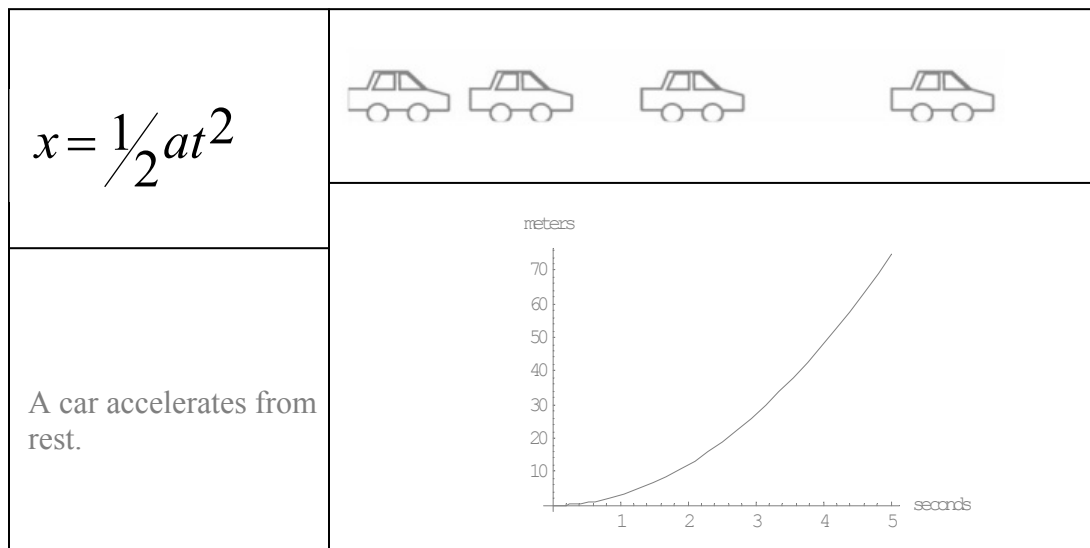


Figure 1: Four different representations of the motion of a car.

Of course, there is no one way to categorize representations. One can productively divide representations into graphical, mathematical, verbal, pictorial,

and/or diagrammatic categories.[9, 10] One can also think in terms of animated versus static representations,[11] or qualitative versus quantitative representations,[3] among many others, depending on one's goals. Scientists can interpret all of these formats effectively and are able to integrate them, translate among them, and assess their usefulness in different situations. Indeed, to a physicist it may seem somewhat artificial (although productive) to draw clean boundaries between categories of representation. There is considerable interest in understanding how physics students use different representations, especially during problem solving.

There have been a number of studies (both within[12-14] and outside of physics education research (PER)) that examined student performances on particular representations of physics problems (see Meltzer for a short overview[14]). In those, it is quite common to see students having difficulties that are bound to the particular representations being used (not being able to interpret a graph of a negative quantity, for example[13]).

There have also been studies that compare student performance on problems that involve multiple representations to performance on problems that involve isolated representations,[15, 16] and studies that investigate student skill at translating between representations.[2, 17] Perhaps most common in PER have been studies that investigate methods of teaching the use of multiple representations during problem solving, generally centering on some kind of multi-step heuristic.[6-8] Heller,[7] for instance, has students build verbal and visual description of a problem, convert that into a "physics" representation, and use that to produce a mathematical representation.

A more recent and less-studied area is the role of meta-representational competence. Meta-representational competence refers to what students know about their own skills and knowledge of representations, and how that meta-knowledge affects problem-solving strategies and successes. Some work has been done on the meta-representational skills that K-12 students bring with them in dealing with physical science,[18, 19] but no work has been conducted (to our knowledge) involving introductory university physics students, one of the major target populations of PER.

The existing research has established that novice physicists (generally introductory physics students) often struggle both with using the canonical representations of physics and with the coordination of multiple representations. Studies also exist that highlight some of the differences between expert and novice problem-solving approaches, with some of these differences involving representation use.[3, 20-22] Indeed, some researchers have argued that facility with multiple representations is a necessary condition for expertise in physics.[1]

It seems clear that if we want our students to succeed in physics, we must attend to their use and understanding of representations. The research base has studies considering the difficulties students have with particular representations such as graphs, studies attempting to teach the use of multiple representations, and early investigations of the role of meta-representational skill. However, a number of gaps remain in our understanding, with four appearing particularly noteworthy to us. First, the extent to which changes in representation can affect performance, and the mechanisms behind that effect, are not well-known. Can we clarify this? Second, the

role of instructional environment in shaping student skills remains largely unstudied. For example, if a course is very rich in representations and multiple representation use, what impact does that have on students' abilities to solve multiple-representations problems, with or without explicit teaching of heuristics? Third, what role do meta-representational skills play at the university level? Finally, how do expert physicists solve multiple representation problems, and what does that tell us about how best to bridge the gap between expert and novice?

These specific points lead into a general problem. To date there have been no comprehensive attempts to understand what factors influence how introductory students succeed or fail when using representations in solving physics problems, and to then organize those results into a coherent description of how students use representations. In this thesis, we develop a simple model (including a set of heuristics for practical use) that lays the foundation for such a description. By addressing the four gaps described above, we provide the necessary background for such a model, while investigating points that are interesting in their own right.

In section I of this thesis (chapters 2-4), we prepare ourselves for experimental studies. Chapter 2 reviews the relevant literature, establishing what is known and what is not known in the field. Chapter 3 describes the primary methods used in the thesis. Chapter 4 discusses the major goals of the thesis in detail.

Section II of this thesis describes the results of our first series of studies. In chapter 5, we confirm that the representation in which students work can have significant effects on their problem-solving success, and that these effects can be quite complex and difficult to predict. To address this complexity, we approach the

question of how and why representation matters on two levels. First, in chapter 6, we consider representations on a macro-scale, investigating how instructional environment can influence student success with different representations and with multiple representations. Here, we find a consistent and positive effect from representation-rich, PER-informed classroom environments. Second, we consider the effect of representation on a micro-scale. In the in-depth problem-solving interviews of chapter 7, we find that representation-dependent cueing is a major factor in student answers and strategy selection, and that student patterns of representation use vary somewhat consistently between experts and novices. We also find a distinct lack of meta-representational skill, at least those skills we measure here.

In section III, we present experiments designed to test our results in more detail. In chapter 9, we investigate the usefulness of separately analyzing the effects of representation from the effects of concept on problem solving. We find that this distinction can be productive, though representation and concept are often intertwined. In chapter 10, we present problems designed to cue student answers in particular ways based on their prior knowledge. In chapters 11 and 12, we explicitly consider multiple representation problems (as opposed to those based primarily on one representation). In chapter 11 we compare two different PER-informed approaches to teaching multiple representations problem solving. One approach centers on specific problem-solving heuristics, while the other infuses all aspects of the course with multiple representations without teaching heuristics. Both approaches are successful in promoting multiple representation use. In chapter 12 we examine expert and novice physics problem solvers individually, establishing key differences



between them. While both experts and novices make significant use of multiple representations, experts make much more productive use of those representations, suggesting greater meta-representational facility.

We thus have four main results corresponding to our four target questions: the complex and context-sensitive nature of representational effects with cueing as a major factor, the effects of representation-rich vs. representation-sparse class environments, the weakness of student meta-representational competence, and specific differences in representation use between experts and novices during problem solving. With these in hand, in section IV (chapters 13 and 14) we sketch a model of representation use during physics problem solving. We introduce a cultural-constructivist perspective, and adapt it to the understanding of representation use. We demonstrate that this model can be useful in understanding and synthesizing our results. Finally, we discuss the instructional implications of our results and possible directions for future experimental or theoretical work.

The work in this thesis has led to four first-author publications, with a fifth in preparation. Chapter 5 is based on the first of these,[24] with Chapter 6 and 7 presenting the results from the second[25] and third[26] papers. Chapter 11 is an extension of a recently-published collaborative effort involving the PER group at Rutgers University in New Jersey.[27] Chapter 12 is being prepared for publication in the near future.

## SECTION I: LITERATURE REVIEW AND METHODS

### Chapter 2: Review of prior work on representations

The body of representations-related research is vast, so we will consider select subcategories that are most relevant to our projects, with emphasis on work in PER and (usually) university-aged students. Some of the earliest work in representations in PER deals with student difficulties with specific kinds of representation, and we begin our review here.

#### *Difficulties with particular representations: Mathematics*

Mathematical representations are ubiquitous and fundamental in physics, to the point that they are not often treated as a class of representations in their own right. However, the importance of mathematics is strongly recognized in PER. Meltzer, for instance, argues that mathematical preparation is among the more important ‘hidden variables’ predicting success in introductory physics.[28] In his study, student gains on an electricity and magnetism assessment are strongly correlated with student ACT mathematics scores or, alternatively, university math placement scores. Student scores were uncorrelated with physics pretest scores, suggesting that deficiency in math skill is harder to overcome than deficient physics preparation. In addition, the E&M assessment used contains a variety of representations, hinting that facility with mathematical representations is linked to facility with others. We will see many examples of the interplay between skills with different representations over the course of this thesis.

Another topic of recent interest in the PER community is that of proportional reasoning, especially when using algebraic equations. In PER, Cohen and Kanim[29] have studied what is known as the algebra reversal error, in which students mistakenly represent a statement such as “There are six students for every professor” as  $6S = P$ . The authors designed an experiment to check a number of possible sources for this error. For instance, students may be translating the statement “There are six students for every professor” directly into an equation in such a way that the number and variable order matches the word order. Alternatively, students may treat algebraic coefficients as labels, so that  $6S$  is literally read to mean six students. The experiment found no clear factors, and the authors concluded that student errors may stem from some fundamental misunderstanding of algebraic representations. Considerable work exists on this topic outside of PER as well,[30-32] though it does not often include the populations of most interest to PER.

Torigoe and Gladding[33] have preliminary results indicating that this kind of proportional reasoning can be a strong predictor of success or failure in introductory physics courses. In their work, they gave introductory physics students a Math Diagnostic Exam (MDE) as a pretest. The exam covered a number of questions involving algebra and geometry. When they compared MDE scores to final grades, two of the three most discriminating questions were algebraic interpretation questions of the form seen above, including the original “ $6S = P$ ” question.

From the above studies, we see the unsurprising result that student facility with mathematics is an excellent predictor of success in physics. Considerably more surprising is the finding that mathematical competence is a better predictor of physics

success than incoming physics domain knowledge. Furthermore, we see at least one specific indicator identified: the ability to translate written language into algebraic representations, consistent with the general notion that skill in translating between representations is an indicator of physics competence.

*Difficulties with particular representations: Vectors*

Physics concerns itself with many quantities that have both magnitude and direction, which we generally represent with vectors. Researchers have documented a number of widespread student struggles with vectors, ranging from the mechanical to the conceptual. Knight[34] and Nguyen and Meltzer[35] have created surveys of introductory students' competence with vectors, and have confirmed that a great many students (both before and after instruction) cannot reliably manipulate vector representations. As many as 50% of Meltzer's subjects were unable to correctly add two 2-dimensional vectors after a full year of introductory physics. As a result of his studies, Meltzer decides to devote more course time specifically to the use of vector representations. The effects of this representation-specific instruction were not studied in Meltzer's paper.

Flores et. al.[36] used pre- and post-tests to address several questions regarding student understanding of vectors. They found that a great many students failed to treat vector quantities such as forces and accelerations as vectors, and were reluctant to use vector representations in solving these problems. These results persisted even in the PER-reformed course that was studied. This paper often (and intentionally) blurs the line between a vector as a kind of representation, and a vector

as a kind of physical entity. While many studies examine difficulties with concepts that are vector in nature ([37] and [38] are typical), far fewer (such as the previous) focus explicitly on the arrow-like representation we refer to as a vector.

*Difficulties with particular representations: Graphs*

In a relatively early work, Beichner[12] created a survey designed to test student facility with graphs in the context of kinematics. Beichner questioned students on a number of specific topics, including their ability to translate between a graph of one concept and that of another, related concept, their ability to find correspondences between graphs and text, and their ability to interpret graphs without making the “graphs as pictures” error. Beichner found a variety of fundamental and persistent mistakes made by students, and concluded that instructors cannot assume fluency in graphs among introductory populations, contributing to the need to study and understand the teaching of representations. In a later study,[39] Beichner began to address teaching technique, and introduced video motion analysis as a tool for promoting graphical facility.

Further studies confirmed these points and others, including student difficulties with the graphing of negative quantities and a tendency towards fulfilling expectations (that a graph should be continuous with no corners, for instance) rather than faithfully representing the situation at hand.[14] Perhaps the most commonly studied class of errors is the aforementioned “graphs as pictures” error, also referred to by mathematics educators as “iconic translation.”[40] Elby refers to this error as a manifestation of WYSIWYG (what you see is what you get), a human tendency to

interpret visual information literally whenever possible. If a graph resembles a hill, it is interpreted as a graph of an object traveling over a hill, even if it is a graph of a less-intuitive quantity such as acceleration. In our thesis, we study several graphical problems, and we will find the WYSIWYG framework to be quite productive.

*Difficulties with particular representations: Language*

Language, either spoken or written, is another way of representing physics concepts or situations, though we often treat it as omnipresent and privileged compared to other representations. Many linguists involve themselves in the study of how we make meaning of written and/or spoken representations, and some have addressed language in physics specifically. Lemke has studied patterns of language particular to the physics classroom, and how sharing or failing to share those patterns leads to productive or unproductive discussions.[41, 42]

Some linguistic work has been done in PER, as well. Williams[43] notes that many of the words that we use to represent physics concepts (force, speed, work) also represent common-language concepts that are much less precisely defined. Thus, students and teachers can be using the same words to represent much different ideas. Brookes[44-46] investigated the role of language in learning physics in much more detail, often in the context of quantum mechanics. He interviewed a number of students and faculty regarding such topics in quantum mechanics as the infinite square well and the Bohr-model, in addition to studying textbook language. He found that much of language use, both expert and novice, takes the form of metaphor and/or analogy. Brookes identified a number of specific metaphorical ideas used (such as

“the potential well step is a physical object”), and noted that much of the difference he observed in success could be attributed to correct or incorrect applications of metaphors. Students have a strong tendency to construct overly literal metaphors, treating potential steps as physical steps, or thinking of a particle as a truly solid and localized object. The physicists studied were capable of applying literal interpretations of the language when appropriate, and ignoring these interpretations otherwise. In short, experts were aware of the limitations of the linguistic representations that they were using, while students were not.

#### *Comparisons across representation type*

The above studies demonstrate that students can struggle with physics material in a number of representation-specific ways (in addition to nonspecific ways). This leads naturally to the question of how strongly student performance can vary across representational format. That is, if a problem is represented in a variety of ways, how strongly and how consistently will student performance vary? We should note that the meaning of the term "problem" has not been completely specified and agreed upon in the PER community. Many researchers implicitly treat problems as always involving some quantitative analysis (including but not limited to Van Heuvelen[5] and Heller[7]). Others do not make such a restriction.[9] We do not wish to debate the proper use of the term problem in this thesis, and we will simply use the term to refer to typical physics tasks given to students, for example those found in homework assignments. This will include questions that do not require calculation.

While much work has been done on student understanding of specific representations, we know of only one paper in PER to directly compare student performance across representation. Meltzer[9] set out to establish that different representations of the same problem would result in different student performance. In this paper, students from large-lecture algebra-based physics courses took quizzes with problems that were represented in up to four different ways (verbal, mathematical, diagrammatic, and graphical). All students received the same problem sets, and the problems of different representation were similar, but not identical. Meltzer found significant performance differences on two sets of questions. On a question regarding the relative force of gravity between the earth and the moon, students answered the verbal question (composed of sentences) and the diagrammatic question (in which relative forces are represented by arrows) significantly differently. Also, there was one question regarding Coulomb's law that was very similar in structure to the gravitation question. Meltzer found that students handling the diagrammatic question differed significantly from the students handling the graphical question. He also found that students were not necessarily consistent in their performance on the same representations across topics, leaving open the question of why these performance differences exist.

Outside of PER, there is at least one similar study. Koedinger and Nathan[47] examined students learning algebra for the first time, and gave them quizzes with questions in three different formats: word problems, strictly symbolic problems, and problems containing a mix of verbal and mathematical representations. The authors also systematically considered a number of other possible variables, such as word



order, type of arithmetic involved, and whole number vs. decimal. They found that students performed significantly better on the verbal representations of their problems than on the strictly symbolic ones. They then analyzed student solution strategies to find that these different problem representations were consistently triggering different approaches to the problems, even if the problems were “the same” by the standards of a mathematician.

These two studies are suggestive, but leave much to investigate further. The Koedinger and Nathan study involves a young population in a non-physics context, with a relatively limited spread of representations. The Meltzer study considers a more PER-relevant population, but the problems studied differ between different representations, and the problems on which students show significant performance differences are arguably ambiguous and hard to interpret. Thus, the question of whether physics student performance varies with representation remains incompletely answered.

### *Multiple representations and problem solving*

So far we have focused on difficulties with or comparisons between specific representations. This categorization is somewhat artificial: We consider it extremely unlikely that anyone ever considers a problem or concept in exactly one representational format. Real problems tend to be at least slightly multiple-representation-based, and a great deal of interest exists in solving physics problems that explicitly use multiple representations, such as equations and free-body diagrams, together. Many in the PER community have long argued that facility with multiple

representations is important and should be a key course goal.[3-6, 8, 21] It has even been suggested that competency with several representations of a concept is a prerequisite for expert-like understanding,[1] and popular research-based physics assessments have implicitly acknowledged this point by including a spread of representations in their questions and by requiring translations among representations to solve problems.[48-50]

Experts and novices differ significantly in their use of multiple representations. Experts tend to use multiple representations in their problem setups more often than novices, who have a tendency to jump directly to mathematics.[3, 10, 20] Thus, use of multiple representations brings student problem-solving procedures more in line with expert procedures. These differences extend beyond problem solving, as research has shown that novices and professional scientists differ significantly in their ability and willingness to use multiple representations productively in more applied settings such as the laboratory or workplace.[2, 51]

This strong association between multiple representation use and level of expertise has prompted a number of studies in which students are explicitly taught how to use multiple representations when solving physics problems. Many of these appear to be rooted in Polya's work in mathematics. Polya advocated the use of a general four-step heuristic for solving difficult mathematical problems, with the steps being to understand the problem, devise a plan, carry out the plan, and look back (checking or reflecting on the answer).[52] For comparison, we have the more representations-oriented problem solving strategy advocated by Heller and Anderson.[7] This strategy has five steps that are taught to students and reinforced

frequently throughout the course. Students begin by constructing a verbal and visual description of the problem. They then convert this into a “physics” representation, and convert that into a mathematical representation. Step four is to execute the solution and calculate an answer, and step five is to check and reflect on that answer. The authors present data suggesting that students who are taught this strategy do in fact use it, and are more successful at solving multiple-representations problems. Several other studies exist in this vein,[6] with some focusing exclusively on certain classes of multiple representation problems, like those involving work and energy bar charts[8] or free-body diagrams.[53]

While much work has been done establishing the relevance of multiple representations skill in physics and in demonstrating that such skill can be taught, very little work has directly studied *how* expert and novice physicists differ in their use of multiple representations.[3, 10] Our early, as well as current, belief is that such an understanding of how physics students use multiple representations, both correctly and incorrectly, will be required for work towards a theoretical understanding of representation use and for refining our teaching of multiple representations problem solving.

### *Meta-representational competence*

Much of what we have reviewed so far could be referred to as cognitive skills, as opposed to metacognitive skills. Metacognition refers to regulatory or self-reflective thinking, or what we know about our own thinking. In some early research into metacognition, Schoenfeld[54] observed mathematicians and mathematics

students and coded their behavior as a function of time into such categories as Analysis, Exploration, Implementation, and Verification. Schoenfeld also tracked when students made reflective remarks about what they were doing and how well it was going. This scheme allowed him to demonstrate stark differences between the higher-level behaviors of experts and novices during problem solving episodes. Experts spent much of their time in preparation and analysis, while novices tended to jump directly into what Schoenfeld calls implementation: solving equations and making calculations. Furthermore, experts frequently step back and explicitly assess how the problem is going, while novices rarely do this.

Little work on metacognition has been done in PER, though in the aforementioned Knight[34] study involving vectors, 86% of the sampled students claimed to have studied vectors before, with an unspecified majority claiming excellent understanding of vectors and their application. However, Knight's data suggested that only 50% of the students had even a marginal understanding of vectors, with considerably fewer being proficient. This brings up the question of how good physics students are at assessing what they know and don't know; a question that previous work cannot fully answer.

DiSessa has begun to study metacognition from a representational perspective, referred to as meta-representational competence. His program traces its roots to a study in which a small group of talented sixth-grade students were allowed to construct and refine their own representations of the motion of a car.[19] They eventually generated something very much like a Cartesian graph, and appeared to have considerable intuitive knowledge regarding what constitutes a useful

representation of physical data. DiSessa took this to mean that student prior knowledge of representations, even at a young age, is rich and generative.

From this base, diSessa has started Project MARC, a program of research designed to investigate what students know about representations and how capable students are of learning new information about representations.[55] Their work so far has focused on student invention and critique of representations of mathematical and physical knowledge, with the expectation of eventually attempting to teach meta-representational competence explicitly. In one such paper, the authors interviewed three middle and high school students from a volunteer-only experimental course on scientific representation.[56] These students were then invited to invent representations of various data, including hurricane wind speed. The students critiqued their representations as a group, and the authors analyzed these critiques, concluding again that students possess a great deal of intrinsic ability to critique and design representations. In a second paper, Azevedo had ninth-graders develop and critique representations of landscapes made of foam and sand. Azevedo noted that these students were aware of basic conventions such as lines representing edges, but found that these students struggled to decide what features of the landscape were productive to attend to, and which were not.[18]

The work done so far on meta-representational competence has a number of common themes. It involves young children, usually from gifted and talented populations. The sample is generally very small to allow interviews. The tasks usually involve creating and critiquing representations and are not part of a larger, background activity (that is, the meta-representational activity is their entire reason

for being there). The results are intriguing, but it is not at all clear what college populations would do on a large scale in a more conventional classroom environment. For our purposes in PER these are important outstanding questions: How effectively can students assess their own knowledge about representations? Do they know enough for us to draw on their knowledge in courses, or for that knowledge to affect their behavior in a consistent way that must be accounted for in models of representation use?

### *Sensitivity to context*

In recent years, there have been many studies showing that student reasoning about physical situations can be very sensitive to the specific context of the problems, greatly increasing the complexity of any attempts to make sense of or predict student behavior. For example, diSessa finds that student reasoning can be described well in terms of basic elements called phenomenological primitives or p-prims[57] that activate or fail to activate in different situations. An example of a p-prim is the expectation that actions (motions, sounds, etc.) die out over time. Elby posits the existence of representational analogs to p-prims, such as the what-you-see-is-what-you-get (WYSIWYG) tendency towards literal interpretations of visual information. Such a p-prim can be either correct or incorrect depending on what it is applied to. Expert physicists build these and other components into coordination classes, conceptual constructs that are robust across contexts as compared to novice student knowledge.[58] Thus, students may appear to understand a particular concept (Newton's 2<sup>nd</sup> law, for instance) in one context and fail to understand the concept in a

different problem context, a context that an expert might find identical to the first. This view contrasts with much early work in which students are taken to have very stable, robust misconceptions regarding the physical world,[59] though even some of these early papers begin to acknowledge the likely role of context.[60]

In a similar vein, Mestre et. al.[61] found that the manner in which students read out information about a physical situation from a representation of that situation (what diSessa calls a readout strategy[58]) varied strongly with very small changes in context. In one experiment, students were shown animations of balls rolling up and down inclined tracks. The students were asked to identify which of the animations were physically reasonable and which were nonphysical. In some cases, one ball rolled along one track and students appeared to base their reasoning regarding this situation on the absolute motion of the ball. In another case, two balls rolled along two superimposed tracks of different shape. In that case, the students appeared to reason based on the relative positions of the two balls, and this difference in readout resulted in students overlooking very nonphysical behavior that they consistently noticed when watching one track or the other in isolation.

The extent to which student responses are robust versus context-sensitive is not fully known, though the studies available suggest that it may be a very large effect. To progress in our understanding of student use of representations, we will need to have some sense of just how strongly the specific context of a problem (for instance, the particular combination of representations, concepts, and course prior knowledge) can affect problem performance.

### Chapter 3: Review and evaluation of available theoretical frameworks

The work just reviewed highlights a number of open questions regarding student use of representations in problem solving, including the variety of representations used in physics, the importance of multiple representation use, and the possible roles of metacognition and context. Any model of representation use in physics should be able to account for the results of previous representations research. Thus, we review a number of candidate frameworks from the literature that may or may not serve as viable starting points.

#### *Dual coding theory*

Many cognitive science studies of representations use dual coding theory[62] (DCT) in making predictions or doing analysis. In this model, human working memory (as opposed to long-term memory) contains two distinct channels: the verbal, and the visual. These can operate to some extent in parallel, and so this may explain how words and pictures used in tandem can result in substantially increased learning gains as compared to one or the other in isolation. If a learner is presented with information in both visual and verbal formats, their effective available working memory is increased and learning is eased. DCT appears to be generally accepted in cognitive research, though there is some disagreement as to its proper use and in how to interpret the numerous cases in which verbal/visual representations used together do not lead to gains. For example, Kirby[63] takes dual coding as a given process and explains his results as instances where dual coding is harmful. In contrast,



Schnotz and Bannert[17] propose a modification of DCT that focuses on the descriptive (verbal-like) and depictive (pictorial-like) mental models that people build. However, the broad result that verbal and visual processing are somewhat distinct is not in dispute.

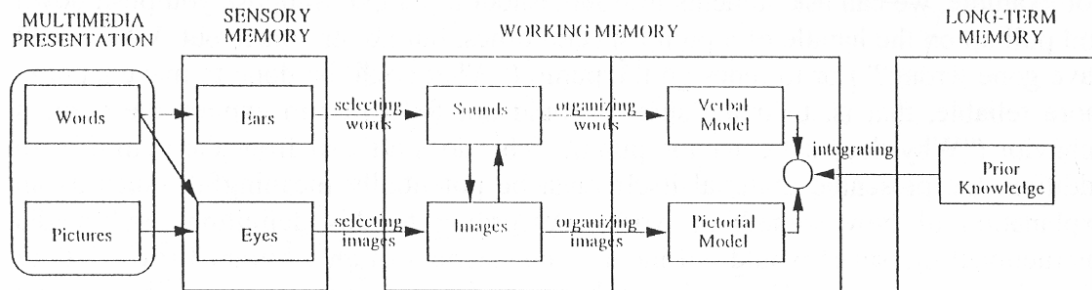


Figure 2. A model of cognition using multiple representations based on Dual Coding Theory. Note the positioning and role of prior knowledge. From Mayer.[16]

While DCT has proven to be useful in cognitive science, we do not consider DCT to be a good candidate for a complete description of representation use in physics. The many studies that make use of DCT[15-17, 51, 64] provide no significant place for contextual variation or inclusion of context in their analysis. In a typical representation of DCT, student prior knowledge (a particular feature of the context in which learning takes place) is placed such that it affects only the later stages of cognition (see Figure 2). The perception and selection of semantic and visual information is not explicitly affected by prior knowledge or any other contextual features here. We note that those authors that do attend to specific context (Kirby,[63] for instance) do not appear to fit this dependence into their guiding theory; rather, the attention to context is an add-on. Furthermore, DCT lacks an

account of meta-level cognition. While this is less of a problem in studies where students are asked to learn simple, mechanistic tasks, physics and higher math problem solving does depend on metacognitive ability[54] and we expect that physics problem solving using multiple representations will be no different. Finally, we believe that the DCT division of representations into only verbal and visual forms will not be adequate for our purposes. Verbal representations are defined to include any symbol strings that are communicated in a speech- or text-based fashion. This would place verbal and mathematical representations on equal footing, and it is well-established that physics students handle mathematics and verbal (conceptual) representations in vastly different ways.[23, 65] Furthermore, it is not clear how one analyzes a graph or diagram in this framework, as these typically have both symbolic and spatial components that are tightly integrated. Quite generally, the representations used in physics are too numerous and rich to be broken into only two categories in a thorough analysis.

### *Cognitive load theory*

Cognitive load theory (CLT)[66] is based on the notion that humans have a limited working memory available for cognitive tasks. The more memory is occupied by one task, the less is available for other, simultaneous tasks. In particular, if a learner is overwhelmed by the immediate demands of solving a problem, they may not have sufficient cognitive resources left over to develop and organize broad problem solving strategies, and as a result their learning will be impeded. This idea has been applied to math and physics problem solving. Ward and Sweller[67] studied

the effect of giving worked examples of homework problems to Australian high school physics students. On exams, these students outperformed students that had no worked examples, even though the students without worked examples solved more problems on their homeworks. The authors hypothesized that the worked examples provided a framework for the students, giving them an attractive alternative to familiar but unproductive problem solving strategies such as searching for equations with the appropriate variables in an iterative fashion. Thus, the students had more working memory available to reflect on that framework and incorporate it into their own methods.

In other work, cognitive load theory has been applied to the use of multiple representations in instruction.[16, 68, 69] In these studies, the authors note several broad features of multiple representation use that they explain in terms of CLT. One of these is the split-attention effect, in which learners solve problems with multiple representations that are separated spatially or temporally. Solving such problems requires the learner to hold parts of each representation in memory for longer, and thus they incur a larger cognitive load than they would if they were solving a problem that makes use of tightly integrated representations. This can result in performance on multiple representation-based learning tasks equal to or lower than the single-representation cases. Another related effect is the expertise reversal effect.[69] Here, researchers have found that an integrated multiple representations approach that works well for novices can be less effective for teaching more experienced learners than a single representation approach. The authors attribute this to the increased cognitive load that multiple representations can generate. If the learner does not need

additional representations to learn a task or solve a problem and if the representations are integrated so that they cannot be considered in isolation, then that learner will have less working memory available for the task at hand.

Cognitive load theory may have applications in a theory of physics representations, and the basic results have been well-established by research. However, we can identify some caveats. First, the implicit goal of most (if not all) related research was to teach students content. If translating between multiple representations hindered students in learning that content, then those multiple representations were considered a liability. Also, the tasks that students were learning in these studies were simple and mechanistic (for example, learning to read a chart in a machine shop or solving "plug and chug" kinematics problems). It is not clear that cognitive load theory would be as useful in analyzing the more complex tasks and skills that we hope to teach our students in physics. Finally, there have been no studies (to my knowledge) where students' learning of cross-representational or meta-representational skills in their own right was analyzed in terms of cognitive load theory, and these are certainly valid goals for instruction. This does not mean that CLT is not a useful tool for us; it simply means that by itself it is likely to be incomplete for our purposes.

#### *Coordination class theory and the resources model*

A third available theory is that of coordination classes.[58, 70] In this theory, diSessa builds on his previous development of the concept of "knowledge in pieces." [57, 71] This perspective challenges a common (though fading) view in PER

in which student knowledge is analyzed at the level of theories or conceptions.[59] There, students are assumed to either understand or fail to understand conceptual elements like Newton's 3<sup>rd</sup> law, force, and acceleration. The transition between not understanding and understanding is not well specified, and the context in which a concept is presented is not generally considered. This misconceptions framework offers few tools for analyzing a case where a student can answer a question related to a particular concept in one context but not in another context. For example, a student may be able to predict the path of a ball that exits a semicircular pipe, but that same student would make an incorrect prediction when asked to predict the path of a ball being twirled on a string when that string is cut.[72] From a physics perspective, these problems use the same principle (Newton's 1<sup>st</sup> law) and are set up the same way (circular motion under the influence of a constraining force). The particular cues available, however, are different and can lead to different student conclusions. In my own anecdotal experience with these questions, students are likely to view the ball on a string example as containing an explicit inward-directed force that persists after the string is cut. In the semicircular pipe example, the students see the pipe as less of a force and more of a constraint, and so the influence disappears as soon as the pipe does.

Using a number of documented examples like these diSessa argues that student knowledge can be analyzed more productively at a finer grain size: the level of resources. Chief among these resources are phenomenological primitives, or p-prims. A p-prim is a basic physical reasoning element that everyone has, such as "more effort begets more result." Such an element goes unquestioned if examined

("that's just the way it is"), and can be applied either correctly or incorrectly to different contexts. For example, this primitive correctly predicts that increased DC voltage applied to a circuit results in increased current, but it also incorrectly predicts that plucking a taut string harder will result in a faster wavepulse (a common enough prediction among introductory students[73]). In diSessa's view, learning physics consists in part of taking primitive reasoning elements and learning which ones to apply in which situations.

With this notion of p-prims in hand, diSessa defines the coordination class. A coordination class is a particular type of conceptual structure that scientists often possess. This coordination class is made up in part of a collection of relevant inferences about the world. This could include physics equations like  $F=ma$ , as they allow for both quantitative and qualitative inferences about some physical quantities given others. Such an inference is in some sense a p-prim bound to a particular context. The coordination class also includes rules for when to apply these inferences, and strategies for obtaining information related to a particular concept (the readout strategies mentioned earlier). This collection is robust and can be applied to many contexts. Not all concepts can be represented by coordination classes (force and mass are examples of physics concepts that may be represented as such), and novice concepts generally do not reach the level of a coordination class. Rather, novice understanding of a concept like force consists of disconnected reasoning fragments that trigger differently in different contexts.

The coordination class approach may be applicable to a model of multiple representations to some extent. The theory is not fully developed,[58] but the focus

on readout strategies and the possibility of representational primitives is enticing. It may be possible to define an analog of a coordination class that describes how experts use physics representations (a contour graph coordination class, for instance). Novice use of representations may be understandable in terms of fragments (particular inferences, readout strategies) and a specification of how they are used in different contexts.

### *Contextual constructivism*

Another theoretical perspective available is that of contextual constructivism.[74, 75] This theory is constructivist in that it focuses on the manner in which learners use tools (physical and conceptual) to construct knowledge and understanding. It is contextual in that it considers the context in which learning is embedded to be inseparable from that learning, and it provides a framework for the careful analysis of that context. This perspective begins with a simple representation of the interaction between a subject and an object, mediated by the use of an artifact or tool (Figure 3, the notion of mediated action). The base of the triangle represents the direct interaction between a subject and an object (say, the observation of a map), and the upper half represents the mediation of that interaction by a tool (for example, a map legend describing the symbols on the map). One can also represent breakdowns in the interaction, a point we shall return to later. Note that depending on what one's analysis focuses on, different things can take the role of subject, object, and tool (a map could just as easily be a tool as a subject). This framework is useful in that it has places for the most relevant components of an interaction and the

relations between those components, while remaining simple enough to be understood by non-experts (for example, physics instructors that do not specialize in PER).

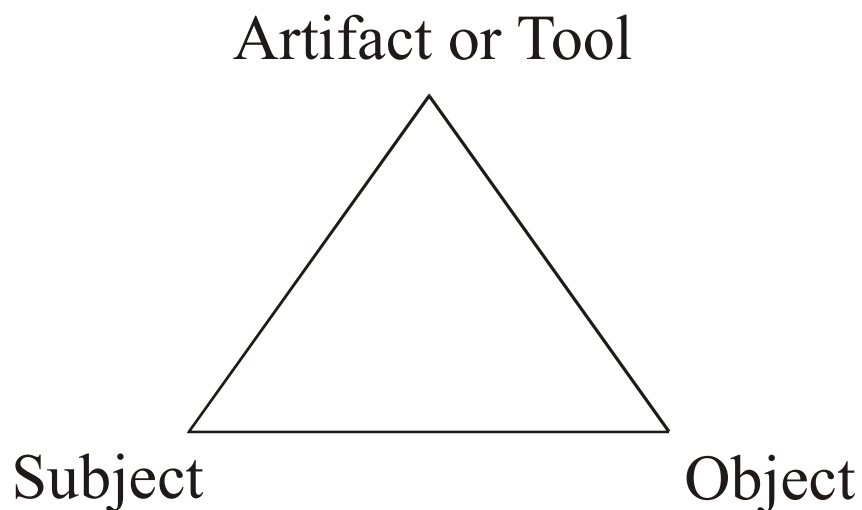


Figure 3. A simple representation of the interaction between a subject and object, mediated by an artifact or tool. From Cole.[74]

By itself, this subject-tool-object model is incomplete, as it divorces an interaction from the context in which it is embedded. Thus, contextual constructivism adds the notion of frames of context.[75] This approach divides the context of an object of study into a number of layers, representable as concentric circles (Figure 4). At the innermost layer depicted here (not necessarily the innermost layer overall), we have context at the level of task. As applied to education, this level describes the particular task that a learner is engaged in, like solving a kinematics problem. This is the level that most physics education research focuses on.



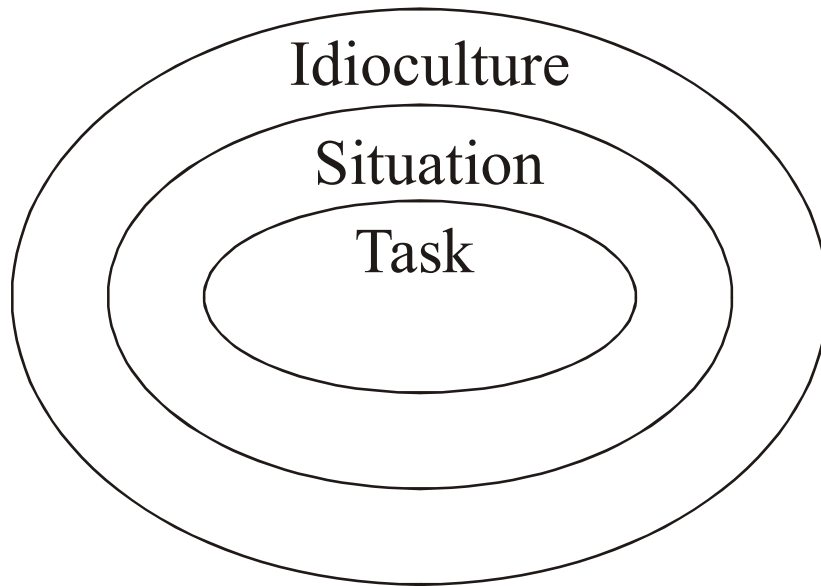


Figure 4. Concentric layers of context present in the contextual constructivism model. From Finkelstein.[75]

The next layer is the level of situation, which includes the immediately available resources, participants, and conditions. Working on a problem set with pencil and paper in the CU Physics Help Room is a particular situation; working on a problem set at home alone is another. Note that the levels of task and situation can influence each other strongly: Solving a kinematics problem in a small group with faculty support is a much different process than solving the same problem alone, and the strategies the learner engages in (and the success of those strategies) will likely change. Similarly, the task can affect the situation: If the students are solving circuit problems, they can make use of a PhET simulation[76] designed to aid in that process. If the students are solving angular momentum problems, no such PhET simulation will be available. In general, the frames of context model contains strong interactions between the various layers of context (which are themselves not rigidly

bounded), though the influences tend to be stronger from the outside-in than from the inside-out.

A third layer is the level of idioculture, which is a collection of situations that are common to a particular group of individuals. The norms and practices that make up a particular course, such as Physics 2010, would constitute an idioculture. This may include an understanding that the professor will talk and the students will listen, or an understanding that the students will be responsible for engaging with the class through a personal response system such as Clickers.[77] It may include the expectation that grades will be based on numerical answers, or the expectation that grades will be based on conceptual reasoning. Again, this level of context will influence and be influenced by other levels. We can extend this analysis to higher levels, but these levels should provide an adequate set to analyze why students learn or fail to learn particular physics concepts in particular environments.[75, 78]

## Chapter 4: Project goals and methods

Our original goal with this thesis was to develop and test a theoretical model of how physics students use representations when solving problems. It eventually became apparent that this was not practical with the current research base. While it appeared to be more-or-less accepted that representation matters, there was not much physics-specific research establishing that different representations do indeed result in different performances. Furthermore, there was no body of work showing how instructional environment can affect representational competence. The effect of environment on general performance is well established; [79] for us to draw instructional conclusions or create a theoretical framework, we need to be able to characterize the size of the effect (and hopefully mechanisms behind the effect) of environment on representational facility. Finally, the moment-to-moment details of exactly how people (both novice and expert) use representations when solving physics problems are unclear, making modeling of this behavior difficult.

Thus, we have conducted a series of experiments designed to fill significant gaps in the body of representations research, with a final goal of creating and evaluating a model and/or set of heuristics that is immediately useful for instruction and can serve as a starting point for future theoretical development. In short our goals are:

- To more firmly establish that representation is a major factor in student performance, and to uncover some of the mechanisms by which representation can affect performance

- To study the effect of different instructional approaches to teaching representation and multiple-representation use
- To evaluate the role of meta-representational skills in solving physics problems at the introductory level
- To characterize the differences in representation use between expert and novice physics problem solvers
- To assemble our most useful results into a concise model describing representation use among novice physics students, facilitating practical analysis

In all of our studies, excepting those on expert problem-solvers, we have focused on introductory algebra-based physics students, as opposed to students at other levels. This limits the applicability of our work somewhat, as it seems quite reasonable that students in other courses would have different levels of skill with different kinds of representation. For example, students in a calculus-based course should be more competent overall with graphs and mathematics. However, we consider this limitation to be a productive trade-off for the depth to which we have been able to study this particular population, a population that tends to be among the largest served by university physics. It remains for future researchers to generalize these results to other course levels.

Our research goals require a two-threaded approach to our studies. We combine large-scale studies of all the students in a class with fine-grained studies in which we examine individual students in interviews. In the large-scale studies, we

give problems (on exams or in recitations) to entire classes, allowing us to make claims regarding averages over large numbers, and allowing us to compare whole classes to each other. In our fine-grained studies, we have the opportunity to validate large-scale conclusions by examining individuals, as well as gaining access to the kinds of data that are not available through whole-class examinations, like details about the specific steps people take when solving problems.

In our large-scale studies, most of our study problems were written for the specific study collaboratively with professors who frequently teach introductory physics at CU. Some of these problems were tested with interviews before administration; others were tested in interviews after administration. A handful of problems were drawn from exams given in past years, or from previous studies in the field.[80] The specifics of problem development and administration will be discussed in the chapters corresponding to the relevant studies.

In our fine-grained studies, our main tool was the ‘think-aloud’ interview,[81] in which students are videotaped while solving study problems. These students are encouraged to think aloud, and are prompted as necessary to explain what they are doing and why. This somewhat changes the nature of the problem-solving episode, as they are essentially being prompted to be reflective about their actions, but the access to student rationale is invaluable. Different interviews have different goals, and our analysis of them varies accordingly as described in the following chapters. All interviewed students were recruited through mass emails to their courses, were paid for their time, and signed informed consent forms.

## SECTION II: INITIAL STUDIES

### Chapter 5: Establishing representational effects on problem solving

(This chapter draws from and extends a paper published in the Physical Review.[24])

In our first study, we hoped to confirm and broaden the idea[9] that representation matters in physics problem solving. That is, we wished to show that presenting problems in different representations, even if those problems were isomorphic from the point of view of a physicist, would provoke substantially different performances from students. The very large class sizes at CU made it possible to divide courses into different groups to receive problems in different representations while retaining good statistics.

We also broadened the examination by investigating whether students can assess their own representational competence, what motives they have for handling a problem in a particular representational format given a choice of formats, and whether providing this choice affects their performance compared to students randomly assigned to particular formats. These questions relate to students' meta-representational competence. Our study differs from the reviewed studies on meta-representational competence in that we ask students to assess fairly standard representations that we have provided rather than ones they have generated themselves. We also have them assess their own skills and preferences regarding these standard physics representations, in part by choosing which representational format they would like to work with on a quiz. There are a number of outcomes that

we might observe here. It may be that students have well-defined learning styles and are aware of them, enabling these students to increase their performance given a choice of representation. It may be that students perform in a relatively consistent way across representational format but are unaware of their strengths, leading to unchanged or even reduced performance when given a choice. Or, it may be that students' performance when given a choice of representations varies and is difficult to predict, with some topics and representations resulting in improved performance and other topics and representations resulting in lower performance. This would suggest a more complicated explanation of how their performance varies as it does, one that must attend carefully to both micro- and macro-level features of the context, and this is in fact what we find.

In short, we have four primary goals in this study:

- To further demonstrate that student performance varies, often strongly, across different representations of physics problems with similar content.
- To investigate why students perform differently on these different representations.
- To show that giving students a choice of representational format will change their performance either for better or for worse, depending on the circumstances.
- To begin to explain how providing a choice of representation results in these performance differences, and to note the possible effects of different instructional techniques.

## *Methods*

We administered our study in recitation to two large (546 and 367 student) algebra-based introductory physics classes at the University of Colorado at Boulder. These courses are composed primarily of students taking the class to satisfy the requirements for life science, social science, and pre-medical programs. These students are typically in their second or third year of study. College algebra is a prerequisite, though in practice student math skills are quite varied. The first course in our study was an on-sequence second-semester class (Physics 202) held in the spring of 2004. The format was mostly traditional, albeit with some in-lecture qualitative and quantitative concept tests using a personal response system.[65, 77] Students had three one-hour lectures per week, and met for two hours each week in either a recitation or a lab. The recitation/lab part of the course was directed by a different professor than the lecture portion. The recitations were generally traditional, with students spending most of their time discussing homework and exam questions with a graduate TA. The labs focused mostly on investigation, testing predictions, and completing open ended tasks (that is, tasks where the students were given a general goal but no specific directions for how to accomplish that goal). Students' grades were based on exams, labs, homework assignments (both online[82] and long answer), and participation in the concept tests.

The second course was an on-sequence first semester class (Physics 201) in the fall of 2004. This course precedes 202 in the standard sequence, but this particular 201 section took place the semester following the 202 class mentioned above, and so each group was being exposed to the study for the first time. The 201



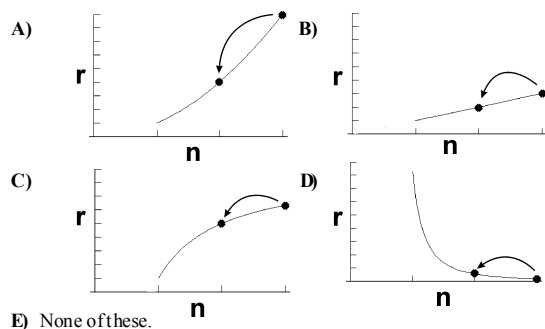
course was taught by a different professor, who is familiar with many of the major results of physics education research. The 201 class was largely reformed, with heavy use of interactive concept tests and an emphasis on tightly integrated lecture demos. The students had the same number of lectures, recitations, and labs as the 202 students. The recitation/lab section was taught by the lecture professor and another professor working together. The recitations focused on working through problems rich in context in small groups, with some demonstrations and some time reserved for homework and exam questions. The labs were a mixture of directed work, open-ended questions, and testing predictions. Students' grades were determined in much the same way as in 202. For the sake of comparison, we videotaped three lectures from the 201 and 202 courses. 57% of the 201 class time was spent on interactive concept tests versus 23% of the 202 class time, supporting the notion that the 201 course had a greater commitment to reform-style student engagement.

For the 202 class, we performed the study in two different subject areas: wave optics and atomic physics. The general subject areas were chosen based on which weeks the recitations were available for study; we attempted to avoid weeks with exams or holidays. The students were assigned four multiple-choice homework questions that covered the same concept in four different representational formats, as well as a one-question multiple-choice quiz given in recitation. We selected specific subtopics that were covered in class and were amenable to representation in a number of different formats. The quiz subtopics were also chosen to match material covered in lab in the hopes that the extra time-on-task from the laboratory would better prepare students to choose between representations. These homeworks were assigned

online as pre-recitation questions and were turned in at the start of the recitation section. Students were expected to turn in pre-recitation homeworks each week and were prepared for the possibility of quizzes, and so these study materials did not represent a significant departure from the norm. The study quizzes were administered by their section TAs. All of the homework and quiz problems are available in Appendix A.

**Question 3 – Graphical**

An electron in a Bohr hydrogen atom jumps from the  $n=3$  orbit to the  $n=2$  orbit. The following graphs show the orbit radius  $r$  as a function of the orbit number  $n$ . Choose the graph that best represents the relative locations of the electron orbits.



**Question 4 – Pictorial**

An electron in a Bohr hydrogen atom jumps from the  $n=3$  orbit to the  $n=1$  orbit. Choose the picture that best represents the relative locations of the electron orbits.

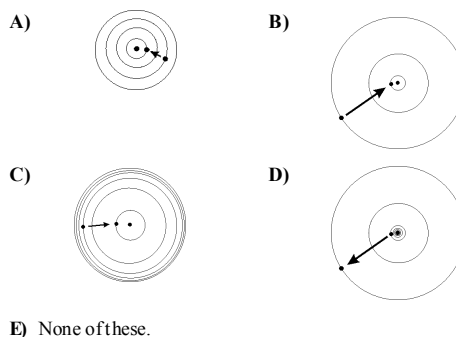


Figure 5: Isomorphic homework problems (in graphical and pictorial/diagrammatic formats) regarding Bohr-model electron orbit radii.

An example of two of four homework problems from one of the two 202 assignments is shown in Figure 5. After turning in the homeworks, the students were given the one-question quiz in one of four representational formats. These quiz problems were isomorphic from format to format, with the answers and distractors mapping from one format to the next. It is worth noting that we use the word 'isomorphic' to mean isomorphic from the point of view of a physicist. A student may have a different view of the similarity (or lack thereof) between these

problems.[10] We also mean isomorphisms between the problem statements and answer choices from representation to representation. We consider it likely that student solution strategies will be considerably less constant across representation (that is, not remotely isomorphic).

Nine of the thirteen 202 recitation sections were allowed to choose from the four representational formats on the quiz without getting to see the problems before they selected. Our intent was for the students to make their choice based on their previous experience with representations in classes and on the homework assignment. The other four sections had quiz formats randomly distributed to the students; these students served as a control group. We provided more of the recitation sections with a choice of format to ensure that a reasonable number of students chose each format. The choice and control sections did not change from one subject area to the next, and the students in the two groups performed similarly on the study homeworks, the course exams, and in the course overall. Both the quizzes and homeworks included a Likert[83] scale survey on which the students could rate the perceived difficulty of the question, and the quizzes included a section where the students were asked to write about why they chose the format they did (if they had a choice) or which format they think they would have performed best at given the choice (if they had a random assignment). Both the quizzes and the homeworks counted towards the students' recitation scores for participation but were not otherwise graded.

The study was conducted in much the same way in the 201 class. We covered two subject areas: energy (in particular, kinetic and potential energies and their connection to motion) and pendulums. For the energy and motion topic, the students

received a four-question pre-recitation homework and an in-recitation quiz. The 201 class was larger (attrition shrinks the 202 class in relation), and so we were able to designate nine of the eighteen recitation sections as control sections, with the remaining nine receiving a choice of quiz format. For the pendulum topic, we gave the students a recitation quiz only (no homework) in order to satisfy schedule constraints. Again, the choice and control groups were the same from one topic to the next, and the two groups performed similarly on homeworks, exams, and the class overall.

In this chapter, we restrict our attention to students who completed a homework (when there was a homework) and the corresponding quiz for a topic, which amounts to roughly 240 and 220 students in the first and second 202 studies, and 330 students in each of the two 201 studies.

### *Data and Results*

In this section, we focus on comparisons of student performances on similar problems in different formats and comparisons of student performance in choice and random-assignment (control) recitation sections. We also examine why students made use of the representations they did and how they used multiple representations when they did.

	Verbal	Math	Graphical	Pictorial
202 Diffraction/Interference HW (N = 241)	0.52	0.61	0.46	0.54
202 Bohr model HW (N = 218)	0.84	0.83	0.76	0.62
201 Mechanics/energy HW (N = 333)	0.54	0.70	0.50	0.49

Table I. Fraction of students answering a homework problem correctly, broken down by representational format and topic. Standard errors vary but are on the order of 0.02.

Table I shows the fraction of students (in both choice and control sections) that answered each of the twelve homework problems (four formats in three different topics) correctly. Table II shows the performances of the students on each format of each in-recitation quiz, grouped by whether they were in a choice or control section. The number of students in each subgroup appears in parentheses.

*Performance across representational format*

All statistical significance tests involving student success rates are two-tailed binomial proportion tests. We shall use to the following terminology: A difference with  $p > 0.10$  is referred to as not significant,  $p$  between 0.10 and 0.05 is marginally significant,  $p$  between 0.05 and 0.01 is significant, and  $p < 0.01$  is highly significant.

Control (random format) group				
	Verbal	Math	Graphical	Pictorial
202 Diff.	0.24 (17)	0.56 (18)	0.25 (16)	0.58 (19)
202 Spec.	0.32 (13)	0.13 (15)	0.53 (17)	0.83 (18)
201 Springs	0.56 (43)	0.41 (39)	0.69 (42)	0.58 (40)
201 Pend.	0.55 (42)	0.30 (40)	0.64 (39)	0.67 (43)

Choice group				
	Verbal	Math	Graphical	Pictorial
202 Diff.	0.35 (17)	0.37 (57)	0.04 (26)	0.82 (59)
202 Spec.	0.81 (21)	0.90 (42)	0.96 (27)	0.39 (58)
201 Springs	0.55 (11)	0.57 (102)	0.88 (17)	0.77 (39)
201 Pend.	0.62 (21)	0.39 (28)	0.65 (40)	0.78 (80)

Table II. Quiz performance of students from the random-format recitation sections (top) and from the recitations sections that had a choice of formats (bottom). The number of students taking a quiz is in parentheses. The quiz topics are diffraction, spectroscopy, springs, and pendulums. Standard errors vary and are not shown.

### *Homework problems*

In examining the homework data, we note that in several cases there were differences in performance from format to format on a particular assignment. When there was a difference in performance between two formats, the mathematical format was often one of the formats involved. This was the only format to require an explicit calculation. The other formats involved conceptual reasoning supported by descriptive language, graphs, or pictures. We see that on average students were most successful with the mathematical homework format, which is consistent with the notion that first-year university physics students are more comfortable with 'plug 'n chug' types of problems than with conceptual problems.[23, 65] We should point out that a mathematical format need not always involve numerical calculation; indeed,

one of the math-format quiz questions (to be described later) was best solved through conceptual reasoning supported by the qualitative use of equations. Nevertheless, in this study the mathematical format usually involved direct calculation.

We also see that there are some noticeable performance differences among the more conceptual formats. For instance, consider the graphical and pictorial problems on the Bohr model assignment, shown in Figure 5. Both require knowledge of how the electron orbit radius varies with the principal quantum number in the Bohr model. The questions differ only in which specific transition is being presented and in whether the problem and solutions are expressed in graphs or pictures/diagrams. Of the 218 students who answered both problems, 76% answered the graphical problem correctly and 62% answered the pictorial problem correctly. This difference is highly significant statistically ( $p = 0.006$ ) and is particularly interesting in that the graphical representation is a rather non-standard one. Students had not seen any graphs of orbital radius versus quantum number, but the pictorial representation of electron orbits should have been somewhat familiar since it is featured in both the textbook and the lectures that preceded this quiz. Further examination of the individual student answers on these two questions indicates that this performance difference can be attributed almost entirely to the 36 students who answered the graphical problem correctly and missed the pictorial problem by choosing the distractor C. This distractor bears a strong resemblance to the energy-level diagrams seen in the Bohr model section of the text and lectures. Since the problems are so similar and the same distractors are present in each problem, it appears that in this case representational

variations may be traceable to a very topic-dependent cueing on visual features of one of the problems.

### *Quiz problems*

We can find another example of performance variation across isomorphic problem presentations in the second 202 quiz, which deals with the emission spectrum of a Bohr-model hydrogen atom. The students were prompted to recall the spectrum of hydrogen, and were asked how that spectrum would change if the binding of the electron to the nucleus were weaker. The questions, answers, and distractors were the same on each quiz except for their representation. Figure 6 shows the problem setups and one distractor for the verbal and pictorial formats (performance data are in Table II). Note that one week previous to the quiz, students completed a lab covering emission spectroscopy, and the quiz images match what students saw through simple spectrometers. Nineteen students in the control group were randomly assigned a verbal format quiz, and 18 were assigned a pictorial format quiz. 32% of the verbal group answered the question correctly, while 83% of the pictorial group answered correctly. This difference is highly significant ( $p = 0.0014$ ). Answer breakdowns indicate that eight of the ten students in the verbal group that missed the question chose the distractor corresponding to the spectral lines moving in the wrong direction (pictured in Figure 6). Only one student from the pictorial group made this error. It is not clear why there would be such a split, especially since the pictorial format shows numerically larger wavelengths as being on the left, opposite the standard number line convention. A possible hypothesis is that students connect



the pictorial format more closely to the lab, giving them additional resources with which to handle the problem. However, as we will see, the students that were given a choice of format performed significantly worse on this pictorial format despite being more likely to cite the lab in making their choice, and so easy identification with the lab cannot be a complete explanation.

Next, consider the performance of control group students on the mathematical formats of the 201 and 202 quizzes. In three of the four quizzes, the average success rate on the math quiz was significantly lower than the average success rate on the other three formats combined. For the spectroscopy quiz, the average verbal/graph/pictorial score was 0.56 versus 0.13 on the math format, a difference significant at the  $p = 0.004$  level. For the 201 spring quiz, the difference was 0.61 vs. 0.41 ( $p = 0.03$ ), and for the 201 pendulum quiz, the difference was 0.62 vs. 0.30 ( $p = 0.0004$ ). The difference between the average verbal/graph/pictorial score and the average math score on the 202 diffraction test was marginally significant ( $p = 0.09$ ). It is somewhat surprising that students were less successful with the randomly assigned math format given their generally higher performance on the equation-based homework problems, though we should note that the students took the quiz in recitation with a time limit (about fifteen minutes) and without access to a textbook, making the environment much different than that in which they would do a homework problem. We should also note that the math problem on the 201 spring

**Spectroscopy Problem – Pictorial Format**

The Balmer series of spectral lines is shown below, as seen through a spectrometer:



Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus and the ionization energy is 13 eV instead of 13.6 eV. Choose the picture that best represents what the new spectrum would look like.



**Spectroscopy Problem – Verbal Format**

Consider the Balmer series of spectral lines from hydrogen gas. Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus. This means that the ionization energy for the electron will be smaller. What will happen to the Balmer lines that we see?

B) The spectral lines will all shift to shorter wavelengths (toward the bluer colors).

Figure 6. Setup and second answer choice for the verbal and pictorial format quizzes given in the second trial. The other distractors match up between the formats as well.

quiz was difficult to solve through explicit calculation, and was more easily handled by using the equations qualitatively. This gives it a different character than the other math-format problems, which is a point we shall return to later.

In closing this subsection, we note that in addition to analyzing homework or quiz problems alone, one can examine whether performance on homeworks is correlated to performance on quizzes in a number of ways. For example, one can ask whether performance on a quiz is correlated to performance on the corresponding homework problem format. Generally, such homework-quiz correlations were very weak, and are not explored further here.

*Effect of student choice of representation*

In Table II, we saw a format-by-format comparison of the students who received a quiz at random and the students who were allowed to choose a quiz format. There were a total of sixteen choice/control comparisons available (four trials with four formats each). Of the eight from the 202 class, six showed a statistically significant difference. These data, along with the significances of the choice/control differences (or lack thereof) in the 201 class, are summarized in Table III.

Quiz subject	Verbal	Math	Graphical	Pictorial
202 Diffraction	<b>X</b>	X	0.04	<b>0.03</b>
202 Spectroscopy	<b>0.002</b>	<b>0.0001</b>	<b>0.0004</b>	0.001
201 Springs	X	<b>0.09</b>	<b>X</b>	<b>0.07</b>
201 Pendulums	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>

Table III. Statistical significance of the quiz performance differences between the format choice and control groups in the 202 and 201 sections. Numbers are p-values using a two-tailed binomial proportion test. X denotes a p-value of greater than 0.10. Bold indicates that the choice group had higher performance than the control group.

These results are notable in that the effects are in some cases quite strong. For instance, 90% of the 42 students in the choice group answered the math format question correctly for the spectroscopy topic, while 13% of the fifteen-student control group answered the same problem correctly. In addition, the direction of the effect can vary. In four of the six cases, giving students a choice of formats significantly increased performance, while in two of the six cases it resulted in a significant decrease. Furthermore, when comparing across content areas we see reversals in the direction of the effect. On the diffraction quiz, students in the choice group do better

than the control group on the pictorial representational format and worse on the graphical representational format, while on the spectroscopy quiz the students in the choice group do worse on the pictorial representation and better on the graphical representation. As we can see, giving students a choice of format does not result in consistently increased or consistently decreased performance relative to the control groups. Rather, the direction of the effect appears to vary strongly across both topic and representation, which suggests two things. First, these students do not have the meta-representational skills necessary to consistently make productive representational choices under these circumstances. Second, a complete explanation of these performance differences will likely be non-trivial and will not be able to rely entirely on broad generalities.

We can further characterize student performance in these cases by considering which distractors they chose. As was mentioned above, the control groups for the pictorial and verbal formats of the 202 spectroscopy quiz (see Figure 6) showed a significant performance difference, with the errors made by the verbal format control group being concentrated almost entirely on the distractor B in which the spectral lines move in the wrong direction (other distractors include the lines compressing, the lines staying the same, and none of the above). The corresponding choice groups did the reverse. The verbal format group had 17 out of 21 people answer correctly, with three choosing the distractor B. The pictorial group had 23 of 58 students answer correctly, with 27 students selecting the distractor B. Thus, we see that the students who chose a verbal-format quiz performed in very nearly the same way as the students who received the pictorial format at random, both in terms of success rate

and choice of distractors. Similarly, the students who chose a pictorial quiz performed in the same way as the students who were randomly assigned a verbal quiz.

In general, of the six statistically significant 202 choice/control comparisons, the performance difference in two of them (spectroscopy verbal and pictorial) was mainly attributable to students focusing on a particular distractor. In the other four, (spectroscopy mathematical and graphical, diffraction mathematical and verbal) the incorrect answers were split among two or more distractors. Note that the quiz distractors map from one format to the others, so this is not simply a case of some of the problems not having any attractive distractors, though apparently (and perhaps not surprisingly) different representations of a problem can make different distractors more or less attractive.

Next, consider the eight choice/control comparisons from the 201 section (Tables II and III). None of the pairs showed different performance at a  $p = 0.05$  significance level. Two were marginally significant (the math format spring quiz at  $p = 0.09$  and the pictorial format spring quiz at  $0.07$ ). There was very little difference in performance between the choice and control groups on the pendulum quiz, which was given four weeks after the spring quiz. The difference between these data and the corresponding 202 data is pronounced. Students in 201 did roughly as well regardless of whether they received their preferred format or a format at random, suggesting that their representational skills are more balanced. That is, they are less likely to have much more trouble with a random representation than with their representation of choice. Since one of the major differences between the 201 and 202

groups was the method of instruction, it may be that the instruction contributed to the effect. We also should note that the 201 and 202 studies involved different topics, which may have contributed to the different performances. In follow up work, we compare two 202 sections with different professors teaching the same topics, allowing us to explore the effects of instructor independent of content.

### *Student self-assessment and assessment of the representations*

In this section we consider data intended to address two related questions. First, how do students assess and value the different representations available here? Second, how (and how successfully) do they assess their own representational competence?

The students in the format choice groups were asked to "Write a few sentences about why you chose the problem format you did." We then coded these responses, separating them into categories that developed as the coding proceeded. In Tables IV and V we present the three dominant categories for each quiz. The complete set of data is in Appendix B. Some remarks regarding our categorization methods: Students in the "visual learners" category have explicitly identified themselves as visual learners or visual people. People that expressed a preference for "plug and chug" problems used language that clearly indicated the insertion of numbers into formulas in a simple fashion, and always used the words plug and/or chug. Students that remarked that they simply found equations or mathematics easier to handle or more straightforward were placed in other categories. In some cases, there is a category for those that chose a format because they were attracted to it and a

separate category for those that chose a format because they were avoiding a different one. Many of the responses were too vague to be useful; "Pictures are pretty," for example. These were discarded.

There are a few notable trends. First, 72% of all the choice group students (including those who did not make comments) selected either a math or pictorial format quiz. We also see that the vast majority of students who cited their lab in explaining their choice chose the pictorial format. This is despite the fact that the recent lab included representations that corresponded to each quiz format.

There are a fair number of students that chose the mathematical format expecting a plug and chug style of problem, except in the case of the 201 pendulum quiz, which followed the 201 quiz on springs. The 201 quiz on springs was unique in that the mathematical format quiz was difficult to handle through explicit calculation alone, and favored qualitative reasoning supported by equations. Eighteen students taking the 201 pendulum quiz mentioned that they didn't like the math format for the earlier spring quiz, with 13 of these choosing the pictorial format the second time. It would appear that in this case there was a mismatch between the students' conception of what constituted a math problem (plugging and chugging) and our conception (either calculation or using equations as a conceptual tool), and the students responded accordingly.

202	Diffraction quiz	202	Spectroscopy quiz
Verbal 10	3: Preference for qualitative analysis	Verbal 12	3: Preference for concepts over math/pictures
	1: Preference for concepts over math/pictures		3: Don't like pictures
	1: Preference for concepts over math		2: The format supports the concepts
Math 24	9: Preference for "plug 'n chug" problems	Math 18	5: Preference for "plug 'n chug" problems
	4: Find equations/numbers easy		4: Preference for mathematics over concepts
	3: Preference for mathematics over pictures		3: Preference for mathematics over pictures
Graphical 12	3: Visual learners/people	Graphical 10	2: Visual learners/people
	2: Like having a visualization provided		2: Preference for visuals over math
	2: Connected it to the pre-recitation HW		2: Connected it to the pre-recitation HW
Pictorial 51	17: Visual learners/people	Pictorial 35	12: Liked the colors/found it attractive
	12: Connected it to lab		8: Like having a visualization provided
	9: Find other formats difficult		6: Connected it to lab

Table IV. Reasons 202 students gave for choosing a particular representation of a quiz. Only the three dominant categories are presented here. The number in each format box is the number of usable responses.



201	Spring quiz	202	Pendulum quiz
Verbal 9	6: Preference for qualitative analysis	Verbal 17	4: Preference for concepts over math/pictures
	3: Preference for concepts over math/pictures		4: Don't like pictures
			3: The format supports the concepts
Math 74	16: Preference for math over concepts	Math 22	7: Preference for "plug 'n chug" problems
	12: Like the straightforward/definite nature		6: Like the straightforward/definite nature
	11: Comfortable handling equations		2: Preference for math over concepts
	7: Preference for "plug 'n chug" problems		
Graphical 11	3: Visual learners/people	Graphical 27	6: Visual learners/people
	2: Like having a visualization provided		6: Like having a visualization provided
	2: Connected it to the pre-recitation HW		3: Didn't like the math format before
Pictorial 26	12: Visual learners/people	Pictorial 62	15: Visual learners/people
	7: Like having a visualization provided		13: Didn't like the math format before
	5: Find other formats difficult (esp. math)		12: Like having a visualization provided

Table V. Reasons 201 students gave for choosing a particular representation of a quiz. Only the 3 dominant categories are presented here. The number in each format box is the number of usable responses.

In the 202 class, 9% of the people who initially chose a verbal format quiz stayed with that format for the second quiz. Twenty-nine percent of the graphical, 42% of the pictorial, and 46% of the mathematical groups stayed with their format. For the 201 section, 73% of the verbal, 25% of the math, 71% of the graphical, and 79% of the pictorial groups stayed with their choice of format from the first quiz to

the second. For all formats but the math (which, for the 201 spring quiz, was different in character from the other math problems in this study), the 201 section was substantially more likely to stay with their choice of format. Of the 76 students in 201 that changed from math on the first quiz to a different format on the second, 11 chose verbal, 22 chose graphical, and 43 chose pictorial. The strong preference for the pictorial format during this switch, the fractions of the class selecting either a math or pictorial quiz, and the student comments overall are all consistent with the notion that students perceive the mathematical and pictorial formats to be dominant and antithetical. That is, when considering the different possible representations of a physics problem, students appear to think primarily of pictorial and mathematical formats (and not so much of others) and to think of these formats as opposites in a sense.

In both the 201 and 202 sections, many of the students who selected a graphical or pictorial format identified themselves as visual learners (15 and 7 of the students on the first and second 202 quizzes, and 15 and 21 students on the first and second 201 quizzes). No students identified themselves as any other type of learner, save one that identified himself as a kinesthetic learner and chose a mathematical format. In the cases of the pictorial formats of the 202 diffraction quiz, the 201 spring quiz, and the 201 pendulum quiz, there were enough self-identified visual learners to compare their performance to the other people choosing the same format. There were 18 self-identified visual learners in the 202 diffraction quiz, who had a success rate of 0.89 as compared to the success rate of 0.78 for the other 41 students. This difference is not statistically significant ( $p = 0.33$ ). For the 201 spring quiz, there were 12 self-

identified visual learners, who had a success rate of 1.00 as compared to 0.67 for the other 27 students. This difference is significant at the  $p = 0.02$  level. For the 201 pendulum quiz, there are 15 self-identified visual learners. These had a success rate of 0.87, as compared to 0.75 for the other 65 students. This difference is not statistically significant ( $p = 0.35$ ). Averaging the above, the self-identified visual learners had a success rate of 0.91 versus 0.75 for the other students, which is significant at the  $p = 0.02$  level. There are a number of confounding factors that leave us hesitant to draw conclusions based on these data alone. Both the students' abilities to assess their own competencies in this fashion and the overall usefulness of categorizing people as different types of learners are somewhat unclear here, and we also note that the students made these identifications (or didn't) after having succeeded or failed at pictorial and/or non-pictorial tasks.

#### *Students' use of multiple representations*

The students in this study were provided with single representations of quiz problems, but in many cases the students' papers showed that they had made explicit use of supplementary representations in solving the quiz. This was most often a picture that they drew in support of a mathematical or verbal format problem. Some students wrote equations in support of non-mathematical formats and a handful wrote out physical principles longhand or drew a graph. As noted previously, it is a common goal in physics education to teach the use of multiple representations,[6, 8] so it is interesting here to compare the performance of students that produce supplementary representations to those that do not. We should emphasize that neither

of the courses studied here made an explicit attempt to teach the use of multiple representations in the style of the just-mentioned references. We should also note that the students who had no explicit supplementary representation may well have used multiple representations in their solution to some extent (it is hard to conceive of a student that can think strictly in terms of one representation and no other), but we shall focus on the students that made these explicit.

On the 202 diffraction quiz, 51% of the 172 choice-group students made explicit use of some supplementary representation. These students had an average success rate of 0.47, compared to 0.54 for the students that did not explicitly use a supplementary representation. This difference is not statistically significant ( $p = 0.12$ ). Breaking it down by format, 35-40% of each of the verbal, graphical, and pictorial groups used a supplementary representation, and in each case the group using such a representation did not do significantly better or worse than the group that did not explicitly use an additional representation. Of the students that chose a math format, 75% (43 out of 57) used a supplementary representation, which was always a picture. These students had a success rate of 0.28, compared to 0.64 for the 14 that did not. This performance difference is significant at the  $p = 0.014$  level.

That the performance difference should favor students that did not draw a picture is surprising, and so we examined the problem in more detail. The math format quiz question is displayed along with a student drawing in Figure 7. This is a question regarding the diffraction pattern coming from two finite-width slits illuminated by monochromatic light, a topic that was featured in a lab but was covered minimally in lecture. The pattern will have a narrow peaks separated by a

distance  $X$  governed by the slit separation  $D$ , and a longer-period envelope with peaks separated by a distance  $x$  governed by the slit width  $d$ . Given the distance from the

**Diffraction Problem – Mathematical Format**

We have a double-slit experiment with incident light of  $\lambda = 633 \text{ nm}$ . On a screen  $3.0 \text{ m}$  from the slit, we see an intensity pattern with small peaks separated by  $0.5 \text{ cm}$ . The first minimum in the overall intensity envelope is at  $2.0 \text{ cm}$  from the center of the pattern. Calculate the separation of the slits,  $d$ . Circle the appropriate letter.

- A)  $D = 3.8 \times 10^{-5} \text{ m}$
- B)  $D = 3.8 \times 10^{-4} \text{ m}$
- C)  $D = 9.5 \times 10^{-5} \text{ m}$
- D)  $D = 9.5 \times 10^{-4} \text{ m}$
- E) None of the above.

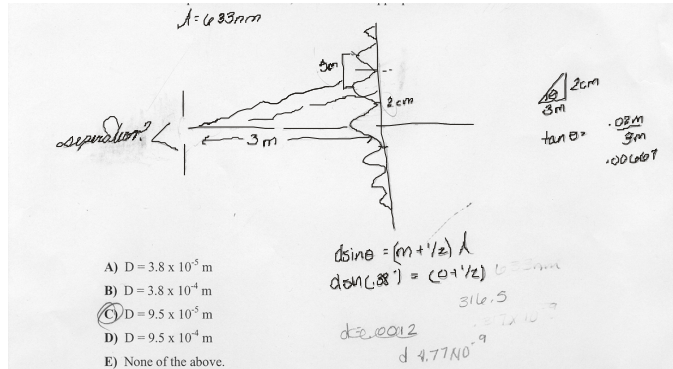


Figure 7. A student's use of a supplementary representation (hybrid graph/picture) to solve the math format 202 diffraction quiz.

slits to the screen  $L$ , the wavelength of the incident light  $\lambda$ , and either  $X$  or  $x$ , one can calculate either  $D$  or  $d$  using  $(d,D)\sin(\theta)=n\lambda$ . Most student errors involved mixing up  $D$  and  $d$ . We examined each student picture (which was often a hybrid of a picture and a graph) and categorized it. Almost no one drew a correct two-slit diffraction pattern, suggesting that this topic was not well understood at this point. Of the 35 students that drew a picture and answered the question wrong, thirteen students drew a picture that represented a single-slit diffraction pattern with peaks separated by  $x$ , which led to a mix-up of  $D$  and  $d$  in the equation. There were also eight students that drew such a picture with peaks separated by  $X$  and then used the equation appropriately, answering the question correctly. Fourteen of the students that drew a picture drew a single-slit diffraction pattern with both  $x$  and  $X$  labeled as follows:  $X$  was marked off between two peaks far from the center, and  $x$  was marked off as the distance from the centerline to the first minimum. These students did appear to notice

that the distance labeled 0.5cm on their paper was roughly twice as wide as the distance labeled 2cm. This drawing was an apparent misinterpretation of the phrase "The first minimum in the overall intensity envelope is at 2.0 cm from the center of the pattern" present in the problem statement. This language is similar to that of the text and of the lab that covered this topic, though this is no guarantee that it would be understood. Of these fourteen students, two answered correctly and twelve answered incorrectly, calculating  $d$  instead of  $D$ . These twelve students account for much of the performance difference between the picture and no-picture groups.

These data recall Chi's[10] finding that in some cases novice problem solvers draw more pictures than experts while making more errors. This suggests that one motivation for using multiple representations is to work through something perceived to be difficult. However, the students that drew a picture rated the problem to be just as difficult as the students that did not draw a picture. On a Likert scale from 1 (easiest) to 5 (hardest), the students that drew a picture gave this problem an average rating of 3.76 while those that did not draw a picture gave a rating of 3.79. It is thus not clear whether the students that struggled with this problem were more likely to draw a picture. There have been other studies in which including multiple representations of a problem resulted in poorer performance than using single representations. This performance difference was interpreted broadly either as an increase in cognitive load when the representations are separated[67, 84] or as an increase in load stemming from an inability to map from one representation to the next.[15] The case here is somewhat different in that it appears to be tied to the specific contextual features of the problem and the problem-solvers. It seems that

here the higher error rate among students using multiple representations is traceable to a particular misunderstanding of the problem statement that was much more likely to be detrimental if it was expressed in a pictorial manner. If the problem or the general background of the students on this topic had been slightly different, this likely would not have occurred.

On the 202 spectroscopy quiz, 10 out of 148 students in the choice group used a supplementary representation. This is too small a sample for analysis. We do find it notable that there would be such a large difference from topic to topic, with 51% of students using an explicit supplementary representation for the diffraction quiz and only 7% using one for the spectroscopy quiz. The average success rate across all students on this quiz (choice and control) was 0.62, which is significantly greater than the 0.42 for the students on the diffraction quiz ( $p = 0.0004$ ). In contrast, the choice and control students gave the spectroscopy quiz a difficulty rating of 3.60 averaged across all formats, compared to the rating of 3.47 for the diffraction quiz. Thus, it appears that the spectroscopy quiz was easier for the students, though they did not rate it as such. It is not clear whether this influenced their decision to use an explicit supplementary representation.

On the 201 spring quiz, 74 of the 169 students in the choice group used a supplementary representation. Sixty-nine of these were students that had chosen a math-format problem. These students had a success rate of 0.55 as compared to 0.61 for the 33 students that chose a math format and did not use a supplementary representation. This difference is statistically insignificant ( $p = 0.60$ ). The use of supplementary representations was less common but somewhat more spread out for

the control group on the same quiz: 9, 23, 9, and 4 students used a supplementary representation on the verbal, math, graphical, and pictorial formats respectively. This variation accounts for 45 students out of 164, or 27%. The 23 students that used a supplementary representation on the math format had a success rate of 0.43, which did not differ significantly from the success rate of 0.38 achieved by the 22 students that did not use a supplementary representation ( $p = 0.62$ ). These data were very similar for the 201 pendulum quiz.

In summary, students that use explicit supplementary representations on these quizzes are roughly as successful as those that do not, with one case in which they are less successful. This finding is consistent with the cognitive science results mentioned previously in which researchers found that multiple representations do not necessarily increase performance. Rather, multiple representations are tools that students can either use productively or not.

### *Discussion and Conclusion*

This study began with a number of goals. First, we wished to know whether student performance on physics problems varies with representational format. We see evidence that it does, often strongly. In the case of the Bohr-model homework problem, the performance difference between the nearly-isomorphic graphical and pictorial problems is due to students selecting a particular distractor. This distractor is one that superficially resembles energy-level diagrams that they have seen associated with this material, but only when it is represented pictorially. We also see students in the random-format groups doing much better on a pictorial format



spectroscopy quiz than on a verbal format of that same quiz. It is less clear what might have triggered this. While the choice groups make it clear that students connect the pictorial format more closely to their laboratory experiences, the lab did not ask them to consider this particular concept (though it did make use of similar representations). This issue is further confounded by the fact that the students that were allowed to choose a pictorial format, in the process perhaps identifying themselves as students connecting more strongly with the laboratory, did significantly worse than the students that chose a verbal format.

We note that students that were randomly assigned a mathematical quiz did significantly worse in three of the four cases than students assigned any other format. This was true in two cases when the math problem involved simple calculation. This is surprising since the selections and comments of the choice groups, in particular the reasons cited for the move away from the math format in response to the first 201 quiz (a problem that was not "plug 'n chug"), suggested that many of the students preferred plug 'n chug problems to other sorts. The poorer performance on the mathematical format was also present on the aforementioned 201 quiz where the mathematical format was more easily handled with conceptual reasoning. In that case, the equations appear to have provoked the students to spend time on unhelpful calculations instead of thinking about the problem. Students calculating without thinking has been observed many times before, and has been attributed in part to a lack of meta-level skills.[54]

Given that students do perform differently on different assigned representations of problems, our second goal was to begin to determine why. The

data suggest that performance on different representations depends on a number of things, including student expectations, prior knowledge, metacognitive skills, and the specific contextual features of the problems and the representations. This dependence on specific, micro-level features also seems to be responsible for the reduced performance of some 202 students that made use of multiple representations, as compared to students that did not. It may also be that different problem representations are prompting different solution strategies, as Koedinger[47] has observed in young algebra students. The strategies of our students cannot be consistently inferred from the data presented here, and so we have interviewed students in-depth as they solve these sorts of problems. The results of these interviews will be part of a later chapter.

Our third goal was to determine whether allowing students to choose which representation they worked in would have an effect on their performance. The data show that giving them this choice for a quiz did indeed result in performance differences as compared to the random-format students; however the direction of that effect turned out to be inconsistent. In some cases, students given a choice of representation did much better than the students that were assigned a format at random; in other cases, they did much worse. Furthermore, whether the choice group did better or worse than the control group for a particular format sometimes varied from one quiz topic to the next, as was the case for both the graphical and pictorial groups in the 202 section. This could possibly be explained by the movement of a group of students that is good at choosing from one format to another, but analysis of the students that switch formats shows that this is not the case. Students that stayed

with these formats did approximately as well on the second quiz as students that switched to these formats.

Finally, we hoped to explain the effect of student choice. To this end, we examined the students' comments. It appears that, correct or otherwise, students generally view mathematical and pictorial representations as dominant and opposite, at least out of the set of representations presented here. Most students selected one of these two formats. Students that switched from a mathematical format typically switched to a pictorial format. Student comments regarding their choice of quizzes frequently pitted mathematics against pictures, with one being favored versus the other. These same comments suggest that students connect pictorial representations quite strongly with "concepts," which students appear to view as unconnected to the mathematics. By itself, this characterization of student motives does not explain why student choice of representation had the effect that it did.

One could suppose that student choices and performances are guided by intrinsic learning styles. While it is the case that self-identified visual learners performed better on one of the pictorial format quizzes than other students (which is cause and which is effect is not clear here), the arrangement of the performance variations seen here suggests that the bulk of the data cannot be explained by a simple alignment of student choices with some individual learning style. For example, the students that chose pictorial format quizzes for both the diffraction and spectroscopy topics had success rates of 0.86 and 0.33 on these quizzes. This is a dramatic difference since it appears that the diffraction quiz was easier overall, as the choice group had respective success rates of 0.48 and 0.70 averaged across all formats. A

learning styles explanation would expect the same students to perform reasonably consistently on the same formats relative to the rest of the class. Considering the complexity of the performance data, it appears likely that an explanation of the effect of giving students a choice of representation will need to carefully attend to the context of the problem, much as we argued above regarding a complete explanation of student performance on different representations.

Another partial explanation is suggested by the fact that the 201 students showed a much smaller performance difference between their choice and control groups than the 202 students. This may be a function of the broader, macro-level features of the context, including the methods of instruction. As we described before, the 201 class included more reforms and the features of the course may have provided students with a more varied set of representational skills. This could have leveled out students' performances on their preferred representations as compared to other representational formats. This might also explain the 201 students' much greater tendencies to remain with a format from quiz to quiz, as students with broader representational skills may be less likely to be dissatisfied with a particular representation. We should caution that this explanation is somewhat speculative, and we note that so far we cannot separate out the effect of instructional differences from the effect of content differences. The 201 and 202 courses are quite different in subject matter and in representational content. To make such an assertion, we would need to perform a repeat study that uses these quizzes and homeworks in a 202 course taught by the same instructor as this 201 course. Such a study is the subject of the next chapter.

The study featured in this chapter has a number of limitations. First, as was just noted, the two courses studied differed in both subject area and instructional method, making cross-class comparisons difficult. Second, while the courses studied had a few hundred students each, it would still be desirable to replicate the study from year to year. Third, physics, including introductory physics, has a very rich collection of subtopics and associated representations. For the sake of this study, we have defined several representational categories, but such a definition cannot be unique or complete. Further, we have only examined a small selection of possible subtopics in this study, though this is perhaps not a severe limitation. One of our conclusions is that specific problems often will have features that are particular to the representation used that have significant impacts on student performance. Thus, there are aspects of this study that we would not expect to be invariant across all subtopics. Finally, this study gives us fairly limited insight as to how the students solved these problems. Such insight will likely be necessary in order to better understand how and why representation affects performance, and so we have performed a number of in-depth student interviews to examine exactly what students do while solving these problems. The results will be seen in a later chapter.

With the above data in hand, it appeared that a complete understanding of student representational competence would need to attend to the specific and general features of the problems, the courses, and the learners, thus provoking the two-threaded micro/macro approach described earlier. In this chapter, we have taken a detailed look at student performance on specific problems. In a later chapter, we will expand this look at the micro-level with a series of problem-solving interviews. We

have also noted the possible effect of instructional method on representational performance. This macro-level effect will be the subject of the next chapter, in which we directly compare the courses studied here with a 202 course taught by the reform-style instructor from the 201 course.

## Chapter 6: Effect of different instructional environments on representational skills

(This chapter is based on a paper published in the PER portion of the Physical Review.[25])

The last chapter[24] showed that physics students' problem solving success depends on the representation of the problem. That is, whether one represents a physics problem in terms of words, equations, graphs, or pictures can have a significant impact on student performance on that problem. We also investigated whether providing students with a choice of representational format would help or harm their problem solving performance. That is, we asked whether students are capable of determining which representations are most helpful to them on problems of a given topic. The results from our study were complicated: Student problem solving performance often depended very strongly on whether or not they had a choice of problem format, but the strength and direction of the effect varied with the representation, the subject matter, and the instructor.

### *Previous study: Summary and questions raised*

Our first study took place in two different large-lecture algebra-based introductory physics courses using the same textbook.[85] Physics 201 covered mechanics and was taught by a reform-style instructor who made heavy use of concept tests, clickers, and well-integrated lecture-demonstrations. Physics 202 covered electromagnetism, optics, and atomic physics, and was taught by a traditional-style professor, though he used clicker questions to a limited extent.

In both courses, we found ample evidence that student problem-solving performance can (but does not necessarily) depend strongly on representation. This finding naturally leads to micro-level questions regarding when and how problem solving depends on representation. In some cases, a problem feature was present in one representation but was either not present or not as attractive in another representation. We also observed in interviews that student problem solving strategies depend sensitively on problem representation, a result that we will consider in depth in the following chapter.

The effect of providing a choice of problem formats was complex. In some cases giving students a choice of format resulted in significantly increased performance with respect to the random-format control sections. In other cases, the choice of formats resulted in significantly decreased performance. These choice/control differences were often quite strong, and also varied by topic area. A group that chose one format would do significantly better than the corresponding control group for one topic, and then would do significantly worse on the next quiz on a different topic but in the same format. As we saw in Table III, the data show that many of the choice/control splits are quite pronounced, and that the direction of the effect can change from topic to topic. Especially interesting is the fact that the 202 course showed much stronger choice/control splits than the 201 course. This led us to macro-level questions: Was this qualitative difference in performance data a result of the different instructional style, the different content area, or some combination? Also, what was it about the difference in content or instruction that was associated with the difference in performance?



In this chapter, we shall address these questions in two parts. As we noted previously, these courses were taught by different professors and covered different material, and so the differences observed could conceivably be explained by differences in instruction, differences in content, or some combination. Our previous hypothesis was that the much different approach of the reformed 201 course resulted in students having a broader set of representational skills. Thus, whether or not they received their "preferred" representation made less of a difference (positive or negative) in performance. The first part of this chapter begins to test our hypothesis by separating out the effect of instruction from the effect of content. We repeated the Physics 202 study in the spring of 2005, when the course was taught by the reform-style professor who had been in charge of the 201 course in the earlier study. We predicted that given the same quizzes, the choice/control splits would be much weaker than they were in the original study with the traditional 202 professor.<sup>i</sup> This prediction held true, leading to the second part of the chapter, in which we analyze the specific differences in representation use in these classes in lectures, exams, and homeworks. The results allow us to conclude that a pervasive use of multiple representations in a physics course can support students learning a broader set of representational skills than students in a representationally-sparse environment.

---

<sup>i</sup> Note that the students in the reformed 202 section were largely those from the reformed 201 section. The students in the traditional 202 course had, for the most part, taken a traditional 201 course as well. Thus, when we compare the two 202 sections, we should be aware that any significant differences may be the cumulative result of two semesters of instruction, not one.

*Methods: Study homeworks and quizzes*

The study was conducted in the same way, using the same homework and quiz problems, as the 202 study in the previous chapter. Note that in all quizzes, the statements and distractors mapped from one quiz representation to the next. The homework problems could not be made completely isomorphic since the students were assigned all of them, but they were designed to be as similar as was reasonable. For instance, some problems asked students how the radius of a Bohr-model electron's orbit varied with the orbital number. The different questions would be the same except for the representation and for which particular transition ( $n = 3$  to  $n = 2$ , for example) was under consideration.

Reform Course, Spring 2005	Verbal	Math	Graph.	Pictorial
202 Diffraction/Interference HW (N = 332)	0.44	0.36	0.39	0.46
202 Spectroscopy/Bohr HW (N = 341)	0.63	0.60	0.55	0.48

Table VI. Fraction of students answering a homework problem correctly, sorted by representational format and topic. Standard errors of the mean vary but are on the order of 0.02.

*Data: Study homeworks and quizzes*

In Table VI, we see the performance of the reformed Physics 202 students on the pre-recitation homeworks. Notably, the fractional and statistical differences between the graphical and pictorial questions on the Bohr model homework (a key comparison in the previous chapter) is smaller than it was in the traditional 202 course (0.55 vs. 0.48 instead of 0.77 vs. 0.62, significant at  $p = 0.05$  versus  $p = 0.006$ ).

In Table VII, we see the performance of the students on the diffraction and spectroscopy recitation quizzes, sorted by representation and by whether the students were in a format choice or control group. Performance variation across representation was generally less statistically significant in this course than it was in the traditional 202 section, including both quizzes and homeworks.[24] We note here that the traditional students noticeably outperformed the reform students on the 202 diffraction quiz. For the sake of replication, the exact questions that were designed for the traditional 202 course were given to the students in the reform 202 course.

Reform Course	Control (random format) group			
Spring 2005	Verbal	Math	Graphical	Pictorial
202 Diffraction	0.19 (46)	0.35 (46)	0.14 (46)	0.18 (44)
202 Spectroscopy	0.59 (46)	0.39 (46)	0.57 (42)	0.54 (46)
202 Diffraction (dist)	0.33	0.45	0.44	0.45

Reform Course	Format choice group			
Spring 2005	Verbal	Math	Graphical	Pictorial
202 Diffraction	0.15 (16)	0.57 (34)	0.13 (37)	0.21 (77)
202 Spectroscopy	0.41 (17)	0.32 (25)	0.49 (37)	0.52 (89)
202 Diffraction (dist)	0.26	0.22	0.41	0.32

Table VII. Quiz performance of students from the random-format recitation sections (top) and from the recitations sections that had a choice of formats (bottom). The number of students taking a quiz is in parentheses. The quiz topics are diffraction and spectroscopy. Standard errors vary and are not shown. The last line indicates how many students chose a particular distractor on the diffraction quiz, as discussed in the text.

The two 202 courses placed different emphases on the different subtopics available, and the reform section spent very little time on double finite-width slit diffraction. Student comments and performance suggest that most students treated

this as a double infinitesimal-width slit problem. One of the distractors is correct for such an interpretation of the problem, and student selection of this distractor is noted in the (dist) line of Table VII (performance is noticeably higher). Because of the different emphases on specific content, comparisons of absolute performances across courses are not likely to be valid, and so we focus on relative student performance across different representations and across choice and control groups. Note, for example, that if absolute performance is too low, relative performances will necessarily be very even regardless of student representational skill. Here, only one of the four reform course quizzes (counting two in 202 presented here and two in 201) shows very low absolute numbers, and these are not present if one considers the likely student misinterpretation described above. Thus, we are confident that the lack of choice/control splits in the reform courses is a genuine effect.

Quiz subject	Verbal	Math	Graphical	Pictorial
202 Diffraction, Reform	X	0.06	X	X
202 Spectroscopy, Reform	X	X	X	X

Table VIII. Statistical significance of the quiz performance differences between the format choice and control groups in the reformed 202 section. Numbers are p-values using a two-tailed binomial proportion test. X denotes a p-value of greater than 0.10. Bold indicates that the choice group had higher performance than the control group.

In Table VIII, we show the statistical significances (p-values using a two-tailed binomial proportion test) of the differences between the choice and control groups on the different topics and formats. Note that these data are essentially the same if one considers the correct diffraction quiz answer to be the distractor mentioned above (the math p-value changes from 0.06 to 0.03; all others remain

insignificant). When comparing these data to those in Table III, we see that they are very similar in character to the reformed 201 course, and much different in character than the traditional 202 course. This suggests that the choice/control splits (or lack thereof) are associated more closely with the instructor and course environment than with the content area. Thus, we analyze these environments in more detail with the goal of finding any differences in representational content and their uses.

### *Study II: Analysis of course environments*

The second part of this chapter involves characterizing the representation use in each of the classes under consideration: the reform 201, the traditional 202, and the reform 202. These courses had many components, including lecture, lab/recitation, exams, and homeworks. The homeworks included problems assigned and graded through a web-based automated system (CAPA[82]), and long-answer hand-graded problems. In comparing the courses, we judged the labs/recitations and CAPA-based homeworks to have very similar representational character. We thus focus our analysis on the lectures, exams, and long-answer homeworks. This approach provides two views of the class. We see how the use of physics representations was modeled for the students (the lectures), and how the students were held responsible for using physics representations themselves (homeworks and exams).

*Methods: Lecture analysis*

We videotaped several lectures from each of the three courses. The lectures covered the material probed by the quizzes and some closely related material. We chose three lectures from each course for analysis, with each set of lectures spread over different topics.

In order to quantify the different uses of representations in the different lectures, we divided each tape into one-minute sections. For each segment, we noted which representations were used significantly according to the following rubric:

- Verbal: Writing sentences expressing an idea or concept on the board; presenting and explicitly referring to a slide with verbal-only content for the sake of the point at hand (words surrounding mathematics are not counted).
- Mathematical: Writing equations; explicitly referring to equations for the sake of the point at hand; doing mathematics. Writing numerical data by themselves is not counted (for example, simply writing  $h = 6.636 \times 10^{-34}$  Js does not count).
- Graphical: Drawing or modifying a graph; explicitly referring to a graph for the sake of the point at hand.
- Pictorial: Drawing or modifying a picture; explicitly referring to a picture for the sake of the point at hand.
- Physical demonstration: Carrying out a physical demonstration.

Note that for lectures, we have added the representational category "Physical demonstration." We also noted which intervals include clicker questions. Finally,

any interval in which more than one representation was used was coded as a Multiple Representations interval (the Clicker category did not count for this purpose).

Because the professor is speaking during nearly every part of a lecture, we did not count spoken words towards the use of verbal representations. This is an example of a broader feature of this study: the privileged position of the verbal representation. Essentially every aspect of the course had some verbal component (even math problems include explanatory text), and so we necessarily have stricter standards as to what counts as verbal representation use compared to the other categories.

Once a lecture was coded, we calculated the fraction of the lecture that showed use of each of the representational categories. We then averaged across the three lectures from each class to obtain an average representational content for those courses' lectures. To test for inter-rater reliability, two individuals (the principal investigator and a researcher unrelated to this study) coded a subset of the lectures working from a written description of the coding standard. Results were extremely similar; the average difference between raters on each category was 1.3%. With approximately 50 one-minute bins per lecture, this represents a typical coding difference of one bin per representation per lecture. This difference is very low, which we attribute to the length and detail of the written coding standard.

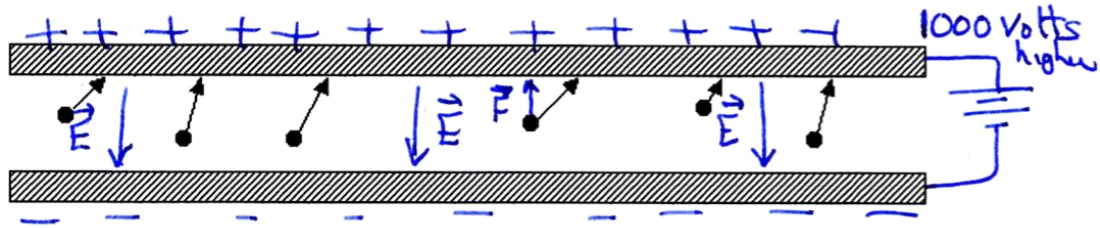
#### *Methods: Exam analysis*

Each of the three courses considered issued three exams, not including the final. The final exam took place after all the recitation quizzes and homeworks used in the study, and thus is not included in the analysis. We quantified the fraction of

each exam that could be described as verbal, mathematical, graphical, and pictorial in representation. We also quantified the fraction of each exam that explicitly required the use of multiple representations. The coding occurred as follows: On an exam, each subproblem is designated as having a verbal, mathematical, graphical, or pictorial component or some combination thereof. The total fraction of the exam composed of a particular representation was defined as the fraction of points possible that came from problems belonging to that representation. Problems with more than one representational component had their points counted in full towards each of the relevant representations; no effort was made to weight the components (for instance, we did not designate a problem as 80% pictorial and 20% mathematical, but rather as 100% of each). Both the problem presentation and the intended solution were considered. For example, a ten-point problem that requires a student to read information off of a provided graph in order to do a numerical calculation is designated mathematical/graphical, and ten points are assigned to both the mathematical and graphical categories. Thus, an exam that is rich in representations can have more than 100 points of representational content assigned to it in this analysis. Any problem that explicitly involves more than one representation has its points counted towards a Multiple Representations category as well. Once we characterized each exam in terms of its representational content, we calculated the average representational content of the exams in each of the courses.



An electrostatic air filter uses electric fields to attract charged particles of dust, pollen, etc. to one of two electrically charged metal plates.



The dust particles are first electrically charged by a small electric discharge and then blown along with the air left to right between the charged plates using a fan. The two plates are wired to a voltage supply so the top plate is 1000 volts higher in voltage than the bottom plate. The two plates are spaced 1 cm apart. The particles are strongly attracted to the top plate and stick there. They can later be removed by just wiping the top plate with a cloth.

a) How are the particles charged, i.e. are they positive or negative? Explain.

*Negatively charged since they are attracted to the positive plate.*

b) Indicate on the figure the direction on the electric field between the two plates and determine the magnitude of that electric field.

*E points downward; away from positive, toward negative*

$$|E| = \left| \frac{\Delta V}{\Delta x} \right| = \frac{1000V}{0.01m} = 100,000 \frac{V}{m} = 1 \times 10^5 \frac{V}{m}$$

Figure 8. Example exam problem with pictorial, mathematical, and verbal components. The problem is from a reform 202 exam, with a handwritten instructor solution

The representational categories are defined as follows:

- Verbal: For multiple choice questions, the available answers are phrases that are conceptual or qualitative in nature. For long-answer questions, there is an "Explain your answer" or similar written component.
- Mathematical: The problem requires numerical or algebraic calculation, manipulation, or interpretation, or other significant and explicitly quantitative reasoning.

- Graphical: A graph presents relevant information, or students have to construct a graph. Diagrams that have labeled axes have a graphical component.
- Pictorial: Students must draw or modify a picture, a picture contains needed information, or a picture meaningfully depicts the relevant physical situation. A picture of a CD accompanying a "plug 'n chug" problem about the wavelength of a CD-reading laser would not fit this standard.

In Figure 8 we see an example exam problem with solution. Part A of this problem was judged by the above standards to have verbal and pictorial components. Part A was worth 11 points (out of 100), and so there was an 11 point verbal component and an 11 point pictorial component for the sake of the averaging. Part B was judged to have mathematical and pictorial components.

*Methods: Homework analysis*

In addition to web-based CAPA problems (formulaic in nature for all classes studied here), students were assigned several additional homeworks requiring more in-depth work. Here, we consider those from the two 202 classes (traditional and reformed) for the sake of direct comparison. We use essentially the same coding scheme as with the exams. The representational content of each assignment is broken down and weighted according to point value. Then, the content of all assignments is averaged together. The reform 202 course had eight homeworks of this sort, and the traditional 202 had fifteen.

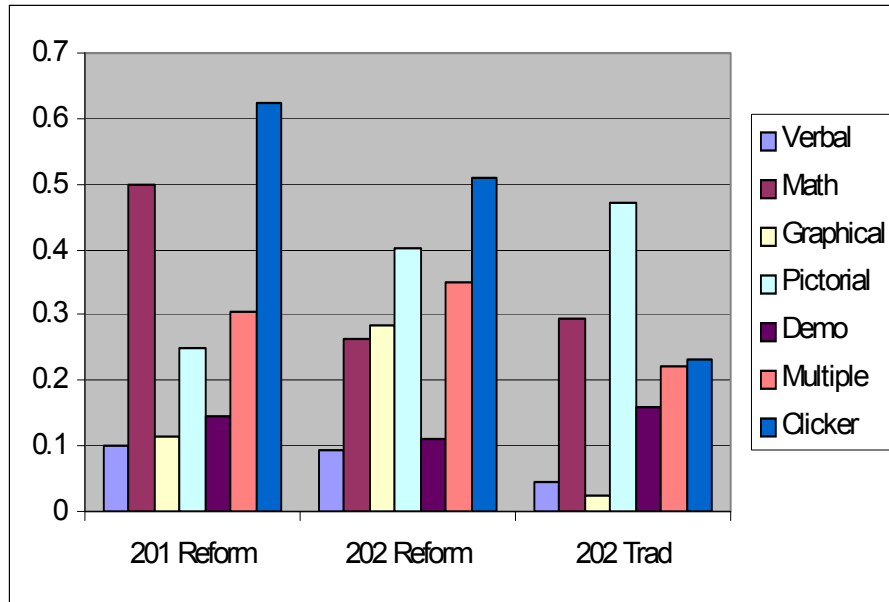


Figure 9: Representational content of the lectures for the reformed 201, reformed 202, and traditional 202 courses. "Multiple" category indicates use of multiple representations. Clicker category indicates fraction of class time involving questions that used a personal response system.

*Data: Lecture content*

In Figure 9, we see the representational content in the reformed Physics 201, reformed 202, and traditional 202 lectures according to the standards described previously. Differences exist between all three sets, suggesting (not surprisingly) that both instructor and content have a significant effect on representation use. That is, a particular instructor will use a different selection of representations in different courses, and in a particular course, different instructors will use different selections of representations. Most relevant to us is the comparison between the reform and traditional sections of 202. The reform section shows a broader selection of representations, with the verbal, math, graphical, and pictorial fractions summing to 1.04 versus 0.83 in the traditional section. We also see more use of multiple

representations (0.35 versus 0.22, significant at the  $p = 0.03$  level using a two-tailed binomial proportion test), and much more use of interactive Clicker questions (0.51 versus 0.23,  $p < 0.0001$ ). The Clicker category does not describe a representation in itself; rather, it tells us something about how representations are used in the course. The resolution of these data are limited by the interrater reliability (on the order of 0.01), and by the division of the lectures into one-minute blocks. Each data set contains three 50-minute lectures, or 150 blocks, suggesting a resulting imprecision on the order of 0.01.<sup>ii</sup>

*Data: Exam content*

In Figure 10, we show the representational content of the exams in the reformed Physics 201, reformed 202, and traditional 202 courses. These data show the average across all exams in each course, excluding the final exam. We see the fraction of the exam problems (weighted according to their point value) that were verbal in nature, mathematical, graphical, and pictorial. We also see the fraction of the exam problems that required explicit use of multiple representations.

---

<sup>ii</sup> We chose one-minute intervals as that seemed a reasonable order of magnitude for discrete portions of a discussion or explanation. Intervals on the order of several seconds result in all intervals containing single representations. Intervals of several minutes tend to include all or nearly all possible representations. As a check, we repeated the analysis with two-minute intervals. The figures for multiple representational use were modestly increased, the other categories increased less, and the relative class-to-class variations (which are of the most interest here) were essentially unchanged.

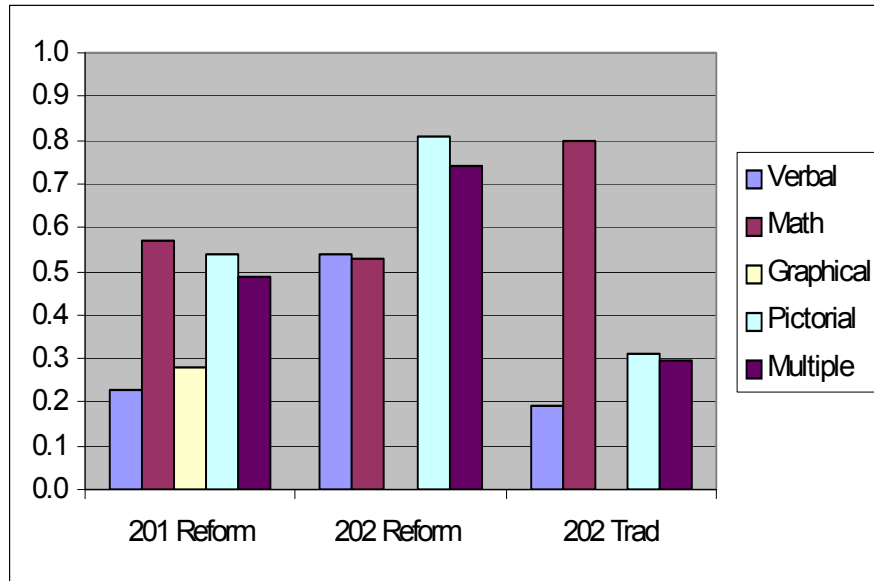


Figure 10. Distribution of representations used in the exams in the three courses studied here. Also includes fraction of the exam problems that required explicit use of multiple representations.

It is clear that the exams from the reform sections of 201 and 202 made use of a broader selection of representations than the traditional 202 section. Perhaps most striking is the difference in the proportion of multiple-representations problems, with 0.49 for the reform 201 course and 0.74 for the reform 202 course versus 0.30 for the traditional course. The difference between the reform 202 and traditional 202 figures is statistically significant ( $p < 0.0001$ , two-tailed binomial proportion test).

*Data: Homework content*

In Figure 11 we see the distribution of representations used in the reformed and traditional 202 course homeworks. The distributions are very similar, with the reform course making somewhat greater use of verbal representations and the

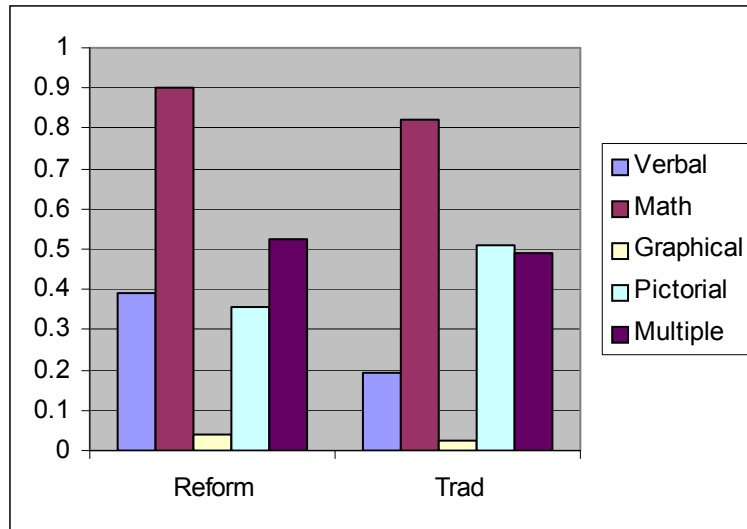


Figure 11. Distribution of representations used in the long-format homeworks in the two Physics 202 courses studied here (reform and traditional). Also includes fraction of the homework assignments that required explicit use of multiple representations.

traditional course making greater use of pictorial representations. The data are similar enough that we can make no claims regarding significant differences in representation quantity.

### *Discussion*

Our original hypothesis appears to be validated. The reform 202 course shows choice/control performance splits that are much more consistent with the reform 201 data than with the traditional 202 data. We thus conclude that these choice/control splits are associated more with instructional environment than content area. The course analysis data demonstrate that major components of the class (in particular the lectures and exams) were strikingly different in their representational character, with the reform content being richer and using multiple representations more frequently. We also believe that the reform course made more productive use

of these representations. For example, half of the reform 202 homeworks require students to estimate, while none of the traditional 202 homeworks do so. Estimation and calculation are very different applications of mathematical (and sometimes graphical) content. The way that the course used representations is not described by the data shown here, which only indicate how often categories of representation were used. We will make no attempt to quantify the productivity of representation use in this chapter, though we suspect that such a quantification would increase the differences shown (for instance, making visible a difference in the character of the long-format homeworks in each class).

This richer use of representations in-class is consistent with the notion that these students are learning a broader set of representational skills, which could explain the choice/control splits or lack thereof. With this broader set, working in a chosen representation as opposed to an assigned one has less impact on performance. It is also perhaps significant that the reform 202 section shows generally smaller performance variations across representation within the choice and control groups. Further, this interpretation allows for the fact that the direction of the choice/control splits can vary across topic. Student quiz comments (examined in the previous study) and student interviews not analyzed here show that students have definite opinions regarding their abilities and the relative worth of different problem representations and that these opinions are generally constant across topics. However, it would appear that their problem-solving skills in different representations are not constant across topics, especially in the traditional course. Thus, students are sometimes correct and sometimes incorrect when they assert that they are better with a particular

representation. This is in itself something of a meta-representational failure, though an understandable one, as physics students are not often asked to engage in such meta-representational tasks. Future work using problem-solving interviews will consider in more detail how accurately students assess their own representational competence.

We do not have evidence to claim that students in the reform sections were necessarily learning better meta-representational skills than the students in the traditional section. It is quite conceivable that these students were no better than those in the traditional 202 course at assessing their own abilities and evaluating the different representations available to them, but that their broader set of representational skills made any meta-representational failures less significant. Of course, neither do the data allow us to conclude that the reform 202 students were not learning better meta-representational skills. To our knowledge, no one has yet undertaken the research task of quantifying meta-representational competence in physics, though some work has characterized student meta-representational skills.[18, 19, 55]

We have suggested that a complete analysis of these courses would require attention to both micro- and macro-level features of the environment. Chapter 5 focused primarily on micro-scale, specific features, demonstrating how particular problem elements, topics, and representations could affect student performance. In this chapter we have taken an explicit look at the effect of instructional environment, considering macro-level influences which can impact the small-scale. For example, instruction can change what resources and skills students bring with them when



solving a particular problem. Here, we have inferred that the reform 202 students were more broadly skilled in using different representations. Thus, the students were less vulnerable to particular variations in problem representation. Also, the course analysis presented here demonstrates that the reform course explicitly valued the use of a variety of representations, often together. Such multiple representation use was demonstrated for the students on a regular basis in the lectures, and they were held accountable for such use in their own work, especially on the exams. It is quite likely that this broad-scale norm helped drive the development of students' specific representational skills, tying together the different contextual levels (micro and macro[75]) discussed here.

Notably, our access to micro-level information has been limited to aggregate data so far. In order to gain a clearer understanding of mechanisms and to validate our conclusions so far, it would be useful to have fine-grained studies of individual students solving these problems. Problem-solving interviews of this sort are the subject of the next chapter.

### *Conclusion*

Our results suggest that instructional environment can play a significant role in developing student representational skills as they apply to problem solving. Pervasive use of different representations and of multiple representations appears to have broadened students' representational skills. Unfortunately, without assessment tasks that are less sensitive to the topic coverage, we are limited in our ability to conclude that student representational skills are stronger in a reform-style class in an

absolute sense; all we are comfortable in asserting is that student problem solving skills varied significantly less from representation to representation. However, we consider it plausible that as these students develop broader skills, they also develop stronger skills overall. We suppose that if instructors wish to make increased representational facility a primary goal of their course, they can do so effectively. In the courses studied, representational richness spanned all (or at least most) aspects of the courses. We suspect that this might be required for a significant effect, but cannot be sure since we have not studied any cases where only some aspects of a course are representationally rich.

## Chapter 7: Fine-grained examination of representational and meta-representational skills

(The material in this chapter is the source for a publication in the Physical Review.[26])

In chapters 5 and 6, we presented the results of a three-semester study investigating student representational and meta-representational skills in large-lecture introductory algebra-based physics courses. Chapter 5 asked two major questions. First, to what extent (and why) does student performance on physics problems depend on the representational format of the problem? We found several instances where student performance was significantly different on problems that were essentially isomorphic, but posed in different representations. This supports previous results in PER,[9] and is consistent with other works in cognitive science (reference [86], for example). In our data, it appeared that subtle features of the problems were cueing students to answer differently, and that those features were specific to the particular problem and representation; that is, it was not simply a case of students being generally more competent with one set of representations than another. Unfortunately, the aggregate nature of the performance data prevented us from making specific inferences about how problem features influenced students' answers in more than a couple of cases.

The second question was whether performance would improve if students were given a choice of representation. This is a meta-representational question: Do students know enough about representations and about their own abilities to make

productive choices? Our results were complex. The first course studied was the second semester of a sequence taught by a traditional professor, Physics 202. In this course, the students who were given a choice of representation (the choice group) often performed either significantly better or significantly worse than students who were assigned a quiz (the control group). Whether the choice group did better or worse varied with representation and topic, but not in a predictable way. In the following two semesters, we performed the study again in Physics 201 and 202 courses taught by a professor who made significant use of PER-based reforms. In these courses, the choice/control splits were nearly nonexistent.

In chapter 6, we attempted to explain the presence or absence of choice/control performance splits and to describe the representational character of the reform and traditional class environments studied.[25] Analysis of the course components (including lectures, exams, and homeworks) suggested that students in the reformed physics courses were being exposed to and held responsible for using a broader variety of representations. In addition, each study's recitation quizzes asked for students to write briefly about which representations they thought that they would perform best with. Student opinions regarding their representational skills appeared to be constant across topics in all three of the courses studied, though performance data suggested that this was not always accurate. These results taken together suggested a possible explanation of the choice/control splits. If the representation-rich reform course environment was leading students to develop broader representational skills, then any meta-representational weaknesses (inaccuracies in assessing one's abilities with regards to a particular quiz, for instance) would have

less impact. That is, whether students received their preferred representation or not would have less effect on their performance than in the traditional course where, perhaps, student skills were less broad.

The above studies made considerable progress towards addressing the questions asked, but the aggregate nature of the available data was a significant limitation. In this chapter, we present data from sixteen student interviews that investigate the above in more detail. These interviews had two goals. First, we hoped to gain a deeper understanding of how problem representation affects performance. We address this in two parts. The original aggregate data suggested that oftentimes student responses would be cued by subtle features of the problem or representation at hand. These cueings are much more accessible in detailed problem-solving interviews, and we find numerous examples of how particular, representation-dependent features of problems can cue students differently, strengthening our earlier conclusions. Next, in watching students solve these problems it became clear that students' solution strategies varied with problem representation. Some prior work in math and science education has investigated the effect of problem representation on student strategies.[47, 87] Koedinger[47] found that young algebra students often chose different problem-solving strategies for problems in different representations (word problems vs. symbolic problems, for example). Such representation-dependent strategy variations could begin to explain the different performances we observe in students solving different physics problems. We quantify the variation of student strategy in these interviews, and discuss the effect strategy variation had on student

performance. The data suggest that students who are more consistent in their choice of strategy perform better.

Our second goal was to validate our conclusions regarding student assessments of different representations and of their own representational skills. These conclusions were key to our arguments and were interesting in their own right. In our previous studies our data included only written student assessments. In these interviews, we could question students more thoroughly, and determine whether our readings of student written remarks were accurate. Since the interviews took place several weeks after the original in-recitation study quizzes, we could also test whether students' opinions of representations and of their own skills were stable. We demonstrate that their opinions are fairly robust and that they do not correlate well with their actual performance, which is consistent with our prior suppositions regarding weak meta-representational abilities.

*Methods: Student interviews*

We recruited eight students from each of the reform 201 and 202 courses, with one student participating in both 201 and 202 interviews, for a total of fifteen unique individuals. Students were solicited for these interviews via course email near the end of the semester and were paid for their time. The student volunteers were primarily from the top two-thirds of the class, with final grades ranging from A to C. Seven of the students were male and eight were female.

The interviews were clinical (in the style of [88]), lasted between 30 and 45 minutes and were videotaped. In the interviews, students were assigned a number of

quiz problems identical to those found in the recitation quizzes and were asked to work aloud. The problem order was varied from interview to interview. The interviewer (the thesis author) did not provide assistance except for problem clarification, and generally allowed the students to work without interruption except for reminders to think out loud and requests for further explanation of what the students did and why. After the students solved the quiz problems, the interviewer asked them a number of questions regarding which representations they found most useful and why. Students were not told whether they answered a question correctly until after the interview.

With two subject areas and four representations, there were eight problems available in total for each interview (see Appendix A for a complete problem listing). Students completed anywhere from two to eight problems in the time allowed, with an average of approximately five problems per interview. This means that students often solved the same problem in different representations in the same sitting. Students rarely expressed explicit awareness of this until prompted at the end of the interview. Nevertheless, this could have led students to be more consistent in their problem-solving approaches than they would have been if they had approached each problem uninfluenced by any others. Students were generally not very consistent in their approaches across representation, so we do not consider this a serious issue.

We analyzed these interviews in three ways. First, we coded each student answer as correct or incorrect. Second, we coded each student's strategy in solving each problem, noting whether it was qualitative or quantitative, which concepts it made use of (energy vs. force, for instance), and flagging any special features

(analogies to other material, for instance). Third, we flagged instances where students expressed a favorable or unfavorable view of a particular representation. More detail on the interview analysis can be found in the Analysis sections, as can sample codings with interview excerpts.

### *Data and Analysis*

In this section, we present and analyze the data from the student interviews. This presentation has two parts, each addressing one of the two research goals identified in the introduction.

#### *Effect of problem representation on performance*

As noted above, we find two major ways in which problem representation can affect performance. Student problem-solving strategies can vary with problem representation, and students can cue on particular, often representation-dependent problem features when selecting their answers. We do not consider it likely that these two dimensions include all the ways in which representation can affect performance, nor do we consider them perfectly distinct (for instance, it appears that student choice of strategy is often cued by representation-dependent problem features). Nevertheless, these categories (which emerged from our analysis) are useful in organizing the available data. We begin by presenting the data on strategy variation, including several examples of student strategies for reference in later sections.



### *Student problem-solving strategies*

When confronted with different representations, students in our interviews ranged from being very diverse in the strategies they used to being very consistent. In this section we examine the selection of strategies employed by students in our interviews.<sup>iii</sup>

As noted above, we coded student strategies according to major problem features such as whether the solution was qualitative or quantitative. The strategy divisions varied from problem to problem; for instance, students' strategies for solving the spectroscopy problems could be binned according to whether they used an energy approach or a wavelength approach (or, rarely, both), and students' strategies for solving the spring problems could be binned according to whether students worked in terms of force or in terms of energy. We also flagged unusual (within this sample) approaches such as analogies to objects not directly related to the problem (planets, flashlights), or the non-quantitative use of equations as conceptual support (which will be discussed more later).<sup>iv</sup>

We begin with interview excerpts of problem solutions from two students whose strategies were varied (one student who is correct across all representations used and one who is correct on only one representation), and one student whose strategies were consistent. Student names are pseudonyms. In the following, we will indicate a student's grade in parentheses the first time we mention them in an

---

<sup>iii</sup> Note that we use the term "strategy" broadly, and use it to include the problem features that a student chooses to focus on in addition to the overall solution plan.

<sup>iv</sup> Strategy categorization is necessarily somewhat subjective, but in nearly all cases, the categories used were unambiguous (either students use quantitative calculations or they do not; either they mention wavelength or they do not). Thus, we had only one researcher code and verify these data.

example. Our analysis focuses primarily on student performance on study tasks, and not on their in-class performance. Nevertheless, this information may be of interest.

*Examples - student strategy selection*

*Adam - Varied strategies, mixed success (course grade: B-)*

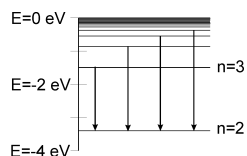
Adam solved the graphical, mathematical, and pictorial versions of the spectroscopy quizzes given in Physics 202. In Figure 12 we show the graphical and pictorial formats of the spectroscopy quiz. Note that the question statements and answer choices map closely from one format to the other. The mathematical format was similar, and asked students to calculate the difference between the  $n = 2$  and  $n = 4$  energy levels given a change in the ionization energy. Adam solved the graphical problem using an analogy to gravity. He stated:

...probably means it's going to be more tightly bound to the nucleus so the, levels are probably going to be lower than they would be if there's a lower ionization level I guess. If we were to reduce the gravity constant so, you know, mass would have less force than it actually does, you'd wind up with wider orbits, I suppose.

Adam chose answer B, which is incorrect and shows an increased spacing, analogous to wider orbits.

**Spectroscopy Problem -- Graphical Format**

The energy level diagram below shows the electron transitions that lead to the Balmer series:

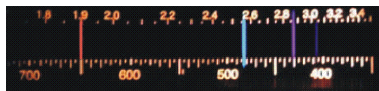


Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus and the ionization energy is 11 eV instead of 13.6 eV. Choose the graph that best represents what the new energy levels would look like.

- A)
- B)
- C)
- D)
- E) Something else.

**Spectroscopy Problem -- Pictorial Format**

The Balmer series of spectral lines is shown below as seen through a spectrometer:



Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus and the ionization energy is 11 eV instead of 13.6 eV. Choose the picture that best represents what the new spectrum would look like.

- A)
- B)
- C)
- D)
- E) Something else.

Figure 12. Graphical and pictorial representations of the spectroscopy quiz given in Physics 202.

Adam then solved the mathematical format. This solution was based on a proportionality argument backed by an equation: The difference in spacing should be that calculable from  $E = -13.6/n^2$ , but scaled by the factor  $11/13.6$  (the ratio of the

hypothetical and real ionization energies). Adam performed the calculation and selected the correct answer, C.

For the pictorial format quiz, Adam used a wavelength picture:

Well, if you've got a lower ionization energy, that means that the photons that are released when it's ionized are gonna have to have lower energy, which means they have to have a longer wavelength, so we're going to see a spectrum that's redshifted slightly compared to what we have normally.

Based on this argument, Adam chose the correct answer, C.

*Betty: Consistent strategies, consistent success (course grade: B)*

Betty solved the verbal, mathematical, and graphical versions of the pendulum quiz (see Figure 13 for the verbal and pictorial problems). In each case, she used the equation for the period of a pendulum to support her reasoning, whether that reasoning involved an exact calculation or not. Regarding the verbal format quiz, Betty said:

I believe the equation for period is like,  $T = 2\pi\sqrt{L/G}$  or something [writes  $T=2\pi\sqrt{(L/G)}$ ]. So if  $L$  is under the square root, um, then it would be the square root of that it would be  $2L$ , so, after one second it should be halfway from where it started.

**Pendulum Problem -- Verbal Format**

I set up a pendulum in front of you and pull it back (to your right), and then let it go. The pendulum takes one second to reach the point opposite from where it started.

Now I lengthen the pendulum's string until it is four times as long as it was, with the mass unchanged. I pull the pendulum back to the right again (far enough that the string is at the same angle as before), and let it go. Where is it after one second? Circle the correct answer.

- A) Straight up and down, and moving left.
- B) Opposite from its starting position.
- C) Straight up and down, and moving right.
- D) Back in its starting position.
- E) Somewhere else.

**Pendulum Problem -- Pictorial Format**

I pull a pendulum back to the position shown below on the left and let it go. It takes one second to swing into the position shown below and on the right.

Start:



After one second:



Now I change the pendulum so that it is four times as long as before, with the same mass. I pull the pendulum back to the same side to the same angle as before and then let it go.

Select the picture that corresponds to the position of the new pendulum after one second. If the pendulum is straight up and down, select the picture that indicates the correct direction of the motion.

A)



B)



C)



D)



E) None of the above.

Figure 13. Verbal and pictorial representations of the pendulum quiz given in Physics 201.

Betty's written calculations coupled with the above indicate that she decided that a quadrupling of the length of the pendulum would halve its period, resulting in the pendulum traveling half as far in a given time interval. She selected A, the correct answer.

Betty began the mathematical quiz by using the same period equation to determine the pendulum's position, giving no obvious attention to the  $x$  vs.  $t$  equation presented. She then used the  $v$  vs.  $t$  equation provided to correctly determine the sign

of the velocity (though in the process she made two offsetting sign errors), and reached the correct answer.

Betty's solution to the graphical pendulum quiz was extremely analogous to the above solutions, once again using the period equation to make a proportionality argument backed by an equation:

...extending the pendulum to four times as long, so since the period is, er T is  $2\pi\sqrt{L/G}$ , four times as long so T is related by being two times as great... If the period is two times as long, at one second it would be half as far, so it would be at this point, zero.

Again, Betty used this reasoning to select the correct answer, A.

*Carmen: Varied strategies, consistent success (course grade: A)*

Carmen solved the mathematical and pictorial versions of the quiz on springs given in Physics 201, shown in Figure 14. Her approach to the mathematical problem involved little calculation, and was based in part on an energy argument:

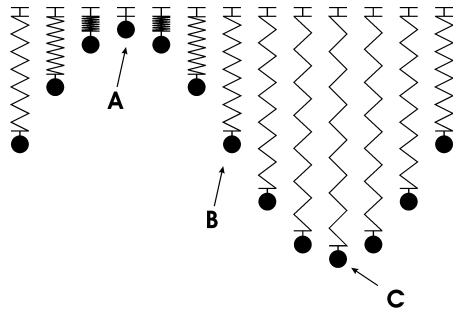
Carmen: ...because, okay, because that's where it's at rest, so... lowest energy... I think it's that. I don't know how to explain it.

Interviewer: What do you mean by 'at rest'?

Carmen: Where it's at rest when the ball is hanging.

**Spring Problem -- Pictorial Format**

A ball on a hanging spring is oscillating up and down as shown in the following snapshots.



At which point is the ball moving the fastest?

- A) The ball is moving fastest at point A.
- B) The ball is moving fastest at point B.
- C) The ball is moving fastest at point C.
- D) The ball is moving fastest at some other point (not A, B, or C).

**Spring Problem -- Mathematical Format**

A ball is hanging from a spring at rest at  $y = 0$  cm. The spring is then compressed until the ball is at  $y = 5$  cm, and is then released so that the ball oscillates. Up is in the positive- $y$  direction. At which point  $y$  is the ball moving fastest? Note that

$$K = \frac{1}{2}mv^2 \quad U_{spring} = \frac{1}{2}k(y - y_0)^2 \quad \text{and} \quad U_{gravity} = mgy$$

where  $y_0$  is the unstretched length of the spring.

- A)  $y = -5$  cm
- B)  $y = 0$  cm
- C)  $y = +5$  cm
- D) The ball is moving fastest at some other point.

Figure 14. Pictorial and mathematical representations of the quiz on springs given in Physics 201.

Carmen's reasoning is difficult to infer precisely, but based on this statement and additional follow-up questions, it appears that she was arguing that the ball will be moving fastest when the spring is at the equilibrium position, where the spring potential energy is lowest. This is the position at which kinetic energy is a maximum

when the spring is moving, though she made no explicit mention of kinetic energy. She selected answer B, which is correct.

Carmen solved the pictorial version using a force and acceleration argument, with no mention of energy:

It makes kind of more sense pictorially, because you know it's stopped here (points at the top of the motion) and that's going to be accelerating, and it'll accelerate until it's at the point it was at rest. And then, the tension of the spring, I don't know what it's called, will start causing it to decelerate, so it's going to be fastest at that point.

Carmen selected the correct answer, B. Notably, the pictorial format of the quiz explicitly depicts stretched and compressed springs, which could perhaps be associated more easily with forces than energies. Also, the mathematical format includes energy equations but not force equations. While Carmen did not make explicit use of these equations, it is plausible that their presence could have cued an energy argument.

Interestingly, no student made reference to both force and energy when solving any one problem. This is consistent with the findings of Sabella and Redish,[89] who demonstrate that even advanced students can struggle when moving from force to energy perspectives and vice versa.



### *Strategy selection and performance*

We next ask whether variation in a student's problem-solving strategies is associated with his or her performance. For this, we need to be able to describe a student as being generally varied or generally consistent in their strategy selection. Each student solved problems in two different topic areas. We consider strategies for quiz problems within a particular topic to be different if they have noticeably different qualitative features. Different features include arguing in terms of energy versus force or using an equation-based argument versus a qualitative proportionality argument. If the number of strategies a student employed was greater than half the total number of problems solved within a topic, we designated that student as varied in strategy choice. For example, a student that solves three representations of each quiz, for a total of six representations, would be designated varied if he or she used four or more strategies in total. This could be either two strategies for each quiz topic, or three for one and one for the other. If the student used only two strategies, (one for all representations of each of the two quizzes), we designate the student as consistent. If the student fits in neither category or uses strategies that are not easily categorized as distinct or similar, we designate the students as mixed-state with respect to strategy.

Of the eight Physics 201 students interviewed, seven solved four or more quiz problems. One solved only two problems, and will not be considered in this analysis (designation by the above standards is impossible). Of these seven students, two (including Betty from the above) were designated as consistent. Five were designated

as varied (including Carmen). None were mixed-state. The two consistent students solved a total of eleven problems, answering nine correctly, for an 82% success rate. The five varied students solved a total of 24 problems, answering 13 correct, for a 56% success rate.

	Varied	Consistent
Physics 201	56% (24)	82% (11)
Physics 202	50% (32)	100% (6)
Overall	52% (56)	88% (17)

Table IX. Success rates for students using a variety of problem solving strategies on different representations of a problem versus students consistently using the same strategy. Parentheses indicate the number of problems solved by students in a category. Data are presented for each class and for both classes averaged together. The performance difference in the Overall category is significant at the  $p = 0.007$  level.

Of the eight Physics 202 students interviewed, all but one solved six or more problems; the other solved four. Of these eight students, one was designated as consistent. Five were designated as varied (including Adam). Two were mixed-state. The consistent student solved six of six problems correctly, for a 100% success rate. The five varied students solved 16 of 32 problems correctly, for a 50% success rate. When we average together the data for the two classes, we see that the consistent students have a success rate of 88% over 17 problems, and the varied students have a 52% success rate over 56 problems. The standard deviations for the averaged data are 11% and 22%, respectively. These average success rates are different at a  $p = 0.007$  significance level using a two-tailed binomial proportion test, and indicate that

of the students studied here, those who are consistent in their problem-solving strategies are outperforming those who are not. We do not consider this an obvious result; one could easily imagine that students who tailor their solutions to the representation at hand would perform better than those that do not. We also cannot be completely certain as to why the more consistent students are performing better (nor can we claim that this pattern would hold in general, outside this sample), though it may be that these students have a better abstract understanding of the problem at hand and thus are less sensitive to representation-dependent problem features. The major data regarding strategy variation are summarized in Table IX.

#### *Representation-dependent student cueing*

In previous chapters, we claimed that subtle, representation-dependent features of a problem can (but do not necessarily) have a significant impact on student success. The aggregate nature of the prior data limited us to only a few examples of this representation-dependent cueing. Here, we bolster our earlier claim with additional examples made possible by the student interviews, and begin to organize these examples into categories that emerged in the analysis. Note that these categories are intentionally narrow. We are not trying to claim that these observed behaviors are ubiquitous; rather we wish to establish their existence for some combinations of representation, topic, and student.

### *Literal interpretation of graphs and language*

In several cases, students appeared to interpret problem features over-literally, drawing inappropriate conclusions or making inappropriate associations that led to incorrect answers. This could include literal reading of actual language, or literal interpretation of other representations (for instance, assuming that a graph feature that is positioned lower on the graph indicates a smaller quantity). Two examples of these literal readings follow. As we saw above, Adam solved three different representations of the spectroscopy quiz problems. He handled the mathematical and pictorial problems expertly (in the opinion of the interviewer), but was incorrect in the case of the graphical problem. Adam's solution to this problem used an analogy to gravity, which was followed by the comparison of energy levels to planetary orbits. We note that the graphical representation was the only representation of this problem that was oriented vertically, with higher-numbered levels being placed physically higher than lower-numbered levels. Adam appeared to cue on this: "probably means it's going to be more tightly bound to the nucleus so the, levels are probably going to be lower than they would be if there's a lower ionization level I guess." We also note that in another part of the interview (not shown) Adam stated that this material connects strongly with what he knows in chemistry about electron "orbitals." We speculate that here, a student who is otherwise quite skilled with the material at hand was prompted to respond incorrectly by a literal interpretation of the vertical arrangement of the levels, which was perhaps reinforced by an association between the word orbitals and the concept of planetary orbits. Student Doug (B-) exhibited a similar pattern. Doug solved the pictorial version of the spectroscopy quiz first. He

expressed uncertainty, but used appropriate reasoning and selected the correct answer. Later, he solved the graphical version, exhibiting correct reasoning at first:

If the electric charges are weaker, and the um electrons aren't held as closely to the nucleus, to the atom, that probably means the distance between... going from  $n, n=1$  to  $n=2$  is shorter. ... In doing this I've been trying to, um, think if it through like chemistry to help me kinda understand ... I've been using  $n=1$   $n=2$  as orbital levels, so as you um increase, your  $n$ , to higher energies. So if it's lower energy, then, the amount of energy to go from  $n=3$  to  $n=2$  or  $2$  to  $1$  would be smaller.

Doug then appeared to make a 'lower means lower' type of error when choosing between the available answers:

I don't think the right answer is on here, to be completely honest (circles none of the above). Because I think that you would have all of your lines slightly smooshed closer and *at a lower energy*. [italics added]

Despite having a fairly complete grasp of the problem, and despite the energy axis being clearly labeled, Doug focused on the physical position of the lines in the representation, and concluded that lower energy lines needed to be below the higher energy lines. Note that there is some potential for confusion here: Since the energy of these bound states is negative, it would be technically possible to use the word

"lower" to refer to states that occupy a lower position on the graph. However, student explanations and answer choices indicate this was not their intent.

These over-literal readings call to mind the "what you see is what you get" (WYSIWYG) knowledge element proposed by Elby,[90] where students interpret a representation in the simplest, most literal way possible (a bump on a graph corresponds to a hill), even if further reflection demonstrates that they 'know' the material well. This WYSIWYG element is a representational analog of the phenomenological primitives (or p-prims) described by diSessa,[57] which include such basic reasoning elements as 'lower means lower'.

#### *Relative versus absolute feature positioning*

Students solving the graphical or pictorial versions of the spectroscopy quiz usually did so in one of two modes. They either viewed the problem in terms of the relative positions of pairs of features (for example, the spacing between particular energy levels or spectral lines), or in terms of the positions of the sets of features as a whole (noting that the entire set of energy levels compressed or expanded, or that the entire set of spectral lines redshifted or blueshifted). We describe this as a focus on feature pairs versus entire feature sets. Students solved a total of five pictorial spectroscopy quiz problems, and seven graphical spectroscopy problems. In four of the five pictorial problems, student language made it clear (for example, "I'd probably pick the same set of spectral lines but at a lower energy") that they were focusing on feature sets, while in one case the student used a mixed approach. In four of the seven graphical problems, students appeared to be focused on feature pairs (for

instance, saying "n2 to n3, it's gonna be, it's gonna be smaller because uh, it's only 11 electron volts instead of 13.6") with one case of a student focusing on feature sets and two cases of students using mixed approaches. These numbers are too small for statistics to be used comfortably, but they suggest that the graphical representation of the problem might cue a different class of strategies than the pictorial representation. While each representation has a set of discrete features, we note that the discrete nature of the energy level diagram is emphasized by the arrows indicating transitions, the lack of background clutter, and the fact that the steps between levels have physical meaning that is discussed in class, while the steps between spectral lines in the pictorial format do not. It is thus perhaps reasonable that students would be likely to focus on a pair of discrete energy levels and their relative positioning while using the graphical representation, while treating the spectral lines as a single feature (or set of features) moving against a background when using the pictorial representation. This is in some ways similar to the results of Mestre et al.,[61] who found that showing students videos of pairs of balls rolling on tracks versus videos of individual balls tended to trigger different readout strategies, in which students considered either absolute or relative ball motions.

#### *Presence or absence of equations*

The presence of equations, not surprisingly, caused some of the interview subjects to attempt calculation-based solutions when they were unnecessary. The quiz problem on springs displayed in Figure 14 could be (and was) solved in a number of ways, and was relatively easy compared to the other problems. Students

Emma (B) and Mindy (A-) solved the pictorial version of this quiz quickly and correctly. Emma inferred the velocity of the ball from the size of the frame-to-frame change in the ball's position. Mindy recalled from lecture that balls on springs are moving fastest as they pass through their equilibrium position. Both Emma and Mindy solved the mathematical version of the spring quiz incorrectly. Mindy did not know what to do with the available equations and gave up after several minutes. Emma set the kinetic and potential energy equations equal to each other to derive  $v = \sqrt{2gy}$ , inferring from that that the ball is moving faster when its y coordinate is higher. She then selected answer choice C. In each case, the student arguably was using equations without thinking about why they were using them, a meta-representational failure frequently observed by Schoenfeld[54] (and, most likely, by any practicing instructor). Also, Sabella[91] notes that students take fundamentally different approaches when engaging in problems they perceive as qualitative or quantitative. From this perspective, Emma and Mindy can be seen as incorrectly judging problems to be strictly quantitative and engaging in inappropriate solution strategies as a result.

In contrast to the above, some students were more expert-like in their handling of equations. Most notable were those that used an equation qualitatively to support the reasoning involved in solving a non-mathematical representation of a problem. By 'qualitatively' we mean that the student used the equation without performing complete calculations. For an example, see Betty's solution to the pendulum problems in the earlier section on strategy selection. There, she used the formula  $T = 2\pi\sqrt{L/G}$  to find that quadrupling the length of a pendulum doubled the period, but



she did not calculate a specific number for T at any time. There were ten instances of students using mathematics qualitatively to support their solutions of non-mathematical representations, spread out over six students and five different combinations of topic and representation, with each topic represented except for the quiz on springs. The students solved the problem correctly in nine of these instances (90% correct). These six students had an overall success rate on all problems of 80%, compared to a success rate of 46% for the other ten students. This difference is statistically significant at the  $p = 0.002$  level. While the samples presented here are small, it appears that the more successful students (at least with regards to these tasks) are the ones capable of using (or willing to use) mathematics as a conceptual support, in addition to any calculation-based uses. This result is again consistent with Sabella[91] and others[3] who have shown that expert problem solvers integrate qualitative and quantitative approaches more often than novices when solving physics problems.

#### *Student assessment of representations and of themselves*

In our original study of this subject (chapter 5), we were able to examine student assessments of their own representational skills and preferences in aggregate, through comments solicited on the study quizzes. We found two notable results. First, students were generally fairly consistent in their representational preferences and assessments. Students that provided comments on both quizzes given over the course of the semester usually claimed to be good at the same representations on each quiz. Second, students' actual performance on the quiz and homework problems

correlated poorly with their assessments of their own skills. The robustness of their opinions and the lack of correlation between self-assessments and performance was a key feature of our previous arguments. The second major goal of the current study is to further validate the conclusions of the earlier chapters using the more detailed information available through interviews. Below, we present data relevant to each of the two results just discussed.

### *Student consistency*

In each interview, we asked students which representations they preferred to work in, and why. We also invited students to make any comments about the representations themselves that they wished. Follow-up questions were posed as needed to clarify student responses. We then compared students' interview assessments with the comments we requested on their original recitation quizzes. Of the fifteen students interviewed (one was interviewed twice), fourteen students had provided comments on the recitation quizzes. Of these, twelve were consistent in their assessments across the quizzes and interviews. We count as consistent any student whose quiz and interview statements (which were separated by several weeks) considered the same representations to be favorable and/or unfavorable. This standard for consistency includes students who favorably rate one representation on one quiz, favorably rate another representation on another quiz, and then describe each of those favorably during the interview. The other two students were somewhat consistent, but not completely. One student stated a preference for one representation (pictorial) on the recitation quizzes and stated a preference for two others (math and

verbal) during the interviews; we classify this stance as inconsistent, though not directly contradictory. Another student was consistent in her assessments of the math, verbal, and graphical representations, but indeterminate in her evaluation of the pictorial representation.

### *Correlation with performance*

We have found that student performances do not improve in general when students are given a choice of representational format in which to work. This result suggests that students' assessments of their own representational skills are not very accurate. In the interviews, we found that most of the students were quite consistent regarding which representations they preferred and/or would do best at. Including interview problems, pre-recitation homeworks, and recitation quizzes, these students have solved a large selection of study problems in different representations (an average of 12 problems per student). To correlate student performances with their assessments, we must be able to define whether a student has rated a particular representation favorably, unfavorably, or neutrally/not at all. Student responses are almost always unambiguously favorable or unfavorable. Two sample quotes (from students Doug (B-) and Tina(B-/C+)) are:

Sometimes the verbal ones are worded in ways which are hard to think about.

Math ones tend to be straightforward.

Given the choice, pictorial and mathematical is preferred. I hate graphical.

We analyzed the responses of thirteen students. The two students whose interview and quiz remarks were inconsistent were discarded, and the student who only provided interview assessments (and not quiz assessments) was included. For each of these students, we noted whether he or she evaluated a representation favorably, unfavorably, or neutrally/not at all. We then divided the set of problems that each student completed into subsets according to whether the student evaluated the representation favorably or otherwise, and found their average performance on each subset. By comparing student performance on each subset, we could describe each student as having done better or worse on problems in their preferred representation. As an example, student Nate (C) rated the mathematical and graphical representations favorably. He rated the pictorial representation unfavorably, and did not discuss the verbal representation in the quizzes or in the interview. We see examples of these ratings in the following subsection, which includes example remarks from Nate. Nate solved eight mathematical and graphical problems and answered three correctly. He solved seven verbal and pictorial questions, and answered three correctly. Thus overall, he performed worse (though only slightly) on problems that were in his preferred formats. Six of the thirteen students analyzed performed better on their favored representations than on other problems. The other seven performed worse. Thus, we see no correlation between their problem-solving success and their representation assessments. This result is consistent with the aggregate performance data in earlier chapters, where providing students a choice of problem representation did not produce a consistently negative or positive effect.

*Sample comments from interviewed students*

In this section we present sample comments from two students, one from Physics 201 and one from Physics 202.

*Betty - Physics 201 (course grade: B)*

Betty was consistent in her assessments of the different representations, and was also accurate in her assessment in that she performed better overall on the representations that she claimed to do better at. Betty was randomly assigned recitation quiz problems. In response to the quiz question "Of the four problem formats you saw on the pre-recitation homework, which do you think you would do best at, given the choice? Why?", Betty wrote "Mathematical - I relate to equations." She expanded on this in the interview:

Well, for me personally, I really um do poorly on conceptual questions, and I'll do a lot better if I have numbers and I can use equations and figure things out in that way.

Later, she said:

I don't like [the verbal quiz] at all because, I don't like it when you make me draw my own picture because sometimes I misunderstand what I'm reading and I draw the picture incorrectly which, affects how I do the whole problem. [the graphical quiz] really appeals to me because, if I'm looking at waves I

like to see this [the graph], and then I really like calculus so, I like to see like if you're telling me take the derivative of a graph, I can like draw it out for myself. ... Ideally, if I was doing a problem, I'd like them to have pictures and equations, because that would probably help me the most.

In two instances, Betty used mathematics to support non-mathematical quiz representations in the manner described earlier, which supports her claims that she is particularly comfortable working with equations. Also notable is the fact that she answered the mathematical version of the spring quiz incorrectly due to an incorrectly labeled picture that she drew (her reasoning was correct), which is also consistent with her assessment.

*Nate - Physics 202 (course grade: C)*

Nate was consistent in his assessments of the different representations, and performed slightly better on those he rated unfavorably or neutrally than on those he preferred. In response to the recitation quiz question, he wrote:

Mathematical because that is all we do with CAPA [a web-based homework system] and I have gotten used to it.

In the beginning of the interview, Nate was asked why he preferred the mathematical formats. His response was:

Kinda what I got used to during physics. Almost all of the tests use math.

Later in the interview, Nate is asked to provide input on all of the representations he has seen:

I guess, the pictures I find hardest because there's a lot goin' on with them. Um, the graphs are a little more manageable, just, I dunno, I guess there's, you know, there's numbers on the page and actual information given to you while this [the pictorial quiz] is kind of an interpret the information kind of thing. When I do like the math best is if you know the equation you just plug it in and you're set.

Betty and Nate were fairly typical in the length and style of their remarks.

### *Discussion*

This chapter set forth two goals. First, we planned to further investigate the means by which problem representation affects performance. We have a number of examples available here and in previous chapters. Second, we hoped to validate our conclusions regarding student assessments of different representations and of their own representational skills.

### *Effect of problem representation on performance*

We have a number of examples available in chapters 5-7 of how representation affects performance. The representation effects are complex and appear to depend on a number of things. Performance can be influenced by the particulars of the representation and how it is implemented, as in the cases where the vertically-oriented graphical spectroscopy problem triggered a 'lower is lower' misinterpretation that was not observed with the horizontally-oriented pictorial spectroscopy problem. Performance can also be influenced by prior student knowledge, including what topics a student has been taught in conjunction with the material at hand, or what other material a student feels comfortable in relating to the problem. An example is the case of Adam, who cited his chemistry experience when he drew an analogy between electron orbitals and planetary orbits. This analogy appears to have reinforced an error of the 'lower is lower' variety. It is also likely that class norms and expectations,[75] such as whether students have been taught to draw pictures in support of particular problems, can play a role. For example, we see here instances of students unnecessarily using equations without a clear understanding of why, and we speculate that this stems from the (not unreasonable) expectation that a mathematically-framed physics question will have a quantitative solution.

The above findings are complicated by the fact that most real situations will involve a combination of the factors discussed. Adam's error probably stems from both a misapplication of his prior chemistry knowledge and a p-prim-like[57] interpretation triggered by the specific features of the graphical problem representation. Emma's inappropriate derivation and use of  $v=\sqrt{2gy}$  for the sake of



solving the mathematical spring problem likely started with a course expectation, and may have been strengthened by the fact that this equation is commonly used in a different mechanics context, making it appear familiar once produced. (This interpretation is speculation, as the interview did not probe this association directly.) This complex dependence of performance on representation, student knowledge, and course norms is consistent with what we have observed before. Of course, not all students exhibited the same sensitivity to changes in problem representation. We found that many students used a variety of solution strategies in their interviews while some used very few. Furthermore, those who used fewer strategies appeared to outperform those who used many. It is possible that students who are more comfortable and competent with a topic are guided more by the topic than the particular representation of that topic; that is, these students may have a better grasp of the abstract concept behind the representation and cue on the deep rather than the surface problem features.[10] This is, of course, speculation at this point.

These results are reminiscent of previous results in PER that show significant differences in student approach and performance on qualitative versus quantitative problems. Such work includes that of Mazur,[65, 91] and the numerous University of Washington studies on student difficulties with various topics.[92] Indeed, some of the mathematical-format problems used in this study were quantitative in nature, in contrast to the more qualitative problems in the other formats (and, sometimes, in other questions using the mathematical format). However, the study here is focused on finer-grained problem divisions: Most of the problem representations were isomorphic, meaning that their intended solutions all included similar blends of

quantitative and qualitative work. Substantial differences in performance and approach emerged in spite of this coarse similarity, which some may find surprising. While it is possible to use the broader characterization of qualitative or quantitative problems, we observe in this study that the focus on representational format provides some insight into some of the factors that influence whether students choose to approach problems qualitatively or quantitatively. We also note that the converse may be true: If students view problems as quantitative or qualitative in nature, that can drive how they use representations (successfully or not).

#### *Student meta-representational skills*

Our second goal was to further support the claim that students assessments are relatively constant over time and across topic, and that these assessments do not generally correlate well with their performance. In comparing student quiz comments to their interview responses, we see that student opinions are indeed fairly consistent, at least over the course of the semester being studied. Furthermore, the performance data support our claim that these opinions are not particularly accurate in general. We must note here that while these student opinions regarding representations appear to be robust over the course of the study, we cannot take this as evidence that students generally come into a physics course with well-formed assessments of representations and of their own representational skills. In some cases they certainly do - Tina expressed such a strong dislike of graphs that the opinion must have existed before she answered the study questions - but we suspect that in some cases students were being asked questions such as these for the first time, and that they were generating

their opinions on the spot. Of course, neither do our data allow the opposite conclusion that students do not generally come into a physics course with well-formed opinions.

Meta-representational activities are not a part of a standard physics course, and we are curious as to what impact such activities could have. Studies have shown that attention to meta-level skills (including an explicit focus on epistemological issues) can have a positive impact on student performance in math[54] and physics[93] courses. Might explicit in-class attention to the uses and drawbacks of different physics representations in different classes improve student representational skills, or at least improve their own self-awareness? Tina was the only student to participate in an interview in both Physics 201 and 202, meaning that she had more formal opportunities to engage in meta-representational reflection than the other students. Perhaps not coincidentally, she was the only student to explicitly challenge her own representational assessments, as seen near the end of the second interview:

Tina: I picked the picture because pictures usually help, they usually make things easier. But not really, ... now that I'm looking at all of them. These two [indicates the mathematical and verbal spectroscopy problems]

Interviewer: Like, in your prior experience, you've found problems that have pictures to

Tina: Yeah, because they're visual, I mean, you could see what was going on, but the information is much more straightforward in these [verbal and mathematical quizzes]. This one [the pictorial problem] I gotta kinda and

figure out that this is wavelength; I need to hear it straightforward like in these. So. Yeah, I dunno, I always picked pictures because I thought it'd be easier and I don't think it ever was.

Interviewer: Really? Even in the past?

Tina: Well, I mean like, until school got hard [laughs].

Tina went on to suppose that pictures are useful when used in conjunction with other representations, such as written descriptions, equations, and graphs. It would appear that she moved from viewing pictures as intrinsically useful to viewing them as one of many tools that need to be used together in order to be most effective. This is, in our opinion, significant meta-representational progress, and we find it satisfying to see that given the opportunity a student can make such gains. Also note that the meta-representational failures observed in this study involve only a subset of student meta-representational skills. Other studies have found evidence of significant meta-representational strengths in students, especially with respect to their ability to generate new representations,[18, 19] so the picture of student meta-representational competence is not strictly negative.

#### *Instructional implications*

While this study did not have the goal of developing or testing new instructional materials or techniques, we can speculate as to the instructional implications of our results. We have found several cases in which student performance was significantly affected by which representation was used, and that

that impact could be tied to surprisingly small representation-dependent problem features that either cued an answer directly or changed the student's overall strategy. Instructor awareness of this sensitivity to representation would likely be productive. This sensitivity would also be relevant in test construction, though we note that PER-based assessments (the FCI,[11, 48] FMCE,[49] and BEMA,[94] for example) generally contain a variety of representations.

We also found that students in our study were not very successful when selecting between different representations to work in, perhaps because of their limited experience with such decisions. This result suggests that if an activity provided students with similar representational freedom (that is, the freedom to select between canonical representations or create new ones), such an activity would probably need to be guided to be more helpful. Or, as an alternative, students would need preparatory meta-representational instruction. We are not aware of any such activities in wide use today, but with the increasing interest in multiple representation use and metacognition, we can envision such an activity becoming more common in the future.

## Chapter 8: Section II summary and positioning for future work

### *Major results*

In the previous three chapters, we have investigated student representational skills at many levels. We have examined the impact of small-scale representation-dependent problem features, the variation of problem-solving strategy with representation, the role of instructional environment (representationally rich vs. sparse), and student meta-representational skills, with consideration of both aggregate data and in-depth interviews. We found four major results. First, student performance on physics problems did depend on problem representation, but the dependence was complex. Particular combinations of representation, topic, and student experience could result in much different performances, often as a result of different strategy selection. This finding suggests that it might be quite difficult to infer whether or not students understand a concept based on an assessment presented in only one representation of that concept.

Second, students (when asked) formed consistent opinions regarding which representations they handle best, but these opinions correlated poorly with their actual performance. Since traditional introductory physics courses usually specify the representations to be used on a problem, this meta-representational failure may not significantly impede student performance. On less-constrained (and more realistic) physics problems, lack of meta-representational skills might be more significant. To our knowledge this dependence has not been studied formally.

Third, we have found some evidence that instructional environment can improve student skills with physics representations. In a PER-informed course that made substantial use of a variety of representations and of multiple representations, students showed much less vulnerability to variations in problem representation than in a corresponding traditionally-taught, representationally-sparse course. Notably, interview subjects from the PER-informed course still showed very weak meta-representational skill, suggesting that such abilities do not “come along for the ride” when they learn to handle multiple representations.

Fourth, we have begun to outline some mechanisms by which representations can drive performance. Most prominent among these is the idea of representation-dependent cueing. This cueing can be fairly direct, which we inferred to be the case with the Bohr-model homework problems from chapter 5, where a distractor from a particular representation of electron orbits superficially resembles another, unrelated representation that students happen to have seen recently. This resemblance cues students to answer incorrectly. Or, a particular representation can cue students to use a problem-solving approach that they might not otherwise have used, as when in chapter 7 Emma correctly answers the pictorial format of the quiz on springs, and minutes later pursues an incorrect algebraic approach to the same problem posed in a mathematical format.

These results partially address the first three of our four thesis questions: We have established that representation matters, and have begun to chart out the way in which representations matters at both broad and fine scales. We can also relate our

results so far to our final goal, that of model development. We show how each of chapters 5, 6, and 7 feed into our four goals in Table X.

	Chapter 5	Chapter 6	Chapter 7
Effect of representation on performance	X	X	X
Dependence of rep. skills on environment		X	
Meta-representational competence	X	X	X
Expert vs. novice behavior			

Table X. Stated thesis goals, and coverage thereof in chapters 5, 6, and 7.

*Relation to model development*

At this point, we shall review how our results so far can support theoretical work. First, we reiterate the dependence of student performance on the whole context of the problem being solved. We observed many cases where student performance depended strongly on the particular representation being used. This happened in aggregate, as in the chapter 5 quizzes and homeworks, where we see significant differences in performance between essentially isomorphic problems in verbal, mathematical, graphical, and pictorial formats of quizzes covering four different topics. We also see this representation dependence in detail, as in the chapter 7 interviews where some students pursue a particular strategy towards a correct answer in some representations, and a different strategy towards an incorrect answer in another representation.

In addition to this dependence on representation, we saw a dependence on topic or concept. In chapters 5 and 6, we saw quiz performance changing from topic to topic, which was not at all surprising. Less expected was the shift in the direction



of choice/control splits. With some topics, certain formats were associated with better performance from the choice groups, while with others the same formats were associated with better performance from the control groups.

Perhaps the most significant theme in these chapters is the sensitivity of student performance to *combinations* of representations, topics, and prior student knowledge.<sup>v</sup> We observed many cases where a student's performance could not be simply attributed to just representation, or just concept, and so forth. This includes the diffraction problem from chapter 5, where students using multiple representations were outperformed by those not using multiple representations, thanks to a misunderstanding enabled by a very particular combination of topic and representation. On the single-student scale, in chapter 7 we saw Adam incorrectly solve the graphical format of the spectroscopy quiz that he had already solved correctly. There, the graphical format of the quiz reminded him of electron orbitals as seen in his chemistry class, leading to an incorrect analogy to planetary orbits. Such a response depended intimately on Adam's recent experiences with chemistry and with gravitation in the first semester of physics, and also required a particular kind of representation to act as a cue. In these and in several other cases, we saw the need to attend to the complete context of the problem in order to understand the outcome. We also saw cause to think of context via analogy to a rope composed of threads: the final product includes the threads themselves as well as the weaving together of those threads.[74, 75, 95] Decomposing the context into its component elements without

---

<sup>v</sup> This is not a comprehensive set of factors that student performance may depend on; it is simply a set that has presented itself as relevant in the studies so far.

attending to the relations between these elements would leave us with an incomplete picture.

Broadly, this suggests that a theoretical understanding of student representation use in physics will have to attend to context, supporting the initial idea that cultural constructivism may become productive. While this is useful to establish, much remained to be done at this point in our work. “Context” is by definition a rather all-encompassing term, and to make practical progress one has to have some specific dimensions along which to work, while choosing to neglect others.

#### *Directions and remaining questions*

With these studies in hand, we were able to clarify the dimensions we would prioritize, and those we would not. Consider the notion of meta-representational competence. Our work and previous work suggested that meta-representational competence could be quite important to eventual student success, but the results of these initial studies showed that novice student beliefs and choices about representations were, on average, very poorly correlated with actual outcomes. Thus, we decided not to include meta-representation in our theoretical work. We consider meta-representation, in particular the teaching of meta-representational competence, to be an important topic for future study. Should it become possible to instill in students some basic level of meta-representational competence through instruction, it may be useful to incorporate this into our theoretical understanding.

One possibility that presented itself was further work along the dimensions of concept and representations. We have already used such dimensions in the previous

chapters, but we have also seen hints that these can be strongly coupled. To us, it became necessary to clarify the extent to which concept and representation could be treated as separate categories. This is the subject of chapter 9, in which we analyze FCI, FMCE, and BEMA data for 1,000-3,000 introductory CU students, testing the usefulness of these categorization schemes.

Another outstanding issue is the limitation of our work so far to single-representation questions.<sup>vi</sup> Generally, facility with multiple-representations use is held up as one of the high goals of physics instruction. From a theoretical perspective, such problems should be more complex to analyze, and so “single representation” problems provide a reasonable starting point. However, we must eventually turn our attention towards multiple-representations problems, which are the focus of the studies in chapters 10 and 11. In chapter 10, we examine multiple representation use on an aggregate scale across two different institutions. In chapter 11, we continue our two-threaded approach by zooming in to a very fine-grained picture, interviewing students during their solutions of multiple-representations problems and studying the patterns of representation use they engage in on a moment-by-moment basis. This fine-grained study also affords us the opportunity to compare novice multiple representation use to the patterns exhibited by physics experts. While the goal of this thesis project is primarily to gain an understanding of novice use of physics representations, comparison to experts provides a useful frame of reference, making it clearer what features of novice representation use distinguish them most clearly from experts, and perhaps suggesting directions for instruction.

---

<sup>vi</sup> As mentioned previously, it is somewhat artificial to consider any question *not* to be multiple-representation in nature, but “multiple representations” has a reasonably well-defined meaning in PER, and this meaning is not satisfied by many of the problems studied in chapters 5-7.

## SECTION III: A DEEPER LOOK

### Chapter 9: Representations, concepts, and context: FMCE and BEMA data

#### *Introduction and goals*

The reader may have noted by now that we have often claimed that introductory student performance is very sensitive to the specific combinations of conditions that we refer to as the problem context. This suggests that trying to break down student performance in terms of *only* concept or *only* representation may not be powerful enough. To help decide this point, we drew on the substantial body of FMCE and BEMA back data available at CU. Nearly 3000 unique students had taken the FMCE[49] by the time of this study, with approximately 1500 students' worth of data available for the BEMA, [94] providing an enormous sample for data mining. These studies also provide a bank of previously validated and studied problems, complementing the study-specific problems we have constructed up until this point. Note that all of the data presented here are from calculus-based introductory physics courses. Until recently, the algebra-based sequences received only the FCI.[48]

We had two major goals. First, we planned to classify the survey problems based on the kinds of activities students needed to perform. In many cases, students needed to be able to translate from one representation to another to solve the problem, for instance, moving from written statements to graphs. In others, they needed to translate from one conceptual domain to another, perhaps thinking in terms of

velocity and then acceleration. We hoped that consistent patterns would emerge in student performances based on what kinds of translations were involved, and patterns did emerge: Students were more successful with problems in which only representational translations were involved as compared to those where conceptual translations or some combination of conceptual and representational translations were involved.

Our second goal was to compare small subsets (usually pairs) of problems that were very similar except for small differences (perhaps only a change of setting or of language). For example, we might identify two problems as asking essentially the same thing with the same distractors available, but with a change in the representation being used or in the setting. Furthermore, we attempted to identify particularly compelling problem features that might cue students to answer in a particular way. Generally, we tried to focus on cueings that might cause students to answer in unexpected ways, as it is not terribly useful to be able to “predict” commonly known student difficulties. In these ways, we can further investigate our claim that very small problem changes can result in significant performance changes. In many instances, problem performance did change significantly between problems.

Both of these goals involve the level of task, or individual problems. We will not be considering the environment level in detail in this study, choosing instead to focus on student performance on these problems averaged across many environments.

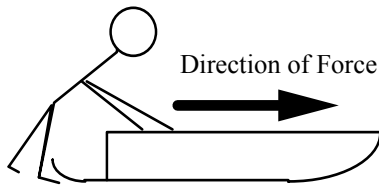
## *Methods*

We began with a brief characterization of each of the problems on the FMCE and the BEMA. In these (not presented), we noted which representations the questions and answers were posed in, and thus identified any translations between representations that were required. We also noted which concepts were involved in the question, for instance, noting whether a kinematics question is framed in terms of velocity or acceleration. Again, it was often the case that the question was posed in terms of one concept and requested an answer in terms of another. There is necessarily some subjectivity here: the primary concept or representation present in a problem is not always concretely defined. Next, we noted any features of the problem that distinguished it from its neighbors. This could include the fact that a problem involved movement to the left when most others involved movement to the right, or a subtle change in language. Finally, we highlighted any representational problem features (language, compelling picture elements, etc.) that we felt might act as particularly strong cues for students.

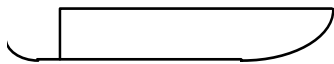
With these characterizations in hand, we made problem groups as appropriate towards our two main goals. The problems on the FMCE were uniquely suited to our first goal. The FMCE problems come in several sets, with each set investigating a particular context and set of representations. Each set often contains a systematic variation of concepts and minor features like mathematical sign or language. In Figure 15, we see the first two problems from the FMCE, which are part of a seven-problem series. These two problems are quite similar, but with one asking about

A sled on ice moves in the ways described in questions 1-7 below. *Friction is so small that it can be ignored.* A person wearing spiked shoes standing on the ice can apply a force to the sled and push it along the ice. Choose the one force (A through G) which would **keep the sled moving** as described in each statement below.

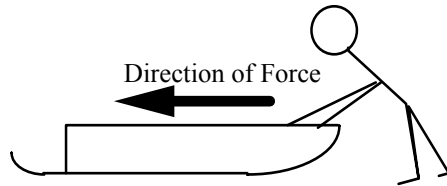
You may use a choice more than once or not at all but choose only one answer for each blank. If you think that none is correct, answer choice J.



- A. The force is toward the **right** and is **increasing** in strength (magnitude).
- B. The force is toward the **right** and is of **constant** strength (magnitude).
- C. The force is toward the **right** and is **decreasing** in strength (magnitude).



- D. No applied force is needed



- E. The force is toward the **left** and is **decreasing** in strength (magnitude).
- F. The force is toward the **left** and is of **constant** strength (magnitude).
- G. The force is toward the **left** and is **increasing** in strength (magnitude).

- \_\_\_\_\_ 1. Which force would keep the sled moving toward the right and speeding up at a steady rate (constant acceleration)?
- \_\_\_\_\_ 2. Which force would keep the sled moving toward the right at a steady (constant) velocity?

Figure 15. FMCE problems 1 and 2.[49] These problems require conceptual translations either from acceleration to force or from velocity to force. Neither requires a representational translation by our standards.

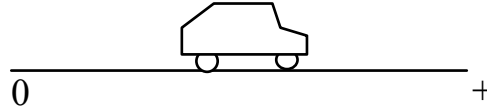
constant *velocity* and the other asking about constant *acceleration*. The answer choices are in terms of forces, so we would characterize these problems as having conceptual translations (either velocity to force or acceleration to force), but no representational translations (they are primarily written language to written language). In Figure 16, we see examples of FMCE problems in which both conceptual and representational translations are present. We also see a minor change in framing: One of the problems describes motion to the left, while another describes motion to the right. Previous work suggests that introductory physics students struggle with graphical representations of negative quantities, making this a potentially useful distinction.[14]

The problem pair in Figure 16 is also suitable for our second goal, where we look at the magnitude of the performance shifts associated with small shifts in problem contexts. In this pair, we see problems that are identical except for the direction of motion. Figure 17 shows us the first three questions from the BEMA exam. These are very similar, making use of the same representation, the same concept, and even the same equation (Coulomb's Law), varying only the quantities present in that equation. The complete FMCE and BEMA exams are reproduced in Appendix C.

These characterizations and groupings were all completed before analyzing any performance data. The large-lecture calculus-based introductory physics courses at CU have given the FMCE five times between 2004 and 2006. The BEMA has been given four times during this same time period in the second semester of this course. In each course, the exam was given before and after instruction, usually the

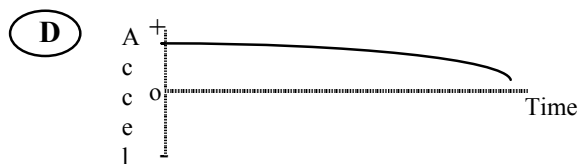
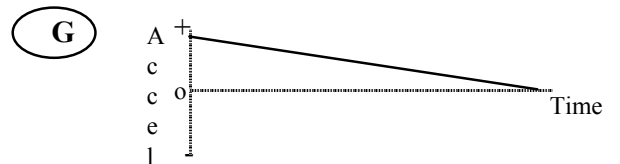
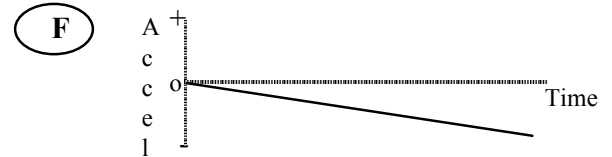
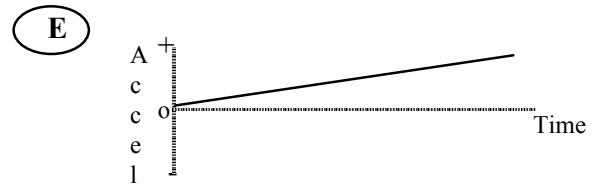


Questions 22-26 refer to a toy car which can move to the right or left on a horizontal surface along a straight line (the + distance axis). The positive direction is to the right.



Different motions of the car are described below. Choose the letter (A to G) of the **acceleration-time** graph which corresponds to the motion of the car described in each statement.

You may use a choice more than once or not at all. If you think that none is correct, answer choice J.



**(J)** None of these graphs is correct.

- \_\_\_\_ 24. The car moves toward the left (toward the origin) at a constant velocity.  
 \_\_\_\_ 26. The car moves toward the right at a constant velocity.

Figure 16. FMCE problems 24 and 26.[49] These problems involve a translation in representation (from graph to written language), and in concept (from force to velocity). These two problems differ in direction of motion, which we would refer to as a subtle change in context.

Two small objects each with a net charge of  $+Q$  exert a force of magnitude  $F$  on each other:



We replace one of the objects with another whose net charge is  $+4Q$ :



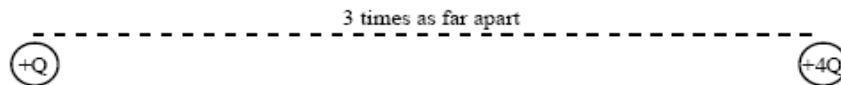
→ Q1 The original magnitude of the force on the  $+Q$  charge was  $F$ ; what is the magnitude of the force on the  $+Q$  charge now?

- (a)  $4F$
- (b)  $5F/2$
- (c)  $3F$
- (d)  $2F$
- (e)  $F$
- (f)  $F/4$
- (g) None of the above

→ Q2 What is the magnitude of the force on the  $+4Q$  charge?

- (a)  $4F$
- (b)  $5F/2$
- (c)  $3F$
- (d)  $2F$
- (e)  $F$
- (f)  $F/4$
- (g) None of the above

Next we move the  $+Q$  and  $+4Q$  charges to be 3 times as far apart as they were:



→ Q3 Now what is the magnitude of the force on the  $+4Q$  charge?

- (a)  $4F/3$
- (b)  $4F/9$
- (c)  $F/3$
- (d)  $5F/18$
- (e)  $2F/9$
- (f)  $F/9$
- (g)  $F/36$
- (h)  $4F$
- (i) None of the above

Figure 17. Questions 1-3 on the BEMA exam. These questions use essentially the same concepts and representations (from the expert perspective), changing only the quantities that would be present in Coulomb's Law.[94]

first and next-to-last weeks of class. In all cases, we restricted our attention to students who completed both the pre and post tests given, making possible comparisons of gains and focusing attention on only those students that completed the course. This matching left us with 1425 students in our FMCE data set, and 1073 in the BEMA data. We averaged these students' scores together to get an average performance on each problem, both pre-instruction and post-instruction. With samples this large, performance differences on the order of 1-2% become *statistically* significant, so we will often have to make judgments regarding how large a performance difference must be to be *educationally* significant. In any analyses involving extremely small thresholds for statistical significance, we will omit p-values for clarity. Any comparisons shown in the Data sections will be restricted to problem pairs for which the performance differences are (in our opinions) educationally interesting as well as statistically significant. To be considered educationally significant, p-values vary, but we look for differences between two performances on the order of a third or more (15 vs. 20 percent correct, for example).

*Data and analysis: Representational and conceptual translations*

As noted before, the FMCE was particularly well-suited to an analysis of conceptual and representational translations. Thus, in this section we restrict ourselves to only the FMCE data. In our opinions, classifications of the BEMA questions with respect to the kinds of translations involved are much less clean.

We flagged 16 of the questions on the FMCE as involving only a translation between concepts (1-13, 27-29). Six represented a translation between only representations (21-23, 25, 40, 42, 43). We considered seven questions to involve significant conceptual and representational translations together (14, 16-19, 24, 26). In Table XI, we see the averaged pre and post test scores for each of these problem categories. For example, the pre-test scores for problems 1-13 and 27-29 average together to 0.25 (the fraction of students answering correctly).<sup>vii</sup>

Translation kind	Pretest avg. (st. dev.)	Posttest avg. (st. dev.)
Conceptual	0.25 (0.10)	0.63 (0.10)
Representational	0.58 (0.29)	0.77 (0.12)
Both	0.24 (0.17)	0.61 (0.12)

Table XI. Average pre- and post-test FMCE scores on questions that involved translations between concepts, translations between representations, or both together.

The conceptual and representational pre-test averages (0.25 and 0.58, respectively) differ at the  $p = 0.01$  significance level using a 2-tailed t-test,[96] meaning that students performed substantially better on the problems that remained within the same concept but shifted in representation than the problems that stayed within the same representation, but shifted concept. A similar performance difference exists between the representation-only problems and those that required translations between concepts and representations, where students had an average score of 0.24. The concept-only and concept/representation performance averages differ with  $p < 0.05$ .

---

<sup>vii</sup> These categorizations were made only by the thesis author; no inter-rater reliability was involved.

On the post-test, students had averages of 0.63, 0.77, and 0.61 on the conceptual, representational, and both categories. The same trends exist as on the pretest: Students score significantly better on the representation translation problems than on the other two categories, with  $p < 0.05$  in either case.

These data suggest either that the problems involving only conceptual translations are no more difficult than those involving both conceptual and representational translations, or that the conceptual translation is the most relevant one in determining difficulty. We can test this surprising result further by comparing four problems from each category that map onto each other very closely, in that the problems regard the same physical setups. For example, problems 1 and 16 ask about the force necessary to keep a frictionless object moving to the right. Problem 16 requires students to choose between answers in a graphical format, adding a representational translation (verbal to graphical) to the conceptual translation already present in both problems (velocity to force). There are four such problem pairs: 1 and 16, 2 and 14, 3 and 18, and 4 and 19. In each case, the higher-numbered problem includes a representational translation. In all four cases, students do better both pre and post on the problems requiring only a conceptual translation. The pretest score pairs are 0.20/0.18, 0.18/0.14, 0.32/0.14, and 0.20/0.14, respectively. The posttest score differences are similar in that they are significant but not especially large.

These four-problem subsets show the pattern that one might expect: Requiring both conceptual and representational translations does increase the problem difficulty as compared to requiring only conceptual translations. Taking the previous problems into consideration, it appears that in this population, the conceptual

translation was the larger of the two factors in determining performance, though both matter.

*Data and analysis: Representational and conceptual translations*

Next, we consider student performance on several pairs and small groups of problems that differ from each other slightly, often involving shifts in context while preserving overall question structure from the point of view of an expert. Generally, our goal is to provide a complement to the data from the previous subsection. Here, the questions identified will often involve the same representations and concepts, with differences existing elsewhere.

*FMCE problems: Association of force with constant velocity*

FMCE problems 2 and 5 ask students what force is necessary to keep a sled moving to the right at constant velocity. Both involve mostly written language, with a supplementary picture of a sled that is not needed to solve the problem.<sup>viii</sup> Problem 5 differs in that the problem statement includes prior history of the sled: “The sled was started from rest and pushed until it reached a steady (constant) velocity toward the right. Which force would keep the sled moving at this velocity?” This is probably intended to cue students into comparing the force necessary to get an object moving with the force needed to keep an object moving. Presumably, this would encourage students to conclude that no force is necessary to maintain motion on a

---

<sup>viii</sup> In some cases a representation will contain information that is necessary for solving the problem and is not present in any other representation. In other cases, a representation will be present but will only duplicate information present in another representation, or will provide information not required for the problem solution.

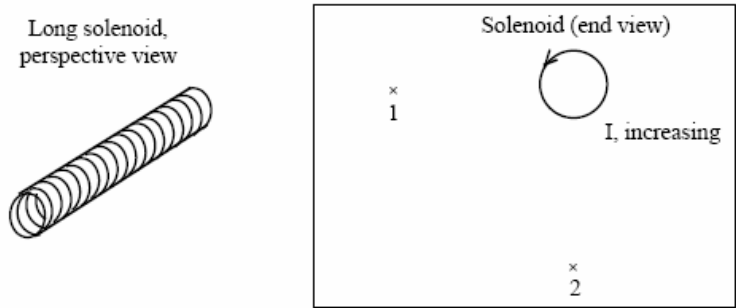
frictionless surface, though an alternative hypothesis is that emphasis on pushing may somehow cue a strong focus on that pushing. The question is successful in calling attention to the cessation of pushing, as students have a success rate of 0.18/0.56 pre/post with question 2, and 0.48/0.75 for question 5. Note that the effect persists post-instruction.

Problems 5 and 21 are not as similar, but merit examination. Problem 21 asks students to identify the graph of the force felt by a frictionless car that is pushed to the right and then released. As in problem 5, the students are asked to identify the force on an object that is moving to the right without friction. In this case, the requirement to interpret a graph may make it more difficult, but our expectation was that the presence of the word “released” would make the problem easier, if anything, since the correct answer is zero force. To our surprise, students had a success rate of 0.48/0.75 pre/post on problem 5 and 0.28/0.63 on problem 21, making problem 21 substantially more difficult. Incorrect answers to these problems did not load onto any particular distractor preferentially, and the difference may be attributable solely to the added difficulty of translating the graphs.

*BEMA: Problems identical under rotation*

Problems 28 and 29 are perhaps the most similar to each other of all the problems analyzed on either exam. Problem 29 is identical to problem 28 after a 90 degree rotation, as seen in Figure 18. Student pre and post scores were not substantially different (0.04 vs. 0.07 pre and 0.15 vs. 0.15 post), though this null result may simply be an artifact of the extremely low performance.

Here is a long solenoid (coils of wire along a long cylinder), and an end view of the solenoid. Conventional current runs counter-clockwise in the solenoid and is increasing with time.



Choose from the following possible directions to answer the questions below:

	⊙ out of page	e
	⊗ into page	f
	zero magnitude	g
	none of the above	h

- Q28 What is the direction (a-h) of the *electric* field at location 1 (marked with an ×)?
- Q29 What is the direction (a-h) of the *electric* field at location 2 (marked with an ×)?

Figure 18. Problems 28 and 29 on the BEMA exam.[94] These problems are identical under a 90 degree rotation, and as expected, student performances are nearly identical.

*FMCE problems: Constant acceleration from gravity*

FMCE problems 8, 9, and 10 ask about the net force on a toy car that is rolling up a hill, coming to a halt, and rolling down. The answer choices refer to the direction of the force (up or down the ramp) and whether it is constant, increasing, or decreasing. Questions 11, 12, and 13 are nearly identical, except that instead of a car, we have a coin that is tossed up into the air, comes to a halt, and falls back down. The answer choices map exactly from the first problem set to the second.



These problems provide a particularly strong example of a shift in setting that preserves the conceptual and representational structure of the problem, save for a picture associated with problems 8-10 whose usefulness is not clear. During analysis, we were unable to identify any problem cues in either set that would lead us to predict that students would do better on one set than another. The ramp problem adds difficulty if students consider the presence of a normal force from the ramp, or the significance of the angle of the ramp, but we did not consider these likely points of interest, especially in pre-test results. Checking the performance data, we found a success rate of 0.13/0.15/0.24 for the 8/9/10 pretest versus 0.21/0.22/0.32 for problems 11/12/13. On the posttest we see scores of 0.57/0.56/0.64 and 0.69/0.67/0.73, respectively. Students are thus significantly more successful, and consistently so, on the problems using the coin context than those using the car context. We have found this result to be particularly difficult to explain, except possibly for the extra complexity associated with the angles and normal forces of the ramp problem. We suspect that it will not be possible to determine the source of these differences with any confidence without conducting problem-solving interviews.

*BEMA problems: Coulomb's law*

BEMA questions 1, 2, and 3 require a conceptual understanding of Coulomb's law, using a diagram with vector elements for support as seen in Figure 17. Questions 1, 2, and 3 involve very nearly the same concepts, representations, and setting from the point of view of the expert. However, some possibilities exist for differences

from the point of view of the novice, in terms of strategy cueing. The ordering of the problems relative to one another may affect performance.[97] Placing problem 2 after problem 1 may cue students to believe that the answer should be different, with a possible response being to multiply the answer to 1 by a factor of 4. Problem 3 has students generate a new answer based on their previous answers in much the same way as problem 1 follows the initial setup, except that now we vary the charge separation instead of the charge, leading to at least a minor conceptual difference. In addition to problem order effects, it is possible that novice students may be using inappropriate strategies to solve these problems, strategies that may be more vulnerable to problem differences than expert strategies. For instance, we suspect that the equal charges in the problem statement could trigger a “balancing” primitive,[57] and changing one of the charges for problems 1 and 2 could thus trigger an “unbalancing” primitive. This would result in reduced performance for problem 2 versus problem 1.

In practice, we do see substantial performance differences, despite the substantial structural similarities between these three problems. The pretest success rates are 0.54/0.39/0.27 on problems 1/2/3, and the posttest scores are 0.78/0.64/0.59, respectively. As with the ramp and coin problems, we cannot confidently attribute these performance differences to any one cause without problem-solving interviews. We can, however, reaffirm that problems that are very similar in structure and intended solution from the point of view of an expert may be quite different to the novice, potentially confounding a priori problem analysis.

*BEMA: Work along different axes*

Problems 14-16 are similar to problems 1-3 in that they all draw on the same representation, the same fundamental concepts, and the same setting. The problems ask students to determine the work done by a constant electric field as a charge is moved between specified points. The questions differ in whether the separations between the points are in the same direction as the field (first problem), perpendicular to the field (second problem), or both (third problem). The first problem requires actual calculation and attention to problem sign, while the second can be solved by recalling a fact: Moving perpendicular to force requires no work. Since this fact is not necessarily known pre-instruction, it may not be a significant factor on the pre-test. The third problem requires the calculation of problem 1 in addition to the factual knowledge associated with problem 2.

Student success on these three versions followed the above reasonably well. On the pretest, problems 1 and 2 showed similar results (0.29 vs. 0.33), whereas on the posttest problem 2 was much easier (0.44 vs. 0.75), possibly reflecting student learning of the fairly straightforward “motion perpendicular to the electric field requires no work” fact. Problem 3 was the most difficult either pre (0.09) or post (0.25). Analyzing the distractors, we find that the distractors with Pythagorean form are much more powerful for problem 3 (58% of answers) than for either problem 1 or 2 (8% of answers).

Since so many students were answering problems 1 and 2 correctly on the post-test, it is surprising that so few could correctly combine the results of problem 1 and 2 to immediately find the answer to problem 3. Instead, they appear to be cueing

on the presence of a hypotenuse, with that cue overriding what they apparently “know” about work and magnetic fields. Thus, we see two major points. First, we see that our analysis of the problem was relatively accurate, but to frame that analysis in terms of conceptual shifts would require a fine-grained standard of how big a change in the problem concept is needed to count as a conceptual shift. Second, we see once again the potentially very powerful effect of a strong cue (the triangular shape present), potentially overriding the concepts that students are intended to draw on.

### *Discussion*

The data address two main questions. First, can we use concept and representation as productive dimensions along which to analyze problems, with a particular eye towards theoretical development? Second, when we identify problem groupings that are very similar except for small but quite identifiable differences, will we continue to see the performance differences we have observed in earlier studies?

The first of these questions can be partially answered simply by attempting to characterize the problems on the FMCE and BEMA. It was immediately clear that the ideas of concept and representation are useful in *describing* these problems. What was less clear was how powerful these dimensions would be for *predictive* analysis. Would problems that depend differently on conceptual or representational translations be associated with consistently different performances?

In these data there are consistent differences between the conceptual translation, representational translation, and “both” categories. Students perform

much better on the set of problems that require only a representational translation than on those requiring only a conceptual translation or both. Surprisingly, problems involving only a conceptual translation were similar in difficulty on average to those requiring both kinds of translation together. However, upon comparing closely-matched subsets of those categories, the “both” problems were in fact more difficult than the concept-only problems.

It is not clear how robust these results would be. The FMCE covers mostly introductory kinematics, with some dynamics, and limits itself primarily to graphical and verbal representations. It may be that the representation-only category was easier because calculus-based introductory physics students are more comfortable interpreting graphs than shifting between different concepts, showing the effect of student backgrounds. Indeed, the knowledge that students bring with them could well have changed their perspective of what constitutes a significant difference between problems, as in the BEMA Coulomb’s law problems, the BEMA work problems, or most strikingly, the FMCE coin vs. ramp problems.

Our second question helps inform the first. We see several problem groups where the problems are extremely similar structurally, sharing the same conceptual and representational framing. These problems differ to varying degrees in other ways, often contextual. We observe that there are often quite significant performance differences between the problems in these groupings, reinforcing our earlier claims that student performance is very sensitive to the specific combination of problem features present.

We should also note that the performance variations in the problems chosen are of comparable or larger size than the category-to-category variations identified in the first section. Thus, from the standpoint of performance, problem-specific effects appear to dominate the effect of whether the problem involves conceptual translation or representational translation overall.

### *Conclusion*

This study set out to reinforce previous results, and to test a possible approach for theoretical development. An ideal theoretical model of representation use in physics problem solving would be simple enough for a practicing instructor to use without enormous time investment, but powerful enough to be somewhat predictive regarding student performance. The simplicity requirement led us to choose few and simple dimensions along which to characterize problems. Our data so far suggest that if it is possible to predict student performance at all on a problem-by-problem basis, it will require a much more detailed analysis than what we have attempted so far, which would likely be too complex for classroom use. Furthermore, the performance effects of context remain strong, even when the contextual shifts are quite minor.

Despite this limitation, it does appear that many problems can be usefully described in terms of the categories used here (concept, representation, and setting). If predictive work is not practical, there may still be considerable use in developing a descriptive language for problem analysis. Indeed, the fact that we can split off categories such as concept and representation and have them reflect performance is useful in itself, as is our identification of likely cueing effects. It would be quite

productive to follow-up this work with interviews on these problems. No plans exist to do this, but chapter 12 includes extensive interviews of other multiple-representations problems with focus on specific student behavior.

After this study, much work remains. We still needed to study the mechanisms by which representation drives performance, and to broaden our investigation to include multiple representation problems. In the next chapter, we examine two such mechanisms: the role of student prior knowledge, and representation-dependent cueing.

## Chapter 10: Representation-dependent cueing and the role of prior knowledge

### *Introduction*

We have seen that the representation of a problem can have a significant effect on student performance on that problem, and that effect can be quite complicated, depending on the particular combination of representations, concepts, prior knowledge, and other context present. In chapters 5, 6, and 7, we saw possible mechanisms by which representation could drive performance. In the quiz problems solved by students in the chapter 7 interviews, students appeared to be cued into particular strategies and answers by features that were compelling in one representation, but not in others. We also saw hints that student prior knowledge and expectations could affect these cueings, such as Adam's understanding of orbitals, or Emma's expectation that a mathematical framing requires a mathematical solution. Generally, for a specific kind of problem, some representations can drive students to answer productively, while other representations of that same problem can drive students to answer unproductively. In this chapter, we shall design new problems with which to investigate the notion of representation-dependent cueing and the effect student prior knowledge will have on this cueing. We will approach this in two parts.

First, we will present a revised version of the FMCE. The original version of the FMCE is usually framed in terms of realistic, everyday contexts such as sleds sliding down hills, coins being tossed, and cars colliding. However, the questions often involve making unrealistic "physics class" assumptions such as frictionless interactions. It is quite possible that the realistic contexts cue students to draw on



their real-world experiences to answer problems that do not necessarily behave in real-world ways. If this were so, it should be possible to rewrite the FMCE in terms of more artificial “physics class” contexts such as steel ramps and air tracks, and to see students answer questions more successfully when the contexts match the questions. This would represent a kind of epistemological cueing, in which students draw on different sets of knowledge and expectations based on whether the context is perceived as realistic or “physics class.”[98] There is some room for debate as to whether this will actually happen. Much work in PER has shown that physics students do not necessarily see real-world knowledge as applying to any physics problems, even those dressed up to appear real-world.[99, 100] On the other hand, there exists a vast body of work inside and outside of PER indicating that students bring with them ideas based on their previous interactions with the world.[101] Indeed, this idea is a foundational one in constructivism. Here, we found that this contextual shift had very little effect on students’ responses, providing a counterpoint to our many observations so far of contextual shifts resulting in strong performance changes.

In the second part of the study, we write several problems chosen according to our expectation that certain representations of these problems will cue students either productively or unproductively as compared to other representations. These problems test for several kinds of cueing, including the WYSIWYG style of cueing previously discussed,[90] and the presence of a balancing p-prim.[57] We also take the opportunity to revisit one problem from our first study,[24] attempting to replicate the results. Some of these problems showed the performance differences predicted by

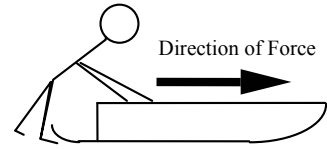
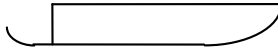
our cueing hypotheses, while a couple showed no significant performance differences.

*Methods: Revised FMCE*

All of the trials described in this chapter took place in the first semester introductory algebra-based large-lecture physics class at CU, taught in the fall of 2006. The instructor was one of the most PER-committed members of the faculty, as evidenced by considerable use of clickers, Tutorials,[102] and substantial use of varied and multiple representations. By our conclusions in chapter 6, we might expect this population to show reduced vulnerability to changes in problem representation, however this may not generalize to changes in context given consistent representations.

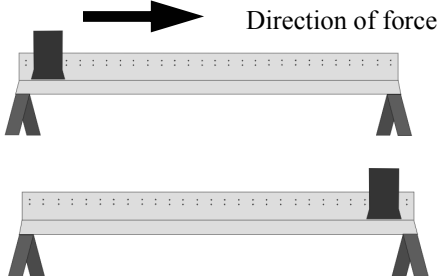
The revised version of the FMCE was developed jointly between the thesis author and the instructor for the course studied. For each problem or group of problems, we revised the context to be more sterile and disconnected from everyday experience, and more strongly connected to the physics classroom. When possible, we made the problems resemble situations from the specific physics course they were taking, as in the case of air cart problems, where the carts are drawn to resemble those actually used. Figures 19 and 20 show the original and revised versions of two FMCE problems. In the first of these, the nominally frictionless sled of the original FMCE is replaced by a more genuinely frictionless air cart, drawn to resemble the one actually used in this course. In the second problem, the two individuals pushing on each other are replaced with two carts that push on each other with a spring. Care

A sled on ice moves in the ways described in questions 1-7 below. *Friction is so small that it can be ignored.* A person wearing spiked shoes standing on the ice can apply a force to the sled and push it along the ice. Choose the one force (A through G) which would **keep the sled moving** as described in each statement below.

	<p><b>A.</b> The force is toward the <b>right</b> and is <b>increasing</b> in strength (magnitude).</p> <p><b>B.</b> The force is toward the <b>right</b> and is of <b>constant</b> strength (magnitude).</p> <p><b>C.</b> The force is toward the <b>right</b> and is <b>decreasing</b> in strength (magnitude).</p>
	<p><b>D.</b> No applied force is needed</p>

\_\_\_\_\_ 1. Which force would keep the sled moving toward the right and speeding up at a steady rate (constant acceleration)?

A cart on a long frictionless air track moves in the ways described in questions 1-7 below. *Friction is so small that it can be ignored.* A force can be applied to the cart (by a string attached to a machine) that pulls the cart along the track. Choose the one force (A through G) which would **keep the cart moving** as described in each statement below. The track is so long that the cart won't reach the end.

	<p>_____ The force is toward the <b>right</b> and is <b>increasing</b> in strength (magnitude).</p> <p><b>B.</b> The force is toward the <b>right</b> and is of <b>constant</b> strength (magnitude).</p> <p><b>C.</b> The force is toward the <b>right</b> and is <b>decreasing</b> in strength (magnitude).</p> <p><b>D.</b> No applied force is needed</p>
--	---

\_\_\_\_\_ 1. Which force would keep the cart moving toward the right and speeding up at a steady rate (constant acceleration)?

Figure 19. Problems 1 on the original and revised versions of the FMCE.[49] Note that part of the figure and answer set has been removed for compactness.

was taken to change only the context and none of the problem structure. For instance, in the second problem shown, the original version portrays one of the two people as actively pushing while the other is passively pushed. The placement of the spring in

the revised version preserves this notion for the two carts. The complete revised FMCE is shown in Appendix D.

- \_\_\_39. Two students sit in identical office chairs facing each other. Bob has a mass of 95 kg, while Jim has a mass of 77 kg. Bob places his bare feet on Jim's knees, as shown to the right. Bob then suddenly pushes outward with his feet, causing both chairs to move. In this situation, while Bob's feet are in contact with Jim's knees,



- A. Neither student exerts a force on the other.
  - B. Bob exerts a force on Jim, but Jim doesn't exert any force on Bob.
  - C. Each student exerts a force on the other, but Jim exerts the larger force.
  - D. Each student exerts a force on the other, but Bob exerts the larger force.
  - E. Each student exerts the same amount of force on the other.
  - J. None of these answers is correct.
- \_\_\_39. Two carts sit on a steel table as shown below. Cart A has a mass of 9.5 kg, while Cart B has a mass of 7.7 kg. Cart A has a compressed spring attached to it, which has a rubber stopper on one side that is pressed up against Cart B. The spring suddenly releases, pushing outward, causing both carts to move. In this situation, while Cart A's plunger is in contact with Cart B,
- A. Neither cart exerts a force on the other.
  - B. Cart A exerts a force on Cart B, but Cart B doesn't exert any force on Cart A.
  - C. Each cart exerts a force on the other, but Cart B exerts the larger force.
  - D. Each cart exerts a force on the other, but Cart A exerts the larger force.
  - E. Each cart exerts the same amount of force on the other.
  - J. None of these answers is correct.

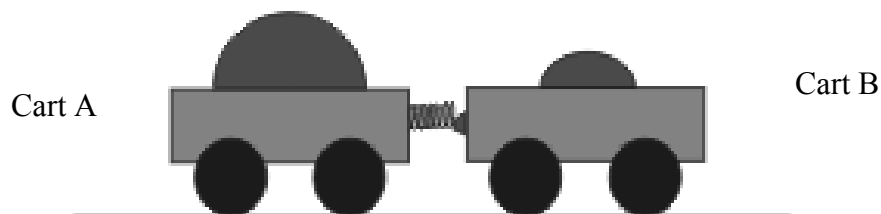


Figure 20. Original and revised versions of problem 39 on the FMCE.[49]

As is common at CU, the FMCE was given in recitation as a pre and post test on the first and last weeks of the course. Half of the recitation sections received the

original FMCE, while the other half received the revised FMCE. We distributed the versions as evenly as possible across the available TAs, days, and times of day. The same sections received the same version pre and post. Since students sometimes change recitation sections over the course of the semester, only those students who completed the same version of the FMCE pre and post were considered for analysis (N = 381 including both versions).

*Data: Revised FMCE*

We had 186 students complete the original version of the FMCE both pre and post, with 195 students completing the revised version pre and post. In Table XII, we show the averaged pre and post scores for each version using both a raw scoring system and the suggested Thornton scoring system.[49] Normalized gains ranged from 0.41 to 0.45 across the different versions and scoring schemes.

	Pre (raw)	Post (raw)	Pre (Thorn)	Post (Thorn)
Original FMCE (184)	10.6	25.9	5.5	16.4
Revised FMCE (195)	11.1	26.8	6.0	16.7

Table XII: Pre and post test scores for the original and revised versions of the FMCE including the raw scores and the scoring system suggested by Thornton. Sample sizes are in parentheses. Raw scores are out of 47, and Thornton scores are out of 33.

The performance difference between the original and revised FMCEs was not statistically significant. Considering individual problems, we can find five on which student pretest scores show a statistically significant difference between the two versions (7, 13, 31, 37, 41) using a two-tailed binomial proportion test and a  $p = 0.05$

threshold. However, note that with 47 questions on the FMCE, using a  $p = 0.05$  threshold for any single question in this post-hoc analysis is inappropriate. A simple and very conservative solution to this problem is to choose the p-value such that if one were to make  $N$  post-hoc comparisons, a difference on any one comparison could be considered significant. We thus choose  $p = 1 - (1-\alpha)^{1/N}$ , where  $\alpha$  is the desired significance level (0.05) and  $N$  is the number of post-hoc comparisons (47).[96] This yields  $p = 0.001$ . This approach is overly conservative in cases with large numbers of problems where clusters of questions may show differences,[96] but in the case of the pre-test, no such clusters were evident. No questions showed differences at a  $p = 0.001$  threshold.

On the post test, five questions showed a performance difference at a 0.05 level (with two at the aforementioned 0.001 level): question 6 ( $p = 0.04$ ), 41 (0.001), 44 (0.001), 45 (0.025), and 47 (0.04). Questions 44, 45, and 47 are part of a four-question block. Question 44 is shown in abbreviated form in Figure 21. On questions 44-47, students are asked to make inferences about the speed of a sled or a steel ball as they slide or roll down two different nearly frictionless hills of the same height. Given that three of the four problems in this set showed a performance difference at the  $p = 0.05$  level, with the difference always involving higher scores on the revised version, this is likely a real effect (and the difference for problem 44 is enough to be significant even considered in isolation). This problem set is one in which we expected a realistic context to trigger more wrong answers than a “physics class” context. A sled on a real hill experiences a great deal of friction,



A sled is pulled up to the top of a hill. The sketch above indicates the shape of the hill. At the top of the hill the sled is released from rest and allowed to coast down the hill. At the bottom of the hill the sled has a speed  $v$  and a kinetic energy  $E$  (the energy due to the sled's motion). Answer the following questions. *In every case friction and air resistance are so small they can be ignored.*

- \_\_44. The sled is pulled up a **steeper** hill of the **same** height as the hill described above. How will the velocity of the sled at the bottom of the hill (after it has slid down) compare to that of the sled at the bottom of the original hill? Choose the best answer below.
- A. The speed at the bottom is greater for the steeper hill.



A steel ball bearing is placed at the top of a steel laboratory ramp. The sketch above indicates the shape of the ramp. At the top of the ramp the bearing is released from rest and allowed to roll down the ramp. At the bottom of the ramp the bearing has a speed  $v$  and a kinetic energy  $E$  (the energy due to the bearing's motion). Answer the following questions. *In every case friction and air resistance are so small they can be ignored.*

- \_\_44. The bearing is put at the top of a **steeper** ramp of the **same** height as the ramp described above. How will the velocity of the bearing at the bottom of the ramp (after it has rolled down) compare to that of the bearing at the bottom of the original ramp? Choose the best answer below.
- A. The speed at the bottom is greater for the steeper ramp.

Figure 21. Problem 44 (abbreviated) from the original and revised FMCEs.[49] Three problems from subset 44-47 showed statistically significant differences in favor of the revised version.

and changing the hill shape will have a significant effect on the final speed. This is likely to be common knowledge for any students that come from someplace with a snowy climate.

Note that there exist other subtle differences between these two problem framings, since one involves rolling and the other involves sliding; however, this does not affect the answer or correct reasoning. Also, the original problem involved a curved surface, while the new problem involves a straight surface. Students often see straight ramps in this course, which may have resulted in better transfer to the straight-ramp problem. Finally, the revised version of the problem presents only the down-slope of the hill, while the original version has irrelevant information in that it displays both sides of the hill. It is not clear without problem-solving interviews whether these differences contributed to the observed effect.

### *Discussion and analysis*

In the first part of this study, we ask whether a changing the problem contexts in the FMCE from realistic to more artificial “physics class” contexts will have a significant effect on student success. We expected that since the problem answers required an idealized, “physics class” approach, that students would do better on the revised version. This was observed in the subset of problems including 44, 46, and 47, but only for the posttest data. No other clear differences existed. The restriction of significant differences to the posttest is reasonable, as it is at that point that students had become most familiar with the “physics class” context.



We found the narrow scope of the FMCE performance differences to be surprising, given the strength of the contextual effects we have observed elsewhere. However, we note that in previous studies, we have attributed performance differences to contextual changes post hoc, and have only focused on those problems that did show an effect. Furthermore, in many of those cases the difference from problem to problem was much larger than the relatively minor revisions seen here. In this study, we focused on one particular contextual shift, which may have been too small to consistently result in performance differences, providing a useful counterpoint to our earlier post-hoc analyses.

*Methods: Representation variation*

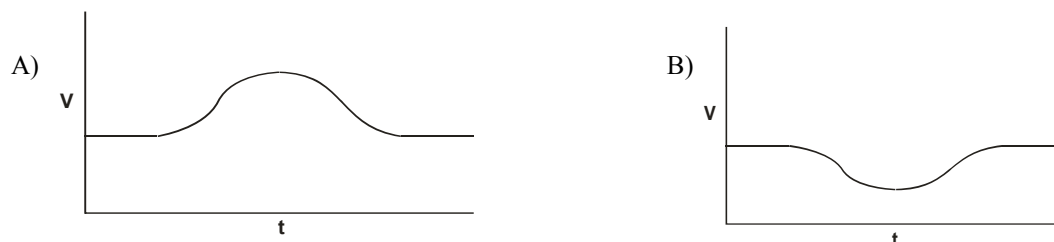
The other problems used in this study were designed to test the variation of student performance under certain changes in representation. These problems were developed jointly between the thesis author, the thesis advisor, and the course instructor, with valuable input from A. Elby. One of these problems had been used in exactly the same form in a previous study.[25] These problems were given as quizzes at the start of recitations. There were three trials on three different weeks. For each trial, a number of versions of the quizzes were given (as many as six), allowing us to control for TA, time of day, day of the week, and problem order in addition to changing the representations of the problems. For example, the first quiz involved three problems, with students receiving one of two representations of those three problems. Varying the problem order resulted in six possible versions, which were distributed amongst the recitation sections in such a way that no one TA, day of the

week, or time of day received a disproportionate share of any representation or problem ordering. The quizzes were always given after lecture on the relevant material. All problems used are shown in Appendix E.

The three questions for the first trial were about the motion of a car traveling over a hill, and were designed to explicitly search for WYSIWYG cueing. Two of the three questions asked students to characterize the velocity and acceleration of the car by choosing one of several graphs or one of several vector sequences. See Figure 22 for examples of the velocity questions, with a partial answer set. The third questions asked students to characterize either the x or y position of the car using a graph. For the graph problems, there is always a distractor that looks like a hill. The vector equivalents of this problem have the same distractors from a physical point of view, but have no superficial resemblance to a hill. In only one case (the graph of vertical position) is the hill-like distractor the correct answer. We expected that students would do especially well on that problem, and would do especially poorly on problems where a hill-like distractor is present but incorrect. The graph and vector problems, and the x and y versions of the position problems, were divided evenly among the recitation sections available.

For the second trial, we gave students one of three representations of a Newton's third law problem. In these problems, students are asked to choose an answer representing the forces acting on a large truck and a small car during a collision. In the three versions, the forces are described in words, are described using bar graphs (with the direction of the force indicated by positive or negative quantities), or are described using vectors (with the direction of the force indicated by

I give a steel ball a quick push along a frictionless track. The following are graphs of that ball's velocity in the x (horizontal) direction as a function of time (after the push). Which graph would be correct if the track went straight, then over a hill (up and back down), and then straight again?



I give a steel ball a quick push along a frictionless track. Below are a few series of vectors showing the velocity of the ball in the x (horizontal) direction at successive times (after the push). Which series would be correct if the track went straight, then over a hill (up and back down), and then straight again?

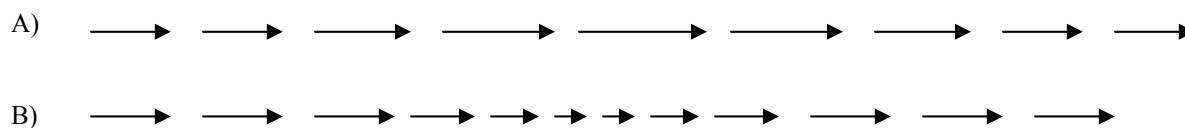


Figure 22. Graphical and vector representations of one of the first-trial problems.

the direction of the vector). The verbal representation of the problem includes the phrase “equal and opposite”, which students usually recognize and strongly associate with Newton’s third law. Thus, we expected this representation to cue students towards a correct answer more often than the other representations.

For the third trial, students answered two questions. The first of these was a ranking task. Students were given four frictionless slides of equal height but substantially different shapes, and were asked to rank the final velocities of people that slid down the slides. Half of the students received a version that represented the slides pictorially, while the other half received a version that described the slides in

words. The ranking was free-response. Students do not need to know anything about the shapes of the slides to answer this question correctly, as they had covered conservation of mechanical energy at this point. Nevertheless, the pictorial version could be said to be more informative, as it is not necessary to draw or otherwise visualize the shapes of the slides if one wants to do so. The second problem was the spring speed problem from chapters 5 and 6. We reissued the mathematical, graphical, and pictorial versions of that quiz, divided evenly among the recitation sections. In the earlier study, we had observed statistically significant performance differences between these three formats.

*Data: Representation variation*

In Table XIII, we see the performance data for the six problems from trial 1, the problems regarding the motion of a ball over a hill or valley shaped track. We see that students answering the vector versions of the velocity and acceleration questions correctly significantly more often than the graphical questions, and answering the y-position question correctly significantly more often than the x-

	Velocity	Acceleration	X position	Y position
Graphical	0.79 (268)	0.44 (267)	0.30 (179)	0.59 (276)
Vector	0.94 (187)	0.65 (187)		

Table XIII. Student performance on the graphical and vector representations of the problems given in trial 1. The performance differences between the velocity questions, the acceleration questions, and the x and y position questions are statistically significant with  $p < 0.0001$ .

position question. The p value is less than 0.0001 in all cases using a two-tailed binomial proportion test. These differences are consistent with our WYSIWYG hypothesis.

In addition to the presence of performance differences, WYSIWYG would predict that the hill-like distractor would account for most of the difference. This was only partially borne out. For the velocity questions, the performance difference was not attributable to any one distractor for the graphical version (though the hill-like distractor was more popular). For the acceleration question, the inverted hill-like distractor (a shape representing a valley instead of the hill described in the problem) accounts for most of the performance difference. For the position graph questions, the hill-like distractor is just as popular as the correct answer for the x-position version, with 30% of the students choosing either.

Trial 2 included three representations of a Newton's third law problem. Students answered the verbal representation of the problem correctly 96% of the time, with  $N = 128$ . Students answered the vector and bar graph representations correctly 91% and 94% of the time, respectively, with  $N = 143$  and  $N = 120$ . We predicted that the verbal representation would be answered most correctly; however, these differences are not statistically significant. Nevertheless, the effect may have been hidden by the abnormally high success rate on these problems.<sup>ix</sup>

Trial 3 included the ranking task about sleds on hills, and the repeat of our spring problem. Students were more successful with the pictorial version of the sled problem than the verbal representation (0.62 vs. 0.5 with  $N = 200$  and 160, giving a

---

<sup>ix</sup> Question 30 on the FMCE is similar. Post-instruction, students averaged 73% correct on this question, which is still quite high given that the FMCE post test came several weeks after Trial 2.

p-value of 0.02). There were no statistically significant performance differences between the three representations of the spring problem, with the math, graphical, and pictorial versions yielding 0.60, 0.62, and 0.65, respectively, with sample sizes of  $N = 116, 141, \text{ and } 103$ .

*Discussion and analysis: Representation variations*

Our first recitation problems behaved largely as expected. Students performed better on the vector versions of the acceleration and velocity problem than on the graphical versions of those problems. Note that the vector representations used are rather non-canonical, and no vector series of that sort had yet been seen by students in the course. It is therefore rather nontrivial from our point of view that students solving that version substantially outperformed students solving the version using a standard representation. Further analysis of the distractors clouded the result: The WYSIWYG distractors (those that superficially resembled hills) were selected slightly more often on the graphical problems, but not to the extent we expected. For the x and y versions of the position problems, results were as expected. Students did much better on the y-position version for which the WYSIWYG answer was the correct answer, and were very likely to select the WYSIWYG distractor for the x-position version. Overall, it appears that WYSIWYG cueing is in fact one of the mechanisms by which these different problem representations can result in significantly different performances.

Our second trial was less successful. Students performed abnormally well on all representations of the Newton's third law question asked as trial 2. We note that

problems 30-39 on the FMCE are Newton's third law questions, and these same students had an average post instruction score of 67% across all of those questions. It may be that despite some previous observations of students immediately and completely forgetting unintuitive results that they are told in physics class,[103] this is one instance in which they were able to retain the "forces are always equal and opposite" fact for long enough to score >90% correct on this quiz. Giving the quiz later in the semester may allow for the absolute performance to drop enough for differences across representation to be visible.

On the third trial, we found a significant performance difference between the verbal and pictorial versions of what is essentially a conservation of energy problem, with students performing better on the pictorial version. This recitation problem is similar to FMCE problems 44-47, where we saw that the less-realistic problem context resulted in increased student performance, possibly because they were less likely to draw on their experiences with systems dominated by friction. One might have expected a pictorial representation to have a similar effect, bringing to mind actual sleds more readily than a verbal description of sledding. This was not observed, though we note that in either representation, the trial 3 problem was in terms of a realistic context. While we have no immediate plans for follow-ups to this experiment, one could imagine another experiment in which we repeat the trial three problem with two additional versions, each involving a "physics class" context. This would enable us to more clearly separate out the possible effects of context and of having to mentally re-represent the verbal representation into something less abstract.

Chapter 9 suggests that we can split off the representation translation component of the task from the shift in context with some expectation of success.

Finally, we were able to repeat a problem trial from the experiment in chapter 5. Here, we found no significant difference between the mathematical, graphical, and pictorial versions of the spring quiz that showed moderate representational effects previously. This may be explainable by the fact that the present course was taught by the most PER-influenced professor in the department, with a great deal more representational richness than is typical. As we have noted in chapters 5 and 6, a course that is representationally rich may result in less significant performance variations across representation. In the next chapter, we will see a more detailed analysis of this professor's approach to representations and that approach's effect on students.

### *Conclusions*

In this study, we set out to elaborate on two points. First, we wished to further explore the notion of context dependence in problem solving by re-representing the FMCE in a less real-world, more physics class way. We found minor effects in the direction expected, but on the mean student performances were very similar on both versions, providing an example of a contextual change that is insufficient to produce *consistent* performance differences.

Second, we wished to revisit possible mechanisms by which representations can drive student performance. We found evidence consistent with the notion that some representations will be prone to what-you-see-is-what-you-get interpretations,



while other representations of the same problems will not, with the positive and null effects being consistent with our predictions. Furthermore, depending on the problem, WYSIWYG can either be productive or unproductive, again consistent with our expectations. As Elby[90] notes in his original paper, WYSIWYG would not be such a fundamental aspect of human cognition if it were not usually useful. In a physics class, where we make use of a number of abstract representations of difficult ideas, WYSIWYG is considerably less likely to always be productive. For instructional purposes, we would be wise to be aware of the WYSIWYG tendency, and the possibility for conflict between this and the representations we use in the classroom.

## Chapter 11: Multiple representation use in different physics course environments

(This chapter is an expanded version of a paper recently accepted for publication in the PER section of the Physical Review.[27])

Up until this point, the problems we have studied have been framed in terms of single representations (graphical, mathematical, or pictorial). As we have noted, no problem can be strictly single-representation, but problems that are referred to as multiple-representation are usually framed so that the different representations involved are explicitly identified. In studying single-representation problems we have hoped to simplify the interpretation of the results, as it seems likely that with multiple representations, student behavior and performance will depend not only on the representations themselves, but on the combination of those representations and interactions between them.

Despite the expected complexity of multiple representations problems, it is crucial that we address them. Instructors and researchers in PER have long argued that students can benefit from solving problems that require the use of multiple representations together.[3-6, 8, 21] Also, these kinds of problems are said to require a more expert-like understanding of the underlying physics.[3, 4] Indeed, experts tend to use multiple representations in their problem setups more often than novices, who have a tendency to jump directly to mathematics[3, 10]. Thus, use of multiple representations brings student problem-solving procedures more in line with expert procedures. These differences extend beyond problem solving, as research has shown that novices and professional scientists differ significantly in their ability and

willingness to use multiple representations productively in more applied settings such as the laboratory or workplace.[2, 51] With this in mind, it seems certain that any useful understanding of how physics students use representations, whether it be formal or heuristic, must include student use of multiple representation. In this chapter and the next, we study problems that are explicitly based in multiple representations. As before, we consider both the level of environment and the level of task. In this chapter, we look at how student solutions of multiple representations problems respond to two different approaches to teaching multiple representations, working collaboratively with Rutgers, the State University of New Jersey. In the next, we interview expert and novice physicists as they solve multiple representations, allowing a moment-to-moment analysis of their strategies.

From an instructional or environmental standpoint, we note that previous work shows that students in traditional physics courses only sometimes use multiple representations,[5] and that efforts specifically focused on increasing student use of multiple representations can be successful, even if students are not graded specifically for multiple representation use.[6-8, 80] In addition, it has been suggested that this multiple representation use can be associated with increased problem-solving performance, though this correlation is far from perfect.[104] To address this, Rosengrant (our collaborator for this work) et. al. have considered the correlation between the quality of multiple representation use and student success.[53, 105] They find that this association is quite strong, a point we return to in the present study.

The courses in the above studies can be described as taking strongly-directed approaches to teaching problem solving with multiple representations. By "strongly

directed," we mean that these approaches teach explicit steps and heuristics for solving multiple-representation physics problems and continue to emphasize these steps throughout the course. Another, less-studied (but perhaps common) approach is to model good multiple-representation problem-solving techniques for students without teaching specific steps. We can refer to this approach as "weakly directed." Arguments can be made in favor of either the strongly or weakly-directed approaches. For example, a strongly-directed approach gives students an easy-to-follow checklist, though it might also result in dependence on algorithms executed with little understanding. A weakly-directed approach may prevent dependence on checklists, but novice students may be incapable of picking up the appropriate skills in the course of an introductory class without such direction. We are unaware of any studies directly comparing strongly and weakly-directed approaches to teaching multiple representation problem solving. In this chapter, we perform such a comparison. Note that in chapter 6 we studied the differences in student success with representations in a PER-based, representation-rich course, and in a non-PER-based, representationally sparse course. Here we consider a finer distinction: that between two different PER-based, representationally-rich approaches to teaching multiple-representations problem solving.

We address these questions regarding the use and learning of representations in two parts. In the first part, we verify that multiple representations aid problem-solving, and ask whether we can begin to understand more specifically how multiple representation use is associated with student performance. In the second part, we ask how multiple-representation use and success with multiple-representation problems

varies with instruction, and examine two multiple-representation rich, PER-based courses that take different approaches to teaching multiple representations: one strongly and one weakly-directed. To this end, we study student performance on five multiple-representation problems in two introductory large-lecture algebra-based physics courses, one taught at Rutgers, the State University of New Jersey, and one taught at the University of Colorado at Boulder (CU). The problems vary in their difficulty and in their framing. For example, one problem hints that a force diagram might be useful, while another makes no such hint. Four of the problems were given in recitation, and a more difficult "challenge problem" was given as a recitation quiz at CU and as part of an exam at Rutgers.

By examining student solutions and performance in detail, we begin to address our first questions. As we have noted, many studies have established that using multiple representations can improve performance. We find, perhaps not surprisingly, that student use of multiple representations does indeed often correlate with success. However, we find that the correlation is nontrivial. Use of multiple representations alone is insufficient for success and can even be associated with lower-than-average performance. Correct use of multiple representations and close coordination of those representations is much more likely to be associated with high success rates. We also find that problem framing can alter student use of multiple representations; for instance, student solutions to problems might show different uses of free-body diagrams (FBDs) depending on whether the problem used the word "force" or not.[106] Notably, this last result regarding framing is tentative: The data here are insufficient to fully characterize the effect, rather the results serve to suggest

that multiple representation use in problem solving may be susceptible to the same kinds of cueing observed in other contexts.[26, 61, 90]

Given these data, the second part of the chapter focuses on a cross-course comparison, and on the question of whether one approach or the other is optimal. Most significant was the overall constancy of the results from the first part of our study across both environments. Both courses were successful in promoting multiple representation use, and student performances were very similar. We note some specific differences that emerged, though we emphasize that the major picture was one of strong similarity. The CU students were slightly more likely to use multiple representations on shorter, easier problems, while Rutgers students were more likely to use complete FBDs on the most difficult problem. These differences and others can be plausibly attributed to the differences in instructional environment, and lead us to suppose that elements of each course might be reasonably combined in the future.

This aggregate-level investigation of two different large-lecture courses at two different institutions also provides us with an opportunity to revisit an earlier point. In chapters 5, 6, and 7, we asked whether students could make productive use of their knowledge about their own abilities with representations. That is, did they have a useful level of meta-representational competence? We found that they did not. In this chapter, we provide students with a short survey in which they are asked various questions regarding their performance in physics class, their opinions regarding different representations, and their assessments of their own skills. We then correlate these responses with student performance.

To summarize, we ask three primary questions, with the associated findings following:

- When and how does the use of multiple representations affect student performance on problems involving free-body diagrams? Here, we find a correlation between FBD use and success, but it is not strict: Poor use of multiple representations is no better and possibly worse than no use thereof.
- What sorts of instructional methods best foster multiple-representation use? We compare two PER-based and representationally-rich approaches that differ significantly in their details, with both yielding very high rates of picture and FBD use among introductory students. Neither is clearly superior by our measures.
- If we test students from these courses for the kinds of meta-representational skills considered in earlier chapters, what will we find? Broadly, we find very weak correlation between success with multiple-representations problems and self-assessments regarding skill with and use of multiple representations, consistent with our previous observation of generally weak meta-representational skills in introductory physics students.

*Methods - Study problems*

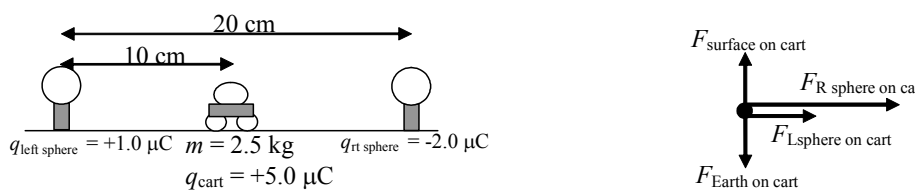
In each course, students received a set of four electrostatics problems in recitation that either required calculation of a force or specified forces in the problems. These problems were given after all lecture coverage of electrostatics and students received recitation credit for significant effort. The problems did not otherwise count towards the course grade. All problems are shown in Figure 23.

1. A small (100 g) metal ball with  $+2.0 \mu\text{C}$  of charge is sitting on a flat frictionless surface. A second identical ball with  $-1.0 \mu\text{C}$  of charge is 3.0 cm to the left of the first ball. What are the magnitudes and directions of the forces that we would have to apply to each ball to keep them 3.0 cm apart?

2. A sphere of 0.3 kg is charged to  $+30 \mu\text{C}$ . It is tied to a second chargeable sphere by a 20 cm rope, and the spheres sit on a frictionless table. If the rope will break at 4.8 N, what charge needs to be on the second sphere to cause the rope to break?

Hint: It may be useful to draw a force diagram.

3. A frictionless metal cart is being held halfway between two stationary charged spheres. The cart's mass is 2.5 kg and its charge is  $+5.0 \mu\text{C}$ . The left sphere has a charge of  $+1 \mu\text{C}$  and the right sphere has a  $-2 \mu\text{C}$  charge. The two spheres are 20 cm apart. At the instant the cart is released, what is the magnitude and direction of the total force on the cart? Refer to the included diagrams for help.



The diagram for problem 3 consists of two parts. On the left, a horizontal line represents a surface. Two spheres are placed on the surface, 20 cm apart. The left sphere has a charge of  $+1.0 \mu\text{C}$  and the right sphere has a charge of  $-2.0 \mu\text{C}$ . A cart with a mass of  $m = 2.5 \text{ kg}$  and a charge of  $q_{\text{cart}} = +5.0 \mu\text{C}$  is positioned halfway between the two spheres, 10 cm from each. On the right, a force diagram shows a central point representing the cart. Four force vectors are shown:  $F_{\text{surface on cart}}$  pointing vertically upwards,  $F_{\text{Earth on cart}}$  pointing vertically downwards,  $F_{\text{R sphere on cart}}$  pointing horizontally to the right, and  $F_{\text{L sphere on cart}}$  pointing horizontally to the left.

4. A 100 gram ball has a charge of  $+40 \mu\text{C}$ . The ball is dropped from a height of 2 m into a  $7000 \text{ N/C}$  electric field pointing up. Draw a diagram showing all the forces involved in the problem, and calculate the magnitude and direction of the net force on the ball.

Figure 23. The four problems given in recitation. Note the different prompts regarding multiple representation use (1: no prompt, 2: hint to draw a force diagram, 3: diagrams included, 4: statement that diagrams are required).



The problems contained a variety of cues regarding the use of multiple representations. The first problem made no mention of multiple representations. The second problem hinted that it may be useful to draw a force diagram. The third problem included both a picture and an FBD as part of the statement. The fourth problem stated that an FBD was required as part of the solution.

Students were also given a more challenging problem, intended to be very difficult to solve without an FBD. This problem was issued with multiple-choice answers on the first exam in the Rutgers course, and as a free-response quiz in recitation just before the first exam in the CU course. This problem and an example solution are shown in Figure 24. We shall refer to this problem as the challenge problem.

A small metal ball with  $Q_1 = +2.0 \mu\text{C}$  of charge hangs at the end of a vertical string. A second, identical ball with  $Q_2 = -2 \mu\text{C}$  of charge hangs at the end of a vertical string. The tops of the strings are brought near each other, and the strings reach an equilibrium orientation (no longer vertical) when the balls are a distance  $d = 3.0 \text{ cm}$  apart. If the force of the Earth on each ball is  $F_1 = F_2 = 30 \text{ N}$ , what is the force  $T$  of the string on each ball? Write your symbolic answer in terms of  $F_1$ ,  $F_2$ ,  $Q_1$ ,  $Q_2$ ,  $d$  and appropriate electrical constants.

$$F_E = \frac{kq_1q_2}{r^2}$$

$$= \frac{9 \times 10^9 \times 2 \times 10^{-6} \times 2 \times 10^{-6}}{(0.03)^2}$$

$$= 40 \text{ N}$$

Symbolic answer here

$$T = \sqrt{(mg)^2 + \left(\frac{kq_1q_2}{r^2}\right)^2}$$

Numerical answer in Newtons here.

50 N

Figure 24. Challenge problem with example solution. The picture drawn shows the common "backwards" picture error, as the balls are supposed to attract each other, leading to inward-facing strings.

Student solutions to these problems were coded in several ways in a scheme that extends methods developed previously.[107] The answers were coded as correct or incorrect. Specific answer features were also tracked; for example, if the answer required a number and a direction, each feature was coded separately. Student use of representations was coded using a more complex scheme. Each solution was coded with respect to any picture used and any free body diagram used. For problems 1, 2, and 4 in the recitation set, the problem was coded either as containing a picture or not containing a picture. A picture was defined as some drawing representing the situation, not to include an isolated free body diagram (coded separately). The expected elements of each picture were then coded as present or absent. For example, in problem 1 (Figure 23) the coders looked for the presence of each of the two charges and for a labeling of the distance between them. The pictures for the challenge problem were coded in more detail: The presence and correctness of the picture was evaluated using a 0-3 rubric, where 3 meant a correct depiction of the physical situation, 2 referred to a common error in which the picture was drawn "backwards," 1 referred to an otherwise incorrect picture, and 0 indicated no picture. The expected elements of the picture were then coded as present or absent, as before in the recitation problem. For problem 3 in the recitation section, a picture was provided, so coders noted whether students made their own marks on the given picture, and whether they re-drew a picture of their own.

Any free-body diagrams were coded in a similar fashion. For each problem, as many as two to four forces could reasonably be present. For each possible force (a gravitational force, a normal force, etc), the force was flagged as being present or not,

being shown in the correct direction or not (ambiguities were also flagged), and being labeled correctly or not. Coding each element of the FBDs and pictures separately facilitated analysis, as most combinations of features of interest were available in the codings.

A researcher from Rutgers University coded all of the data from that institution, and the thesis from CU coded all the data from CU. These authors then both coded two sections of the data chosen at random from their counterpart's data set and compared codings. Agreement varied from 91% to 100%, depending on the category.

*Methods: Course descriptions*

The study involved second-semester large-lecture algebra-based physics courses from CU and Rutgers, taught in the spring of 2006. The instructors for these courses had also taught the first semester of the sequence, and have been involved in PER for many years. Both courses can be described as reformed in nature, making use of many common tools and practices from PER. The courses each had one recitation/lab meeting per week, with two or three full-class meetings. Each course was four credit hours. Lecture sections had approximately 300 students each. Each school is a large state university, with similar standardized test scores for incoming students. As these were life-science track courses, the backgrounds and performances of the students within each of the classes varied considerably.

### *Rutgers University*

The Rutgers course uses the ISLE curriculum, which is inquiry-based and spends considerable time on the use of multiple representations.[108] The instructors use the Active Learning Guide workbook in lecture and in recitation, which includes many tasks designed to teach multiple representation use.[109] The recitations have research-based design elements,[110] and also use ActivPhysics computer simulations, which emphasize conceptual development, problem-solving, and multiple representations.[111] The lectures also use personal response systems (clickers).

For both mechanics and electrostatics problems in the ISLE curriculum, the instructor teaches students an explicit problem-solving heuristic with five main steps, which emphasizes multiple representations and is described elsewhere.[109] Note that this five-step procedure includes within it a sub-procedure for drawing free-body diagrams. These procedures are emphasized whenever multiple representations problems are discussed, though rigid adherence to each step of the procedure is not required, and students were never graded specifically on following the steps.

### *University of Colorado*

The CU course features such reforms such as clickers and PER-based labs and recitation activities,[112] and includes the PhET computer simulations.[52] It also includes substantial multiple representation use in lecture and in homework and exam tasks, but little explicit instruction in multiple representation use is given. The

instructor taught no specific problem-solving heuristics. In Figure 25, we see an example of an exam question from the CU course. Such multiple-representation questions were common. With substantial multiple representation use in lecture and on exams, students were held accountable for using multiple-representations effectively as well as having such use modeled for them, the combination of which was identified as useful in chapter 6.

A small block with mass  $M$  and a net charge  $+Q$  is on a frictionless horizontal table. There is a constant, uniform electric field with magnitude  $E$ , pointing to the right, as shown. There is gravity in this problem. Which statement most accurately describes the motion of the block?

A) The block moves to the right with constant velocity of magnitude  $v = E / (QM)$ .  
 B) The block has a constant acceleration to the right with magnitude  $a = QE/M$ .  
 C) The block is stationary.  
 D) The block has a constant acceleration to the right for a short while but soon acquires a final constant speed of  $v = QE/M$ .  
 E) None of these statements is an accurate description of the motion of the block.

Figure 25. A multiple-representations problem from a University of Colorado exam. The free-body diagram is not part of the problem statement.

*Methods: Environment evaluation*

The multiple-representations reforms present in the Rutgers lectures are well-documented.[108, 109] To establish the representational richness of the CU environment, we analyzed the representational content of their lectures using a procedure developed and validated in chapter 6. To help compare the two course environments, we performed a similar analysis on the exams from each course. We will only summarize the procedure here. To characterize the lecture content, we take

a series of videotaped lectures and break them into one minute intervals. We code each minute according to whether it includes use of verbal, mathematical, graphical, or pictorial representations, with "verbal" including written physics principles, but not spoken language (since spoken language is almost always present). Any interval that has more than one representation is also coded as having "multiple representations." We then average over all lectures to come up with an average fraction of lecture time spent on each category. We videotaped eight CU lectures in between the first day of class and the first exam (the lectures covering the material used in this study).

For the exam content, we focus on all the exams that lead up to the study material, as these would be the only ones likely to influence student behavior in the study. This means we consider all the exams from the first semester of the course, keeping in mind that all student data presented comes from students who took both semesters consecutively (approximately three-quarters of the class). We quantified the fraction of each exam that could be described as verbal, mathematical, graphical, and pictorial in representation on a problem-by-problem basis using the standard from chapter 6. We also quantified the fraction of each exam that explicitly required the use of multiple representations.

*Methods: Multiple representations survey*

Our multiple representations survey asked a variety of questions. A few of these asked general questions about skills with physics and mathematics. Some asked questions about specific representations, such as free-body diagrams. Others asked questions about using multiple representations together. The full survey is

**1. I am usually good at learning physics on my own, without any help from others.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**5. I am either good at physics or bad at physics, and there's nothing I can do to change that.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**8. I often use multiple representations (drawing pictures, diagrams, graphs, etc) when solving physics problems.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**9. When I use multiple representations, I do so because it makes a problem easier to understand.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**15. On a scale of 1-5, rate how much each of the following factors affects your performance in physics class (5 being the highest):**

\_\_\_ Your Effort      \_\_\_ Your Ability      \_\_\_ Teacher/TAs      \_\_\_ Textbook

**16. On a scale of 1-5, rate how often you use the following (when applicable) in solving physics problems, and how comfortable you feel when doing so (5 being the highest):**

Free-body diagrams      \_\_\_ How often      \_\_\_ How comfortable

Figure 26. Sample questions from the multiple representations survey, spanning the categories used.

reproduced in Appendix F, and sample questions from the different categories are presented in Figure 26.

As seen in the figure, questions were either on a five-point agree/disagree scale or required students to make a numerical rating. These questions were written with input from the thesis author, the collaborator from Rutgers, and the thesis advisor. They were then validated over the course of ten student interviews conducted by the thesis author, where the major focus was to ensure consistent and

expected interpretation of the question language (including the phrase “multiple representations”). With each interview, potentially problematic language was identified and changed. Over the last three interviews, no language changes were deemed necessary.

These surveys were given near (but not at) the end of the semester in the recitation sections of both courses studied. For analysis purposes, we averaged across each course, but did not average both courses together, as it was reasonable to expect variations between the two courses. Most of the analysis involved the calculation of Pearson correlation coefficients[96] between average student responses for particular questions and their performance on study problems or some subset thereof.

### *Data and Analysis*

We present the data and analysis in four parts. First, we compare the representational content of the courses studied. Second, we examine student performance and representation use on a problem-by-problem basis, comparing across courses when appropriate. Third, we focus more closely on cross-course analysis. Finally, we review student meta-representational skills, as inferred from the multiple representations survey.

#### *Part I: Environment data*

We have claimed that both the CU and Rutgers courses are representation-rich, noting the various curriculum reforms present in each. In particular, we saw that the Rutgers course made use of specific curricula intended to promote the use of



multiple representations in lecture. Since CU used no such documented curricula, we present data on the representational richness of CUs lectures here. For the sake of cross-course comparisons, we also analyze the representational content of the exams in each class.

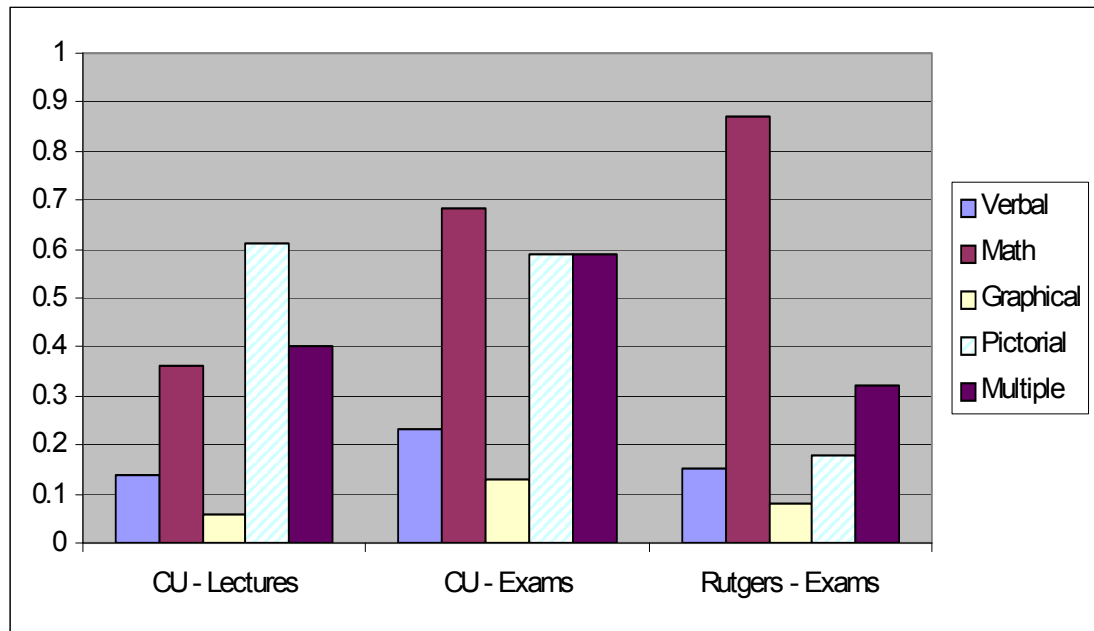


Figure 27. Fraction of lectures and exams at CU and exams at Rutgers using verbal, math, graphical, pictorial, and multiple representations.

In Figure 27, we see the fraction of the sampled lectures that contained verbal, mathematical, graphical, and pictorial representations. The CU data show more representations being used more often than in the similar, traditionally taught class studied in chapter 6, supporting the claim that this environment is representationally rich, much like the Rutgers lecture environment. Such richness is consistent with the

strong, broad similarities observed in representation use among CU and Rutgers students.

In Figure 27 we also see the fraction of the first-semester exams (those leading up to this study) using each of these four representations and using multiple representations. In these, we see a difference between the courses. The CU exams tended to use more representations more often, and used multiple representations more often, while the Rutgers exams focused more on mathematical representations. We emphasize that this is an evaluation of the representations contained within the problem, not necessarily an evaluation of all the representations used by the students. Students were intended to (and often did) use FBDs and pictures in many of their solutions of Rutgers exam problems that were strictly mathematical in presentation.

## *Part II: Performance and representation use*

### *Recitation problems*

In Table XIV we see the fraction of the students in each course answering each of the four recitation problems correctly. The numbers in parentheses indicate the number of students sampled for each problem. Problems 1, 2, and 3 had the same sample size.<sup>x</sup>

Problems 1 and 2 are similar in that both require single applications of Coulomb's Law, but with different variables to be solved for (force in problem 1, and charge in problem 2). Students in both courses performed significantly worse on 1

---

<sup>x</sup> We had initially planned to give problem 4 separately, but changed this partway through the study, resulting in some recitation sections not receiving problem 4.

than on 2, with an average fraction correct across courses of 0.37 for problem 1 and 0.53 for problem 2. These differ at a  $p < 0.0001$  level using a two-tailed binomial proportion test, but this is likely a result of the extra information requested by problem 1. Problem 1 asks students to note the direction of the force calculated, and examination of student solutions shows that many students simply overlooked or ignored this directive. Thus, we also include in Table XIV the fraction of students answering the scalar portion of problem 1 correctly, which does not differ significantly from the fraction answering problem 2 correctly. We consider problems 3 and 4 to be less directly comparable to the others since their solutions were substantially different, as were their treatments of multiple representations in the setup.

	Prob. 1	1 (scalar)	Prob. 2	Prob. 3	Prob. 4
Rutgers	0.36 (296)	0.51	0.49	0.29	0.38 (155)
CU	0.38 (314)	0.60	0.56	0.43	0.40 (269)

Table XIV. Fraction of students answering the four recitation problems correctly at Rutgers and CU. Parentheses indicate sample sizes. Samples for problems 1, 2, and 3 are the same. Standard errors vary but are on the order of 0.03. The 1 (scalar) category refers to the scalar portion of the answer for problem 1.

In Table XV we see the fraction of students in each course that drew a picture with their problem solution. Since problem 3 provided a picture, we instead show the fraction of students in each course that re-drew their own picture. Students were quite likely to draw pictures in all cases, with 90% or more of students drawing a picture in four of six cases (not counting problem 3). Students were equally likely to

draw pictures for problems 1 and 2. Table XV also shows the fraction of students identifying any forces correctly in their solution, using some kind of vector representation. Since problem 3 provided an FBD, we show the fraction of students who re-drew some force information on their own. Two data features are notable: First, the vast majority of students drew a complete and correct FBD for problem 4 (almost all who identified at least one force identified both possible forces). Since this problem asked students to draw an FBD as part of their answer and since it was the last problem in the set, we consider this an indication that students were taking the problems seriously throughout the set. Second, students were much more likely to draw some kind of FBD for problem 1 than for problem 2. Forty-five percent of students identified some forces correctly for problem 1, compared to 29% for problem 2 ( $p < 0.0001$ ). Mathematically, these problems were very similar, and it is possible that this difference resulted from some difference in the problem framing, a point we will return to in the discussion.

Pictures	Prob. 1	Prob. 2	Prob. 3	Prob. 4
Rutgers	0.90	0.85	0.00	0.73
CU	0.92	0.91	0.13	0.90

Forces	Prob. 1	Prob. 2	Prob. 3	Prob. 4
Rutgers	0.36	0.26	0.03	0.90
CU	0.53	0.32	0.11	0.92

Table XV. Fraction of students drawing a picture for each of the four recitation problems, and fraction of students identifying any forces correctly in their solution. Standard errors vary but are on the order of 0.03.

### *Challenge problem*

In the first column of Table XVI, we see the fraction of students answering the challenge problem correctly in each course. Because the Rutgers problem was given as a five-answer multiple-choice question and the CU problem was given as a free-response question, we do not consider the difference in performance between Rutgers and CU to be significant or useful for further analysis. In the next three columns, we see the fraction of students identifying exactly 1, 2, or 3 forces correctly in their solution. Note again that an FBD was not requested by the problem. More than 98% of students drew a picture. The last two columns show what we refer to as type 2 and type 3 picture use (from the previously described picture rubric). Picture type 3 is complete and correct. Picture type 2 was a common misinterpretation of the problem statement, where students drew the balls as if they were repelling (an example is shown in Figure 3). Students drawing picture type 1 were otherwise incorrect, are not shown in the table, and will not be considered further.

Exam	Correct	1 force	2 forces	3 forces	Picture 2	Picture3
Rutgers	0.56* (283)	0.09	0.22	0.51	0.34	0.41
CU	0.29* (280)	0.23	0.31	0.32	0.39	0.45

Table XVI. Fraction of students answering the challenge problem correctly, fraction correctly identifying 1, 2, or 3 of the possible forces, and fraction drawing either picture type 2 or type 3. Parentheses indicate sample size. \*Note that the Rutgers version was multiple choice, while the CU quiz was free-response.

Once again, we note the very frequent use of multiple representations, with nearly all students in either course drawing a picture, and 83% of students identifying at least one force correctly, despite no request for a picture or FBD in the problem.

*Relation of performance to representation use*

We next consider student performance as a function of multiple representation use. That is, we ask whether students that used pictures and FBDs performed better. We cannot compare problem-by-problem performance between picture-drawing students and non-picture-drawing students since nearly all drew a picture. Instead, we begin by examining student success as a function of correct FBD use. Previous work has shown that students who construct a correct FBD to help them solve problems do significantly better than students who do not construct diagrams or who construct incorrect diagrams.[53] In Figure 28, we see the success rate for students correctly identifying 0, 1, 2, or 3+ (3 or 4) forces per problem on the challenge problem. Since the CU and Rutgers problems differed in format (CU being free response and Rutgers multiple-choice), we have normalized the data to reflect this: Each CU data point has been renormalized by a constant factor so that the CU and Rutgers overall mean scores are identical, allowing for easier trend comparison.

This scaling does not change the shape of the curve observably. Overlap is very thorough. Student performance drops from 0 to 1 forces identified, and increases to 2 and finally to 3. Uncertainties are relatively large for 0, 1, and 2 forces, but we consider the fact that both schools' curves overlap so closely to make the observed trend more likely to be real. Averaging the CU and Rutgers data sets results

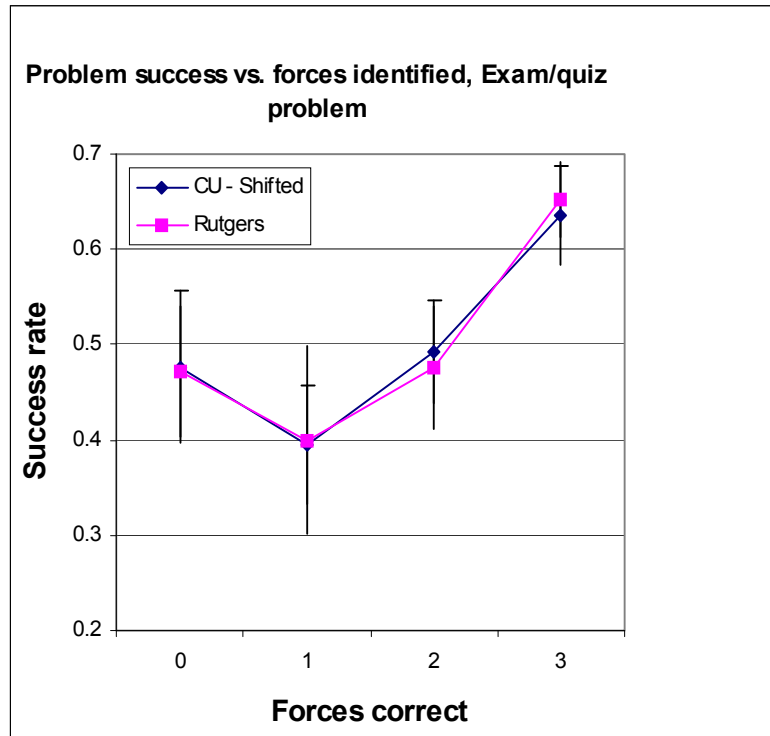


Figure 28. Challenge question performance as a function of number of forces identified correctly. Note that CU scores have been shifted to account for free-response/multiple-choice difference.

in error bars of approximately 70% of this size (not shown). We can perform a similar analysis for problems 1 and 2. For those data (not shown), we cannot conclusively claim that more correct use of multiple representations leads to higher performance (and neither can we claim that it does not): the trend is more-or-less flat. We note here that this analysis would be inappropriate for problems 3 and 4, as problem 3 provides students with a complete FBD already, and problem 4 tells students explicitly to draw an FBD as part of their answer. In the challenge problem and in problems 1 and 2, a free-body diagram is potentially useful but is neither provided nor required.

The above data suggest to us a finer-grained analysis here. Is success associated with any more specific pattern of representation use? In Table XVII, we show student performance versus the identification of each force, the correct representation of each force, and the correct, labeled representation of each force present in the exam problem. Rutgers and CU data are very similar, so we display only CU data. Student performance is flat along the vertical dimension (which would show a dependence on correctness or labeling), and mostly flat along the horizontal dimension (which would show a dependence on force type). There is a minor excess in the second column, corresponding to the electrostatic force. Notably, this is the force whose correctness can be most easily impacted by drawing a type 2 versus a type 3 picture, so this excess might be more reflective of picture type than anything else. Generally, the weak dependence on any one factor suggests that only correct coordination across all of the forces will be associated with success.

Forces	mg	F (electrostatic)	Tension
Present	0.30 (243)	0.38 (186)	0.32 (222)
Correct	0.30 (243)	0.38 (114)	0.36 (198)
Correct and labeled	0.31 (240)	0.39 (108)	0.34 (191)

Table XVII. Fraction of students answering the challenge problem correctly, shown for each force available in the problem and broken down according to whether that force was present, drawn correctly, or drawn correctly and labeled. Data for CU and Rutgers are very similar, so we display only that for CU.



### *Part III: Cross-class comparison*

The above data are interesting when viewed from a problem-by-problem perspective. We see that multiple representation use can significantly influence success, especially on more difficult problems, and that complete, correct multiple representation use is associated with high performance. We also note that the data, when viewed from a cross-course perspective, show similarities and differences. Performances overall are quite similar, as is the dependence of performance on representation use. In this section, we investigate those differences and similarities in more detail, to work towards an understanding of their source.

#### *Cross-class performance*

First, we compare Rutgers and CU performances on problems 1-4 (Table XIV). We can compare the performances pairwise, but since this analysis is post-hoc we must modify the p-value considered significant (or use an appropriate post-hoc test). A simple and very conservative approach is to choose the p-value such that if one were to make N post-hoc comparisons, a difference on any one comparison could be considered significant. We thus choose  $p = 1 - (1-\alpha)^{1/N}$ , where  $\alpha$  is the desired significance level (0.05) and N is the number of post-hoc comparisons (4).[96] This yields  $p = 0.013$ . The CU/Rutgers performances on problem 3 differ at a  $p=0.0002$  level using a two-tailed binomial proportion test, but no other pair differs significantly. Averaged across all recitation problems, the two courses do not differ significantly in problem performance.

### *Cross-class representation use*

Perhaps the most noticeable result is the very large fraction of both course types that made use of pictures and free-body diagrams, despite the significant differences in instruction. Student performance is also very constant across courses, as the performances for problems 1, 2, and 4 are statistically indistinguishable, with the challenge problem performances also similar after accounting for the format differences.<sup>xi</sup> Thus, we have a significant performance difference on only one of the five problems studied. However, some differences emerge in representation use. On the recitation problems that neither demand nor provide a free-body diagram (problems 1 and 2), the CU students identify at least one force correctly significantly more often (43% vs. 31%,  $p = 0.002$ , Table XIV). In contrast, on the exam problem (where the vast majority of students in both courses draw some forces), the Rutgers students are significantly more likely than the CU students to identify all three forces correctly, generating a complete and correct FBD (51% vs 32%,  $p < 0.0001$ , Table XVI). Picture use is comparable on problems 1 and 2, but on problem 4 CU students were more likely to draw a picture (90% vs. 73%,  $p < 0.0001$ ).

### *Cross-class performance versus representation use*

As noted, the dependence of performance on representation use is similar in both classes. The trends of correctness vs. FBD use in Figure 28 are nearly identical,

---

<sup>xi</sup> We can coarsely correct for the format differences by assuming that  $\frac{1}{4}$  of the CU students that answered the question incorrectly would have been able to answer correctly by chance in a multiple-choice format. Very few of the incorrect free-response answers matched the Rutgers multiple-choice distractors, removing a possible confounding factor.

and in neither class does the performance difference for the challenge problem depend on which specific force was identified (as opposed to how many). Since the data suggest that complete and coordinated use of multiple representations is most relevant, we can continue along these lines by breaking down student challenge problem performance by both FBD use and picture use. While nearly all students drew a picture, not all students drew the same picture. In Table XVIII, we show student performance as a function of picture type drawn (2 or 3) and as a function of the number of forces correctly identified (2 or 3). We note that for Rutgers students, the performance difference between using two forces and three forces was minimal, while the difference between using picture 2 and picture 3 was large. Conversely, for CU students the difference between using 2 forces and 3 forces was large, while the difference associated with the picture types was small.

Rutgers	2 forces	3 forces	CU	2 forces	3 forces
Picture 2	0.44 (39)	0.46 (37)		0.28 (47)	0.38 (21)
Picture 3	0.75 (8)	0.71 (87)		0.22 (27)	0.42 (47)

Table XVIII. Fraction of students answering the challenge problem correctly, broken down by whether they drew picture type 2 or 3 and whether they identified 2 or 3 forces correctly.

From the above, we see some specific differences between the two classes: the CU students appear to be more likely to use multiple representations on the simpler problems (specifically, problems 1, 2, and 4), while the Rutgers students are more successful with FBDs on the more difficult challenge problem. Furthermore,

the correctness of the picture seems to be a more significant factor for Rutgers students, while correctness of the FBD appears to be the most significant factor for CU students.

*Part IV: Meta-representational competence*

With student performance on the study problems in-hand, as well as their answers to the multiple representations survey, we were able to correlate their question responses with their actual performance. For CU, we had  $N = 204$ , and for Rutgers, we had  $N = 169$ . There were a total of 26 items on the survey. For this post-hoc analysis, we need the statistical significance of a result involving any individual item to be better than  $p = 0.002$ . For a correlation between a performance parameter and student responses to a survey question, this corresponds to a correlation coefficient of 0.23 for the Rutgers sample, or 0.21 for the CU sample.

Three of the multiple representations questions showed a significant and positive correlation with student performance on the problems studied in the CU sample. Two of these were agree/disagree statements: “I am good at finding and fixing my conceptual mistakes”, and “I am good at figuring out how closely related different representations are (words, equations, pictures, free body diagrams, etc.)”. A fourth was a rating question: “On a scale of 1-5, rate how often you use the following (when applicable) in solving physics problems, and how comfortable you feel when doing so (5 being the highest)”, with the relevant subquestion involving comfort with free-body diagrams. The correlation coefficients were 0.26, 0.21, and 0.21, respectively.

For the Rutgers students, the same three questions showed positive, significant correlations (0.31, 0.25, 0.28). In addition, a fourth question showed a correlation with performance: "I am usually good at learning physics on my own, without any help from others" (0.29).

### *Discussion*

Our first goal was to ask whether multiple representation use mattered, and if so, how. The challenge problem data confirm what has been observed previously: Students that use free-body diagrams correctly significantly outperform those who do not.[80]<sup>xii</sup>

However, the trends were less clear for the recitation data, especially for problems 1 and 2. There, the quality of a student's FBD is not clearly associated with their success, which may be due in part to the relative simplicity of these problems. A student with a good grasp of the material could reasonably solve both of these problems in a "plug 'n chug" fashion, without any additional representations, leading to a less-straightforward dependence.

The challenge problem was more difficult, and perhaps benefits more from the use of a picture and free-body diagram. This is consistent with the fact that many more students used both pictures and FBDs for the challenge problem than for the recitation problems, and with the fact that the dependence of performance on representation use was much clearer for the challenge problem. Thus, these data

---

<sup>xii</sup> We will not attempt to argue causation here. It could well be that using multiple representations leads to better performance directly. Alternatively, it could be that better students overall are more likely to use multiple representations. Or, most likely (in our opinions), these are both true, and are intertwined.

suggest the somewhat intuitive result that for difficult problems in this style, multiple representations can be especially helpful. There are no guarantees, of course: From Figure 28, we can see that students who drew an FBD that was only partially correct were no more likely to answer the problem correctly than those who drew no FBD at all. Indeed, it is possible (though not conclusive from these data) that the students who drew only one force correctly did worse than those who drew none, which is reasonable if we assume that the "no forces" group includes some students who are extremely comfortable with the material and skip diagrams, or keep track of information with mental representations rather than external representations. Along these lines, we note in Table XVII that no one type of force in student solutions was a driving factor in student success. Only the successful coordination of all three forces was associated with better-than-average performance.

On a problem-by-problem basis, we note that problems 1 and 2 are very similar in their solution. Each requires a single use of Coulomb's Law with one variable missing. From an expert perspective, it is not obvious that one is more difficult than the other, or that one is more likely to benefit from a picture and/or FBD than the other. Students did, in fact, perform very similarly on the scalar-only parts of problems 1 and 2. Yet, many more students drew an FBD for 1 than for 2 (45% vs. 29%). This difference occurs despite the fact that problem 2 hinted directly for students to use a force diagram, and problem 1 did not. We speculate that this variation might have been a result of the problem framing. Problem 1 asked students to calculate a force using Coulomb's Law, perhaps suggesting an FBD, while problem 2 asked students to calculate a charge. This is potentially significant if true: A

change in framing (in this case, a language cue) had a significant effect on multiple representation use where an explicit statement did not. In future work, we will vary the problems slightly to look for influences on representation use. For instance, problem 2 could be changed to provide charge and request force magnitude. If the framing is in fact responsible for the difference in representation use between problems 1 and 2, we should expect this change to problem 2 to result in increased use of FBDs.

Our second goal was to compare the two courses directly. Both courses were representationally rich, but with a significant difference. The Rutgers course strongly directed student use of multiple representations, providing specific problem-solving procedures that were emphasized throughout the course. The CU course was representationally rich, presenting a variety of representations in lecture, recitation, and on exams, but did not teach specific procedures. Despite this difference, both courses were very successful in fostering multiple representation use. On all five problems in both recitation and exam environments, students were extremely likely (typically >90%) to use supplementary representations like pictures and FBDs. For comparison, van Heuvelen observes much less frequent multiple representation use in traditional courses.[5] Performance was also quite similar across the courses, and only problem 3 showed a significant difference. The main feature distinguishing problem 3 from the others was the fact that a picture and FBD were included with the problem statement, so students were less dependent on their own supplementary representations.

While we consider the major result to be the strong similarities between the results for both courses, some aspects of the data did differ from course to course. The CU students were more likely to use multiple representations on the shorter, easier recitation problems (particularly problems 1 and 2). The Rutgers students were more likely to use complete and correct FBDs on the challenge problem. This suggests a possible explanation. Since the Rutgers students are being taught (but not graded on) a multi-step problem solving process using multiple representations, they may be less willing to engage in that process in the easier, lower stakes recitation problems, and more willing to engage in that process for a high-stakes exam problem. In comparison, the CU students have learned to use multiple representations, but without specific procedures or guidelines for their use. This could result in more willingness to use them on lower-stakes problems, and in relatively less success with them on higher-stakes problems (though those that do succeed in using multiple representations appear to succeed similarly in solving the problem). If this is the case, we might expect there to be less performance dependence on representation use for Rutgers students on high stakes problems: most students in that case would be using the problem-solving procedure, and the ability and willingness to draw a complete FBD might be less of a discriminator than it would be for CU students. In one sense, we do not observe this. Student performance as a function of FBD correctness is nearly identical in both courses. In another sense, we do see this. Picture correctness is a much more powerful discriminator for Rutgers students than for CU students, which is consistent with the notion that most students, strong or weak, are drawing fairly good FBDs (73% identifying two or more forces correctly),



so that some other factor could present itself as a strong discriminator. Either way, these differences in representation use should not detract from the striking broad similarities observed in the data from the two courses.

There is another possible contributor to the surprising observation that CU students solved problems using multiple representations as often as the Rutgers students, whose course appears more likely upon initial inspection to promote multiple representation use. In chapter 6, we suggested that facility with multiple representations might best be promoted by infusing all aspects of a course with multiple representation use. Here, we saw that the Rutgers and CU environments differed in another way. The CU exams were richer in representations and in multiple representations use, whereas the Rutgers exams were more focused on mathematical representations. The fact that the CU exams were more likely to hold students accountable for being able to interpret a variety of representations might have offset some of the effect of the more detailed Rutgers multiple-representations curriculum. Note that the data do not demonstrate this effect clearly; we mention it only as a possible confounding factor. Nor is this a value judgment: The extent to which course exams should focus on non-mathematical representations is dependent on course goals and upon which aspects of the course are meant to promote which goals.

Our final goal was to test the meta-representational skills of these students by testing whether their responses to questions about their skills with different representations and their uses thereof correlated with their actual performance. We found weak correlations. Of the 26 survey questions used, four showed some

statistically significant correlation with student performance in one or the other course. Two of these questions were unrelated to specific representations, asking about general ability to learn physics and to find and fix conceptual mistakes. The other two asked about comfort with free-body diagrams and student self-perception of their ability to figure out how closely related different physics representations. These questions were not especially noteworthy amongst the broader set. There were a total of twelve questions involving multiple representation use, with only two showing a marginally statistically significant correlation with performance.

Given the exceptionally poor correlations between student self-assessments and performance in our earlier studies, these results are almost surprising. However, we note that only two of the questions specifically regarding representation use showed any correlation with performance, and both of those were near the limit of statistical significance, leaving our previous conclusion intact: Introductory physics students are rarely capable of predicting their own skills with different and/or multiple representations when solving problems.

### *Conclusions*

We can draw two main conclusions from our results, each addressing one of our two chapter goals. First, we confirm that multiple representation use is important in successful physics problem solving as seen in previous work, but find that the dependence is not trivial. Coordinated and correct use of multiple representations on challenging problems can be very helpful, but multiple representation use on simple problems, or poor use of multiple representations, might not have a positive impact

on student success. This dependence of performance on representation use was very similar across two different courses. Second, we find that multiple representation use can be taught, and in more than one way. One of the physics courses studied took a strongly-directed approach to teaching physics problem solving with multiple representations, while the other took a weakly-directed approach. Both courses were very successful in promoting multiple-representation use across a variety of problems, and student performances were generally comparable. Notably, both courses were heavily PER-influenced. We observed some minor differences between the two courses. The CU students were more likely to use multiple representations on some of the easier problems, while the Rutgers students were more likely to use multiple representations correctly on the more difficult and higher-stakes challenge problem. This is consistent with the idea that Rutgers students are learning approaches to using multiple representations which, while successful, might not be drawn on in lower-stakes situations.

In addition to the above, we observe that problem framing may have a powerful effect on student use of representations, possibly a more powerful effect than explicit references to multiple representations in the problem. This observation is tentative, but it reinforces previous results of this nature,[106] and we plan to investigate this more thoroughly in future work.

For instruction, we note that multiple representation use can be taught successfully. Furthermore, an instructor can do so in either strongly or weakly-directed manners. Neither of these approaches was clearly superior for this purpose, so the instructor has some freedom in choosing between them according to other

course goals. Alternatively, one might adapt elements of each. Also, neither approach resulted in particularly strong meta-representational skills, consistent with our earlier results.

From the above, we can draw some broader conclusions. First, we see our earlier suspicion confirmed: successful multiple representation use in these problems depended less on any one representation than on the coordinated use of all the representations present. Thus, at the level of task, it remains likely that we will not be able to make sense of student use of multiple representations when problem solving by treating the various representations in isolation. At the level of environment, we see two different methods of teaching multiple representation use leading to broad similarities in performance. The similarities are consistent with our chapter 6 hypothesis, that representationally rich course environments will lead to greater facility with a variety of representations, used singly or together in problem solving.

## Chapter 12: Task-level analysis of multiple representation use by experts and novices

### *Introduction*

In the previous chapter, we examined problems that were explicitly based on the use of multiple representations, in contrast to the generally single-representation problems from our earlier studies. Much of the analysis was focused on the level of environment, where we characterized the two different kinds of classes and approaches to teaching multiple representation problem solving and the effects that those approaches had on aggregate data. We found that multiple-representation rich instruction does foster use of multiple representations during problem solving (pictures and free-body diagrams in this case), and that this can happen in substantially different environments. We also found that using multiple representations during problem solving was not a guarantee of success. Success was much more strongly associated with a correct, coordinated use of the representations present than with just use. This finding is consistent with the picture that has emerged so far: Student performance is a function of representation use, but that function is complex, depending not just on the features of the problem and the problem context, but on the relations among those features. This pattern was true even in the first chapters of this thesis, where we studied relatively simple, single-representation problem, first using aggregate data, and then using detailed interviews.

As we have noted, skill in using and coordinating multiple representations is often considered to be a prerequisite for expertise in physics.[1] Thus, to develop a practical understanding of how physics students use representations in problem

solving, we must examine multiple representations problem solving episodes, with a particular eye towards how expert and novice problem solvers differ in their use. In this chapter we present our last major study, in which we examine the most representationally-complex problems present in this thesis. All are explicitly based on the use of multiple representations, with some presenting those representations to the students, and others requiring the students to generate several representations in order to solve the problem. Our focus here will be on the level of individual students solving individual problems on a moment-to-moment basis, providing the highest level of detail that is reasonably available. We have interviewed novice problem solvers taken from introductory algebra-based physics students. We have also interviewed graduate students who serve as examples of expert problem solvers. Very little prior work exists in PER on multiple-representations problem solving at the scale of single problems,[105, 113] (though see Kozma's work comparing chemistry experts in the workplace to novices engaged in academic tasks[2, 51]) meaning that for us to make such an analysis, we will also have to develop the tools to do so.

We have three major goals in this study. First, we need to create tools for characterizing and analyzing the use of multiple-representations problem-solving episodes at this fine-grained scale. We will have three categories of tools. The first of these is timing data that takes into account the amount of time spent on various tasks. For example, we will measure how often and when students use representations during problem solving episodes. Second, we find it useful attend to sequential data, or data that focuses explicitly on the order in which students use

particular representations, with less attention on the time spent on each one. Third, we will see that in addition to describing which representations are in use and for how long, we need to characterize the kinds of activities students are using these representations for. For this, we adapt a classification scheme used by Schoenfeld[54] (with such activity categories as analysis, implementation, and verification), resulting in what we shall refer to as Schoenfeld diagrams.

With these tools in hand, we can apply them to our second goal. In order to work towards a model of how students handle representations, we would like to have very thorough characterizations of problem-solving episodes available. We will present three case studies of individual students, who we classify as a weak novice, a strong novice, and an expert, respectively. These students serve as a summary of the twelve novice and five expert students that we have analyzed. In these case studies, we will see the themes previously established in this thesis and will examine them in more depth.

Our third goal is to look for general patterns in multiple representation use, especially when they allow us to make generalizations about the differences that exist between experts and novices. We find three results. First, the experts solve their problems<sup>xiii</sup> more quickly, often using the same set of representations as novices but in a shorter time. Surprisingly, the novices interviewed do not show the reluctance to use a variety of representations that one might expect. Second, novices are more concrete in their representation use, engaging in simple linear sequences of representations when they are successful, and usually being consistent in their choice

---

<sup>xiii</sup> One might at this point ask whether it is appropriate to speak of these tasks as “problems” from the point of view of the experts, a point we will return to in this chapter.

of representations as starting points for all problems. Experts are more likely to vary their starting points, and to visit and revisit the different representations available even when they are not having apparent difficulties. Third, experts differ quite noticeably from novices regarding their applications of representations, with more careful analysis and self-checking, and less weakly-directed, unplanned work. These results together allow us to paint a picture of what distinguishes a novice from an expert in terms of representation use, a picture likely to have consequences for instruction.

#### *Methods – General*

Our problem-solving interviews were in the style of those previously conducted in this thesis (chapter 7, Ref. [26]). Six of our novices were drawn from the first-semester large lecture introductory algebra-based physics course from the fall of 2005, which we refer to as physics 201. The other six novices (for a total of twelve novice interviewees) were drawn from the second semester of this course in the spring of 2006, physics 202. Our five expert problem-solvers were physics graduate students, usually in the first year or two of their program. All participants volunteered after solicitation through mass emails, and were paid for their time. Interviews typically lasted about an hour.

The three groups of students solved three sets of problems. The 201 novices solved what we will refer to as the car problem. The students were given sets of representations of the motion of a car. This included a set of graphs of position versus time, a set of graphs of velocity versus time, a set of Flash animations



depicting a moving car, and a set of written descriptions of a moving car. The students were instructed to make as many groups as possible of members from the various sets; that is, they were told to select position graphs, velocity graphs, animations, and written descriptions that all corresponded to one another. They were also told that not all members of each set would be used in all groups, and that it was possible to find partial groups of less than four elements. In general, we took care to make it difficult to construct groups based on elimination or other such strategies. There were six flash movies, providing a rough span of the motions present in this problem. In Movie A, a car enters from the left of the screen, slows down, reverses direction, and exits, showing constant acceleration. In Movie B, the car starts on the left side at rest, and accelerates constantly until it goes off-screen to the right. In Movie C, the car drives from left to right at a constant speed. In Movie D, the car comes in from the left, stops suddenly, remains motionless for a moment, and then exits to the right with constant acceleration. In Movie E, the car is motionless. In Movie F, the car moves from left to right, slowing down without coming to a halt. Corresponding motions (and others) are represented in the graphs of position and velocity and in the written descriptions.

The 202 novices solved the five electrostatics problems seen in chapter 11 and in Ref. [27]. All of these problems involved numerically calculating either a force or a charge. One explicitly required the production of a free-body diagram as part of the answer, while all five (especially the fifth, or challenge problem) were made easier by drawing a picture and/or an FBD.

The expert problem solvers solved all of the problems given to the 201 and 202 novices, as well as one problem designed to be challenging for an expert. The expert problem is the pulley problem used by Larkin in Ref. [3]. The problem statement follows:

We have three pulleys, two weights, and some ropes, arranged as follows:

- 1) The first weight ( $W_1$ ) is suspended from the left end of a rope over Pulley A. The right end of this rope is attached to, and partially supports, the second weight.
  
- 2) Pulley A is suspended from the left end of a rope that runs over Pulley B, and under Pulley C. Pulley B is suspended from the ceiling. The right end of the rope that runs under Pulley C is attached to the ceiling.
  
- 3) Pulley C is attached to the second weight ( $W_2$ ), supporting it jointly with the right end of the first rope.

Find the ratio of  $W_1$  to  $W_2$ .

This problem is nearly insoluble without constructing a picture and a free-body diagram, though we encourage the reader to attempt the solution with only a mental representation of the problem. The answer is available in a footnote in the

data section of this chapter. All of the problems discussed above are reproduced in Appendix G.

### *Methods – Coding*

All of these interviews were coded in two main ways. First, we coded representation use as a function of time. We divided the interview episode into ten-second blocks and noted which of the available representations (pictures, FBDs, written language, math, movies, position graphs, velocity graphs) students made use of or reference to in each block. It was possible for more than one representation to be present in more than one block. This procedure was quite similar to that demonstrated in chapter 6, and we performed no additional reliability tests.

The parts of the interviews corresponding to the solution of the FBD challenge problem were coded in a second way, as well. Here, we coded for the kinds of activities students engaged in, such as planning and implementation. We adapted the activity categories and rubric from Schoenfeld,[54, 114] with input from the thesis author, the thesis advisor, and member of the CU PER group unrelated to the project. The activity rubric follows:

Reading: Reading the problem statement, either out loud or quietly. This includes silence following the initial reading when accompanied by a gaze or gesture in the direction of the problem statement.

Translation: Taking information directly from the problem statement and re-representing it. This includes writing numerical data or the quick construction

of a diagram on which to place data from the problem statement. This does not include substantial work independent of the problem statement, but only back-and-forth from the problem statement to a representation of given data.

**Analysis:** Represents a directed, systematic attempt to more fully understand the problem. It can include constructing supplementary representations like free-body diagrams or pictures once the reading and translation phases are over. It can include talking out loud about their understanding of the problem. Generally, analysis has an end in mind, either explicit or implicit.

**Exploration:** A less-structured version of analysis. The student is searching for options or trying things out with little direction or expectation of moving forward. Examples include a student searching through equations in the book, or cycling through their previous work out of apparent inability to proceed further.

**Planning:** Explicitly planning future moves (this usually means saying it out loud). A student will say “Ok, I need to draw a free-body diagram, then I’ll balance forces, then I’ll solve for my variable.” When they begin implementing their plan or they return to analysis, this phase ends.

**Implementation:** Differs from analysis and exploration in that it is strictly process-oriented or “crank-turning”, generally mathematical, but including any algorithmic crank-turning with little clear awareness of what students are doing or why (for instance, drawing a picture because “that’s what you do”, even if a student never refers to it again). Note that this does not require

students to be turning the correct cranks; if a student is engaged in a mathematical wild-goose chase, this counts as implementation.

Verification follows the discovery of an answer (right or wrong), and involves taking some kind of step to check the correctness of the answer (a final or intermediate answer). If a student decides outright that an answer is wrong and begins work again, this is analysis/exploration/implementation rather than verification.

As before, we divided the interview into ten-second segments and coded each block according to which activity the students were involved in. No block was coded as having more than one kind of activity unless that block represented a transition from one activity period to another. To test for inter-rater reliability, the thesis author and one unrelated researcher each coded the same problem-solving session. The coders agreed 86% of the time. An example Schoenfeld plot is shown in Figure 29. The episode displayed is a novice attempt at solving the electrostatics challenge problem. The student's attempt involved a lot of equation-hunting and generally unfocused behavior, and was ultimately unsuccessful. During our analysis, we created Schoenfeld diagrams for each student's solution of the electrostatics challenge problem.

#### *Methods – Timing data*

With representation use coded in time, we then calculated a number of numerical parameters. These included time spent per problem, number of

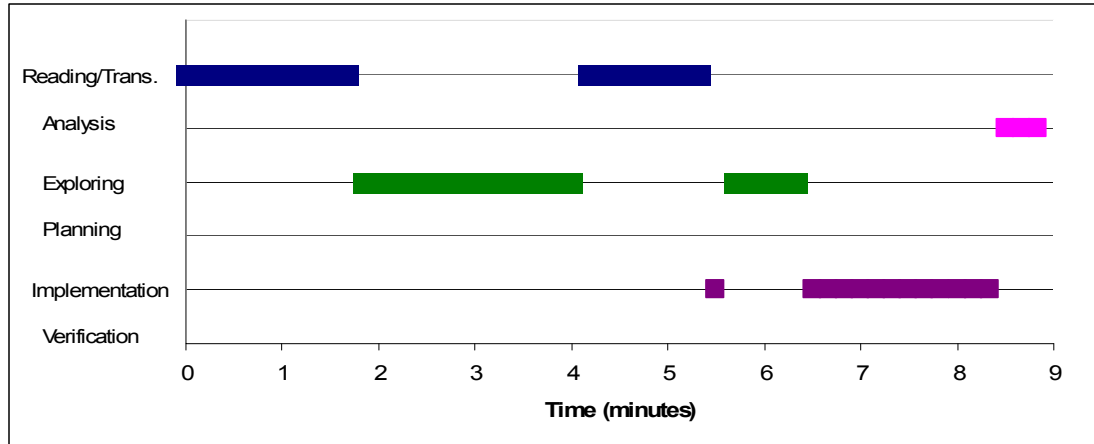


Figure 29. Schoenfeld diagram for a novice student's unsuccessful attempt at the electrostatics challenge problem. Note the lack of planning or analysis.

representations used per problem, the density of different representations used per unit of time, and the number and density of transitions between representations. Transitions were determined via a manual examination of the coded data, as an automated search tended to produce spurious results.

We also produced graphs of the representations being used as a function of time. In Figure 30, we see an example graph, showing a 202 novice's solution of the electrostatic challenge problem. The bars show usage of each of the indicated representations, with time in minutes displayed on the bottom axis.

#### *Methods – Sequential data*

Our timing data are useful for seeing representation use as a function of time, but is less useful at capturing the character of a problem-solving episode at a glance. An alternative is to depict the representations being used in the order that they are

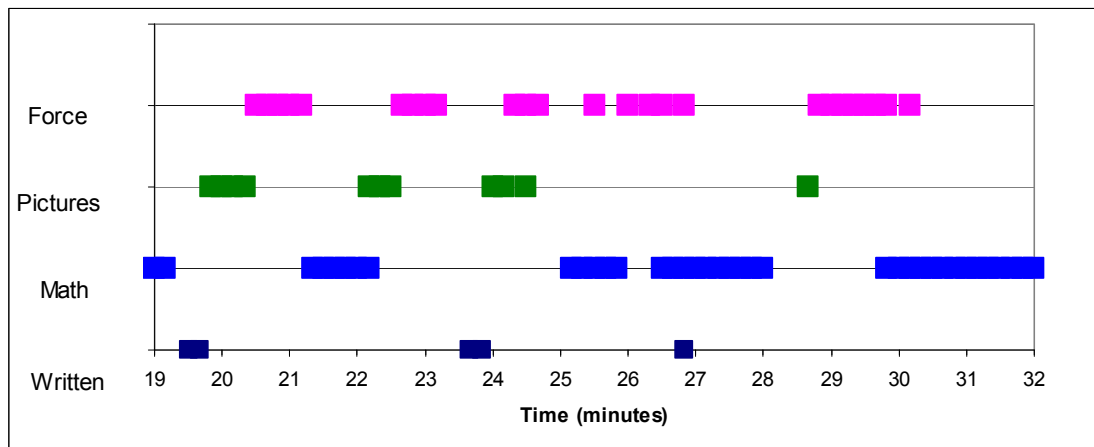


Figure 30. Graph of a 202 novice’s representation use during their solution of the electrostatics challenge problem.

used. Figure 31 shows two such examples. In the first of these, we see a novice student’s correct solution to the first of the five electrostatics problems from chapter 11. The student reads the written description of the problem, works with a set of equations, and draws a picture to clarify the direction of the force calculated. Once they have moved on from a particular representation, they do not return to it. In the lower diagram, we see a student’s unsuccessful attempt to solve the same problem. The double-headed arrow represents a very close coordination between two representations, in this case, the problem statement and a picture that they draw while continuously referring to the problem statement. The student then draws an FBD,

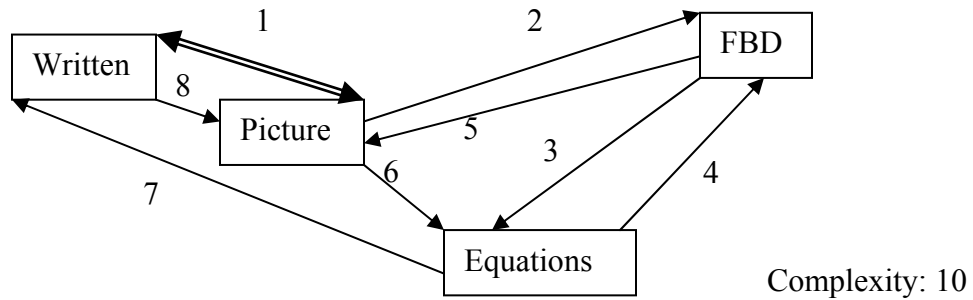
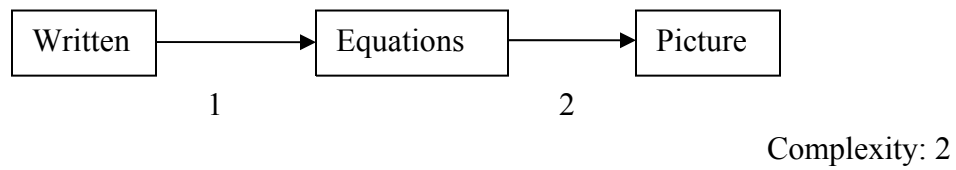


Figure 31. Sequence diagrams showing the order in which students used representations in two of their problem solutions. Above: a correct solution to the first of the electrostatics problems from chapter 11. Below: a solution where the student failed to generate the correct answer. The complexity parameter is defined in the text.

writes some equations, refers back to the FBD, draws a picture, does more math, returns to the problem statement, and takes a look at their picture before giving up.

This visualization, which we will refer to as a sequence diagram, provides a quick and easy sense of the complexity or linearity of a student’s solution. We created a sequence diagram for each of the problems solved by each of the students in this study.

For the purposes of comparing many students, it is useful to define a numerical parameter that expresses the complexity of the diagram. An obvious first choice for this complexity parameter would be the number of transitions indicated by the diagram. However, this choice has two problems. First, it treats close coordination of representations the same way as a switch between representations,



and second, it over-represents quick, repeated transitions between the same two representations. To solve this, we count the number of transitions between representations, counting only one transition per ten-second block for periods of quick back-and-forth, and counting one transition for a move into a period of close coordination, and one transition for a move out of a period of close coordination. This complexity parameter is displayed for each of the problem solutions shown in Figure 31.

### *Data*

The tools described in the methods section provided an enormous data set. As a result, it will be necessary for us to be selective in our presentation. The data we show here have been chosen to be both concise and representative.

#### *Data: Correctness and timing*

In Table XIX, we see student performance on all of the problems solved. One of the students solving the electrostatics problems did not finish due to external factors, and will not be considered here. The Car (avg) column shows the average number of correct and incorrect representation groupings made by the expert and novice students, in a correct/incorrect format. Most of the difference between experts and novices here can be traced to the fact that only one of the novices attempted to make a group that didn't include a movie, whereas all of the experts made at least one subgroup that didn't include movie, using a position or velocity graph as a starting point.

	Cars	Electrostatics					Pulley
		Prob. 1	2	3	4	Challenge	
Experts	7.6/0	4/1	4/1	5/0	5/0	5/0	3/2
Novices	5.0/1.0	3/2	3/2	1/4	0/5	1/4	

Table XIX. Student performance on the study problems. The Car column shows the number of correct/incorrect groups and subgroups made. The columns for problem 1-4 show the number of students answering correctly/incorrectly for the first four electrostatics problem. The Challenge and Pulley columns show the number of students answering the challenge and pulley (expert only) problems correctly/incorrectly.

The columns for Prob. 1-4 and the challenge problem show the number of students answering correctly/incorrectly on the first four electrostatics problems and the electrostatics challenge problem. Note that the novices solving the electrostatics problems are not the same novices that solved the car problems. The last category shows the number of experts answering correctly/incorrectly on the expert-only pulley problem.

Overall, the experts were very successful. One expert missed problems 1 and 2 by incorrectly recalling Coulomb's law, writing it with a  $1/r$  dependency instead of  $1/r^2$ . All other expert solutions were correct until they reached the pulley problem, at which point three were successful and two were unsuccessful.

The novices were reasonably successful on the car problem, forming an average of one incorrect grouping per person. This was frequently a grouping involving Movie F, in which a car slowed down constantly without coming to rest. Despite the speed of the car upon exiting the screen being approximately half of the

initial speed, many of the novices perceived it to be constant speed, or had trouble deciding. In contrast, only one of the experts spent significant time considering this point, and that expert resolved the difficulty correctly. The fact that experts and novices differed so strongly in terms of their qualitative perception of the motion is surprising, but is not unprecedented. This is likely related to the differences in readout strategies[70] observed by Mestre between more and less expert students.[61]

In the electrostatics data, novice performance trends downward as we move across the table from problem 1 to the challenge problem. No students solved problem 4 correctly, and one out of five solved the challenge problem correctly. This was expected, and is consistent with the free-response data from chapter 11.

In Table XX, we see the average time taken by the experts and novices to solve the various problems. For the car problem, both experts and novices attempted groupings based on all of the movies, and so we display the time taken to make that set of groupings (5.8 minutes for experts, and 14.6 minutes for novices). The other columns show the time taken for problems 1-4 considered together, the challenge problem, and the expert-only pulley problem. Generally, experts took 40-60% of the time required by novices to solve the same problems. Note that on the pulley problem, one expert had a very fast solution (4 minutes), while all others took considerably longer.

	Cars, A-F	Prob. 1-4	Challenge	Pulley
Experts	5.8 min.	12.6	4.7	9.2
Novices	14.6	22.0	10.4	

Table XX. Average time for experts and novices to complete the interview tasks, in minutes. “Cars, A-F” column shows the time to complete all groupings involving movies A-F. “Prob. 1-4” refers to the first four electrostatics problems.

*Data: Representation use*

In Table XXI, we see a summary of the overall representation use by the students solving the car problems. There were a total of 9 groups of representations (collections of associated graphs, movies, and descriptions) made by students. Six of these involved movies A-F, with the movie C subgroup only having two elements (the movie and a written description). Three of these groups contained no movie, and were typically centered around graphs of velocity (graphs B, D, and H). Only one novice identified any subgroups that did not include movies, and we do not include those in our averaged data.

Performance data and timing data (not all shown) suggest the approximate difficulty scale shown in Table XXI. The representation grouping centered around Movie E was the easiest (taking the least time and resulting in the fewest errors), followed by that involving Movies D and B, that involving movies A, F, and C, and finally those involving no movies, which were only attempted by experts.  $V_B$ ,  $V_D$ , and  $V_H$  refer to three of the velocity graphs labeled B, D, and H (seen in Appendix G), as these were elements common to the experts’ extra subgroups.

Car problems	Complexity, Experts	Complexity, Novices	T. Density, Experts	T. Density, Novices
E	4.0	5.3	4.8	4.8
D, B	4.8/4.5	6.2/7.5	4.5/4.5	2.9/3.1
A, F, C	6.4/4.0/7.7	8.5/9.2/7.5	4.7/5.6/3.4	3.1/3.2/3.0
V <sub>B,D,H</sub>			3.4/4.7/5.1	

Table XXI. Average number of transitions between representations and density of those transitions (per minute) during solutions. Organized by expert vs. novice and by representation grouping. Groups A-F are those representation groups that include movies A-F. and groups V<sub>B</sub>, V<sub>D</sub>, and V<sub>H</sub> are groupings involving velocity graphs B, D, and H (see Appendix G) but no movie. V<sub>B,D,H</sub> involved no movies and were only found consistently by the experts.

In the first two columns, we see the average complexity parameter associated with the sequence diagrams for each group. Since all the representations needed for the problem were provided to the students, these diagrams show student progression between these various representations. Low parameters represent a solution in which students used only the minimum set of representations in sequence, without revisiting representations during the solution. High parameters represent a more complex, iterative approach in which students moved back and forth between the representations available. We show no parameter for the V<sub>B,D,H</sub> subgroups, as those had only two or three representations in them (rather than four), and so comparing this parameter with that from the other subgroups is not meaningful. For both novices and experts, solution complexity trends roughly upwards with problem difficulty. Complexity is also generally higher for the novice solutions overall.

In the second two columns, we have an alternative method of characterizing representation use. There, we see the time density of transitions between

representations, in transitions per minute. That is, we see a measure of how quickly the problem solver was moving back and forth between the representations available. Experts show no clear and consistent variation in transition density across the tiers of difficulty. Novices show a sharp break between their transition density for the first groups (the trivial case of a motionless car) and all other groups. Notably, the transition density for both experts and novices is the same for the trivial group. The data appear consistent with the interpretation that group E represented a simple exercise for all problem solvers, while the other group represented simple exercises for the experts, but not for the novices.

In Table XXII, we see the same data for the set of electrostatics problems and the pulley problem. As in Table XXI, we have arranged the problems in terms of apparent problem difficulty, defined as before in terms of student success and time to finish. We see no significant differences in solution complexity except for problem 4 and the challenge problem. The novice solutions to the challenge problem are approximately as complex as the expert solutions to the pulley problems.

The representation transition data are intriguing: We see a higher density of transitions for experts solving all problems than for novices, consistent with the expectation that experts will be using multiple representations for these problems, and solving them quickly.<sup>xiv</sup> We also note that transition density is fairly constant for problem to problem, with a spike present for both groups at problem 4. This result

---

<sup>xiv</sup> At this point one could point out, correctly, that experts would not need to create other representations (pictures and free body diagrams) to make sense of the novice tasks. However, it has been our observation in this study that these expert physicists make use of multiple representations even when they don't need to.

likely reflects the fact that this was the only problem explicitly requiring students to draw extra representations as part of their answer.

Electrostatics problems	Complexity, Experts	Complexity, Novices	T. Density, Experts	T. Density, Novices
Prob. 1/2	7.3/5.8	7.8/5.0	2.2/1.8	1.4/1.3
Prob. 3/4	8.5/6.0	6.5/10.3	2.3/2.7	1.3/2.3
Challenge	11.8	15.8	2.4	1.5
Pulley	17.0		2.0	

Table XXII. Complexity parameter describing the representational structure of student solutions, and density of transitions between representations (per minute), organized by problem and expert vs. novice.

*Data: Starting points*

On the problems studied, there was some flexibility as to which representations the students could start with on each problem. For instance, after reading the electrostatics problem statements, students could either draw a picture, a free body diagram, or write equations. On the car problems, we took care not to present any one of the four available representations as the intended starting point, so that students were forced to choose their own starting point. These starting points, and the different choices made by experts and novices, are potentially quite informative.

On the electrostatics problems, experts and novices were quite consistent. Of the five novices that completed all five problems, we see only two instances where the novice started with an equation (one on problem one, and one on problem two, not the same student). In all other cases, the student started with a picture. Experts were similar: On nearly every problem, including the pulley problem, the expert

began with a picture immediately after reading the problem statement. One expert began problem one with an equation, and another started with an equation on problems 1-3.

On the car problems, novices were very self-consistent. Three of the novices began every grouping by picking out a movie to use as a centerpiece, while the other three began nearly every grouping with the written description. Among the students starting with written descriptions, the only exception was one student who started five groups with a written description and began the sixth with a movie. In short, novices perceived either movies or the written descriptions to be the natural starting points and did not deviate from their choice.

Experts were considerably more flexible. Experts started their problems using all four of the available representations (written, position graphs, velocity graphs, and movies). Only one of the experts chose a starting representation consistently (the written one), and none of the other four deviated less than twice from their primary choice. In total, we have 13 expert groupings that began with movies, 13 that began with written descriptions, 7 that began with velocity graphs, and 5 that began with position graphs.

We can compare expert and novice flexibility in at least one other way. When solving the car problems, novices always finished one group before moving on to another, and very rarely returned to a finished group, doing so only when a later group directly conflicted with a previous group. One of the experts departed radically from this pattern, making many of his groups simultaneously, in parallel. A second expert completed as many three-representation groups as possible before going back



and adding in the movies last. The other three experts generally finished groups in sequence, but were relatively comfortable with returning to a completed group or looking ahead to a group yet to be made.

### *Data and analysis: Case studies*

With an understanding of the tools available and some aggregate data in hand, we will now provide examples of thorough characterizations of problem-solving episodes. We will consider the solutions of the electrostatics challenge problem by three example students. The first of these, Carrie, we consider to be one of the weaker novices interviewed, and she is unsuccessful in solving the problem. The second of these, Sam, is the strongest of the novices interviewed, and is successful in her attempt. Finally, we consider Jim, the strongest expert in our opinion. He solves the problem very quickly and correctly. In each case, we will begin by describing the student's performance overall and our impression of the student. We will then examine their challenge problem solution in detail.

#### *Case study: Carrie*

Carrie struggled with the electrostatics problem set, getting none correct. In general, Carrie was not completely ignorant of the appropriate solution techniques, mostly remembering such things as Coulomb's Law and  $F = qE$ , but her recollections were often piecemeal, and her attempts to fill in gaps in her understanding were not successful. Among other things, she remembered Coulomb's Law as having a  $1/r$  dependency, contributing to her errors on the first three problems. In some ways, we

would describe her as the stereotypical novice, applying learned procedures and equations with little understanding.

In Figure 32, we see three representations of Carrie's solution to the electrostatics challenge problems. One of these shows the representations she used as a function of time in minutes. The next shows a Schoenfeld diagram of the activities she engaged in on the same timescale. Last, we see a sequence diagram showing the sequence of particular representations used.

Carrie begins by reading the problem and translating the description into two sketches and a free-body diagram. She then arranges the charge and distance information into Coulomb's law without an intermediate symbolic step, and remembers out loud that the functional form of Coulomb's law is like that of the law of gravity. Next, she draws revised versions of her picture and free-body diagram, adding more labels and information. Importantly, her picture is of the backwards, type two sort identified in chapter 11, with positive and negative charges hanging on strings as if they were repelling.

At 4:30 into the problem, Carrie re-reads the problem and re-examines her picture and free body diagram. She says that she is confused as to why the charges aren't touching if one is positive and the other is negative. Unable to resolve this, she dismisses the issue and returns to Coulomb's law, beginning a long episode in which she tries to decide which elements of her picture fit into Coulomb's law. She eventually gives up.

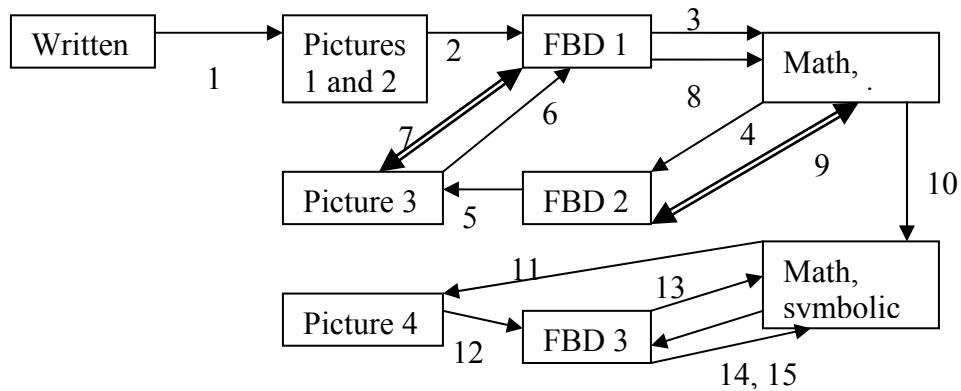
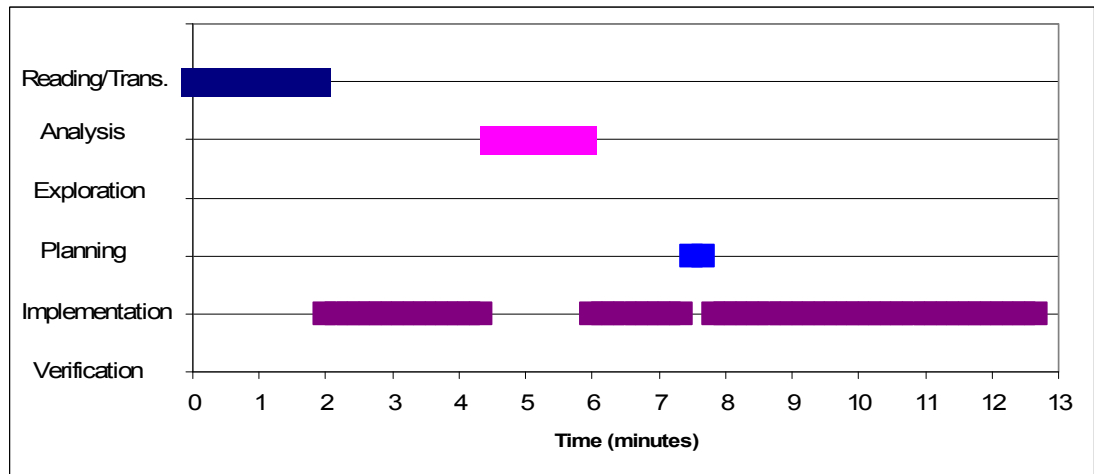
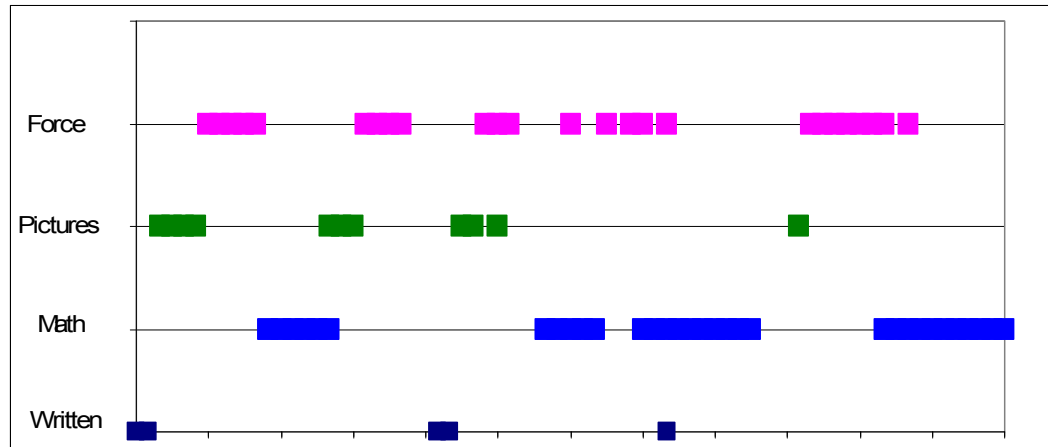


Figure 32. Three representations of novice Carrie's attempted solution of the electrostatics challenge problem. Above, we have the representations used as a function of time. In the middle, we have the kinds of activities she was engaged in. On the bottom, we have the sequence of representations used.

We should revisit our remark about Carrie being a stereotypical novice, as this label has many possible interpretations. Carrie is not hesitant to use non-mathematical representations, and is at times quite careful in coordinating between them. However, this is true of all students studied, and may be a result of the representation-rich course she was a member of (the one featured in chapter 11). Our anecdotal impression was that Carrie was drawing pictures and free-body diagrams because that is the norm, and that only once did she turn these representations towards sense-making and analysis (4:30-6 minutes on the charts). Indeed, most of her time spent was coded as “implementation”, during which she appeared to be manipulating pictures, diagrams, and equations with no obvious direction. Ultimately, the problem was too complex to be solved given her inability or unwillingness to engage in deep analysis/sense-making.

*Case study: Sam*

Sam was the strongest overall of the novices interviewed with the electrostatics problems. Sam answered problems 1, 2, and the challenge problem correctly, and missed problems 3 and 4 due to misplaced factors of ten. Sam appeared comfortable working symbolically, usually made sound inferences, and even used unit analysis to aid in solving problem 2.

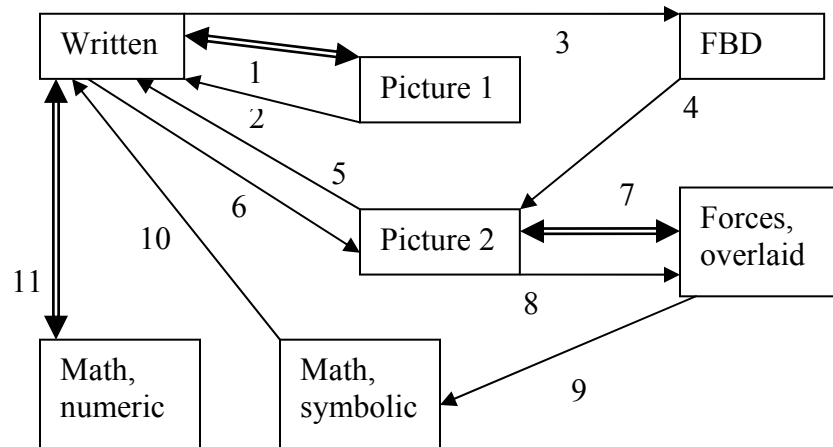
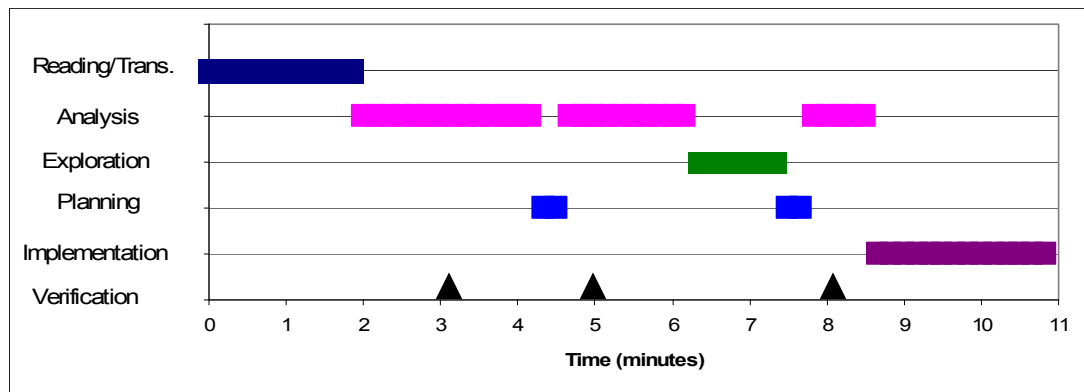
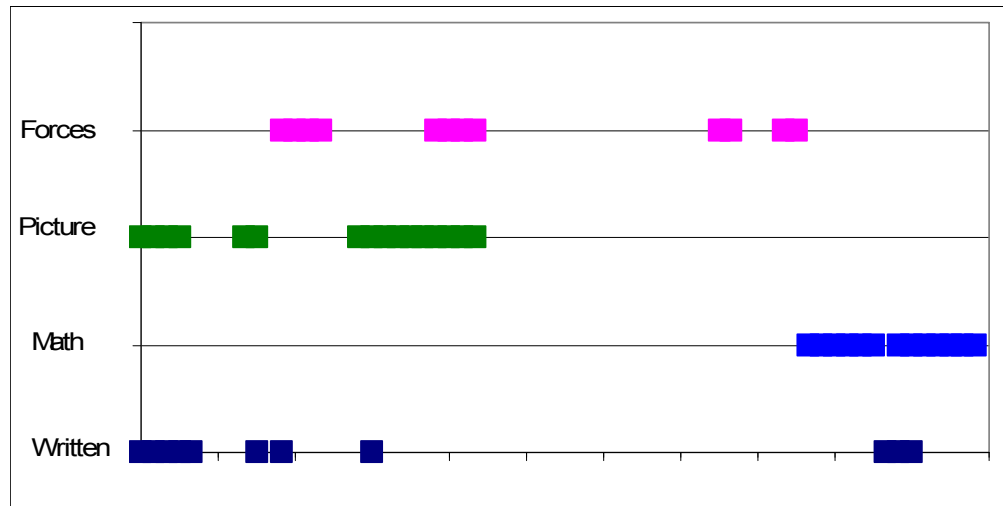


Figure 33. Three representations of novice Sam's successful solution of the electrostatics challenge problem. The triangles on the Schoenfeld plot represent out-loud assessments of how the problem is going so far.

In Figure 33, we diagram Sam's solution to the challenge problem. She began by drawing a picture based on the problem statement, coordinating extremely closely between the two (looking quickly between them, marking points of correspondence with her fingers, and so forth). She briefly revises her picture, and then begins drawing a free-body diagram and thinking out loud about which forces should go where (with this, in part, prompting us to code a period of analysis).

Three minutes into the problem, Sam looks back at the problem statement and notes that the picture she has drawn shows the charges repelling, when they are of opposite sign. She explicitly wonders how this could be correct, and whether she's made a mistake. The tick mark at this point on the graph indicates this explicit and audible self-checking remark. After more analysis and reference to her picture and to her force diagram, Sam draws the correct picture, and draws an updated free-body diagram on top of it. She makes a brief statement about her intentions (coded as planning), before pulling up short.

Near minute 5, Sam says that she does not know which of the available forces she should use to get the tension, or how to combine them if necessary. She begins a period of thinking out loud without clear reference to representations on the paper (but with considerable gesture, a kind of representation not considered in this thesis). The period of exploration indicates an ultimately fruitless examination of her class notes. At 8 minutes, Sam spends more time gesturing and updating her free-body diagrams while thinking about which forces apply. She eventually realizes the solution, which she spends the last two and a half minutes implementing, carefully

referring back to the problem statement to verify the numbers she is using in her symbolic equation

*Case study: Jim*

Jim was a first-year graduate student, but was noticeably stronger than any other expert interviewed. He solved all the problems quickly and correctly, and then solved an additional reserve problem asking him to estimate the number of times per second a bee needs to flap its wings. Despite the apparent triviality of many of the novice tasks (from his perspective), he still drew a picture on all but problem 2 of the electrostatics set.

In Figure 34, we see representations of Jim's solution to the challenge problem. After thoroughly reading the problem (nearly a minute of reading before taking any other actions), he drew a correct picture of the charges on strings. He was nearly unique in this regard, as some of the experts still had some difficulty in translating the problem statement into a picture without getting it temporarily backwards. He then set up a block of equations before re-reading the problem. Just after the 2 minute mark, he said that he was trying to figure out how far apart the strings should be. After some deliberation, he concluded that it did not matter. He then set up a free-body diagram, checked it against his first block of equations, wrote a second block, and solved for the necessary force.

Our primary impression from Jim was one of efficiency with algorithms and mastery of the relevant concepts. Still, the problem was complex enough that his use

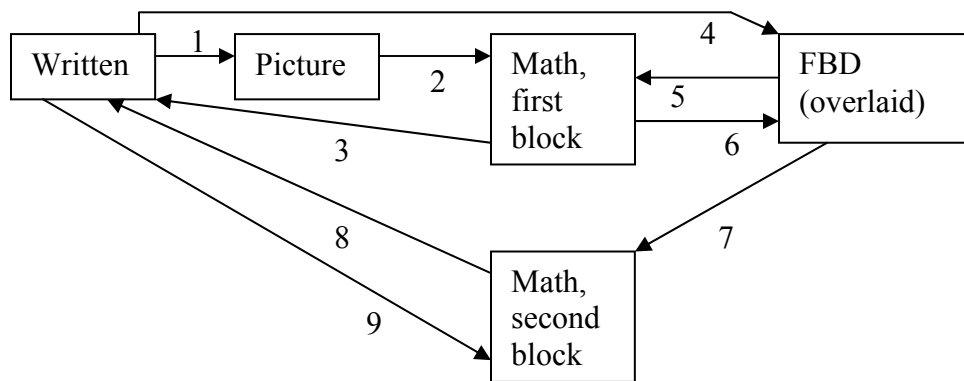
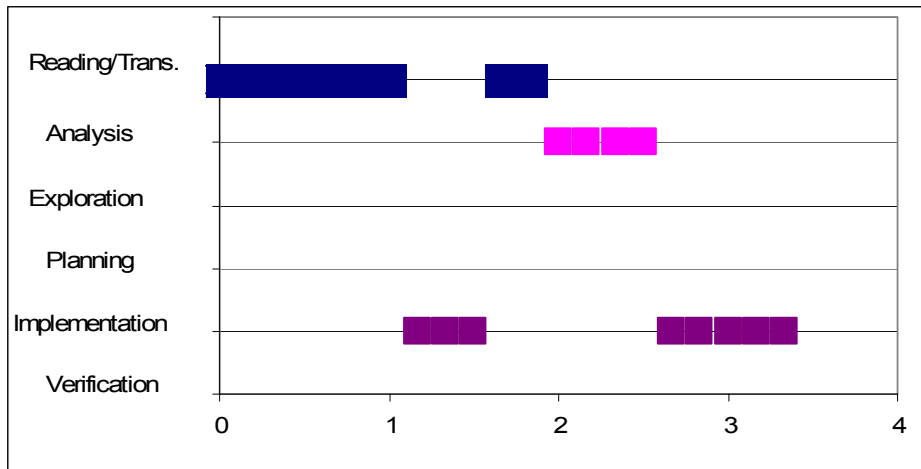
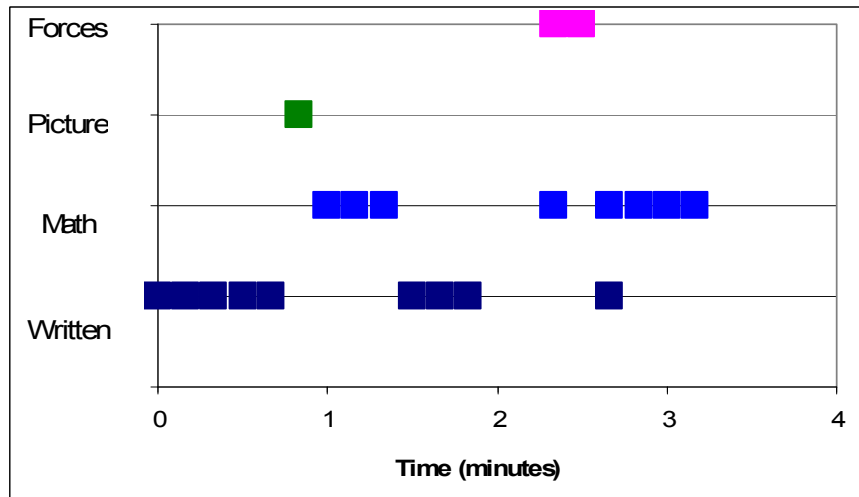


Figure 34. Three representations of expert Jim's successful solution of the electrostatics challenge problem.



of pictures and free-body diagrams did not appear to be token or out of habit; rather, he appeared to be using both of these in making sense of the problem and possibly in non-verbalized self-checks. Note that there was no period coded as verification, as we never perceived him to be engaged only in checking his answer.

*Case studies: Comparisons*

These three students represent the full spread of subjects interviewed, including a weak novice, a successful novice, and the most expert problem-solver interviewed. Despite that, at least one similarity presented itself. All three of these subjects, much like all the others, made considerable use of multiple representations in their problem solutions. This is counter to early evidence of introductory students using multiple representations only rarely,[5] but is consistent with newer observations of students using pictures and free-body diagrams quite often in representation-rich PER-informed courses.[27] Indeed, without checking on the total time elapsed, it would be difficult to identify the weak novice, strong novice, and expert using only the displays of what representations were being used and in what order (the sequence diagrams and representations vs. time chart). These students become most distinguishable when we consider how they applied their representations. The expert and successful novice spent time using the representations to make sense of the physics, while the unsuccessful novice appeared to be drawing pictures and free body diagrams out of a sense of requirement, and not towards any particular purpose.

This sample is representative of the data as a whole. Examination of solution sequence diagrams and representation vs. time charts is not nearly as useful in separating novices from experts as is combining that data with Schoenfeld plots representing the kinds of activities students engaged in. We close this section with a report on an interesting informal test of these results. The thesis author presented the advisor with five sets of the above representations, with no identifying information, and gave the advisor ten minutes to classify the students without being told the categories or how many representatives of each there were. The advisor was able to correctly identify students at the level of weak novice vs. strong novice vs. expert, and was also able to offer some more specific comments regarding their likely solution strategies. The advisor had not yet seen any data regarding these students. It appears that these representations allow very thorough characterization of multiple-representations problem-solving episodes in a fairly compact format.

### *Discussion*

In this chapter, we had three major goals. First, we wished to develop useful tools and procedures for characterizing multiple-representations problem solving at a fine level of detail. To this end, we have settled on coding schemes, diagrams, and tabled data that, as far as we can tell, provide a reliable and detailed picture of a student's solution processes that is reasonably concise.

Second, we planned to provide analyses of representative problem-solving episodes, giving us a sense of how different kinds of students solve physics problems with multiple representations. Our interviews with Carrie, Sam, and Jim are

representative of all the interviews conducted. We see, as in previous chapters, that even unsuccessful novices often attempt to use multiple representations. This may be because doing so has become the norm in these physics classes, and so they draw pictures and free-body diagrams without significant understanding of why. We also see hints of the general lack of meta-representational skills inferred from chapters 5-7. Sam was an exceptional novice, and was one of the only ones that was clearly using multiple representations for sense-making, with some apparent idea of why a physicist would want to use multiple representations. Our observations do, however, provide an interesting caveat to our previous claim that one may be able to neglect meta-representational competence in an analysis of representation use in introductory physics problem solving. It still appears likely that intro students bring very little in terms of understanding why or how to use representations. However, it also appears that meta-level skills regarding applying these representations (as seen in the Schoenfeld plots) are among the biggest we have observed between experts and novices.

Our third goal was to look for generalizations about representation use in physics problem solving, especially when those generalizations differentiate between experts and novices. We have already mentioned two such generalizations: ubiquitous multiple representation use in all groups and the difference in metaskill observed between experts and novices.

Some of our other observations were less surprising: Experts solved the novice-level problems in much less time, with much more success. Expert solutions were also slightly simpler, as they did not move back and forth between the

representations they created or had available as often as did novices (see the complexity parameters in Tables XXI and XXII, for example). In general, solutions from both groups became more complex as the problem difficulty increased and as student success decreased, with expert solutions of the expert-only pulley problem containing more complex representation use than any other problems. In contrast, successful novice student solutions for the simpler problem were usually very linear, in that the representations used were each only used once, and in sequence.

The electrostatics problems were of a style familiar to the students, and they tended to rely on a standard pattern of representation use (read the problem, draw a picture and/or free body diagram, and solve equations). The car problems were designed to be less familiar and potentially much more flexible. Here, the experts and novices differed in their representational flexibility. Novices always worked from the same starting representations in solving the car problems, and, across all novices interviewed, only used two of the available representations as starting points. Novices also solved one group fully before progressing to another, almost never revisiting old answers or leaving a partial group early. Experts used all four available representations as starting points, used different representations as starting points within the same task, and were more likely to work in a piecemeal or iterative fashion when it suited them, with no associated performance cost.

### *Conclusions*

In this chapter, we successfully developed and used a set of tools that allows us to characterize individual multiple-representations problem solving episodes with

considerable detail. These tools, when applied to interviews of successful and unsuccessful novice and expert problem solvers, allow us to begin to identify the major features that distinguish novices from experts when solving physics problems that are representation-heavy. Surprisingly, students at all levels were very willing to use multiple representations, which is perhaps another example of the effects of the broader environments. For reasonably complex problems, their solutions also looked very similar from the point of view of the representations used and the complexity of the sequence of those representations.

There were at least two major differences between the experts and novices seen here. One was the apparent flexibility of the experts in the less-constrained problems, with the other being simple facility with the various representations. The experts were unlikely to make a mistake in representing or re-representing problem information, even in the case of the enormously complex pulley problem. The second major difference was perhaps the clearest indicator of expert/novice status without access to performance information. Experts showed consistently superior meta-level problem solving skills, engaging in considerably more analysis and planning. Novices were much more likely to behave mechanically or algorithmically, and to produce multiple representations without being able to make much use of them.

If this is widespread and repeatable, it is quite significant. As we have seen in this thesis, our PER-informed courses are successful at getting students to use a variety of canonical representations while solving physics problems. However, they may not be learning why they are using these representations, or how to use them to maximum effect. This is reasonable, since meta-level skill tends to be slow to

develop.[84] Furthermore, classes almost never teach meta-representational competence or other meta-level problem solving skills explicitly, meaning that if they are learned, they are picked up informally over many courses. The significance of these skills in our observations raises the question of whether these skills can be taught formally, and whether they can be consistently taught to novices at all in the short time available to introductory physics instructors.

## SECTION IV: SYNTHESIS AND IMPLICATIONS

### Chapter 13: Model and heuristics

This thesis has been concerned with filling in some of the major gaps in our knowledge of how physics students, especially novices, use representations when solving problems. Ultimately, we hoped that this foundation would allow us to lay down a tentative model and/or set of heuristics for describing representation use in problem solving. In chapter 4, we suggested starting points for such a model drawn from the existing literature. Here, we review some of the candidate ideas and how they relate to the studies we have performed.

#### *Coordination classes and p-prims*

In his theoretical framework,[58, 70] diSessa defines the coordination class to be a conceptual structure usually associated with an expert. This structure contains a collection of inferences potentially both quantitative and qualitative, readout strategies, and sets of rules regarding the contexts in which the different inferences and readout strategies apply. It is robust and useful in a variety of contexts, though different contexts can trigger different aspects of the coordination class. We have not often used the concept of a coordination class to interpret the results of our experiments, and we did not expect to. The coordination class represents the sort of compiled physics knowledge that few introductory students possess.<sup>xv</sup>

---

<sup>xv</sup> It is possible that introductory physics students possess knowledge in other domains that fits the definition of a coordination class.

A precursor to and component of coordination class theory is the idea of “knowledge in pieces”, or p-prims.[57] These p-prims represent basic elements of our reasoning about the world, such as the idea that “more X implies more Y”, or the idea that under certain circumstances, things balance. Once triggered by a particular context, these p-prims go unchallenged by those using them. DiSessa describes novice knowledge as fragmentary, with mistakes in reasoning often attributable to the activation of p-prims in inappropriate contexts (such as association of the “dying away” p-prim with a frictionless physical situation).

The p-prims framework carries two additional, intertwined principles with it: the ideas of cueing and the context sensitivity of novice knowledge. For a novice, the same idea (by the expert’s standard) in different contexts can be associated with much different p-prims. Oftentimes, differences in representation can cue different responses, as with Elby’s WYSIWYG p-prim.[90]

The ideas of p-prims, knowledge in pieces, cueing, and context sensitivity, when turned towards making sense of representation use, have been enormously productive in this thesis. In chapters 5-7, 9, and 10, we saw strong examples of similar questions posed in different representations resulting in different student responses and student strategies, sometimes including what could be considered readout strategies (for example, Doug’s interpretation of the graphical spectroscopy quiz in chapter 6). In our detailed problem-solving interviews, we have seen p-prim-type arguments in action, like “lower is lower” or WYSIWYG. This was often representation-specific, as the problem features responsible for cueing students would



only be present or compelling in one of the available representations (the Bohr model homework of chapter 5, or the hill problems of chapter 10).

Note that we do not always make use of the full range of coordination class theory in this thesis. We have only briefly touched on readout strategies (in chapter 7 as noted above, and in chapter 12 involving expert and novice perceptions of car motion). While we believe that the concept of readout strategies may be quite powerful in making sense of representation use generally, we have found that our needs are often met by a focus on p-prims.

### *Contextual constructivism*

Another major framework from chapter 4 was contextual constructivism.[74, 75] Contextual constructivism makes heavy use of a triangular representation of tool use, with a subject, object, and tool making up the vertices of the triangle, and the lines joining these vertices representing the interactions between any two elements. When applied to cognition, this perspective places much of the cognitive work in the coordination of the two paths to the object (the one mediated by the tool, and the direct path), often by offsetting the two paths at the final vertex, calling specific attention to the coordination that must occur (see Figure 35).

Contextual constructivism treats context by dividing it into a number of layers, representable as concentric circles. Three such layers refer to the particular task a learner is engaged in, the immediately surrounding situation, and the broader

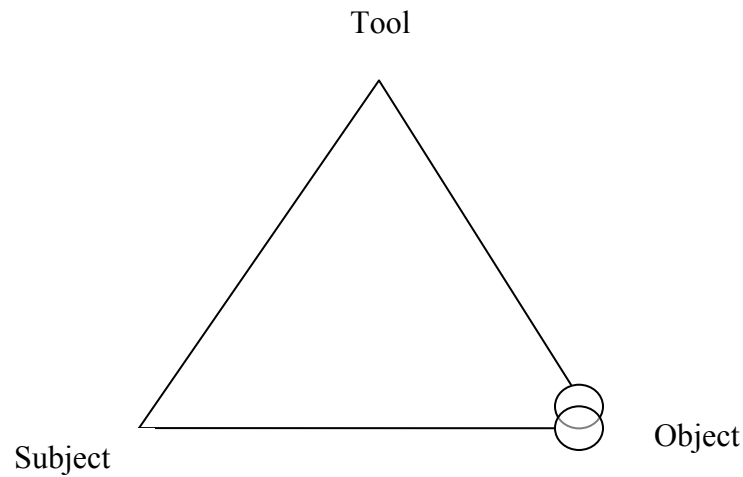


Figure 35. Contextual constructivist triangle. The coordination of the direct and tool-mediated paths to the object is often highlighted in applications of this model to cognition, and is represented here by the two incompletely overlapped circles.

idioculture. The frames of context model also treats interactions between the various layers of context, though these interactions are not modeled as explicitly as the interactions between subject, object, and tool.

In our experiments, we found that student performance when solving a problem depended on many variables, many of which could be concisely represented by an adapted version of the cultural constructivist triangle. In Figure 36, we show this triangle applied to our purposes. The vertices explicitly represent the student, the conceptual knowledge addressed by the problem at hand, and the representation in which the problem is cast. The knowledge and background of the student, the conceptual domain of the problem, and the specific representation of the problem

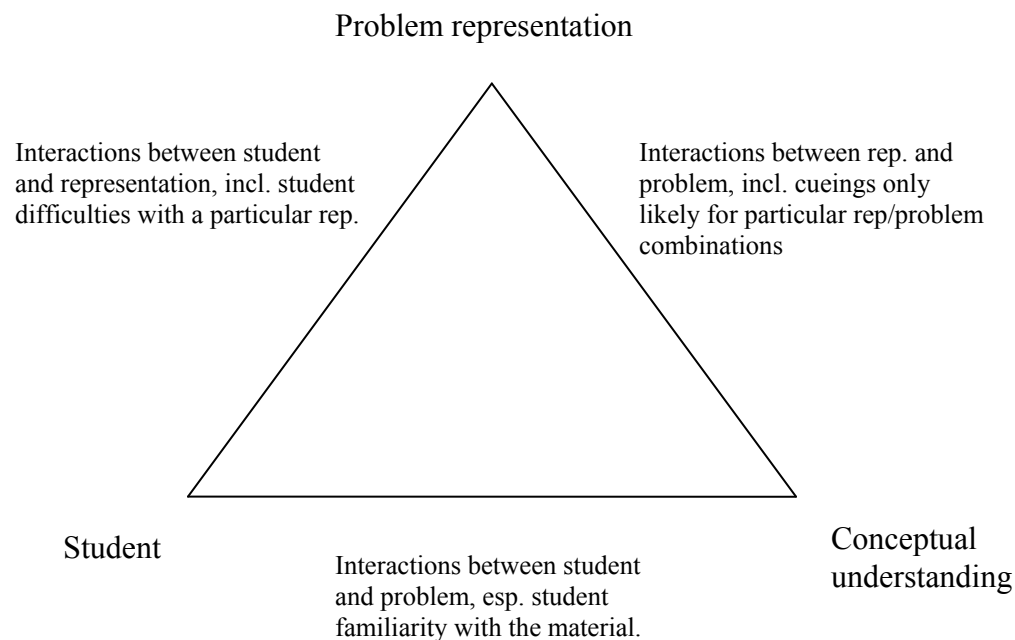


Figure 36. Adaptation of the contextual constructivist subject/object/tool triangle to student use of representations during physics problem solving. Each of the vertices and the interactions between the vertices represent a major factor in student performance.

have all played a primary role in our problem analyses so far. The lines connecting the vertices represent the interactions between the student and conceptual understanding, the student and representation, and the conceptual understanding and representation. These, too, have been major features of our analyses. In making sense of student performance, we have seen the significance of student understanding or misunderstanding of a concept or of a representation. We have also seen that a representation can be particularly well- or ill-suited to a particular task, or that it can have certain compelling features that depend in the combination of representation and concept. All of these factors are located in the triangle seen in Figure 36.

We have not, in our analyses, made substantial use of frames of context. We have, rather, focused on the levels of task and environment without making any further subdivisions. This coarse delineation has been adequate for our purposes, since we were usually either discussing the specifics of a task or the effects of the broader environment taken as a whole.

### *Synthesis*

In this subsection we describe the set of theory elements that we believe are powerful enough to allow for description and analysis of the kinds of results produced by our studies. This is not intended to be a comprehensive theory of representation use. Rather, we are aiming for a model that is as simple as possible while still retaining enough power to make sense of student representation use. We organize this model as four major elements. Three of these are guiding principles, and the fourth is the adapted triangle, which we use as an organizational framework. The principles are as follows:

- The strong sensitivity of novice performance to context. That is, we should be prepared for the possibility that very small changes in representation, problem framing, or student background might result in very large changes in performance, and that these changes might be quite difficult to predict.
- The mechanisms by which representations can drive performance. We have seen that students can cue strongly on certain problem features that may only

be particularly compelling in one of many possible representations. This cueing can affect either answer selection or strategy selection. Among the most compelling kinds of cues we have observed is the inappropriately literal interpretation of representations, referred to as WYSIWYG cueing.

- The mechanisms by which instructional environment can affect the task level. In our experiments, we have seen that extensive and varied representation use in problem solving, whether promoted explicitly or implicitly, will affect the extent to which students will use representations in their solutions, and will also broaden their representational skill set. This principle, while coarse, appears to be fairly robust.

In addition to these principles, we have the triangle framework described above on which we can organize the specifics of a problem-solving episode for descriptive or analytic purposes. Note that in addition to providing an explicit location for many of the features we have identified as relevant in problem-solving, this framework is fairly easily modifiable. For instance, students can fail to solve a problem for many reasons. They could struggle with the representation itself, or with the concept, or the representation may be ill-suited to the concept at hand. We can differentiate these by dashing the appropriate line of the triangle. We can also represent more complicated scenarios by adding vertices, or tagging the vertices with pertinent additional information (highlighting the presence of a compelling cue, for

instance). This triangle tool is not intended to encompass the three principles just identified; it is simply a useful representation of problem solving episodes.

In Figure 37, we show representations of three archetypal uses of representation during problem solving. In the first, we have a student using a single representation to solve a problem, as in the Bohr model homework sets of chapter 5 that required very little translation. In the second, we see two concentric triangles, where one is dashed to represent the student's weak grasp of that representation and its application. With this, we represent what we refer to as representational scaffolding, where a student takes a representation that they understand to support and/or construct a second representation that they are not as comfortable with. The students that solved the challenge problem in the interviews of chapter 12 usually drew free-body diagrams, and those that were successful (expert and novice alike) often wrote their equations based on their diagram, making point-by-point correspondences between the two representations. In the third depiction, we see an approach wherein students use two representations to solve a problem, but both are comfortable and are turned towards reasonably separate ends, so that the two representations can be considered somewhat independently.<sup>xvi</sup> We see this pattern when a picture is used to organize information that is then plugged into an equation, rather than being used to generate the form of that equation. This was the case for the multiple-representations approach to the diffraction quiz in chapter 5, which we highlight in the next subsection.

---

<sup>xvi</sup> The different representations in a multiple-representations problem are never completely separate and independent, and so we still show a link between them in the third part of Figure 37. The relative independence of the representations is expressed by the separation of the corresponding vertices.

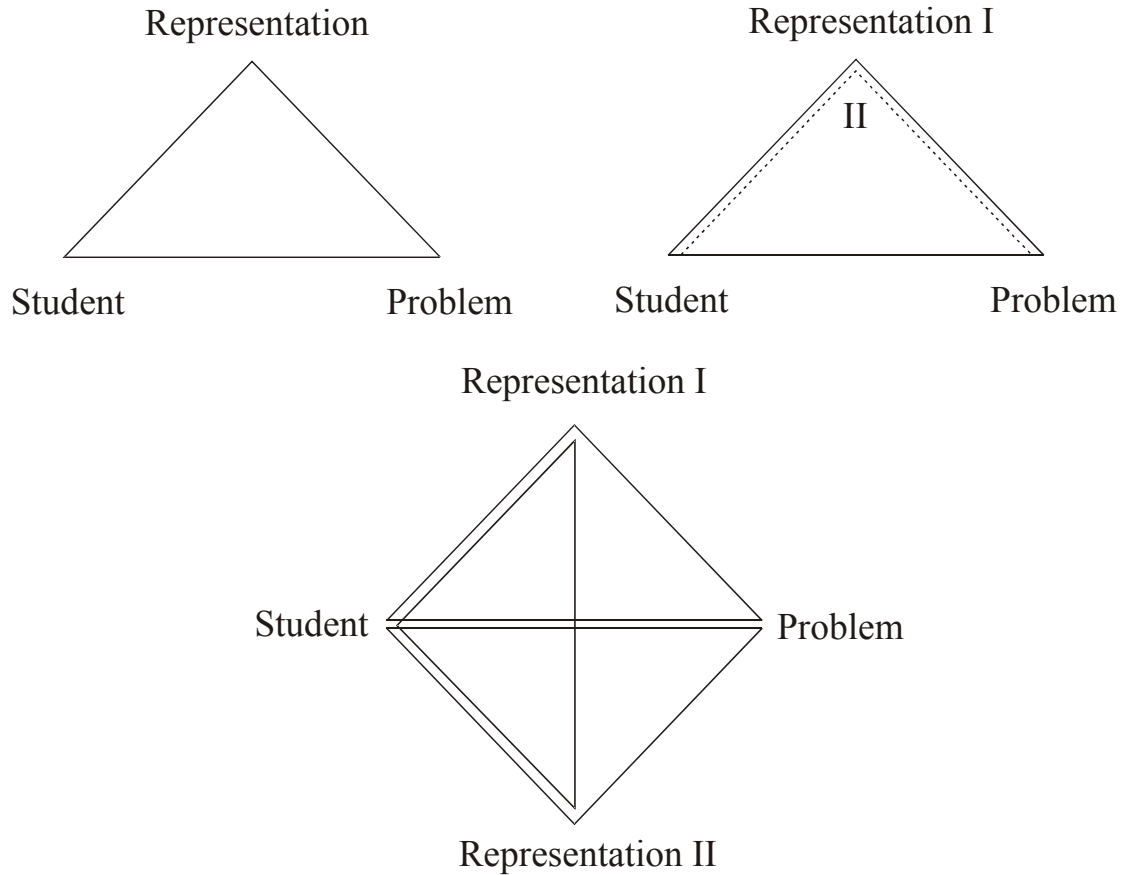


Figure 37. Three example categories of representation use while solving problems. The second displays representational scaffolding, in which one familiar representation is used to construct or make sense of a second, unfamiliar representation on a point-by-point basis. The third displays use of multiple representations that are relatively independent.

*Examples: Analyses of observed representation use*

In this subsection, we show applications of our framework towards making sense of example observations from our data. First, we look at the multiple-representations approach from chapter 4's diffraction quiz. Note that in this case our data were counterintuitive, in that students that used multiple representations (a

picture of their own construction and equations) were actually less successful than those that did not. In Figure 38 we reproduce the problem statement and one example solution. Recall that students who drew a picture and used that to help decipher the language of the problem were more likely to misinterpret it and ultimately use the wrong version of the  $(d,D)\sin(\theta)=n\lambda$  diffraction equation.

In Figure 39 we show an analysis of the student approach that leads to the performance difference between the single and multiple representation groups, which is the approach shown in Figure 38. On the left, we represent the initial student-problem interaction, mediated by the relevant equations and the problem statement (the verbal representation). We have omitted the abstract student-problem interactions here for clarity.

In this analysis, we assume that the student knows the form of the diffraction equation, can execute a calculation, and can read the problem statement. Thus, we have solid lines connecting the student to each representation and the student to the problem via the equation.(i) However, the student cannot correctly interpret the language of the problem statement, and does not know how to map the given information onto the known equation. We represent these breakdowns with dotted lines.(ii) The student then draws a diagram, presumably to help with these difficulties. On the right we show the three-representation interaction that results. The student has



**Diffraction Problem -- Mathematical Format**

We have a double-slit experiment with incident light of  $\lambda = 633 \text{ nm}$ . On a screen  $3.0 \text{ m}$  from the slit, we see an intensity pattern with small peaks separated by  $0.5 \text{ cm}$ . The first minimum in the overall intensity envelope is at  $2.0 \text{ cm}$  from the center of the pattern. Calculate the separation of the slits.

- d. Circle the appropriate letter.  
 A)  $D = 3.8 \times 10^{-5} \text{ m}$   
 B)  $D = 3.8 \times 10^{-4} \text{ m}$   
 C)  $D = 9.5 \times 10^{-5} \text{ m}$   
 D)  $D = 9.5 \times 10^{-4} \text{ m}$   
 E) None of the above.

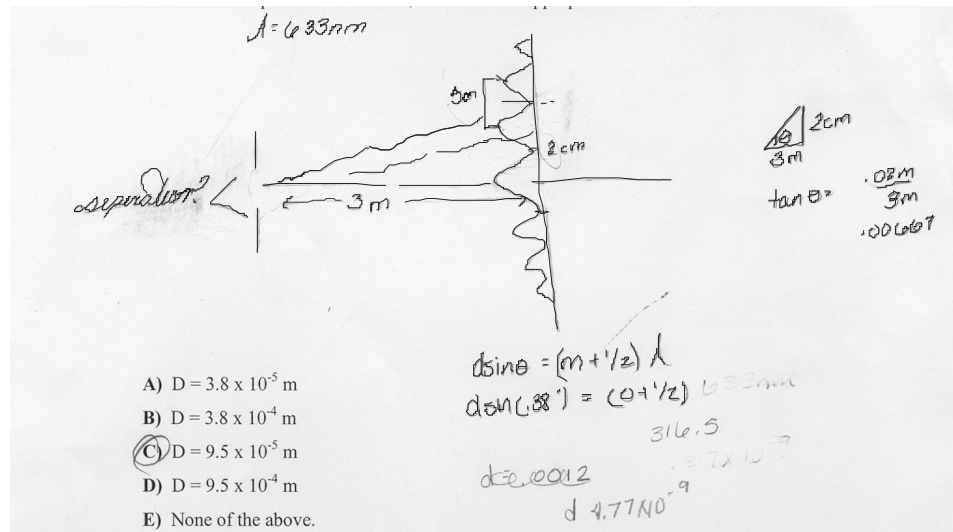


Figure 38. Incorrect student solution to the diffraction quiz of chapter 5, making use of multiple representations.

no trouble drawing a diffraction-related diagram,(iii) but misinterprets the problem statement and draws the wrong diffraction diagram.(iv) At this point, they could still arrive at the correct answer depending on how they map this diagram onto the equations, but they do this incorrectly as well.(v) These failures are again represented by dotted lines.

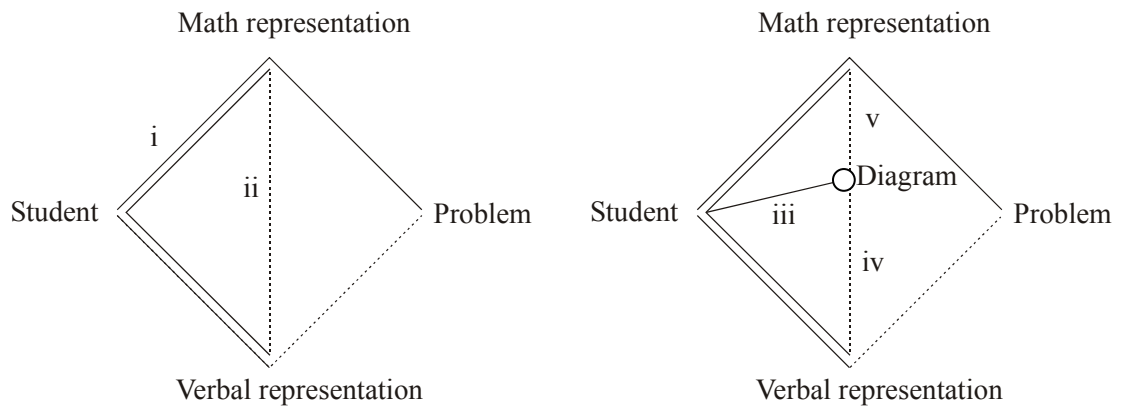


Figure 39. Triangle representation of student incorrect multiple representation solutions to the diffraction problem of chapter 5.

In this example, the triangle formulation provides a compact display of the representations being used and student success and/or failure. The notions of representational scaffolding and sensitivity to context provide directions for analysis that are powerful and relatively straightforward.

In a second example, we consider an expert's solution to the pulley problem of chapter 12. This expert is the same expert seen solving other problems there, who we refer to as Mike. In Figure 40, we show a graph of the representations Mike uses as a function of time. Figure 41 shows a scan of his solution. In Figure 42, we show a triangle representation of Mike's solution to the pulley problem. The solution proceeds in two stages. First, Mike spends a couple minutes constructing the pictures labeled 1 and 2 (labels added during analysis). He builds these pictures piecewise, stepping through the problem statement one step at a time. Thus, the statement

scaffolds the pictures very closely, though we note that Mike has no significant difficulty with either representation, and so we do not use dotted lines in this stage. Second, Mike begins a fairly complex coordination of the written problem statement and his picture in order to construct and constantly check a set of equations. He continues writing equations in this fashion (labeled M1-M4) until he arrives at the answer.

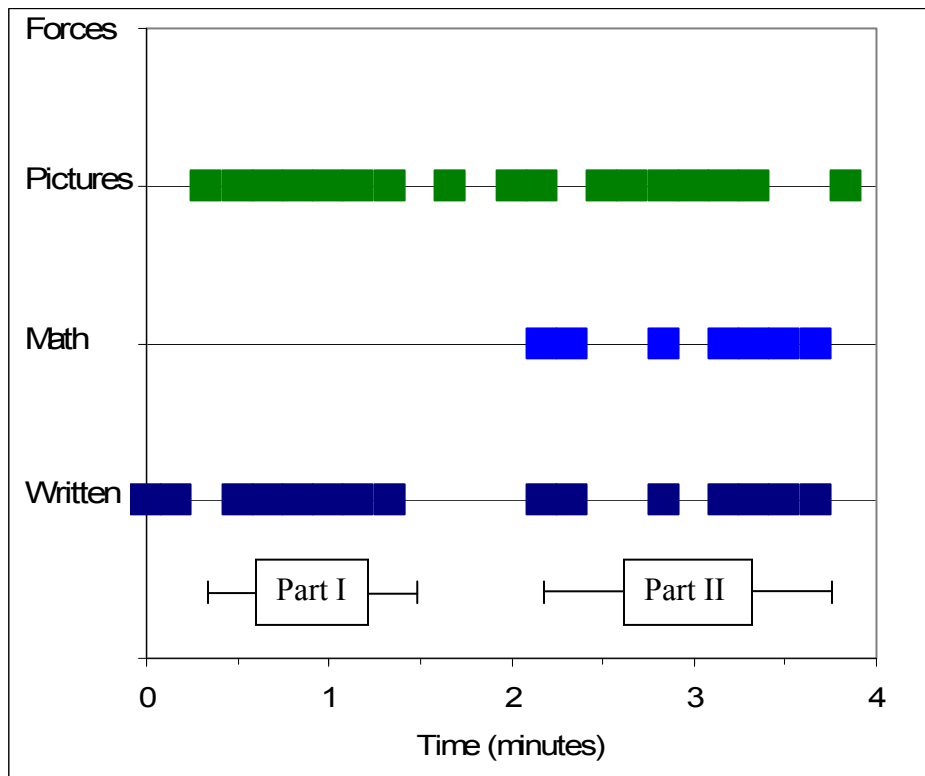


Figure 40. Expert Mike's use of representations during his successful solution of the pulley problem from Chapter 12. Note the periods of extremely close coordination of different representations.

We have three pulleys, two weights, and some ropes, arranged as follows:

- 1) The first weight ( $W_1$ ) is suspended from the left end of a rope over Pulley A. The right end of this rope is attached to, and partially supports, the second weight.
- 2) Pulley A is suspended from the left end of a rope that runs over Pulley B, and under Pulley C. Pulley B is suspended from the ceiling. The right end of the rope that runs under Pulley C is attached to the ceiling.
- 3) Pulley C is attached to the second weight ( $W_2$ ), supporting it jointly with the right end of the first rope.

Find the ratio of  $W_1$  to  $W_2$

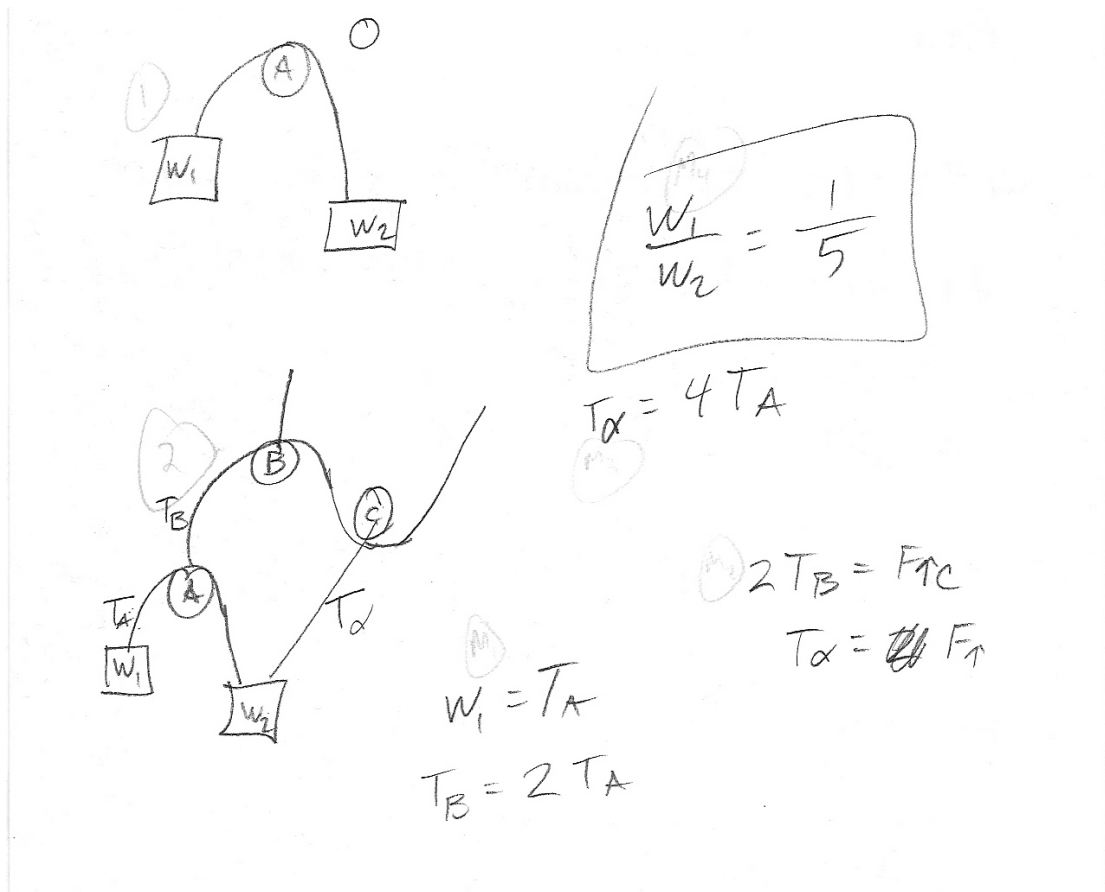


Figure 41. Expert Mike's solution of the pulley problem.

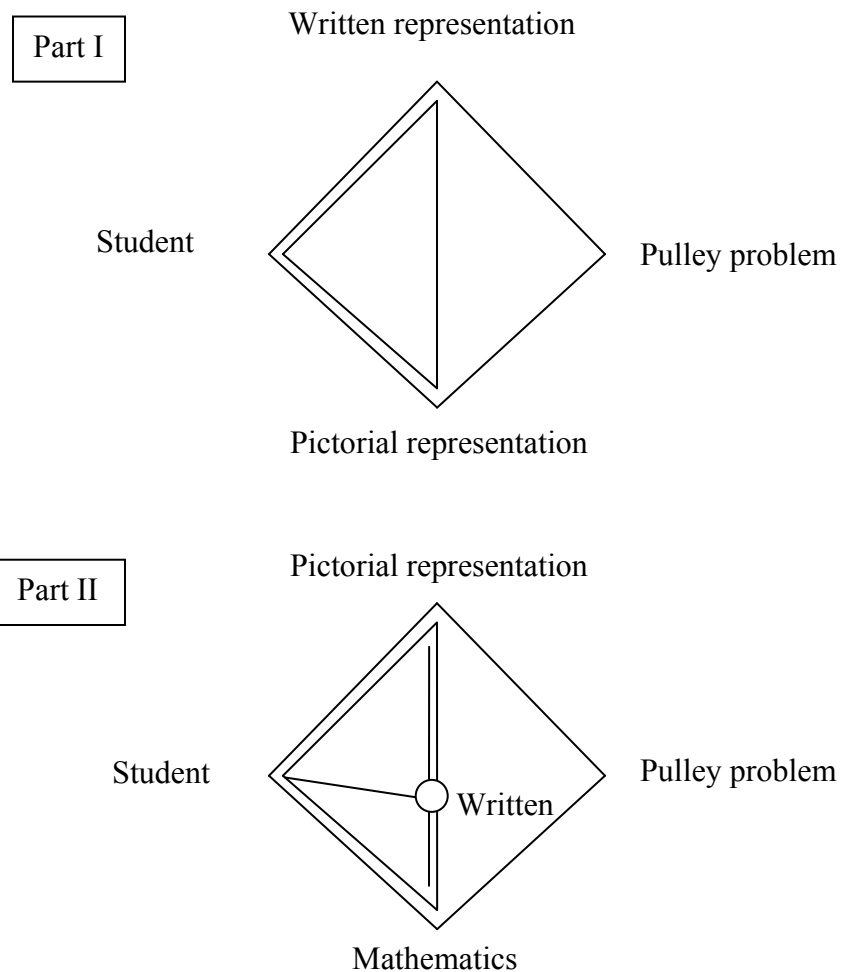


Figure 42. Representation of the two parts of Mike's solution to the pulley problem, focusing on the very close coordination of the different representations he uses.

The representation of Figure 42 emphasizes the most notable feature of Mike's solution: the exceptionally close coordination of the different representations involved. We do not see this level of coordination from any novices, or from experts on the simpler problems. We do not claim that this kind of representational

coordination will be present with any expert solution of a difficult problem, as this problem had an especially complex physical situation to sort out. However, we do believe based on our observations in this thesis that only an expert physics problem solver will be able to generate such a solution, providing a target for instruction.

### *Discussion and Conclusions*

The triangle representation of student problem-solving episodes combines strengths of some of the tools developed in chapter 12. Much like the Schoenfeld plots, this representation calls attention not just to which representation is being used, but also shows the purposes those representations are being served. Unlike the Schoenfeld diagrams that were mostly about representation nonspecific categories, this analysis tool focus more on which representations are being used and their uses with respect to one another. This triangle representation also resembles the sequence diagrams of chapter 12 in that it allows us to show the sequence of representations being used, though it allows considerably more room for detail.

The triangle analysis tool, coupled with the three principles described previously (sensitivity to context, the role of cueing, and the effects of instructional environment) provide a practical and concise package for making sense of student use of representations during physics problem solving. This package, while useful in practice, does have limitations from a theoretical standpoint. We are not sure that we have fully described and incorporated the effects of context (a notoriously slippery target). We also have not incorporated our results regarding meta-representational competence.

Overall, the principles and tools described in this chapter represent a useful contribution towards our understanding of physics problem solving. We believe it also provides a base for a more complete theoretical description of representation use, as well as a step towards bridging the gap between the resources model[57, 70] and the contextual constructivist model,[74, 75] two major theoretical frames of the day.

## Chapter 14: Review and conclusions

In this thesis, we had four major research goals, designed to lay a foundation for the development of a model of introductory physics student use of representations during problem solving. We also hoped to be able to turn our major results into implications for instruction. In this chapter, we will review our four research goals, our model and heuristics, and the implications our results have for teaching physics.

### *Major goals*

Our thesis laid out four primary goals, which we review here.

- We wished to clarify the extent to which changes in representation can lead to changes in performance, and to identify mechanisms that drive this effect.
- We needed to determine the role of instructional environment in shaping student representational competence. Specifically, will a course rich in representations and multiple representations lead to stronger and/or broader student skills?
- We wanted to examine the role of meta-representational skill at the university level. Do students know enough about their own knowledge of representations for us to tap that knowledge productively in the classroom?



- In order to better bridge the gap between novice and expert physicists, we need to characterize the differences in expert and novice representation use during physics problem solving.

In Table XXIII, we list these four goals and the chapters that most strongly address each goal. We then review each of these goals and our results in detail.

Research goal	Chapters providing major coverage
Effect of representation on performance	5, 6, 7, 9, 10
Dependence of rep. skills on environment	6, 11
Meta-representational competence	5, 6, 7, 11
Expert vs. novice behavior	7, 11, 12

Table XXIII. Stated thesis goals, and the primary chapters addressing each. Minor support of each of the four goals may be found in other chapters.

*Goal 1: Effect of representation on performance*

We found repeated evidence of the major effect of representation on student performance. In chapters 5, 6, and 10, we split large-lecture introductory algebra-based physics courses into multiple groups, giving each group quiz problems posed in different representations. We made the problems as similar as possible, to the extent that the problems would be perceived as isomorphic by many physics experts. In a great many of these cases, there were very significant performance differences between the different problem representations, often on the order of 50% or more.

With these representational effects established, it was natural to ask how, exactly, representation drove performance. While we do not doubt that there are

quite a few specific ways that problem representation can affect success, we observed two primary mechanisms. The first of these was representation-dependent cueing. By this, we refer to the fact that some representations of a problem will contain features that are particularly compelling, either linguistically, visually, or otherwise. A number of examples can be found in our data, but one category of cueing was most common. In chapters 5 (the Bohr model homework problem), 7 (interviews involving the graphical format of the spectroscopy quiz), and 10 (the ball on a hill/valley problem), the results were consistent with what-you-see-is-what-you-get (WYSIWYG) cueing. There, students interpret visual representations of the problem in a very literal way (lower on the page indicates lower in energy, for instance). This is often productive, and in such cases we do not even notice it. However, in physics we use a number of very abstract representations, where a WYSIWYG interpretation can fail, and this is what we see in these chapters. Similarly, Podolefsky finds that representation-dependent WYSIWYG cueing is a major factor in his studies of reasoning through analogy in physics.[115-117] He also finds that appropriate representation choice during instruction can influence whether WYSIWYG is applied by the students productively or unproductively.

A second mechanism was the change in problem-solving strategy provoked by the different representations. In chapter 7 we saw this especially clearly: Of the students interviewed, some were quite consistent in their approach to problems regardless of the representation used. Others took significantly different approaches depending on the representation of the problem, including using force vs. energy

pictures, or using quantitative vs. qualitative approaches. The consistent students appeared to outperform those whose strategies varied with representation.

In addition to these mechanisms, we should note another general principle that is relevant the effects of representation of performance: the major role of context. By context we refer to all the elements of a problem and its setting and, perhaps most importantly, the relations among those elements. In many cases, including but not limited to the multiple representations diffraction question from chapter 5 and Jim's interview in chapter 7, we found that minor and easy to overlook (from our perspective) interactions between concept, representation, setting, and student knowledge resulted in significant performance differences. This result must be taken into account in any attempt to make sense of student representation use in general.

When we began this thesis, there existed very little in the way of direct comparisons across representation, with fairly tentative results.[9] We can now say with confidence that problem representation can frequently have a profound impact on student performance, and that we have some sense of *how* representation matters, paving the way for the rest of our work.

*Goal 2: Effect of instructional environment on representational competence*

The environment associated with a course is enormously complex, including all the structural features (material covered, types and frequencies of assessments, size) and all the cultural features (which we mean to include attitudes, norms, expectations, power relationships, and so forth). In this thesis, we have mostly

focused on the representational components of the instructional environment and the means by which that representational content was communicated to students.

In chapters 6 and 11 we focused explicitly on the representational context of five different courses from four different professors. One of these courses (Physics 202 from chapter 5) was traditionally taught, and in terms of lectures and exams was fairly sparse in representations. Two of the courses (Physics 201 and 202 from chapters 5 and 6) were taught by a PER-influenced professor whose course was much richer in representations and multiple representations. The last two courses (the 202-equivalent courses at Rutgers and CU from chapter 10) were taught by PER faculty, and were again very rich in representations. The latter courses differed in a significant way: One course taught step-by-step procedures for handling multiple representations during problem solving very explicitly, while the other taught procedures implicitly, by example.

We saw these differences in environment manifest themselves in two ways. First, in chapters 5 and 6, students in the more representationally-rich courses showed much less performance variation with respect to representation than those in the representationally-sparse course. We interpret this plus the reduced effect of representation choice as evidence that the students are developing competency with a broader set of representations in the richer course.

Second, we found in chapter 11 that two fundamentally different approaches to teaching multiple representation use resulted in surprisingly similar success rates and frequencies of multiple representation use. We did observe minor differences in the specifics of multiple representation use, with the strongly-directed students using

free-body diagrams more often on high-stakes, high difficulty problems, and the weakly-directed students using pictures and free-body diagrams more often on the easier problems.

These results serve two purposes for us. Not only do they provide useful guidelines for instruction, but they also establish means by which the features of the course at the level of environment can affect student performances at the level of task.

### *Goal 3: Relevance of meta-representational competence*

Meta-representational competence can include many things, including the ability to generate or critique completely new kinds of representations.[19] We have focused our attention on meta-representational skills that are more practical for the university classroom, such as the ability to productively choose between standard physics representations, or the ability to accurately assess one's own skill with different representations.

In chapters 5-7, we found evidence that students' meta-representational skills were not well-developed, which is not surprising since such skills are not encouraged or tested in a typical physics course. Giving students a choice between quiz representations resulted in a statistically significant performance decrease as compared to a control group nearly as often as it resulted in a significant performance increase. This was not a result of students having no opinions on the matter: Student beliefs regarding their facility with different representations appeared to be well-formed and consistent over the semester. These beliefs had little correlation with their actual performance. This was confirmed in a series of problem-solving

interviews involving individual students. These students gave very similar self-assessments to those given earlier in the semester, and those assessments were almost completely uncorrelated with their performance on the study tasks.

In chapter 11, we gave a survey on representations and representation use to students in two large-lecture introductory courses. We checked the correlation between student responses on this survey with their performance on five multiple-representations problems, with generally null results. This again suggests that introductory physics students are not good at accurately assessing or describing how they use representations in problem solving.

For the sake of model development, this limited meta-representational skill, coupled with the fact that meta-representational tasks are not part of a standard curriculum, suggests to us that the abilities to assess one's own representational competence and/or choose between representations might not be among the most important points when analyzing novice behavior. However, this does not mean that meta-representational skill is not important to problem-solving. In chapter 12, we did not examine these meta-representational skills, but we did study other aspects of meta-representational competence. In particular, we looked at the sorts of uses towards which novices and experts applied the representations they were using. It was here that we saw the sharpest differences between experts and novices. Novices tended to use physics representations mechanically or algorithmically, while experts turned these same representations towards sense-making and analysis. If experts and novices show such significant differences in other meta-representational skills as well

(and we believe that they would), then we are left with the question of how best to bridge that particular gap between expert and novice.

We should also acknowledge that while students do not appear to be bringing with them the kinds of meta-representational skills that we had hoped for, that does not mean that they possess no such skills. Since students are rarely asked to engage in explicit meta-representational tasks, we could consider it a productive decision on their part to neglect such skills, at least in the short-term.

*Goal 4: Expert vs. novice problem-solving behavior*

In many of our thesis studies, we implicitly characterized novice problem solving behavior, involving both single and multiple representations problems. Students were often cued into different answers or strategies depending on the problem representation. On the multiple representations problems of chapter 11, use of supplementary representations like pictures and free-body diagrams was associated with improved performance, but only when those representations were used correctly. We observed the best performance when all the representations present were correct and properly coordinated.

In chapter 12, we focused explicitly on comparing the problem-solving procedures of novices and experts on multiple representations problems, with those problems either providing multiple representations or a practical requirement to produce additional representations. Experts and novices were both quite willing to create pictures and free-body diagrams to accompany their solutions. The major differences were in their patterns of use. Experts were faster and more flexible in

terms of which representations they would use and the order in which they would use them. They were also capable of translating between the different representations more quickly, even when very difficult problems were involved. Most significant was the novice tendency to use representations with little direction or plan, in an algorithmic way. Experts used multiple representations for problem analysis.

### *Model of representation use*

In chapter 13 we presented a concise and powerful set of principles and a tool for analysis and description of representation use during problem solving. This package is straightforward enough for non-experts to use in practice, while still being capable of making sense of a wide variety of problem solving episodes.

We also believe that we have succeeded in providing a proper foundation for future theoretical development. The major results of this thesis clarify which aspects of problem-solving episodes are most pertinent, and provide a base of observational data. In addition, we have identified major differences between expert and novice use of representations during problem solving. While some aspects of physics representation use during physics problem solving remain to be fully characterized (the role of context and of meta-representational skill), our understanding is now much deeper than when we began.

### *Instructional implications*

At least one major implication for instruction presented itself very early on in this thesis (chapters 5 and 6). Instructors can foster broader representational skills by



infusing all aspects of their course (exams, lectures, recitations) with a variety of representations. This approach can succeed even without an explicit plan for teaching representational facility. In chapter 11, we confirmed this result, and generalized it to include the use of multiple representations together, in this case pictures, free-body diagrams, and mathematics. In that study, only one of the courses examined made a significant attempt to explicitly teach skill with multiple representations. Effective though that may be, it appears that simply making a variety of representations present in the course is sufficient to provide a significant positive benefit.

These early chapters, along with chapter 12, introduced another possibility for instruction. We observed that introductory physics students had little in the way of knowledge about their own skills with representations, and what representations are useful for. Experts, on the other hand, were very focused and productive in their application of representations to problems both simple and difficult. This suggests that experts have some knowledge, explicit or implicit, of what representations are for, and of their own capabilities. Introductory physics courses do not routinely contain any activities that might foster this kind of meta-representational skill. It may be that such instruction would help close the gap between expert and novice abilities to solve physics problems with multiple representations.

### *Conclusion*

In this thesis we have significantly expanded our understanding of how novice physicists use representations and multiple representations in solving problems. Our

efforts have been focused towards four major points: the magnitude of the effect of representation shift and the mechanisms by which representations drive performance, the role of instructional environment in determining student representational skills, the role of meta-representational competence, and the main differences in representation use between expert and novice physicists. Each of these points complements the representational research that already exists in PER, resulting in a much more complete picture.

Our work is significant for instruction both in terms of immediate impact, and in terms of direction for future work. We have seen that physics courses can foster skill with a variety of representations and with multiple representations, along with a willingness to use those representations. What is less clear is whether introductory courses are teaching students why they are using these representations. Since our results indicate that using representations productively is a major factor separating experts and novices during problem solving, this suggests to us that future work should determine the feasibility of teaching this kind of meta-representational skill to introductory physics students. This is potentially a tall order, as metaskills may take a very long time to develop. Nevertheless, we believe that our students will become much more expert-like in their multiple-representations problem solving if they can learn at least a basic set of these skills.

## Bibliography

1. McDermott, L., *A view from physics*, in *Toward a scientific practice of science education*, M. Gardner, et al., Editors. 2004, IOS: Amsterdam.
2. Kozma, R.B., *The material features of multiple representations and their cognitive and social affordances for science understanding*. Learning and Instruction, 2003. **13**: p. 205.
3. Larkin, J.H., *The role of problem representation in physics*, in *Mental Models*, D. Gentner and A. Stevens, Editors. 1983, Lawrence Erlbaum: Mahwah, New Jersey.
4. Reif, F., *Understanding and teaching important scientific thought processes*. Am. J. Phys., 1995. **63**: p. 17.
5. Heuvelen, A.V., *Learning to think like a physicist: A review of research-based instructional strategies*. Am. J. Phys., 1991. **59**: p. 981.
6. Dufresne, R.J., R.J. Gerace, and W.J. Leonard, *Solving physics problems with multiple representations*. The Physics Teacher, 1997. **35**: p. 270.
7. Heller, P., R. Keith, and S. Anderson, *Teaching problem solving through cooperative grouping. Part I and II*. Am. J. Phys., 1992. **60**: p. 627.
8. Heuvelen, A.V. and X. Zou, *Multiple representations of work-energy processes*. Am. J. Phys., 2001. **69**: p. 184.
9. Meltzer, D.E., *Relation between students' problem-solving performance and representational mode*. Am. J. Phys., 2005. **73**: p. 463.
10. Chi, M., P. Feltovich, and R. Glaser, *Categorization and representation of physics problems by experts and novices*. Cognitive Science, 1980. **5**: p. 121.
11. Dancy, M. and R. Beichner, *Impact of animation on assessment of conceptual understanding in physics*. Phys. Rev. ST Phys. Educ. Res., 2006. **2**(010104).
12. Beichner, R.J., *Testing student interpretation of kinematics graphs*. Am. J. Phys., 1994. **62**: p. 750.
13. Goldberg, F.M. and J.H. Anderson, *Student difficulties with graphical representations of negative values of velocity*. The Physics Teacher, 1989. **27**: p. 254.
14. McDermott, L., M.L. Rosenquist, and E.H.V. Zee, *Student difficulties in connecting graphs and physics: Examples from kinematics*. Am. J. Phys., 1987. **55**: p. 503.
15. Ainsworth, S.E., P.A. Bibby, and D.J. Wood, *Analyzing the costs and benefits of multi-representational learning environments*, in *Learning with multiple representations*, M.W.v. Someren, et al., Editors. 1998, Pergamon.
16. Mayer, R.E., *The promise of multimedia learning: using the same instructional design methods across different media*. Learning and Instruction, 2003. **13**: p. 125.
17. Schnotz, W. and M. Bannert, *Construction and interference in learning from multiple representation*. Learning and Instruction, 2003. **13**: p. 141.
18. Azevedo, F.S., *Designing representations of terrain: A study in meta-representational competence*. Journal of Mathematical Behavior, 2000. **19**: p. 443.

19. diSessa, A.A., et al., *Inventing graphing: Meta-representational expertise in children*. Journal of Mathematical Behavior, 1991. **10**: p. 117.
20. Chi, M., R. Glaser, and E. Rees, *Expertise in problem solving*, in *Advances in the psychology of human intelligence Vol. 1*, R.J. Sternberg, Editor. 1992, Lawrence Erlbaum: Hillsdale, NJ.
21. Maloney, D.P., *Research on problem solving: Physics*, in *Handbook of Research on Science Teaching and Learning*, D. Gabel, Editor. 1994, MacMillan Publishing Co: New York, NY.
22. Hsu, L., et al., *Resource Letter RPS-1: Research in problem solving*. Am. J. Phys., 2004. **72**: p. 1147.
23. Redish, E.F., *Teaching physics with The Physics Suite*. 2003, New York, NY: Wiley.
24. Kohl, P.B. and N.D. Finkelstein, *Student representational competence and self-assessment when solving physics problems*. Phys. Rev. ST Phys. Educ. Res., 2005. **1**(010104).
25. Kohl, P.B. and N.D. Finkelstein, *The effect of instructional environment on physics students representational skills*. Phys. Rev. ST Phys. Educ. Res., 2006. **2**(010102).
26. Kohl, P.B. and N.D. Finkelstein, *The effects of representation on students solving physics problems: a fine-grained characterization*. Phys. Rev. ST Phys. Educ. Res., 2006. **2**(010106).
27. Kohl, P.B., D. Rosengrant, and N.D. Finkelstein, *Strongly and weakly directed approaches to teaching multiple representation use in physics*. Phys. Rev. ST Phys. Educ. Res., accepted for publication.
28. Meltzer, D.E., *The relationship between mathematics preparation and conceptual learning gains in physics: A possible "hidden variable" in diagnostic pretest scores*. Am. J. Phys., 2002. **70**(12): p. 1259-1268.
29. Cohen, E. and S.E. Kanim, *Factors influencing the algebra "reversal error."* Am. J. Phys., 2005. **73**: p. 1072-1078.
30. Lobato, J. and E. Thanheiser, *Developing understanding of ratio and measure as a foundation for slope.*, in *Making sense of fractions, ratios, and proportions: 2002 yearbook*, B. Litwiller and G. Bright, Editors. 2002, National Council of Teachers of Mathematics: Reston, VA. p. 162-175.
31. Kaput, J.J. and M.M. West, *Missing-value proportional reasoning problems: Factors affecting informal reasoning patterns.*, in *The development of multiplicative reasoning in the learning of mathematics*, G. Harel and J. Confrey, Editors. 1994, State University of New York press. p. 237-287.
32. Tourniaire, F. and S. Pulos, *Proportional reasoning: A review of the literature*. Educational studies in mathematics, 1985. **16**: p. 181-204.
33. Torigoe, E.T. and G.E. Gladding, *Mathematical communication in physics*. p. (unpublished).
34. Knight, R.D., *The vector knowledge of beginning physics students*. The Physics Teacher, 1995. **33**: p. 74-77.
35. Nguyen, N. and D.E. Meltzer, *Initial understanding of vector concepts among students in introductory physics courses*. Am. J. Phys., 2003. **71**(6): p. 630-638.

36. Flores, S., S.E. Kanim, and C.H. Kautz, *Student use of vectors in introductory mechanics*. Am. J. Phys., 2004. **72**(4): p. 460-468.
37. Trowbridge, D.E. and L.C. McDermott, *Investigation of student understanding of the concept of velocity in one dimension*. Am. J. Phys., 1980. **48**: p. 1020.
38. Trowbridge, D.E. and L.C. McDermott, *Investigation of student understanding of the concept of acceleration in one dimension*. Am. J. Phys., 1981. **49**: p. 242.
39. Beichner, R.J., *The impact of video motion analysis on kinematics graph interpretation skills*. Am. J. Phys., 1996. **64**: p. 1272-1277.
40. Monk, S., *Students' understanding of a function given by a physical model. The concept of function: Aspects of epistemology and pedagogy*. MAA Notes, 1992. **25**: p. 175-194.
41. Lemke, J.L., *Talking physics*. Phys. Educ., 1982. **17**: p. 263-267.
42. Lemke, J.L., *Multiplying meaning: Visual and verbal semiotics in scientific text*, in *Reading science*, J.R. Martin and R. Veel, Editors. 1998, Routledge: London. p. 87-113.
43. Williams, T.H., *Semantics in teaching introductory physics*. Am. J. Phys., 1999. **67**(8): p. 670-680.
44. Brookes, D.T. and E. Etkina, *Do our words really matter? Case studies from quantum mechanics*. PERC Proceedings 2005, 2005: p. 57.
45. Brookes, D.T., *The role of language in learning physics*. Ph.D. Thesis, Rutgers, the State University of New Jersey, 2006.
46. Brookes, D.T. and E. Etkina, *Using conceptual metaphor and functional grammar to explore how language used in physics affects student learning*. Phys. Rev. ST Phys. Educ. Res., 2007. **3**(010105).
47. Koedinger, K.R. and M.J. Nathan, *The real story behind story problems: Effects of representations on quantitative reasoning*. J. of the Learning Sciences, 2004. **13**: p. 129.
48. Hestenes, D., M. Wells, and G. Swackhamer, *Force Concept Inventory*. The Physics Teacher, 1992. **30**: p. 141.
49. Thornton, R.K. and D.R. Sokoloff, *Assessing student learning of Newton's laws: The force and motion conceptual evaluation*. Am. J. Phys., 1998. **66**: p. 228.
50. Ding, L., et al., *Evaluating an electricity and magnetism assessment tool: Brief Electricity and Magnetism Assessment*. Phys. Rev. ST Phys. Educ. Res., 2006. **1**(010105).
51. Kozma, R.B. and J. Russell, *Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena*. J. of Research in Science Teaching, 1997. **34**: p. 949.
52. *Physics Education Technology Project*. <http://phet.colorado.edu> (retrieved August 29, 2006).
53. Rosengrant, D., E. Etkina, and A.V. Heuvelen, *Multiple representations in problem solving: Does a free-body diagram help?* J. of Research in Science Teaching, submitted.

54. Schoenfeld, A.H., *What's all the fuss about metacognition?*, in *Cognitive science and mathematics education*, A.H. Schoenfeld, Editor. 1987, Lawrence Erlbaum Associates: Hillsdale, NJ. p. 189.
55. diSessa, A.A. and B.L. Sherin, *Meta-representation: an introduction*. *Journal of Mathematical Behavior*, 2000. **19**: p. 385.
56. diSessa, A.A., *Students' criteria for representational adequacy*, in *Symbolizing, modeling and tool use in mathematics education*, K. Gravemeijer, et al., Editors. 2003, Kluwer Academic Publishers: Dordrecht, the Netherlands.
57. diSessa, A.A., *Knowledge in pieces*, in *Constructivism in the Computer Age*, G. Forman and P.B. Pufall, Editors. 1988, Lawrence Erlbaum: Hillsdale, NJ.
58. diSessa, A.A., *What coordination has to say about transfer*, in *Transfer of Learning: Research and Perspectives*, J. Mestre, Editor. forthcoming, Information Age Publishing: Greenwich, CT.
59. McCloskey, M., *Naive theories of motion.*, in *Mental Models*, D. Gentner and A. Stevens, Editors. 1983, Erlbaum: Hillsdale.
60. McDermott, L.C., *How we teach and how students learn - A mismatch?* *Am. J. Phys.*, 1993. **61**: p. 295.
61. Mestre, J.P., et al., *The dependence of knowledge deployment on context among physics novices*, in *Proceedings of the International School of Physics Enrico Fermi Course CLVI: Research on Physics Education*, E.F. Redish and M. Vicentini, Editors. 2004, IOS: Amsterdam.
62. Clark, J.M. and A. Paivio, *Dual coding theory and education*. *Educ. Psych. Review*, 1991. **3**: p. 149.
63. Kirby, J.R., *Collaborative and competitive effects of verbal and spatial processes*. *Learning and Instruction*, 1993. **3**: p. 201.
64. Seufert, T., *Supporting coherence formation in learning from multiple representations*. *Learning and Instruction*, 2003. **13**: p. 227.
65. Mazur, E., *Peer Instruction: A user's manual*. 1996, Upper Saddle River, NJ: Prentice Hall.
66. Chandler, P. and J. Sweller, *Cognitive load theory and the format of instruction*. *Cognition and Instruction*, 1991. **8**: p. 293.
67. Ward, M. and J. Sweller, *Structuring effective worked examples*. *Cognition and Instruction*, 1990. **7**: p. 1.
68. Kalyuga, S., P. Chandler, and J. Sweller, *Incorporating learner experience into the design of multimedia instruction*. *J. of Educational Psych.*, 2000. **92**: p. 126.
69. Kalyuga, S., et al., *The expertise reversal effect*. *Educational Psychologist*, 2003. **38**: p. 23.
70. diSessa, A.A., *Coordination and contextuality in conceptual change*, in *Proceedings of the International School of Physics Enrico Fermi Course CLVI: Research on Physics Education*, E.F. Redish and M. Vicentini, Editors. 2003, Italian Physical Society: Bologna, Italy.
71. diSessa, A.A., *Changing Minds*. 2000, Cambridge, MA: MIT Press.

72. Smith, J.P., A.A. diSessa, and J. Roschelle, *Misconceptions reconceived: A constructivist analysis of knowledge in transition*. J. of the Learning Sciences, 1993. **3**: p. 115.
73. Wittmann, M.C., *The object coordination class applied to wavepulses: Analyzing student reasoning in wave physics*. Int. J. of Science Education, 2002. **24**(1): p. 97.
74. Cole, M., *Cultural psychology: A once and future discipline*. 2001, Cambridge, MA: Harvard University Press.
75. Finkelstein, N.D., *Learning physics in context: A study of student learning about electricity and magnetism*. Int. J. of Science Education, 2005. **27**: p. 1187.
76. Perkins, K.K., et al., *PhET: Interactive simulations for teaching and learning physics*. The Physics Teacher, 2006. **44**: p. 18.
77. [www.h-itt.com](http://www.h-itt.com).
78. Finkelstein, N.D. and S.J. Pollock, *Replicating and understanding successful innovations: Implementing tutorials in introductory physics*. Phys. Rev. ST Phys. Educ. Res., 2005. **1**(010101).
79. Hake, R.R., *Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses*. Am. J. Phys., 1998. **66**: p. 64.
80. Rosengrant, D., A.V. Heuvelen, and E. Etkina. *Two-year study of students' use of free-body diagrams*. in *Proceedings of the NARST 2006 Annual Meeting*. 2006.
81. Lewis, C.H. and J. Rieman, *Task-centered user interface design: A practical guide*, G. Perlman, Editor. 1993, HCI Bibliography.
82. *The LearningOnline Network with CAPA*. <http://www.lon-capa.org> (retrieved July 4, 2005).
83. Likert, R., *A technique for the measurement of attitudes*. Archives of Psychology, 1932. **140**(55).
84. Mayer, R.E., *Multimedia learning: Are we asking the right questions?* Educational Psychologist, 1997. **32**: p. 1.
85. Giancoli, D., *Physics (Fifth Edition)*. 1998, Upper Saddle River, NJ: Prentice Hall.
86. Zhang, J. and D. Norman, *Representations in distributed cognitive tasks*. Cognitive Science, 1994. **18**: p. 87.
87. Mayer, R.E., *Different problem-solving strategies for algebra, word, and equation problems*. J. of Exp. Psych.: Learning, Memory, and Cognition, 1982. **8**: p. 448.
88. diSessa, A.A., A. Elby, and D. Hammer, *J's epistemological stance and strategies*, in *Intentional conceptual change*, G. Sinatra and P. Pintrich, Editors. 2003, Lawrence Erlbaum: Mahwah, NJ.
89. Sabella, M.S. and E.F. Redish, *Knowledge organization and activation in physics problem-solving*. Am. J. Phys. PER Supp., submitted.
90. Elby, A., *What students' learning of representations tells us about constructivism*. Journal of Mathematical Behavior, 2000. **19**: p. 481.

91. Sabella, M.S., *Using the context of physics problem solving to evaluate the coherence of student knowledge*, in *Department of Physics*. 1999, University of Maryland.
92. *University of Washington PER group publications*.  
<http://www.phys.washington.edu/groups/peg/pubs.html>.
93. Elby, A., *Helping physics students learn how to learn*. Am. J. Phys. PER Supp., 2001. **69**: p. 54.
94. Sherwood, B. and R. Chabay, *A Brief E&M Assessment*. unpublished.
95. McDermott, R., ed. *The acquisition of a child by a learning disability*. Understanding practice: Perspectives on activity and context, ed. S. Chaiklin and J. Lave. 1993, Cambridge University Press: New York.
96. Sheshkin, D., *Handbook of parametric and nonparametric statistical procedures, Third edition*. 2004, Boca Raton, FL: CRC Press.
97. Gray, K.E., *The effect of question order on student responses to multiple choice physics questions*, in *Physics*. 2004, Kansas State University.
98. Hammer, D., et al., *Resources, framing, and transfer*, in *Transfer of learning: Research and perspectives*, J. Mestre, Editor. 2005, Information Age Publishing: Greenwich, CT.
99. Adams, W.K., et al., *New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey*. Phys. Rev. ST Phys. Educ. Res., 2006. **2**(1): p. 010101.
100. Redish, E.F., J.M. Saul, and R.N. Steinberg, *Student expectations in introductory physics*. Am. J. Phys., 1998. **66**: p. 212.
101. McDermott, L.C. and E.F. Redish, *Resource letter: PER-1: Physics education research*. Am. J. Phys., 1999. **67**: p. 755.
102. McDermott, L.C. and P.S. Schaffer, *Tutorials in Introductory Physics*. 1998, Upper Saddle River, NJ: Prentice Hall.
103. Wieman, C.E. and K.K. Perkins, *Transforming physics education*. Physics Today, 2005. **58**(11): p. 36-41.
104. DeLeone, C. and E. Gire. *The effect of representation use on student problem solving*. in *Proceedings of the 2005 PERC*. 2005: AIP Conference Proceedings.
105. Rosengrant, D., E. Etkina, and A.V. Heuvelen. *Case study: Students' use of multiple representations in problem solving*. in *Proceedings of the 2005 PERC*. 2005: AIP Conference Proceedings.
106. Rosengrant, D., E. Etkina, and A.V. Heuvelen. *An overview of recent research in multiple representations*. in *Proceedings of the 2006 PERC*. 2006: AIP Conference Proceedings.
107. Etkina, E., et al., *Scientific abilities and their assessment*. Phys. Rev. ST Phys. Educ. Res., 2006. **2**(020103).
108. Etkina, E. and A.V. Heuvelen. *Investigative Science Learning Environment: Using the processes of science and cognitive strategies to learn physics*. in *Proceedings of the 2001 PERC*. 2001: AIP Conference Proceedings.
109. Heuvelen, A.V. and E. Etkina, *Active Learning Guide*. 2006, San Francisco, CA: Addison Wesley.



110. Etkina, E. and S. Murthy. *Design labs: Students' expectations and reality*. in *Proceedings of the 2005 PERC*. 2005: AIP Conference Proceedings.
111. *Active Physics*. <http://www.physics.ohio-state.edu/>.
112. Finkelstein, N.D., et al., *When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment*. Phys. Rev. ST Phys. Educ. Res., 2005. **1**(010103).
113. Singh, C., *When physical intuition fails*. Am. J. Phys., 2002. **70**(11): p. 1103.
114. Schoenfeld, A.H., *Mathematical problem solving*. 1985, Orlando, FL: Academic Press. 409.
115. Podolefsky, N.S. and N.D. Finkelstein, *Use of analogy in learning physics: The role of representations*. Phys. Rev. ST Phys. Educ. Res., 2006. **2**(020101).
116. Podolefsky, N.S. and N.D. Finkelstein, *Analogical scaffolding and the learning of abstract ideas in physics: Empirical studies*. Phys. Rev. ST Phys. Educ. Res., (in review).
117. Podolefsky, N.S. and N.D. Finkelstein, *Analogical scaffolding and the learning of abstract ideas in physics: An example from electromagnetic waves*. Phys. Rev. ST Phys. Educ. Res., (in review).

## Appendix A: Homework and Quiz problems for chapters 5-7

Physics 2010 HW – Energy and motion

### Question 1 - Verbal Format

A professor drops a ball from the top of an eight-story physics building. At what point has the ball reached half of the speed it has just before it hits the ground? Neglect air resistance.

- A) The ball has reached half of its final speed when it has fallen two stories.
- B) The ball has reached half of its final speed when it has fallen four stories.
- C) The ball has reached half of its final speed when it has fallen six stories.
- D) The ball has reached half of its final speed at some other point.

How difficult did you consider this question? (Circle the best number)

Easy    1            2            3            4            5            Hard

### Question 2 - Mathematical Format

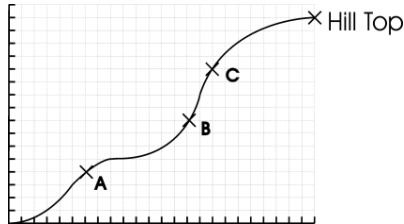
A dumbbell ( $m = 15 \text{ kg}$ ) is dropped from a height of 12 meters. Calculate the speed of the dumbbell when it has fallen to a height of 3 meters. Neglect air resistance, and round to the nearest 0.1 m/s. Note that  $g = 9.8 \text{ m/s}^2$ .

- A) 9.4 m/s
- B) 13.3 m/s
- C) 7.7 m/s
- D) None of the above.

### Question 3 - Graphical Format

A roller coaster car approaches a hill going just fast enough to reach the top and stop. A graph of the hill's height versus its horizontal coordinate is shown.

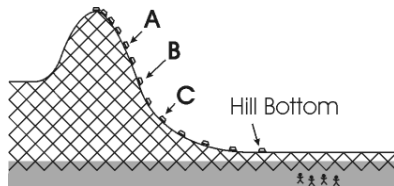
At what point has the car slowed down to half of its original speed? Neglect friction.



- A)      Point A            B)      Point B    C)      Point C
- D)      Somewhere else.

### Question 4 - Pictorial Format

A roller coaster car comes to rest at the top of a hill before starting down the other side. At what point on the track is the car moving at one half of the speed it has at the bottom of the hill? Ignore friction.



- A)      Point A    B)      Point B    C)      Point C
- D)      Somewhere else.

**Question 1 - Verbal Format**

We have two very flat pieces of glass on top of each other. The left side of the top piece of glass is propped up very slightly, so that there is a thin layer of air between the pieces. When we shine light on the glass from above, we observe light and dark interference fringes evenly spaced across the glass.

Now we prop up the right side of the top piece of glass by the same amount, so that both sides of the top piece of glass are held up by the same amount. What do we observe?

- A) Now that both sides are propped, we will see twice the effect, and so there will be twice as many fringes.
- B) Now that both sides are propped, we will see half the effect, and so there will be half as many fringes.
- C) The change we made was symmetric, and so we will see no change in the fringe pattern.
- D) The spacing between the pieces is constant, and so we will see no fringes at all.
- E) The new setup is exactly out of phase with the old one, so we will see light where before there was dark, and dark where before there was light.

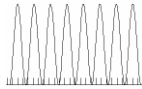
**Question 2 - Mathematical Format**

We have light of wavelength  $\lambda = 600$  nm shining on two very flat pieces of glass on top of each other. The top piece is propped up on one side, and we see bright and dark interference fringes across the glass. By how much does the thickness of the air gap change as you go from one bright fringe to the next bright fringe?

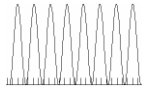
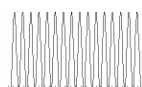

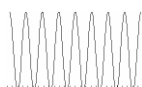

- A)  $t = 150$  nm
- B)  $t = 300$  nm
- C)  $t = 450$  nm
- D)  $t = 600$  nm
- E)  $t = 1200$  nm

**Question 3 - Graphical Format**

We have two very flat pieces of glass on top of each other. One side of the top piece of glass is propped up very slightly, and we shine light on the glass from above. We observe fringes whose intensity vs. position can be graphed as:



Now we change things by halving the distance by which the side of the top piece is propped up. Choose the graph corresponding to the intensity of the fringes we now see:






- A) 
- B) 
- C) 
- D) 
- E) 

**Question 4 - Pictorial Format**

We have two very flat pieces of glass on top of each other. One side of the top piece of glass is propped up very slightly, and we observe the following interference fringes:



Next, we prop up the side of the top piece twice as far as before. Choose the picture representing the resulting fringe pattern.

- A) 
- B) 
- C) 
- D) 
- E) 

**Question 1 - Verbal Format**

An electron in a Bohr-model hydrogen atom is in the 'orbit' with the lowest possible energy. How does the radius of the electron orbit change if it moves up to the third energy level?

- A) The radius of the new orbit will be three times the original radius.
- B) The radius of the new orbit will be nine times the original radius.
- C) The radius of the new orbit will be one-third the original radius.
- D) The radius of the new orbit will be one-ninth the original radius.
- E) None of these.

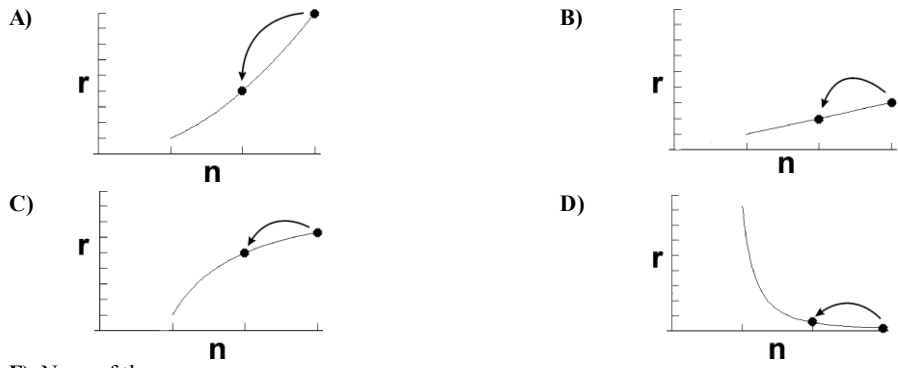
**Question 2 - Mathematical Format**

The Bohr radius for an electron is  $r_1=0.529 \times 10^{-10}$  m. Calculate the radius of the  $n=4$  energy level.

- A)  $r_4=0.033 \times 10^{-10}$  m
- B)  $r_4=0.132 \times 10^{-10}$  m
- C)  $r_4=2.116 \times 10^{-10}$  m
- D)  $r_4=8.464 \times 10^{-10}$  m
- E) None of the above.

**Question 3 – Graphical**

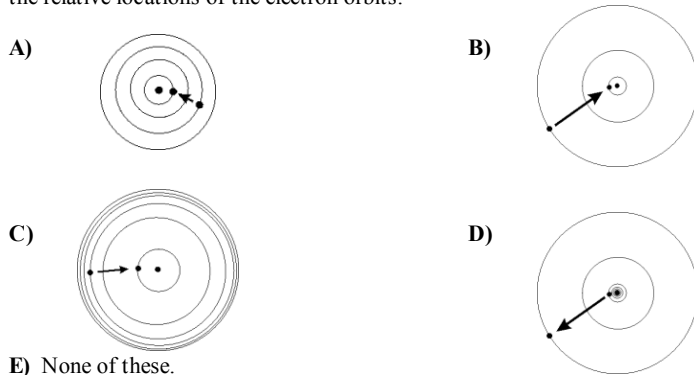
An electron in a Bohr hydrogen atom jumps from the  $n=3$  orbit to the  $n=2$  orbit. The following graphs show the orbit radius  $r$  as a function of the orbit number  $n$ . Choose the graph that best represents the relative locations of the electron orbits.



- E) None of these.

**Question 4 – Pictorial**

An electron in a Bohr hydrogen atom jumps from the  $n=3$  orbit to the  $n=1$  orbit. Choose the picture that best represents the relative locations of the electron orbits.



- E) None of these.

**Diffraction Problem -- Verbal Format**

We have a double-slit experiment set up. A helium-neon laser is shining on a pair of finite-width slits, and we see a corresponding intensity pattern on a screen. The pattern consists of narrow, closely spaced spots that get brighter and dimmer as you look across the screen, periodically dropping to nothing. I take the slits away and replace them with a pair of slits that are the same width, but are twice as far apart. What happens to the intensity pattern? Circle the appropriate letter.

- A) The entire pattern squishes together so that it is half as wide.
- B) The narrow peaks are half as far apart, and the rest of the pattern is unchanged.
- C) The narrow peaks are the same distance apart, but the places where the peaks drop away to nothing are twice as far apart.
- D) The narrow peaks are the same distance apart, but the places where the peaks drop away to nothing are half as far apart.
- E) The narrow peaks are twice as wide, and the rest of the pattern is unchanged.

How difficult did you consider this question? (Circle the appropriate number)

Easy    1        2        3        4        5        Hard

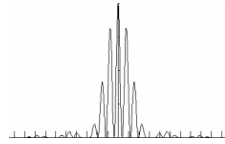
**Diffraction Problem -- Mathematical Format**

We have a double-slit experiment with incident light of  $\lambda = 633 \text{ nm}$ . On a screen  $3.0 \text{ m}$  from the slit, we see an intensity pattern with small peaks separated by  $0.5 \text{ cm}$ . The first minimum in the overall intensity envelope is at  $2.0 \text{ cm}$  from the center of the pattern. Calculate the separation of the slits,  $d$ . Circle the appropriate letter.

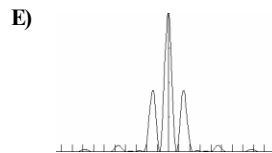
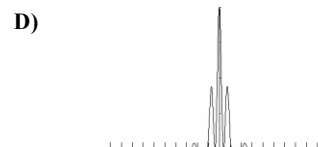
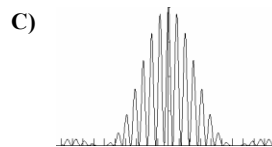
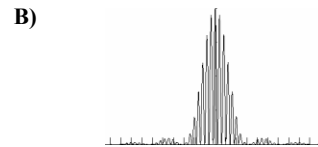
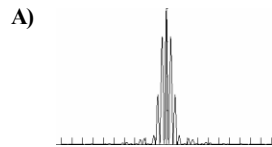
- A)  $D = 3.8 \times 10^{-5} \text{ m}$
- B)  $D = 3.8 \times 10^{-4} \text{ m}$
- C)  $D = 9.5 \times 10^{-5} \text{ m}$
- D)  $D = 9.5 \times 10^{-4} \text{ m}$
- E) None of the above.

**Diffraction Problem -- Graphical Format**

We have a double-slit experiment setup. A helium-neon laser is shining on a pair of finite-width slits, and we see a corresponding intensity pattern on a screen. The graph of the intensity of this pattern versus position is:

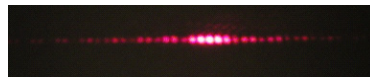


I take the slits away and replace them with a pair that has slits that are of the same width, but twice as far apart. What happens to the pattern on the screen? Circle the appropriate letter.

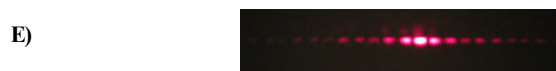
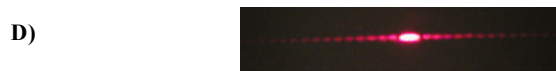
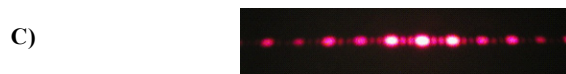
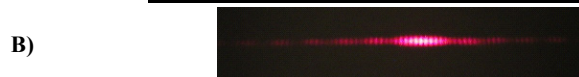
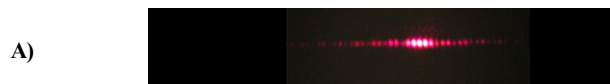


**Diffraction Problem -- Pictorial Format**

Suppose we have a double-slit experiment set up. A helium-neon laser is shining on two slits of finite width. On the screen behind the slits we see an intensity pattern that looks like:



Now we change the double slit setup so that the slits are twice as far apart. Which of the following intensity patterns will we see? Circle the appropriate letter.



**Spectroscopy Problem -- Verbal Format**

Consider the Balmer series of spectral lines from hydrogen gas. Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus. This means that the ionization energy for the electron will be smaller. What will happen to the Balmer lines that we see?

- A) The spectral lines will remain the same.
- B) The spectral lines will all shift to shorter wavelengths (toward the bluer colors).
- C) The spectral lines will all shift to longer wavelengths (toward the redder colors).
- D) The spectral lines will all shift toward the center of the visible spectrum.
- E) Something else.

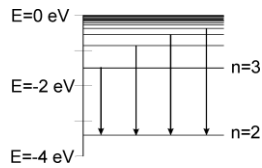
**Spectroscopy Problem -- Mathematical Format**

Suppose that we change the hydrogen atom so that the ionization energy for the electron is 11 eV instead of 13.6 eV. Calculate the energy of the photon emitted when the electron moves from the  $n = 4$  to the  $n = 2$  orbit.

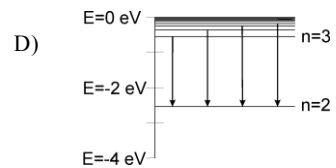
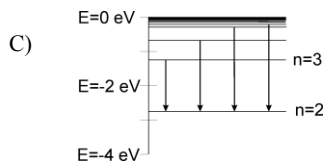
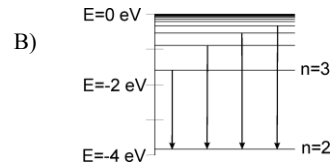
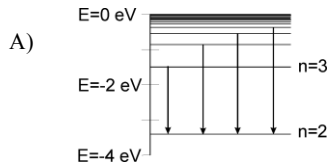
- A) 3.15 eV
- B) 2.75 eV
- C) 2.06 eV
- D) 2.55 eV
- E) None of the above.

**Spectroscopy Problem -- Graphical Format**

The energy level diagram below shows the electron transitions that lead to the Balmer series:



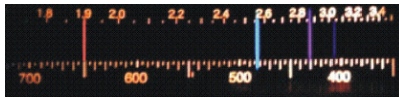
Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus and the ionization energy is 11 eV instead of 13.6 eV. Choose the graph that best represents what the new energy levels would look like.



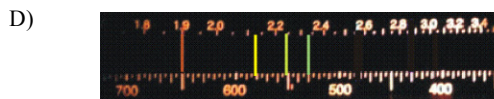
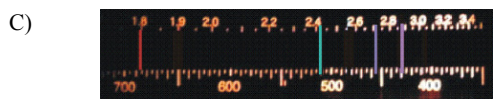
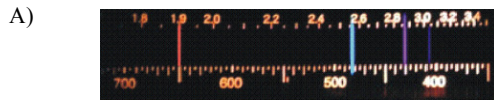
E) Something else.

**Spectroscopy Problem -- Pictorial Format**

The Balmer series of spectral lines is shown below, as seen through a spectrometer:



Now suppose we are in a world where electric charges are weaker, so the electron is not held as tightly by the nucleus and the ionization energy is 13 eV instead of 13.6 eV. Choose the picture that best represents what the new spectrum would look like.



E) Something else.



**Spring Problem -- Verbal Format**

A ball is hanging on a spring, and is oscillating up and down. At which point is the ball moving fastest?

- A) The ball is moving fastest when it is at its highest point.
- B) The ball is moving fastest when it is at the midpoint of its motion and is moving down.
- C) The ball is moving fastest when it is at its lowest point.
- D) None of the above are true.

**Spring Problem -- Mathematical Format**

A ball is hanging from a spring at rest at  $y = 0$  cm. The spring is then compressed until the ball is at  $y = 5$  cm, and is then released so that the ball oscillates. Up is in the positive- $y$  direction. At which point  $y$  is the ball moving fastest? Note that

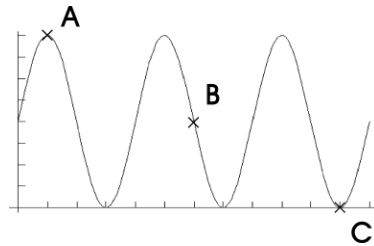
$$K = \frac{1}{2}mv^2 \quad U_{spring} = \frac{1}{2}k(y - y_0)^2 \quad \text{and} \quad U_{gravity} = mgy$$

where  $y_0$  is the unstretched length of the spring.

- A)  $y = -5$  cm
- B)  $y = 0$  cm
- C)  $y = +5$  cm
- D) None of the above.

**Spring Problem -- Graphical Format**

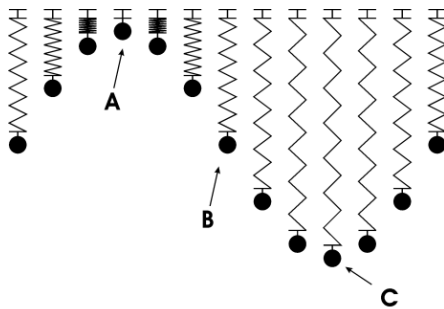
A ball is hanging on a spring and oscillating up and down. The height of the ball as a function of time is graphed below. At which point is the ball moving fastest?



- A) The ball is moving fastest at point A.
- B) The ball is moving fastest at point B.
- C) The ball is moving fastest at point C.
- D) None of the above are true.

**Spring Problem -- Pictorial Format**

A ball on a hanging spring is oscillating up and down as shown in the following snapshots.



At which point is the ball moving the fastest?

- A) The ball is moving fastest at point A.
- B) The ball is moving fastest at point B.
- C) The ball is moving fastest at point C.
- D) None of the above are true.

**Pendulum Problem -- Verbal Format**

I set up a pendulum in front of you and pull it back (to your right), and then let it go. The pendulum takes one second to reach the point opposite from where it started.

Now I lengthen the pendulum's string until it is four times as long as it was, with the mass unchanged. I pull the pendulum back to the right again (far enough that the string is at the same angle as before), and let it go. Where is it after one second? Circle the correct answer.

- A) Straight up and down, and moving left.
- B) Opposite from its starting position.
- C) Straight up and down, and moving right.
- D) Back in its starting position.
- E) Somewhere else.

**Pendulum Problem -- Mathematical Format**

A pendulum of length  $L = 1$  m starts at  $x = +5$  cm and is released at  $t = 0$ . At  $t = 1$  s, it is at almost exactly  $x = -5$  cm. Now suppose we change the length of the pendulum to  $L = 4$  m without changing the mass. We pull it back and release it from  $x = +20$  cm at  $t = 0$ . Find  $x$  and the sign of the pendulum's velocity at  $t = 1$  s.

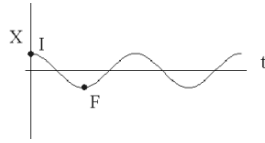
Possibly useful equations:

$$x = A \cos\left(\frac{2\pi t}{T}\right) \quad v = -A \frac{2\pi}{T} \sin\left(\frac{2\pi t}{T}\right) \quad T = 2\pi \sqrt{\frac{L}{g}}$$

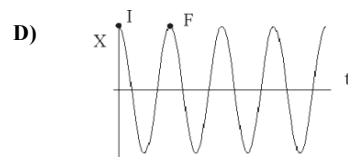
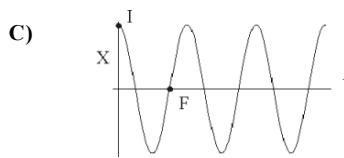
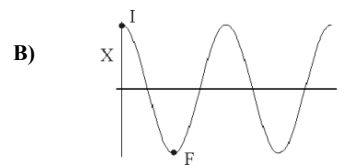
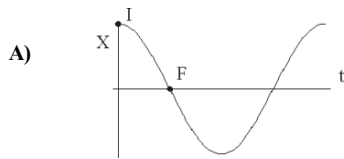
- A)  $x = 0$  cm,  $v$  is negative.
- B)  $x = -20$  cm,  $v$  is zero.
- C)  $x = 0$ ,  $v$  is positive.
- D)  $x = +20$  cm,  $v$  is zero.
- E) None of the above.

**Pendulum Problem – Graphical Format**

The following graph shows the motion of a pendulum after we let it go at  $t = 0$ . The horizontal axis shows the time  $t$  and the vertical axis shows the position of the pendulum  $X$ . It starts at point I and is at point F at  $t = 1$ s.



Now I change the pendulum so that it is four times as long as before, but the same mass. I pull it back to the same angle it was at before, and let it go. The graphs below show several possible motions of the pendulum. Let point F again represent  $t = 1$ s. Circle the graph that correctly describes the motion.



E) None of the above.

**Pendulum Problem – Pictorial Format**

I pull a pendulum back to the position shown below on the left and let it go. It takes one second to swing into the position shown below and on the right.

Start:

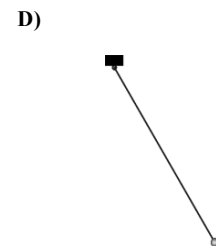
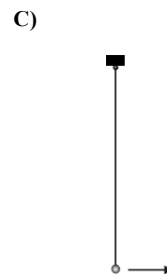
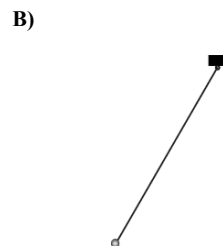
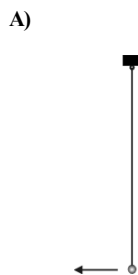


After one second:



Now I change the pendulum so that it is four times as long as before, with the same mass. I pull the pendulum back to the same side to the same angle as before and then let it go.

Select the picture that corresponds to the position of the new pendulum after one second. If the pendulum is straight up and down, select the picture that indicates the correct direction of the motion.



E) None of the above.

## Appendix B: Student reasons for choosing quiz representations (Chapter 5)

2020 Diffraction quiz	
Verbal 10 usable	3: Preference for qualitative analysis. 1: Connected it to the pre-recitation homework. 1: Prefers concepts to math. 1: Not good at math. 1: Thought it would be good practice.
Math 24	9: Preference for “plug ‘n chug” problems. 4: Find equations/numbers easy to work with. 3: Preference for mathematics over pictures. 2: Connected it to the lab. 1: Connected it to the pre-recitation homework.
Graphical 12	3: Visual learners/people. 2: Like having a visualization provided. 2: Connected it to the pre-recitation homework. 1: Connected it to the lab. 1: Like having a qualitative/quantitative hybrid
Pictorial 51	17: Visual learners/people. 12: Connected it to lab. 9: Find other formats (esp. math/words) difficult. 7: Like having a visualization provided. 5: Preference for concepts/concepts over math.

2020 Spectroscopy quiz	
Verbal 12	3: Preference for concepts/words over math/pictures. 3: Don't like pictures. 2: The format supports the concepts.
Math 18	5: Preference for “plug ‘n chug” problems. 4: Preference for mathematics over concepts. 3: Preference for mathematics over pictures. 2: Like the definite/straightforward nature. 1: The format supports the concepts.
Graphical 10	2: Visual learners/people. 2: Preference for visuals over math. 2: Connected it to the pre-recitation homework. 1: The format supports the concepts. 1: Like having a qualitative/quantitative hybrid
Pictorial 35	12: Liked the colors/found it attractive 8: Like having a visualization provided. 6: Connected it to lab. 5: Visual learners/people. 2: Preference for concepts/concepts over math. 2: Thought it would be good practice.

2010 Spring quiz	
Verbal 9	6: Preference for concepts/intuition 3: Other formats are difficult.
Math 74	16: Preference for mathematics over concepts. 12: Like the definite/straightforward nature. 11: Comfortable handling equations. 7: Preference for “plug ‘n chug” problems. 7: Find other formats difficult. 7: Connected it to the pre-recitation homework
Graphical 11	3: Visual learners/people. 2: Like having a visualization provided. 2: Find other formats difficult. 2: Connected it to the pre-recitation homework. 1: The format supports the concepts. 1: Like having a qualitative/quantitative hybrid
Pictorial 26	12: Visual learners/people. 7: Like having a visualization provided. 5: Find other formats difficult (esp. math). 3: Preference for pictures over equations/numbers.

2010 Pendulum quiz	
Verbal 17	4: Preference for concepts/words over math/pictures. 4: Clear/ordered presentation. 3: Other formats are difficult. 2: Eases visualization of the problem. 2: Didn't like the math format before.
Math 22	7: Comfortable handling equations. 6: Like the definite/straightforward nature. 2: Preference for mathematics over concepts. 2: The format supports the concepts. 1: Preference for “plug ‘n chug” problems.
Graphical 27	6: Visual learners/people. 6: Like having a visualization provided. 3: Like having a qualitative/quantitative hybrid 3: Didn't like the math format before. 2: Suits the topic well. 1: The format supports the concepts.
Pictorial 62	15: Visual learners/people. 13: Didn't like the math format before. 12: Like having a visualization provided. 5: Don't like math generally. 5: Preference for pictures over equations/numbers. 3: Connection to real life. 3: Connection to lab and lecture demos. 2: The format supports the concepts.

## Appendix C: Standard-format BEMA and FMCE exams (chapter 8/9)

---

Two small objects each with a net charge of  $+Q$  exert a force of magnitude  $F$  on each other:



We replace one of the objects with another whose net charge is  $+4Q$ :



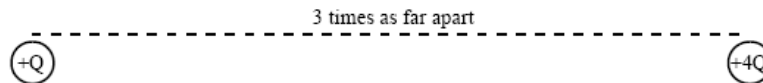
→ Q1 The original magnitude of the force on the  $+Q$  charge was  $F$ ; what is the magnitude of the force on the  $+Q$  charge now?

- (a)  $4F$
- (b)  $5F/2$
- (c)  $3F$
- (d)  $2F$
- (e)  $F$
- (f)  $F/4$
- (g) None of the above

→ Q2 What is the magnitude of the force on the  $+4Q$  charge?

- (a)  $4F$
- (b)  $5F/2$
- (c)  $3F$
- (d)  $2F$
- (e)  $F$
- (f)  $F/4$
- (g) None of the above

Next we move the  $+Q$  and  $+4Q$  charges to be 3 times as far apart as they were:



→ Q3 Now what is the magnitude of the force on the  $+4Q$  charge?

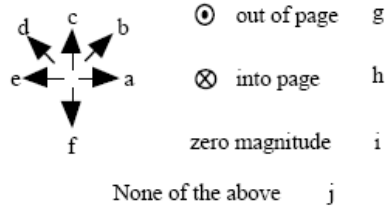
- (a)  $4F/3$
- (b)  $4F/9$
- (c)  $F/3$
- (d)  $5F/18$
- (e)  $2F/9$
- (f)  $F/9$
- (g)  $F/36$
- (h)  $4F$
- (i) None of the above

Here are two charges of equal magnitude but opposite sign, separated by a distance  $s$ :



2  
x

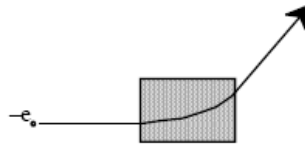
Choose from the following possible directions to answer the questions below:



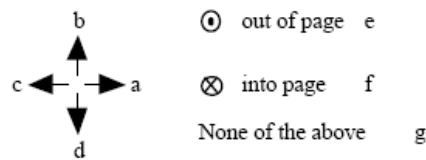
→ Q4 What is the direction (a-j) of the electric field at location 1 (marked with an x)?

→ Q5 What is the direction (a-j) of the electric field at location 2 (marked with an x)?

A moving electron with charge  $-e$  travels along the path shown, and passes through a region of electric field. There are no other charges present. The electric field is zero everywhere except in the gray region.



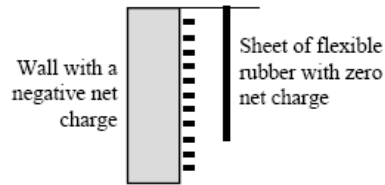
Choose from the following possible directions to answer the question below:



→ Q6 What is a possible direction (a-g) of the electric field in the region where the field is non-zero?

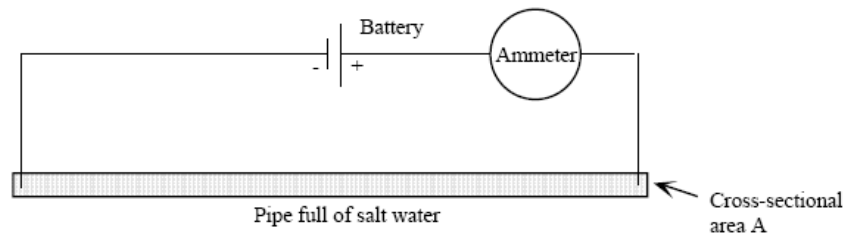


A non-conducting wall is given a negative net charge. Next, a sheet of very flexible rubber with zero net charge is suspended from the ceiling near the charged wall as shown below.



- Q7 The rubber sheet will:
- not be affected by the charges on the wall since rubber is an insulator.
  - not be affected by the charged wall because the rubber sheet has zero net charge.
  - bend away from the wall due to the electrical repulsion between the electrons in the rubber and the charges on the wall.
  - bend away from the wall due to the polarization of the rubber molecules by the charged wall.
  - bend toward the wall due to the polarization of the rubber molecules by the charged wall.
  - none of the above.

Salt water contains  $n$  sodium ions ( $\text{Na}^+$ ) per cubic meter and  $n$  chloride ions ( $\text{Cl}^-$ ) per cubic meter. A battery is connected to metal rods that dip into a narrow pipe full of salt water. The cross-sectional area of the pipe is  $A$ :

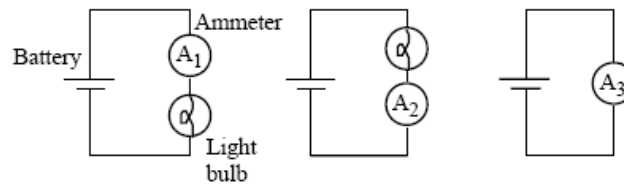


- Q8 What is the direction of conventional current flow in the salt water?
- To the right.
  - To the left.
  - There is no conventional current, because the motions of the positive and negative ions cancel each other out.

The magnitude of the drift velocity of the sodium ions is  $v_{\text{Na}^+}$  and the magnitude of the drift velocity of the chloride ions is  $v_{\text{Cl}^-}$ . Assume that  $v_{\text{Na}^+} > v_{\text{Cl}^-}$ . ( $+e$  is the charge of a proton.)

- Q9 What is the magnitude of the ammeter reading?
- $enAv_{\text{Na}^+} - enAv_{\text{Cl}^-}$
  - $enAv_{\text{Na}^+} + enAv_{\text{Cl}^-}$
  - $enAv_{\text{Na}^+}$
  - $enAv_{\text{Cl}^-}$
  - zero

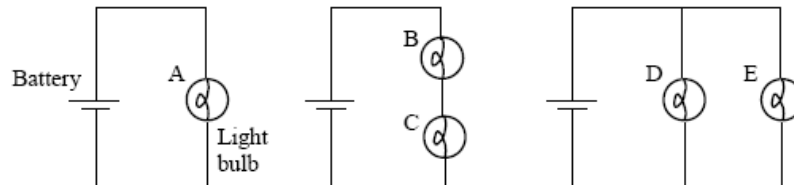
A student has set up the three circuits shown. The light bulbs and the batteries are identical.



→ Q10 Rank all 3 ammeters ( $A_1$ ,  $A_2$ , and  $A_3$ ) in order of their current measurements from greatest to smallest.

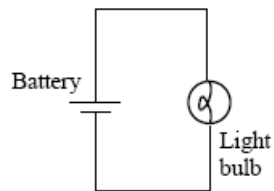
- |                       |                       |
|-----------------------|-----------------------|
| (a) $A_1 = A_2 = A_3$ | (e) $A_2 = A_3 > A_1$ |
| (b) $A_1 = A_2 > A_3$ | (f) $A_3 > A_1 = A_2$ |
| (c) $A_1 = A_3 > A_2$ | (g) $A_3 > A_1 > A_2$ |
| (d) $A_2 = A_1 > A_3$ | (h) $A_3 > A_2 > A_1$ |
| (i) None of the above |                       |

In these three circuits, all the batteries are identical and have negligible internal resistance, and all the light bulbs are identical.

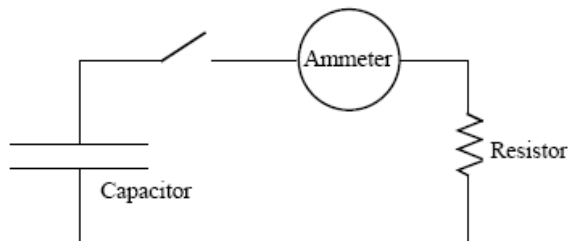


→ Q11 Rank all 5 light bulbs (A, B, C, D, E) in order of brightness from brightest to dimmest.

- |                         |                         |
|-------------------------|-------------------------|
| (a) $A = B = C > D = E$ | (e) $A = D = E > B = C$ |
| (b) $A > B = C = D = E$ | (f) $A = D = E > B > C$ |
| (c) $A > B = C > D = E$ | (g) $A > D = E > B = C$ |
| (d) $A > B > C > D = E$ | (h) $D = E > A > B = C$ |
| (i) None of the above   |                         |



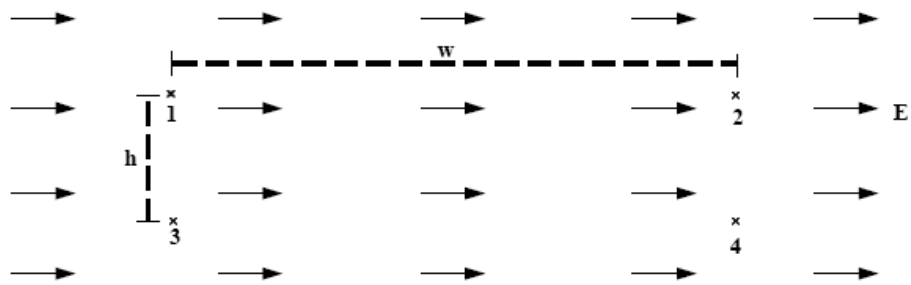
- Q12 Which of the following statements is true about the electric field inside the bulb filament?
- (a) The field must be zero because the filament is made of metal.
  - (b) The field must be zero because a current is flowing.
  - (c) The field must be zero because any excess charges are on the surface of the filament.
  - (d) The field must be non-zero because the flowing current produces an electric field.
  - (e) The field must be non-zero because no current will flow without an applied field.
  - (f) The field must be zero for reasons not given above.
  - (g) The field must be non-zero for reasons not given above.



- Q13 The capacitor is originally charged. How does the current  $I$  in the ammeter behave as a function of time after the switch is closed?
- (a)  $I = 0$  always.
  - (b)  $I = \text{constant} \neq 0$
  - (c)  $I$  increases, then is constant.
  - (d)  $I$  instantly jumps up, then slowly decreases.
  - (e) None of the above.

---

In a certain region of space there is a uniform electric field of magnitude  $E$ :



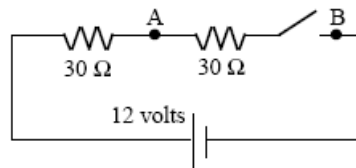
Choose from the following possible values to answer the three questions below:

- (a)  $+Ew$
- (b)  $-Ew$
- (c)  $+Eh$
- (d)  $-Eh$
- (e)  $+E\sqrt{(h^2 + w^2)}$
- (f)  $-E\sqrt{(h^2 + w^2)}$
- (g) zero

→ Q14 The potential difference  $V_2 - V_1 = ?$

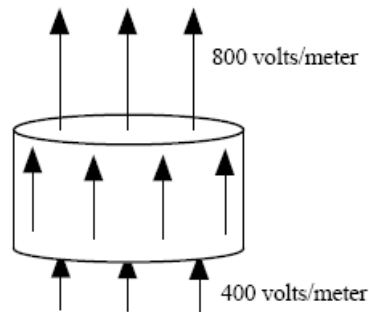
→ Q15 The potential difference  $V_3 - V_1 = ?$

→ Q16 The potential difference  $V_4 - V_1 = ?$



- Q17 What is the magnitude of the potential difference between points A and B on the circuit, while the switch is open?
- (a) 0 volts.
  - (b) 3 volts.
  - (c) 6 volts.
  - (d) 12 volts.
  - (e) None of the above.
- 

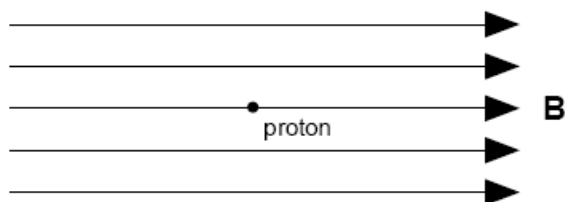
Here is a cylinder on whose surfaces there is a vertical electric field of varying magnitude as shown. The electric field is uniform on the top face, and also uniform on the bottom face.



- Q18 This cylinder encloses
- (a) no net charge.
  - (b) net positive charge.
  - (c) net negative charge.
  - (d) There is not enough information available to determine whether or not there is net charge inside the cylinder.

- 
- Q19 In static equilibrium, the potential difference between two points inside a solid piece of metal
- (a) is zero because metals block electric interactions.
  - (b) is zero because the electric field is zero inside the metal.
  - (c) is non-zero if the piece of metal is not spherical.
  - (d) is non-zero if there are charges on the surface of the metal.
  - (e) is non-zero for reasons not given above.
  - (f) is zero for reasons not given above.
- 

A proton is initially at rest in a region of constant magnetic field (shown below). There are no other charges present.

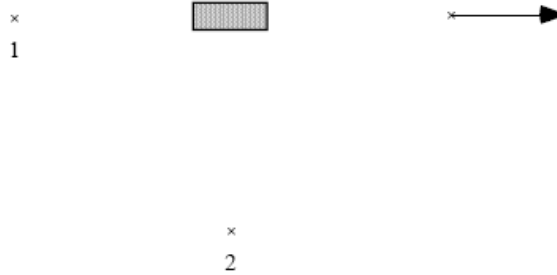


Choose from the following possible directions to answer the question below:

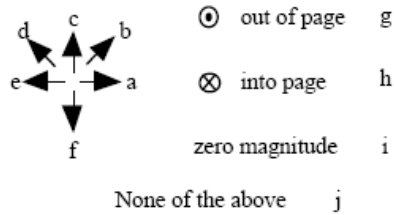
- |        |                   |   |
|--------|-------------------|---|
| b<br>↑ | ⊙ out of page     | e |
| ← a →  | ⊗ into page       | f |
| ↓      | zero force        | g |
| c      | none of the above | h |

- Q20 What is the direction (a-h) of the initial magnetic force on the proton?

Here is a bar magnet. The magnetic field made by the bar magnet at one location is shown on the diagram:



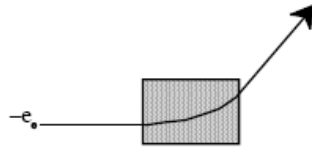
Choose from the following possible directions to answer the questions below:



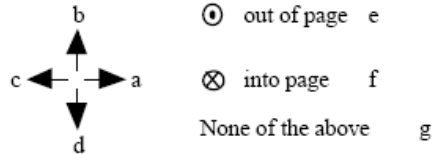
→ Q21 What is the direction (a-j) of the magnetic field of the bar magnet at location 1 (marked with x)?

→ Q22 What is the direction (a-j) of the magnetic field of the bar magnet at location 2 (marked with x)?

A moving electron travels along the path shown, and passes through a region of magnetic field. There are no other charges present. The magnetic field is zero everywhere except in the gray region.

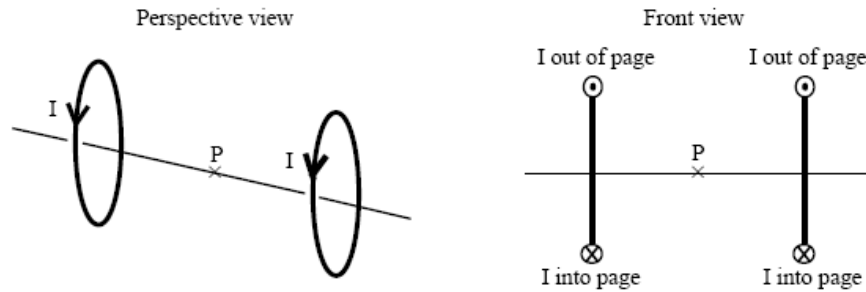


Choose from the following possible directions to answer the question below:

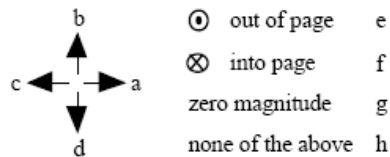


→ Q23 What is a possible direction (a-g) of the magnetic field in the region where the field is non-zero?

Two identical circular loops of wire, perpendicular to the page, carry the same conventional current  $I$ :



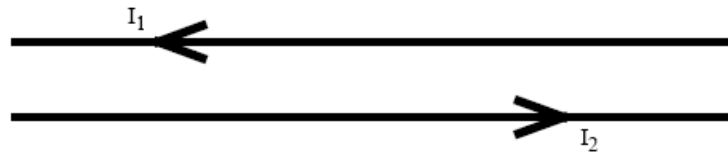
Choose from the following possible directions to answer the question below:



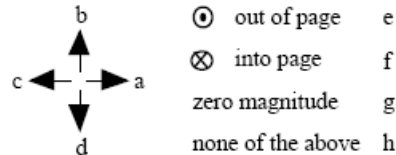
→ Q24 In the front view, what is the direction (a-h) of the magnetic field due to the loops at location P, midway between the loops?



Two wires lie in the plane of the page. Wire 1 carries conventional current to the left, and wire 2 carries conventional current to the right:

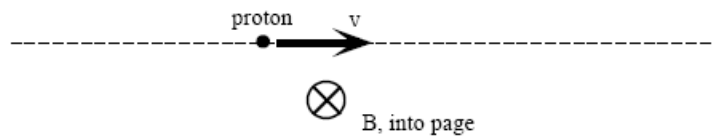


Choose from the following possible directions to answer the question below:

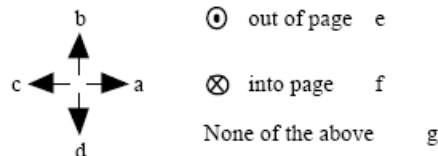


→ Q25 What is the direction (a-h) of the magnetic force that wire 1 exerts on wire 2?

A proton moves with constant velocity  $v$  to the right through a region where there is a uniform magnetic field of magnitude  $B$  that points into the page. There is also an electric field in this region. The magnetic field and electric field are produced by devices not shown on the diagram.



Choose from the following possible directions to answer the question below:

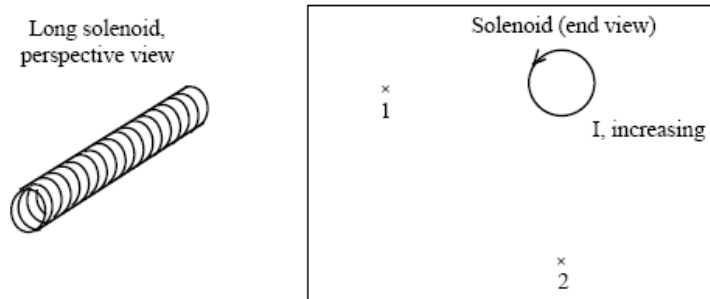


→ Q26 What is the direction (a-g) of the electric field in this region?

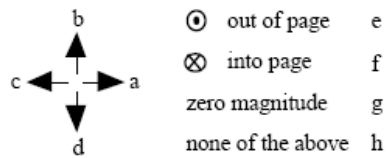
→ Q27 What is the magnitude of the electric field?

- |                              |                               |
|------------------------------|-------------------------------|
| (a) $evB$                    | (e) $e\vec{v} \times \vec{B}$ |
| (b) $\vec{v} \times \vec{B}$ | (f) $vB/e$                    |
| (c) $vB$                     | (g) $ev$                      |
| (d) $B$                      | (h) None of the above         |

Here is a long solenoid (coils of wire along a long cylinder), and an end view of the solenoid. Conventional current runs counter-clockwise in the solenoid and is increasing with time.

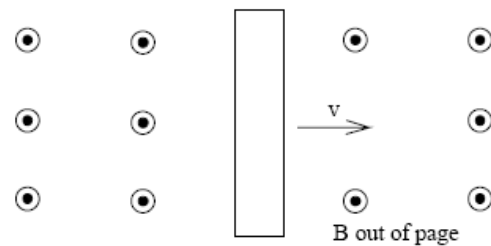


Choose from the following possible directions to answer the questions below:

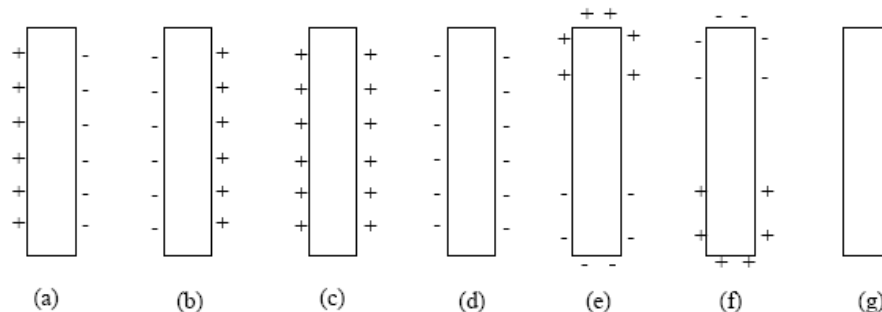


- Q28 What is the direction (a-h) of the *electric* field at location 1 (marked with an ×)?
- Q29 What is the direction (a-h) of the *electric* field at location 2 (marked with an ×)?

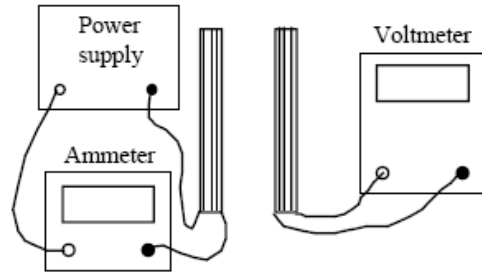
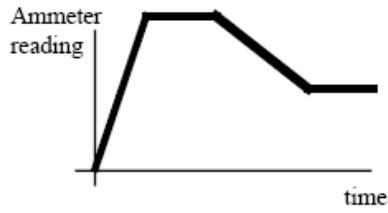
A neutral metal bar is moving at constant velocity  $v$  to the right through a region where there is a uniform magnetic field pointing out of the page. The magnetic field is produced by some large coils which are not shown on the diagram.



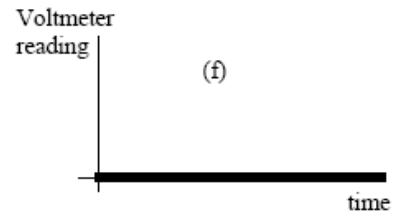
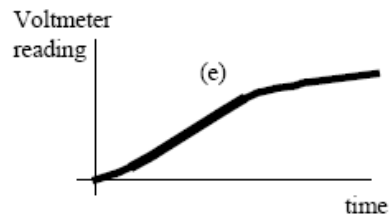
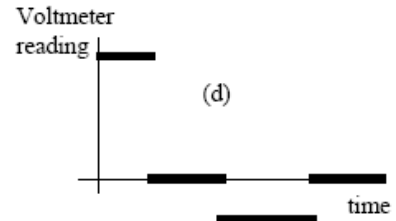
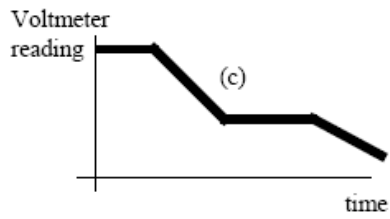
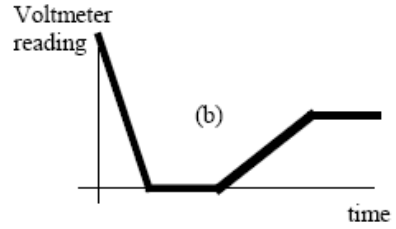
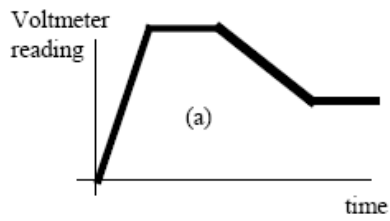
- Q30 Which of the following diagrams best describes the state of the metal bar?



A variable power supply is connected to a coil and an ammeter, and the time dependence of the ammeter reading is shown. A nearby coil is connected to a voltmeter.



→Q31 Which of the following graphs correctly shows the time dependence of the voltmeter reading?

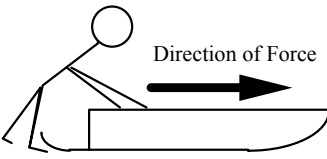


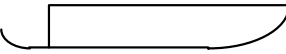
**FORCE AND MOTION CONCEPTUAL EVALUATION**

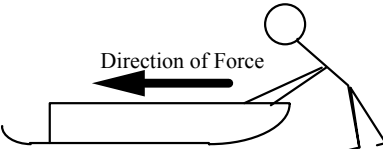
**Directions:** Answer questions 1-47 in spaces on the answer sheet. Be sure your name is on the answer sheet. Answer question 46a also on the answer sheet. Hand in the questions and the answer sheet.

A sled on ice moves in the ways described in questions 1-7 below. *Friction is so small that it can be ignored.* A person wearing spiked shoes standing on the ice can apply a force to the sled and push it along the ice. Choose the one force (A through G) which would **keep the sled moving** as described in each statement below.

You may use a choice more than once or not at all but choose only one answer for each blank. If you think that none is correct, answer choice J.

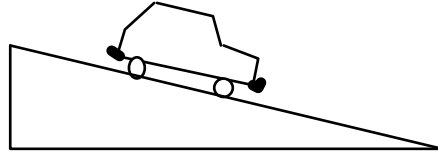
	<p><b>A.</b> The force is toward the <b>right</b> and is <b>increasing</b> in strength (magnitude).</p> <p><b>B.</b> The force is toward the <b>right</b> and is of <b>constant</b> strength (magnitude).</p> <p><b>C.</b> The force is toward the <b>right</b> and is <b>decreasing</b> in strength (magnitude).</p>
---	---

	<p><b>D.</b> No applied force is needed</p>
---	---

	<p><b>E.</b> The force is toward the <b>left</b> and is <b>decreasing</b> in strength (magnitude).</p> <p><b>F.</b> The force is toward the <b>left</b> and is of <b>constant</b> strength (magnitude).</p> <p><b>G.</b> The force is toward the <b>left</b> and is <b>increasing</b> in strength (magnitude).</p>
---	--

- \_\_\_\_\_ 1. Which force would keep the sled moving toward the right and speeding up at a steady rate (constant acceleration)?
- \_\_\_\_\_ 2. Which force would keep the sled moving toward the right at a steady (constant) velocity?
- \_\_\_\_\_ 3. The sled is moving toward the right. Which force would slow it down at a steady rate (constant acceleration)?
- \_\_\_\_\_ 4. Which force would keep the sled moving toward the left and speeding up at a steady rate (constant acceleration)?
- \_\_\_\_\_ 5. The sled was started from rest and pushed until it reached a steady (constant) velocity toward the right. Which force would keep the sled moving at this velocity?
- \_\_\_\_\_ 6. The sled is slowing down at a steady rate and has an acceleration to the right. Which force would account for this motion?
- \_\_\_\_\_ 7. The sled is moving toward the left. Which force would slow it down at a steady rate (constant acceleration)?

Questions 8-10 refer to a toy car which is given a quick push so that it rolls up an inclined ramp. After it is released, it rolls up, reaches its highest point and rolls back down again. *Friction is so small it can be ignored.*



Use one of the following choices (A through G) to indicate the **net force** acting on the car for each of the cases described below. Answer choice **J** if you think that none is correct.

- |   |   |
|---|---|
| <input type="radio"/> <b>A</b> Net <b>constant</b> force <b>down</b> ramp   | <input type="radio"/> <b>E</b> Net <b>constant</b> force <b>up</b> ramp   |
| <input type="radio"/> <b>B</b> Net <b>increasing</b> force <b>down</b> ramp | <input type="radio"/> <b>D</b> Net force zero                             |
| <input type="radio"/> <b>C</b> Net <b>decreasing</b> force <b>down</b> ramp | <input type="radio"/> <b>F</b> Net <b>increasing</b> force <b>up</b> ramp |
|   | <input type="radio"/> <b>G</b> Net <b>decreasing</b> force <b>up</b> ramp |

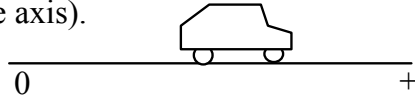
- \_\_\_\_\_ 8. The car is moving up the ramp after it is released.
- \_\_\_\_\_ 9. The car is at its highest point.
- \_\_\_\_\_ 10. The car is moving down the ramp.

Questions 11-13 refer to a coin which is tossed straight up into the air. After it is released it moves upward, reaches its highest point and falls back down again. Use one of the following choices (A through G) to indicate the force acting on the coin for each of the cases described below. Answer choice **J** if you think that none is correct. **Ignore any effects of air resistance.**

- A.** The force is **down** and constant.
- B.** The force is **down** and increasing
- C.** The force is **down** and decreasing
- D.** The force is zero.
- E.** The force is **up** and constant.
- F.** The force is **up** and increasing
- G.** The force is **up** and decreasing

- \_\_\_\_\_ 11. The coin is moving upward after it is released.
- \_\_\_\_\_ 12. The coin is at its highest point.
- \_\_\_\_\_ 13. The coin is moving downward.

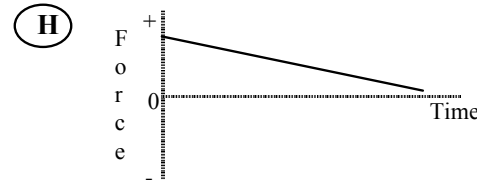
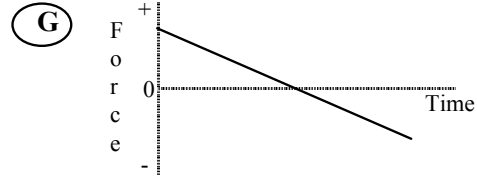
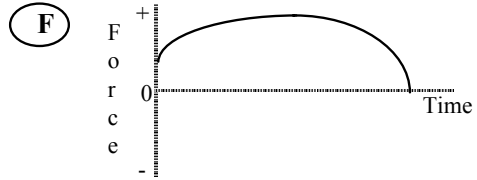
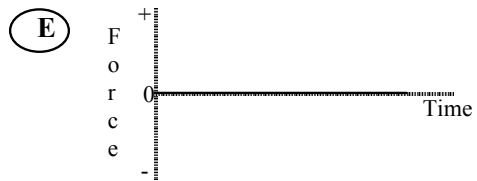
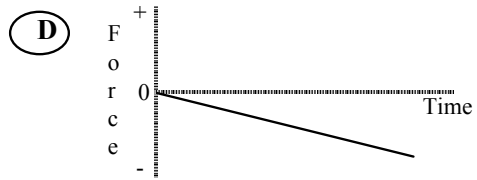
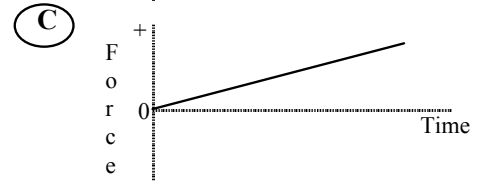
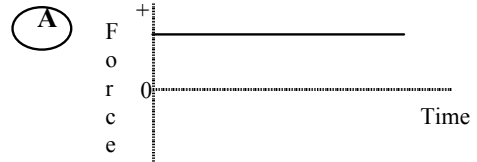
Questions 14-21 refer to a toy car which can move to the right or left along a horizontal line (the positive part of the distance axis).



Assume that friction is so small that it can be ignored.

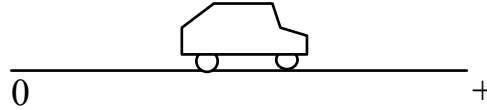
A force is applied to the car. Choose the one force graph (A through H) for each statement below which could allow the described motion of the car to continue. You may use a choice more than once or not at all. If you think that none is correct, answer choice J

- \_\_ 14. The car moves toward the right (away from the origin) with a steady (constant) velocity.
- \_\_ 15. The car is at rest.
- \_\_ 16. The car moves toward the right and is speeding up at a steady rate (constant acceleration).
- \_\_ 17. The car moves toward the left (toward the origin) with a steady (constant) velocity.
- \_\_ 18. The car moves toward the right and is slowing down at a steady rate (constant acceleration).
- \_\_ 19. The car moves toward the left and is speeding up at a steady rate (constant acceleration).
- \_\_ 20. The car moves toward the right, speeds up and then slows down.
- \_\_ 21. The car was pushed toward the right and then released. Which graph describes the force after the car is released.



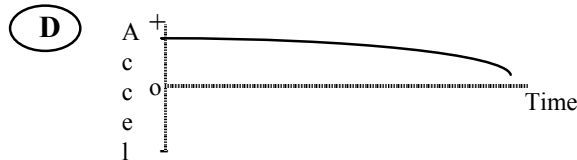
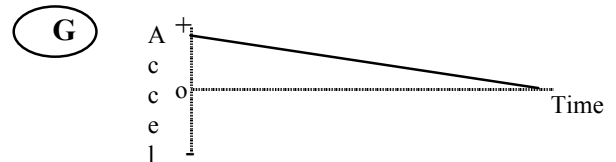
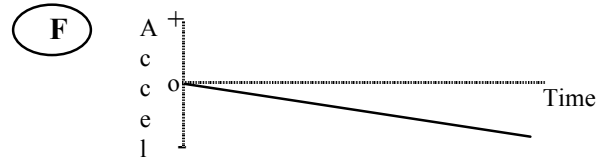
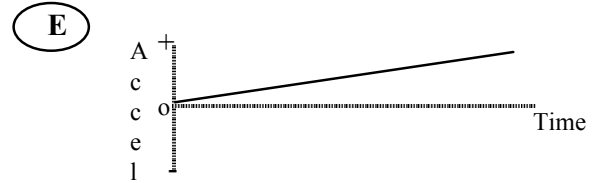
(J) None of these graphs is correct.

Questions 22-26 refer to a toy car which can move to the right or left on a horizontal surface along a straight line (the + distance axis). The positive direction is to the right.



Different motions of the car are described below. Choose the letter (A to G) of the **acceleration-time** graph which corresponds to the motion of the car described in each statement.

You may use a choice more than once or not at all. If you think that none is correct, answer choice **J**.



**(J)** None of these graphs is correct.

- \_\_\_ 22. The car moves toward the right (away from the origin), speeding up at a steady rate.  
 \_\_\_ 23. The car moves toward the right, slowing down at a steady rate.  
 \_\_\_ 24. The car moves toward the left (toward the origin) at a constant velocity.  
 \_\_\_ 25. The car moves toward the left, speeding up at a steady rate.  
 \_\_\_ 26. The car moves toward the right at a constant velocity.

Questions 27-29 refer to a coin which is tossed straight up into the air. After it is released it moves upward, reaches its highest point and falls back down again. Use one of the following choices (A through G) to indicate the acceleration of the coin during each of the stages of the coin's motion described below. Take **up** to be the **positive** direction. Answer choice **J** if you think that none is correct.

- A.** The acceleration is in the negative direction and constant.  
**B.** The acceleration is in the negative direction and increasing  
**C.** The acceleration is in the negative direction and decreasing  
**D.** The acceleration is zero.  
**E.** The acceleration is in the positive direction and constant.  
**F.** The acceleration is in the positive direction and increasing  
**G.** The acceleration is in the positive direction and decreasing

- \_\_\_ 27. The coin is moving upward after it is released.  
 \_\_\_ 28. The coin is at its highest point.  
 \_\_\_ 29. The coin is moving downward.

Questions 30-34 refer to collisions between a car and trucks. For each description of a collision (30-34) below, choose the one answer from the possibilities **A** through **J** that best describes the forces between the car and the truck.

- A.** The truck exerts a greater amount of force on the car than the car exerts on the truck.

- B. The car exerts a greater amount of force on the truck than the truck exerts on the car.
- C. Neither exerts a force on the other; the car gets smashed simply because it is in the way of the truck.
- D. The truck exerts a force on the car but the car doesn't exert a force on the truck.
- E. The truck exerts the same amount of force on the car as the car exerts on the truck.
- F. Not enough information is given to pick one of the answers above.
- J. None of the answers above describes the situation correctly.

In questions 30 through 32 the truck is **much heavier** than the car.



- \_\_\_\_\_30. They are both moving at the same speed when they collide. Which choice describes the forces?
- \_\_\_\_\_31. The car is moving much faster than the heavier truck when they collide. Which choice describes the forces?
- \_\_\_\_\_32. The heavier truck is standing still when the car hits it. Which choice describes the forces?

In questions 33 and 34 the truck is a **small pickup** and is the **same weight** as the car.



- \_\_\_\_\_33. Both the truck and the car are moving at the same speed when they collide. Which choice describes the forces?
- \_\_\_\_\_34. The truck is standing still when the car hits it. Which choice describes the forces?

Questions 35-38 refer to a large truck which breaks down out on the road and receives a push back to town by a small compact car.



Pick one of the choices A through J below which correctly describes the forces between the car and the truck for each of the descriptions (35-38).

- A. The force of the car pushing against the truck is equal to that of the truck pushing back against the car.
  - B. The force of the car pushing against the truck is less than that of the truck pushing back against the car.
  - C. The force of the car pushing against the truck is greater than that of the truck pushing back against the car.
  - D. The car's engine is running so it applies a force as it pushes against the truck, but the truck's engine isn't running so it can't push back with a force against the car.
  - E. Neither the car nor the truck exert any force on each other. The truck is pushed forward simply because it is in the way of the car.
  - J. None of these descriptions is correct.
- \_\_\_\_\_35. The car is pushing on the truck, but not hard enough to make the truck move.
  - \_\_\_\_\_36. The car, still pushing the truck, is **speeding up** to get to cruising speed.
  - \_\_\_\_\_37. The car, still pushing the truck, is at cruising speed and continues to travel at the **same speed**.
  - \_\_\_\_\_38. The car, still pushing the truck, is at cruising speed when the truck puts on its brakes and causes the car to **slow down**.

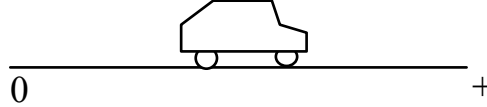




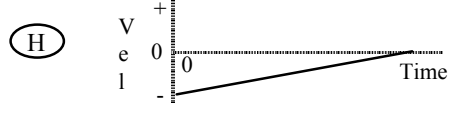
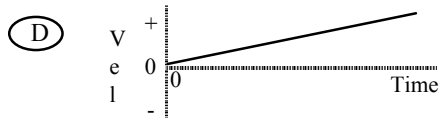
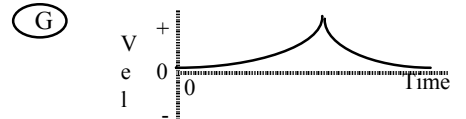
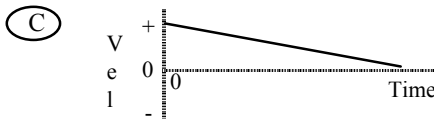
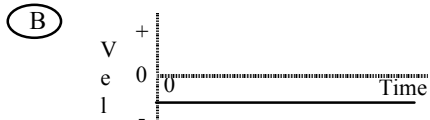
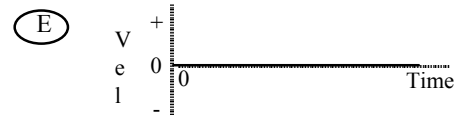
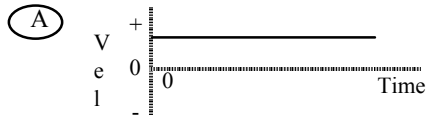
39. Two students sit in identical office chairs facing each other. Bob has a mass of 95 kg, while Jim has a mass of 77 kg. Bob places his bare feet on Jim's knees, as shown to the right. Bob then suddenly pushes outward with his feet, causing both chairs to move. In this situation, while Bob's feet are in contact with Jim's knees,

- A. Neither student exerts a force on the other.
- B. Bob exerts a force on Jim, but Jim doesn't exert any force on Bob.
- C. Each student exerts a force on the other, but Jim exerts the larger force.
- D. Each student exerts a force on the other, but Bob exerts the larger force.
- E. Each student exerts the same amount of force on the other.
- J. None of these answers is correct.

Questions 40-43 refer to a toy car which can move to the right or left along a horizontal line (the positive portion of the distance axis). The positive direction is to the right.



Choose the correct velocity-time graph (A - G) for each of the following questions. You may use a graph more than once or not at all. If you think that none is correct, answer choice J.



(J) None of these graphs is correct.

- 40. Which velocity graph shows the car moving toward the right (away from the origin) at a steady (constant) velocity?
- 41. Which velocity graph shows the car reversing direction?
- 42. Which velocity graph shows the car moving toward the left (toward the origin) at a steady (constant) velocity?
- 43. Which velocity graph shows the car increasing its *speed* at a steady (constant) rate?



A sled is pulled up to the top of a hill. The sketch above indicates the shape of the hill. At the top of the hill the sled is released from rest and allowed to coast down the hill. At the bottom of the hill the sled has a speed  $v$  and a kinetic energy  $E$  (the energy due to the sled's motion). Answer the following questions. *In every case friction and air resistance are so small they can be ignored.*

- \_\_44. The sled is pulled up a **steeper** hill of the **same** height as the hill described above. How will the velocity of the sled at the bottom of the hill (after it has slid down) compare to that of the sled at the bottom of the original hill? Choose the best answer below.
- A. The speed at the bottom is greater for the steeper hill.
  - B. The speed at the bottom is the same for both hills.
  - C. The speed at the bottom is greater for the original hill because the sled travels further.
  - D. There is not enough information given to say which speed at the bottom is faster.
  - J. None of these descriptions is correct.
- \_\_45. Compare the kinetic energy (energy of motion) of the sled at the bottom for the original hill and the steeper hill in the previous problem. Choose the best answer below.
- A. The kinetic energy of the sled at the bottom is greater for the steeper hill.
  - B. The kinetic energy of the sled at the bottom is the same for both hills.
  - C. The kinetic energy at the bottom is greater for the original hill.
  - D. There is not enough information given to say which kinetic energy is greater.
  - J. None of these descriptions is correct.
- \_\_46. The sled is pulled up a **higher** hill that is **less** steep than the original hill described before question 44. How does the speed of the sled at the bottom of the hill (after it has slid down) compare to that of the sled at the bottom of the original hill?
- A. The speed at the bottom is greater for the higher but less steep hill than for the original.
  - B. The speed at the bottom is the same for both hills.
  - C. The speed at the bottom is greater for the original hill.
  - D. There is not enough information given to say which speed at the bottom is faster.
  - J. None of these descriptions is correct.
- 46a. Describe in words your reasoning in reaching your answer to question 46. (**Answer on the answer sheet** and use as much space as you need)
- \_\_47. For the higher hill that is less steep, how does the kinetic energy of the sled at the bottom of the hill after it has slid down compare to that of the original hill?
- A. The kinetic energy of the sled at the bottom is greater for the higher but less steep hill.
  - B. The kinetic energy of the sled at the bottom is the same for both hills.
  - C. The kinetic energy at the bottom is greater for the original hill.
  - D. There is not enough information given to say which kinetic energy is greater.
  - J. None of these descriptions is correct.

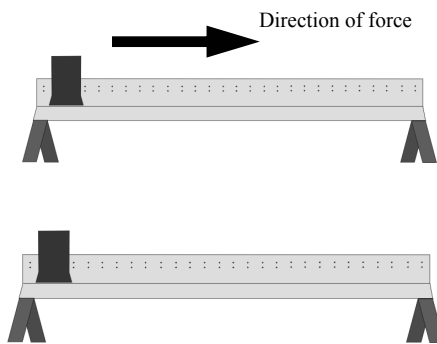
## Appendix D: Revised FMCE (chapter 10)

### FORCE AND MOTION CONCEPTUAL EVALUATION

**Directions:** Answer questions 1-47 in spaces on the answer sheet. Be sure your name is on the answer sheet. Hand in the questions and the answer sheet.

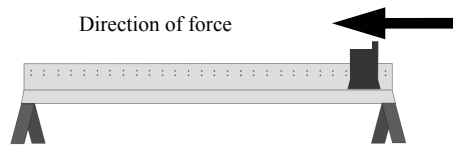
A cart on a long frictionless air track moves in the ways described in questions 1-7 below. *Friction is so small that it can be ignored.* A force can be applied to the cart (by a string attached to a machine) that pulls the cart along the track. Choose the one force (A through G) which would **keep the cart moving** as described in each statement below. The track is so long that the cart won't reach the end.

You may use a choice more than once or not at all but choose only one answer for each blank. If you think that none is correct, answer choice J.



- A. The force is toward the **right** and is **increasing** in strength (magnitude).
- B. The force is toward the **right** and is of **constant** strength (magnitude).
- C. The force is toward the **right** and is **decreasing** in strength (magnitude).

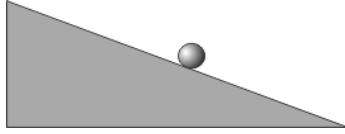
D. No applied force is needed



- E. The force is toward the **left** and is **decreasing** in strength (magnitude).
- F. The force is toward the **left** and is of **constant** strength (magnitude).
- G. The force is toward the **left** and is **increasing** in strength (magnitude).

- \_\_\_\_\_ 1. Which force would keep the cart moving toward the right and speeding up at a steady rate (constant acceleration)?
- \_\_\_\_\_ 2. Which force would keep the cart moving toward the right at a steady (constant) velocity?
- \_\_\_\_\_ 3. The cart is moving toward the right. Which force would slow it down at a steady rate (constant acceleration)?
- \_\_\_\_\_ 4. Which force would keep the cart moving toward the left and speeding up at a steady rate (constant acceleration)?
- \_\_\_\_\_ 5. The cart was started from rest and pushed until it reached a steady (constant) velocity toward the right. Which force would keep the cart moving at this velocity?
- \_\_\_\_\_ 6. The cart is slowing down at a steady rate and has an acceleration to the right. Which force would account for this motion?
- \_\_\_\_\_ 7. The cart is moving toward the left. Which force would slow it down at a steady rate (constant acceleration)?

Questions 8-10 refer to a steel ball bearing which is given a quick push so that it rolls up an inclined laboratory ramp. After it is released, it rolls up, reaches its highest point and rolls back down again. *Friction is so small it can be ignored.*



Use one of the following choices (**A** through **G**) to indicate the **net force** acting on the bearing for each of the cases described below. Answer choice **J** if you think that none is correct.

- |   |   |
|---|---|
| <input type="radio"/> <b>A</b> Net <b>constant</b> force <b>down</b> ramp   | <input type="radio"/> <b>E</b> Net <b>constant</b> force <b>up</b> ramp   |
| <input type="radio"/> <b>B</b> Net <b>increasing</b> force <b>down</b> ramp | <input type="radio"/> <b>D</b> Net force zero                             |
| <input type="radio"/> <b>C</b> Net <b>decreasing</b> force <b>down</b> ramp | <input type="radio"/> <b>F</b> Net <b>increasing</b> force <b>up</b> ramp |
|   | <input type="radio"/> <b>G</b> Net <b>decreasing</b> force <b>up</b> ramp |

\_\_\_\_\_ 8. The bearing is moving up the ramp after it is pushed.

\_\_\_\_\_ 9. The bearing is at its highest point.

\_\_\_\_\_ 10. The bearing is moving down the ramp.

Questions 11-13 refer to a coin which is tossed straight up into the air. After it is released it moves upward, reaches its highest point and falls back down again. Use one of the following choices (**A** through **G**) to indicate the force acting on the coin for each of the cases described below. Answer choice **J** if you think that none is correct. **Ignore any effects of air resistance.**

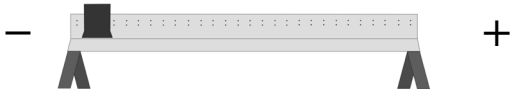
- A.** The force is **down** and constant.
- B.** The force is **down** and increasing
- C.** The force is **down** and decreasing
- D.** The force is zero.
- E.** The force is **up** and constant.
- F.** The force is **up** and increasing
- G.** The force is **up** and decreasing

\_\_\_\_\_ 11. The coin is moving upward after it is released.

\_\_\_\_\_ 12. The coin is at its highest point.

\_\_\_\_\_ 13. The coin is moving downward.

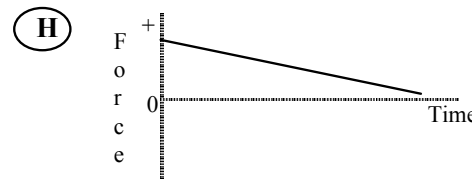
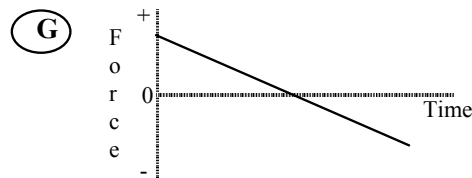
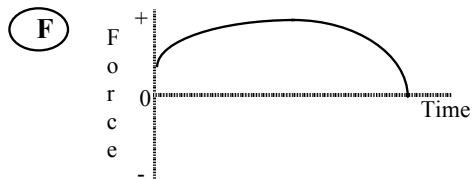
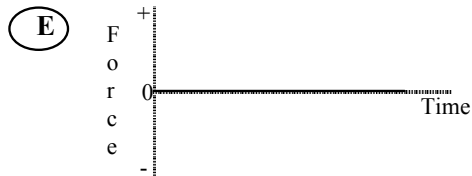
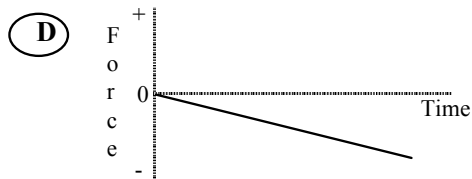
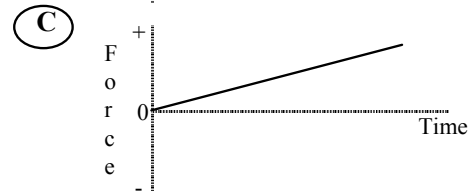
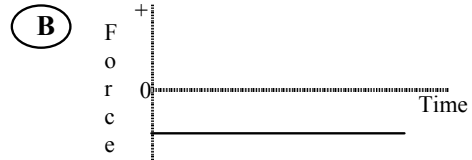
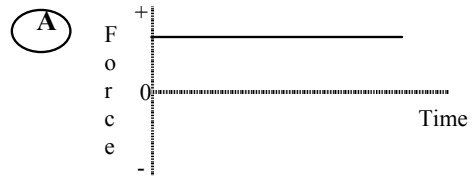
Questions 14-21 refer to an air cart that can move to the right or left along a very long horizontal track (with the right being the positive direction).



A force is applied to the cart. Choose the one force graph (A through H) for each statement below which could allow the described motion of the cart to continue.

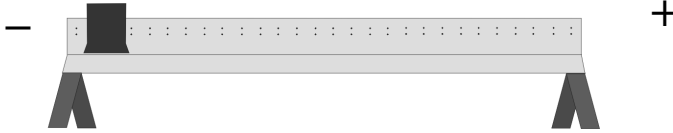
You may use a choice more than once or not at all. If you think that none is correct, answer choice J.

- \_\_ 14. The cart moves toward the right (away from the origin) with a steady (constant) velocity.
- \_\_ 15. The cart is at rest.
- \_\_ 16. The cart moves toward the right and is speeding up at a steady rate (constant acceleration).
- \_\_ 17. The cart moves toward the left (toward the origin) with a steady (constant) velocity.
- \_\_ 18. The cart moves toward the right and is slowing down at a steady rate (constant acceleration).
- \_\_ 19. The cart moves toward the left and is speeding up at a steady rate (constant acceleration).
- \_\_ 20. The cart moves toward the right, speeds up and then slows down.
- \_\_ 21. The cart was pushed toward the right and then released. Which graph describes the force after the car is released?



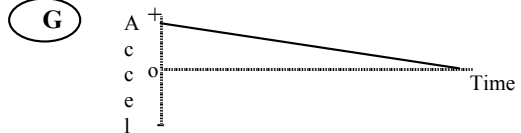
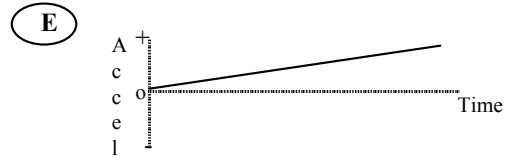
(J) None of these graphs is correct.

Questions 22-26 refer to an air cart which can move to the right or left on a horizontal surface along a straight line (the + distance axis). The positive direction is to the right.



Different motions of the cart are described below. Choose the letter (A to G) of the **acceleration-time** graph which corresponds to the motion of the cart described in each statement.

You may use a choice more than once or not at all. If you think that none is correct, answer choice J.



(J) None of these graphs is correct.

- \_\_\_\_ 22. The cart moves toward the right (away from the origin), speeding up at a steady rate.  
 \_\_\_\_ 23. The cart moves toward the right, slowing down at a steady rate.  
 \_\_\_\_ 24. The cart moves toward the left (toward the origin) at a constant velocity.  
 \_\_\_\_ 25. The cart moves toward the left, speeding up at a steady rate.  
 \_\_\_\_ 26. The cart moves toward the right at a constant velocity.

Questions 27-29 refer to a coin which is tossed straight up into the air. After it is released it moves upward, reaches its highest point and falls back down again. Use one of the following choices (A through G) to indicate the acceleration of the coin during each of the stages of the coin's motion described below. Take **up** to be the **positive** direction. Answer choice **J** if you think that none is correct.

- A. The acceleration is in the negative direction and constant.
- B. The acceleration is in the negative direction and increasing
- C. The acceleration is in the negative direction and decreasing
- D. The acceleration is zero.
- E. The acceleration is in the positive direction and constant.
- F. The acceleration is in the positive direction and increasing
- G. The acceleration is in the positive direction and decreasing

\_\_\_27. The coin is moving upward after it is released.

\_\_\_28. The coin is at its highest point.

\_\_\_29. The coin is moving downward.

Questions 30-34 refer to collisions between masses  $M_1$  and  $M_2$ . For each description of a collision (30-34) below, choose the one answer from the possibilities A through J that best describes the forces between  $M_1$  and  $M_2$ .

- A.  $M_1$  exerts a greater amount of force on  $M_2$  than  $M_2$  exerts on  $M_1$ .
- B.  $M_2$  exerts a greater amount of force on  $M_1$  than  $M_1$  exerts on  $M_2$ .
- C. Neither exerts a force on the other;  $M_2$  gets smashed simply because it is in the way of  $M_1$ .
- D.  $M_1$  exerts a force on  $M_2$  but  $M_2$  doesn't exert a force on  $M_1$ .
- E.  $M_1$  exerts the same amount of force on  $M_2$  as  $M_2$  exerts on  $M_1$ .
- F. Not enough information is given to pick one of the answers above.
- J. None of the answers above describes the situation correctly.

In questions 30-32,  $M_1$  is much more massive than  $M_2$ .

\_\_\_\_\_30. They are both moving at the same speed when they collide. Which choice describes the forces?

\_\_\_\_\_31.  $M_2$  is moving much faster than the more massive  $M_1$  when they collide. Which choice describes the forces?

\_\_\_\_\_32. The more massive  $M_1$  is not moving when  $M_2$  hits it. Which choice describes the forces?

In questions 33 and 34,  $M_1$  and  $M_2$  have the same mass.

\_\_\_\_\_33. Both the truck and the car are moving at the same speed when they collide. Which choice describes the forces?

\_\_\_\_\_34. The truck is standing still when the car hits it. Which choice describes the forces?

Questions 35-38 refer to a very massive freighter ship which breaks down in the ocean and receives a push back to dock by a tugboat.

Pick one of the choices A through J below which correctly describes the forces between the tugboat and the freighter for each of the descriptions (35-38).

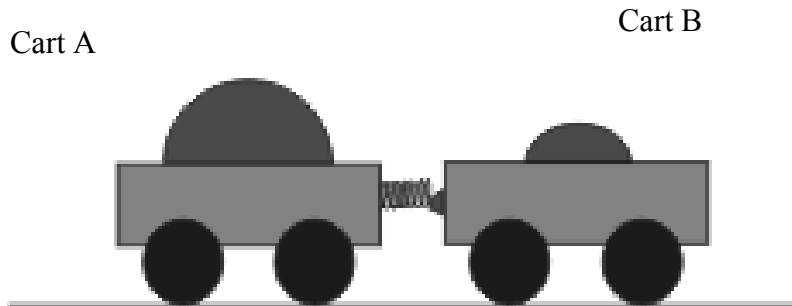
- A. The force of the tugboat pushing against the freighter is equal to that of the freighter pushing back against the tugboat.
- B. The force of the tugboat pushing against the freighter is less than that of the freighter pushing back against the tugboat.

- C. The force of the tugboat pushing against the freighter is greater than that of the freighter pushing back against the tugboat.
- D. The tugboat's engine is running so it applies a force as it pushes against the freighter, but the freighter's engine isn't running so it can't push back with a force against the tugboat.
- E. Neither the tugboat nor the freighter exert any force on each other. The freighter is pushed forward simply because it is in the way of the tugboat.
- J. None of these descriptions is correct.

- \_\_\_\_\_35. The tugboat is pushing on the freighter, but not hard enough to make the freighter move.
- \_\_\_\_\_36. The tugboat, still pushing the freighter, is **speeding up** to get to cruising speed.
- \_\_\_\_\_37. The tugboat, still pushing the freighter, is at cruising speed and continues to travel at the **same speed**.
- \_\_\_\_\_38. The tugboat, still pushing the freighter, is at cruising speed when the freighter drops its anchor, which drags along and causes the tugboat to **slow down**.

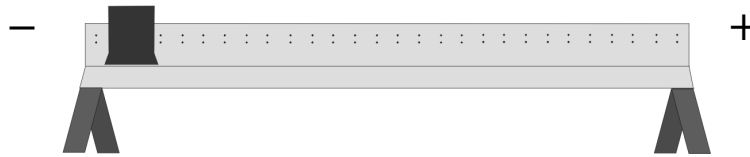
- \_\_\_\_\_39. Two carts sit on a steel table as shown below. Cart A has a mass of 9.5 kg, while Cart B has a mass of 7.7 kg. Cart A has a compressed spring attached to it, which has a rubber stopper on one side that is pressed up against Cart B. The spring suddenly releases, pushing outward, causing both carts to move. In this situation, while Cart A's plunger is in contact with Cart B,

- A. Neither cart exerts a force on the other.
- B. Cart A exerts a force on Cart B, but Cart B doesn't exert any force on Cart A.
- C. Each cart exerts a force on the other, but Cart B exerts the larger force.
- D. Each cart exerts a force on the other, but Cart A exerts the larger force.
- E. Each cart exerts the same amount of force on the other.
- J. None of these answers is correct.

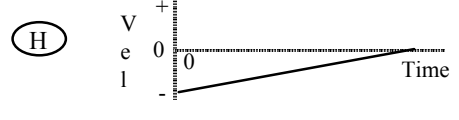
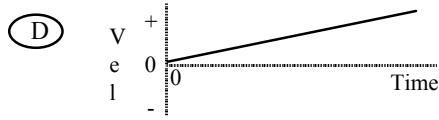
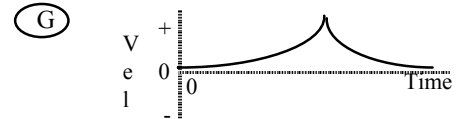
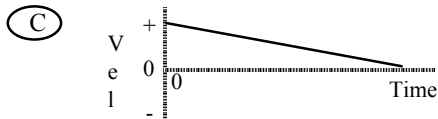
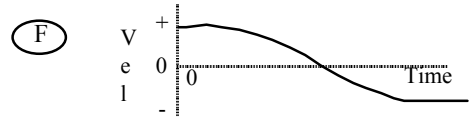
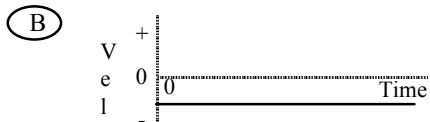
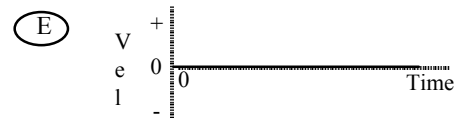
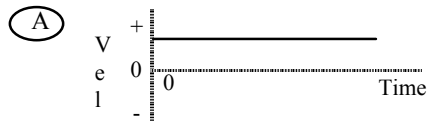




Questions 40-43 refer to an air cart which can move to the right or left along a horizontal track. The positive direction is to the right. The track is long enough that the cart won't fall off.



Choose the correct velocity-time graph (A - G) for each of the following questions. You may use a graph more than once or not at all. If you think that none is correct, answer choice J.



(J) None of these graphs is correct.

- \_\_\_40. Which velocity graph shows the cart moving toward the right (away from the origin) at a steady (constant) velocity?
- \_\_\_41. Which velocity graph shows the cart reversing direction?
- \_\_\_42. Which velocity graph shows the cart moving toward the left (toward the origin) at a steady (constant) velocity?
- \_\_\_43. Which velocity graph shows the cart increasing its *speed* at a steady (constant) rate?

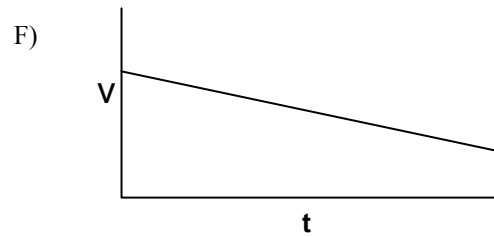
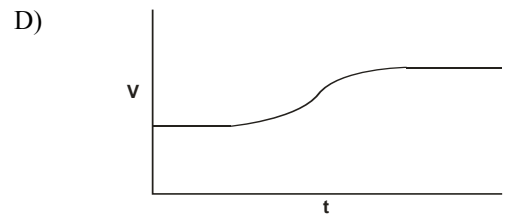
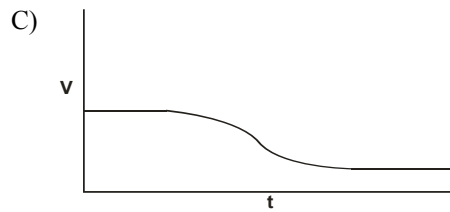
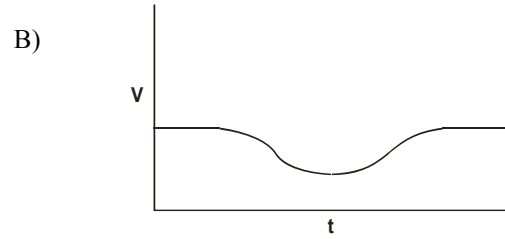
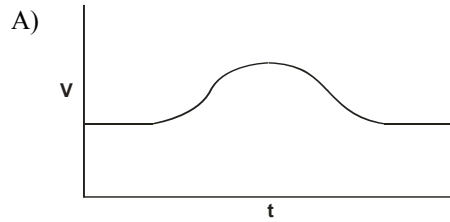


A steel ball bearing is placed at the top of a steel laboratory ramp. The sketch above indicates the shape of the ramp. At the top of the ramp the bearing is released from rest and allowed to roll down the ramp. At the bottom of the ramp the bearing has a speed  $v$  and a kinetic energy  $E$  (the energy due to the bearing's motion). Answer the following questions. *In every case friction and air resistance are so small they can be ignored.*

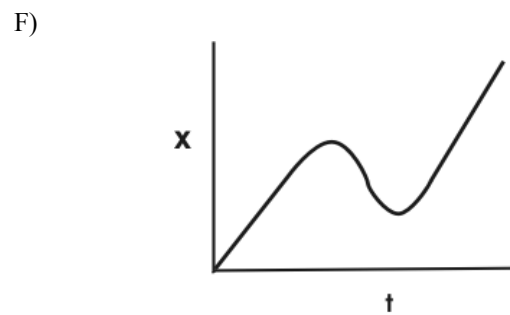
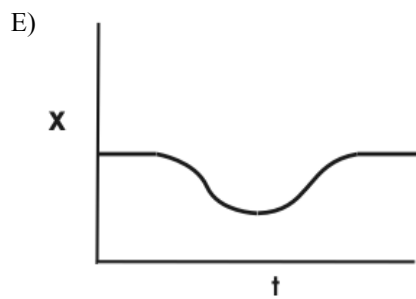
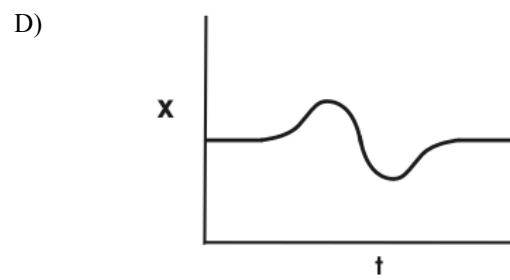
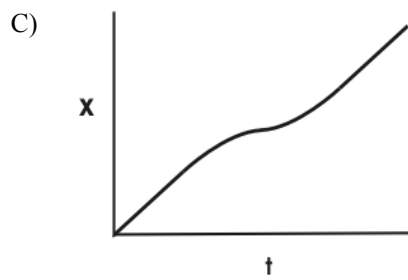
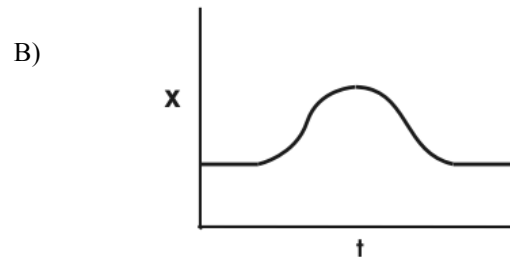
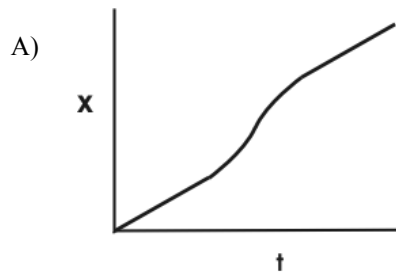
- \_\_44. The bearing is put at the top of a **steeper** ramp of the **same** height as the ramp described above. How will the velocity of the bearing at the bottom of the ramp (after it has rolled down) compare to that of the bearing at the bottom of the original ramp? Choose the best answer below.
- A. The speed at the bottom is greater for the steeper ramp.
  - B. The speed at the bottom is the same for both ramps.
  - C. The speed at the bottom is greater for the original ramp because the bearing rolls farther.
  - D. There is not enough information given to say which speed at the bottom is faster.
  - J. None of these descriptions is correct.
- \_\_45. Compare the kinetic energy (energy of motion) of the bearing at the bottom for the original ramp and the steeper ramp in the previous problem. Choose the best answer below.
- A. The kinetic energy of the bearing at the bottom is greater for the steeper ramp.
  - B. The kinetic energy of the bearing at the bottom is the same for both ramps.
  - C. The kinetic energy at the bottom is greater for the original ramp.
  - D. There is not enough information given to say which kinetic energy is greater.
  - J. None of these descriptions is correct.
- \_\_46. The bearing is placed at the top of a **higher** ramp that is **less** steep than the original ramp described before question 44 above. How does the speed of the bearing at the bottom of the ramp (after it has rolled down) compare to that of the bearing at the bottom of the original ramp?
- A. The speed at the bottom is greater for the higher but less steep ramp than for the original.
  - B. The speed at the bottom is the same for both ramps.
  - C. The speed at the bottom is greater for the original ramps.
  - D. There is not enough information given to say which speed at the bottom is faster.
  - J. None of these descriptions is correct.
- \_\_47. For the higher ramp that is less steep, how does the kinetic energy of the bearing at the bottom of the ramp after it has rolled down compare to that of the original ramp?
- A. The kinetic energy of the bearing at the bottom is greater for the higher but less steep ramp.
  - B. The kinetic energy of the bearing at the bottom is the same for both ramps.
  - C. The kinetic energy at the bottom is greater for the original ramp.
  - D. There is not enough information given to say which kinetic energy is greater.
  - J. None of these descriptions is correct.

## Appendix E: Recitation problems used in chapter 10

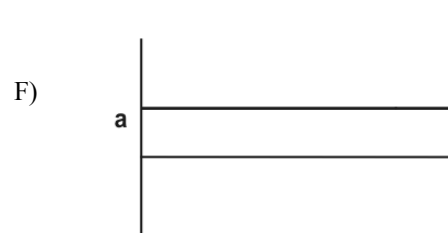
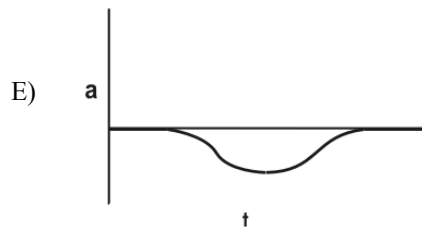
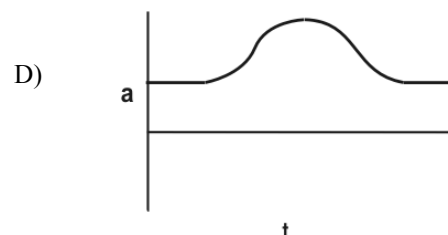
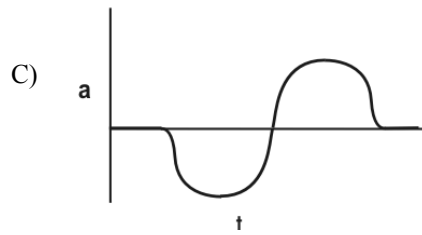
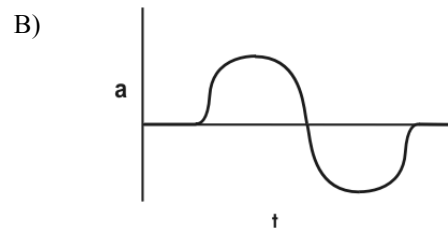
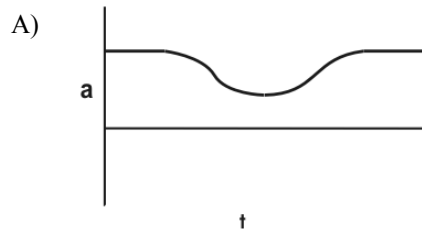
I give a steel ball a quick push along a frictionless track. The following are graphs of that ball's velocity in the x-direction as a function of time (after the push). Which graph would be correct if the track went straight, then over a hill (up and back down), and then straight again?



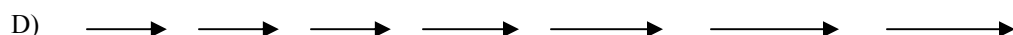
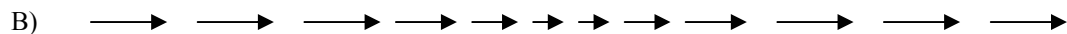
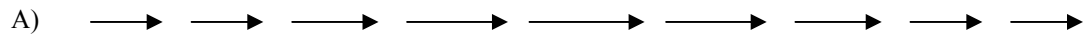
I give a steel ball a quick push along a frictionless track. The following are graphs of that ball's position in the  $x$ -direction as a function of time (after the push). Which graph would be correct if the track went straight, then across a valley (down, and then back up), and then straight again?



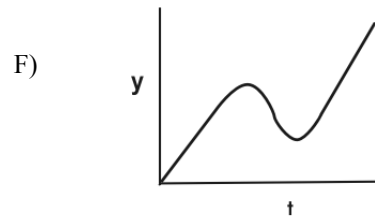
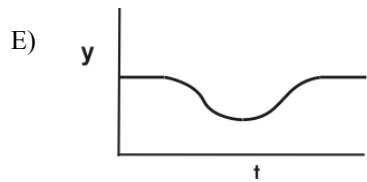
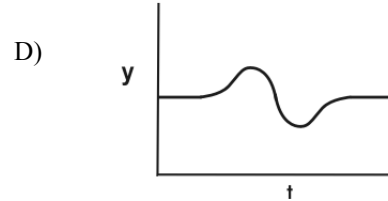
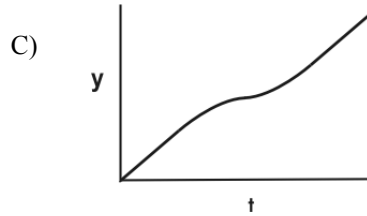
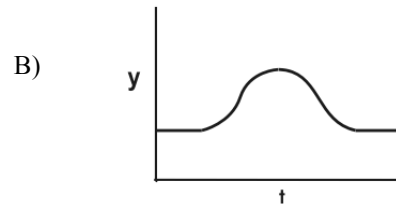
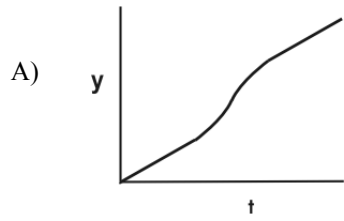
I give a steel ball a quick push along a frictionless track. The following are graphs of that ball's acceleration in the x (horizontal) direction as a function of time (after the push). Which graph would be correct if the track went straight, then over a hill (up and back down), and then straight again?



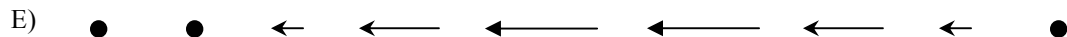
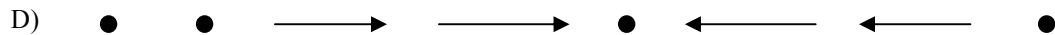
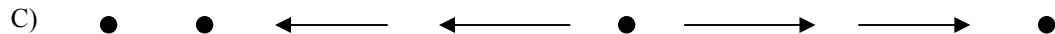
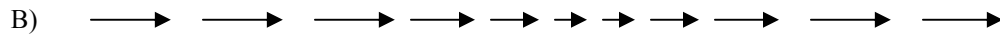
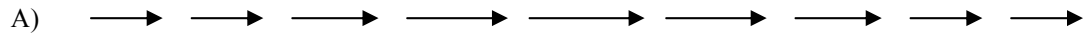
I give a steel ball a quick push along a frictionless track. Below are a few series of vectors showing the velocity of the ball in the x-direction at successive times (after the push). Which series would be correct if the track went straight, then over a hill (up and back down), and then straight again?



I give a steel ball a quick push along a frictionless track. The following are graphs of that ball's position in the  $y$  direction as a function of time (after the push). Which graph would be correct if the track went straight, then across a valley (down, and then back up), and then straight again?

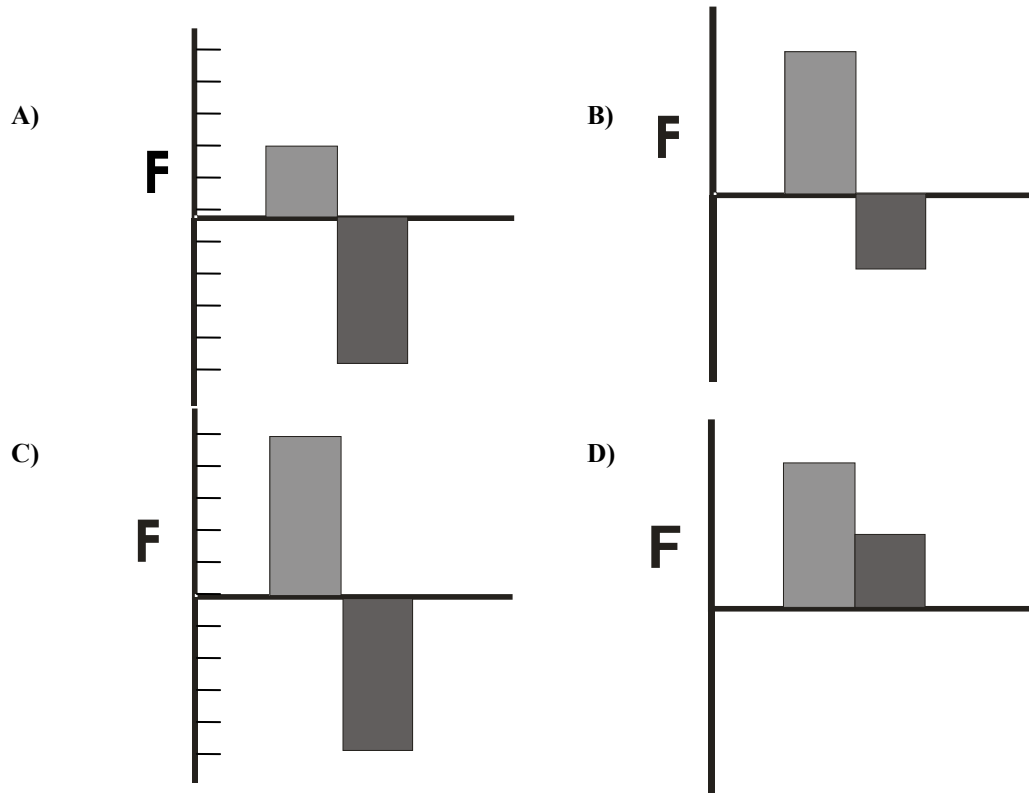


I give a steel ball a quick push along a frictionless track. Below are a few series of vectors showing the acceleration of the ball in the  $x$  (horizontal) direction at successive times (after the push). Which series would be correct if the track went straight, then over a hill (up and back down), and then straight again? A dot indicates a vector of length zero.



A large truck and a small car collide head-on, with the truck coming in from the left and the car from the right. Which of the following bar graphs best represents the magnitude and direction of the forces exerted by the truck on the car and by the car on the truck? Down on the graph is a force to the left, and up on the graph is a force to the right.

- Force on car by truck
- Force on truck by car

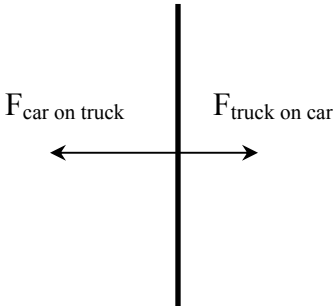
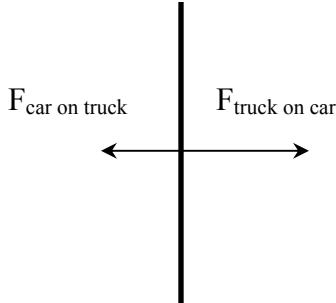
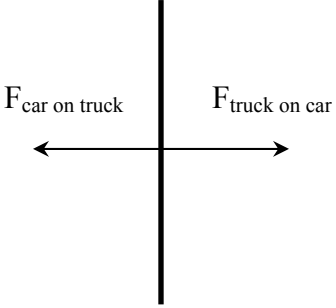
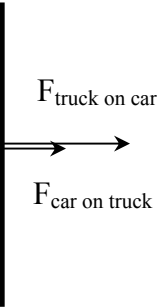


E) None of these accurately describe the collision.

A large truck and a small car collide head-on, with the truck coming in from the left and the car from the right. Which of the following best describes the magnitude and direction of the forces exerted by the truck on the car and by the car on the truck?

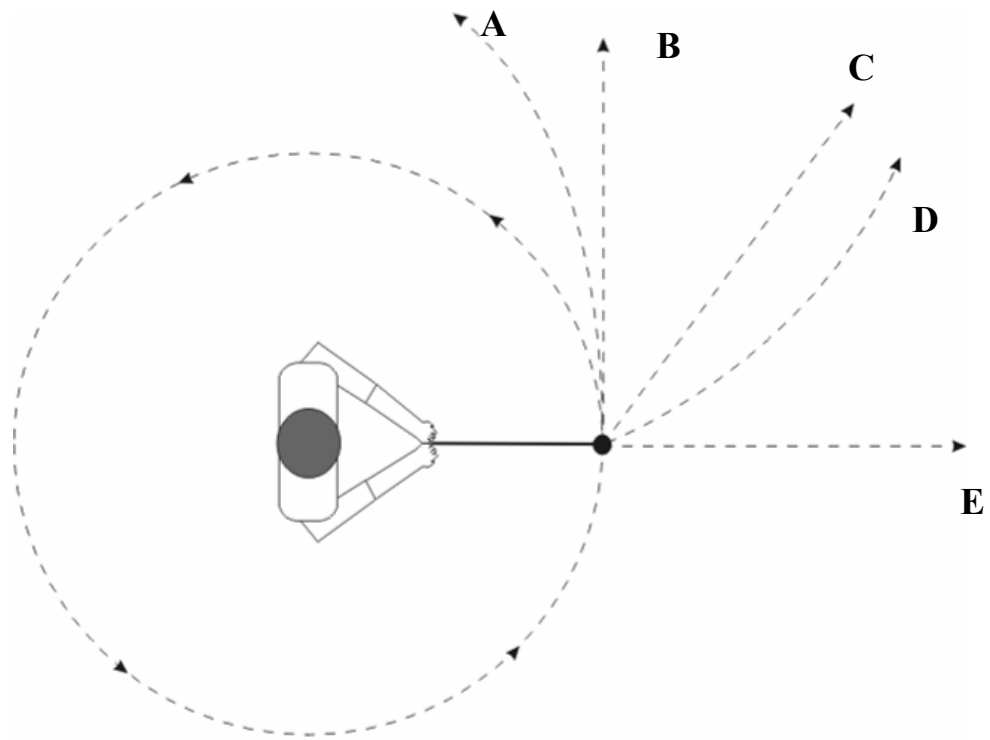
- A) The forces are in opposite directions, and the car exerts more force on the truck than the truck exerts on the car.
- B) The forces are in opposite directions, and the truck exerts more force on the car than the car exerts on the truck.
- C) The car and truck exert equal and opposite forces on each other.
- D) The forces are in the same direction, and the truck exerts more force on the car than the car exerts on the truck.
- E) None of these accurately describe the collision.

A large truck and a small car collide head-on, with the truck coming in from the left and the car from the right. Which of the following diagrams best describes the magnitude and direction of the forces exerted by the truck on the car and by the car on the truck?

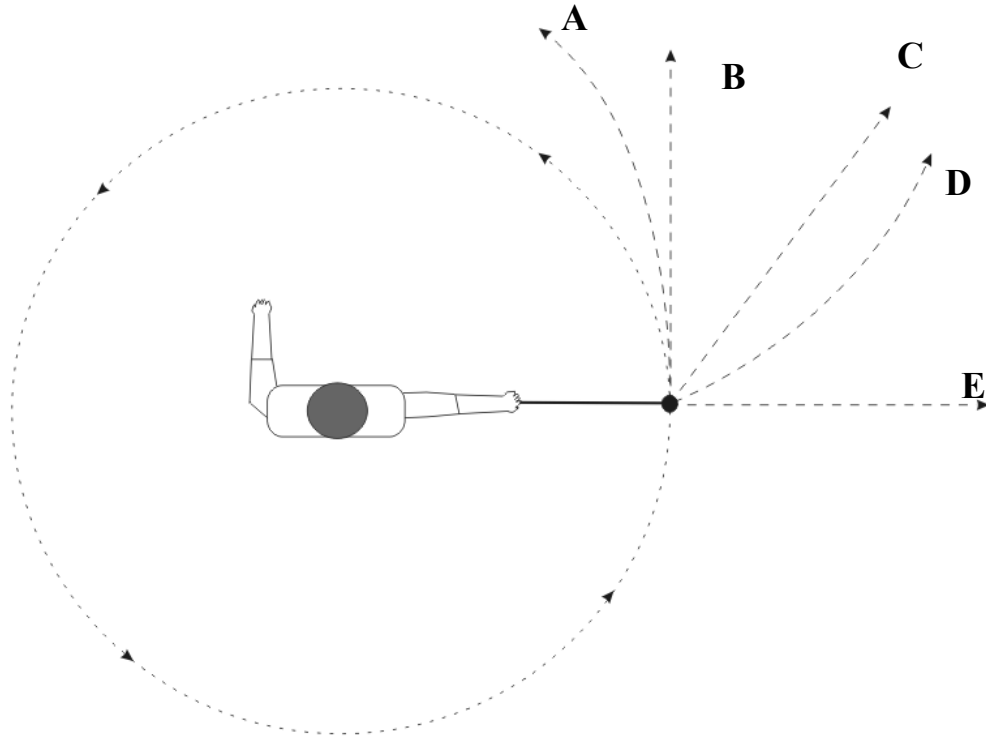
- A) 
- B) 
- C) 
- D) 
- E) None of these accurately describe the collision.



An athlete is swinging a heavy ball on a chain in a circle in the horizontal plane as shown from above. At the moment shown in the diagram, the ball is released. Choose the correct path showing where the ball goes after release (circle the appropriate letter).

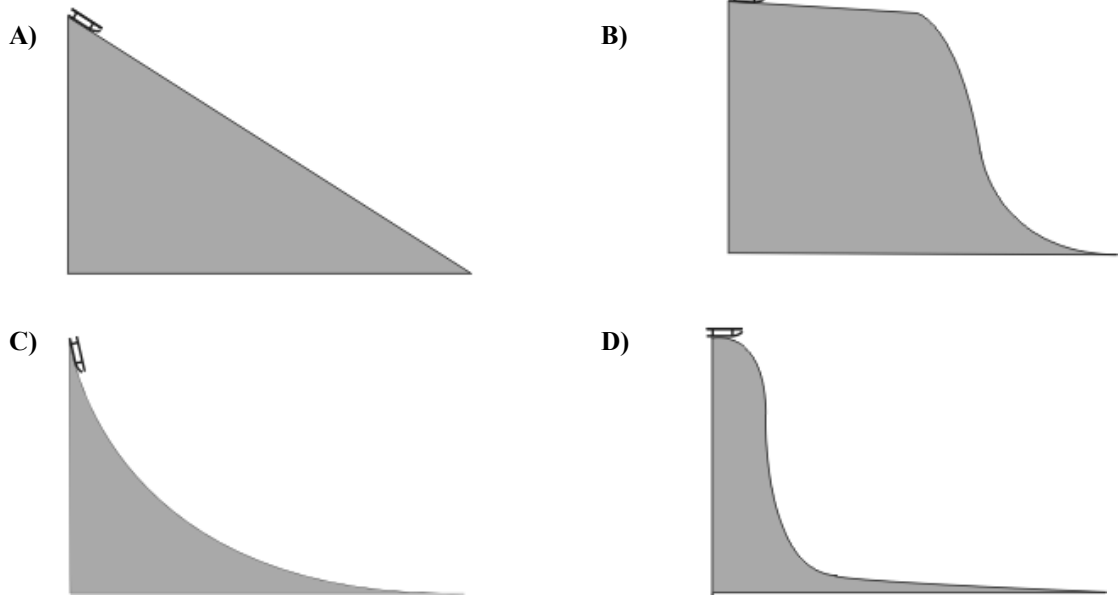


An athlete is swinging a heavy ball on a chain in a circle in the horizontal plane as shown from above. At the moment shown in the diagram, the ball is released. Choose the correct path showing where the ball goes after release (circle the appropriate letter).



In this problem, we're going sledding on **frictionless** snow. There are a total of four hills to sled down, **all of the same height**, shown below. Rank the four hills by how fast the sled will be going at the bottom.

For example, if you think the sled on hill A will be going faster than the one on hill B, which will be going as fast as the one on hill C, which will be going faster than the one on hill D, write  $A > B = C > D$



In this problem, we're going sledding on **frictionless** snow. There are a total of four hills to sled down, **all of the same height**, described below. Rank the four hills by how fast the sled will be going at the bottom.

For example, if you think the sled on hill A will be going faster than the one on hill B, which will be going as fast as the one on hill C, which will be going faster than the one on hill D, write  $A > B = C > D$

- A) Hill A has no curves. It goes down straight at about a 30 degree angle.
- B) Hill B starts with a long, almost flat section that's barely downhill, and then drops off very steeply until it reaches the bottom.
- C) Hill C starts steep and gradually becomes less steep until it is nearly level at the bottom.
- D) Hill D has an immediate and sharp dropoff that goes nearly to the bottom, with a long, very gentle slope after that.

## Appendix F: Multiple representations usage survey (chapter 11)

**1. I am usually good at learning physics on my own, without any help from others.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**2. I am usually good at solving physics problems on my own, without any help from others.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**3. I am good at finding and fixing my conceptual mistakes**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**4. I am good at finding and fixing my mathematical mistakes.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**5. I am either good at physics or bad at physics, and there's nothing I can do to change that.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**6. I feel motivated to learn physics.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**7. I feel motivated to learn in general.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**8. I often use multiple representations (drawing pictures, diagrams, graphs, etc) when solving physics problems.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**9. When I use multiple representations, I do so because it makes a problem easier to understand.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**10. When I use multiple representations, I do so because I will be more likely to get the right answer.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**11. When I use multiple representations, I do so because the instructor (or the book or the TA) tells me that I should.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**12. When I am drawing free body diagrams (or force diagrams) that include numbers and equations, I check to make sure that the diagram and the math match well.**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**13. I am good at representing information in multiple ways (words, equations, pictures, free body diagrams, etc.).**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**14. I am good at figuring out how closely related different representations are (words, equations, pictures, free body diagrams, etc.).**

Strongly Disagree      Disagree      Neutral      Agree      Strongly Agree

**15. On a scale of 1-5, rate how much each of the following factors affects your performance in physics class (5 being the highest):**

\_\_\_ Your Effort      \_\_\_ Your Ability      \_\_\_ Teacher/TAs      \_\_\_ Textbook

**16. On a scale of 1-5, rate how often you use the following (when applicable) in solving physics problems, and how comfortable you feel when doing so (5 being the highest):**

Free-body diagrams	_____ How often	_____ How comfortable
Equations and numbers	_____ How often	_____ How comfortable
Graphs	_____ How often	_____ How comfortable
Written explanations	_____ How often	_____ How comfortable

## Appendix G: Car motion representation choices (chapter 12)

You will be given four animations of moving cars. For each one, match the animation to the correct description of the motion, the correct position versus time graph, and the correct velocity versus time graph.

- A) The car accelerates in the  $+x$  direction, and then accelerates in the  $-x$  direction.
- B) The car is initially moving in the  $+x$  direction, stops suddenly, and then accelerates in the  $+x$  direction.
- C) The car starts at rest and then accelerates in the  $+x$  direction.
- D) The car starts at rest and then accelerates in the  $-x$  direction.
- E) The car shows zero acceleration during the movie.
- F) The car starts in motion and undergoes continuous acceleration in the  $+x$  direction during the movie
- G) The car starts in motion and undergoes continuous acceleration in the  $-x$  direction during the movie.

