ROLE OF MULTIPLE REPRESENTATIONS IN PHYSICS PROBLEM SOLVING

by

Alexandru Maries

B.S., Ramapo College of New Jersey, 2009

M.S., University of Pittsburgh, 2011

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

DEPARTMENT OF PHYSICS AND ASTRONOMY

This dissertation was presented

by

Alexandru Maries

It was defended on

December 3rd, 2013

and approved by

Dr. Robert Devaty, Associate Professor, Department of Physics and Astronomy

Dr. Arthur Kosowsky, Professor, Department of Physics and Astronomy

Dr. Russell Clark, Senior Lecturer, Department of Physics and Astronomy

Dr. Larry Shuman, Professor, Department of Industrial Engineering

Dissertation Advisor: Dr. Chandralekha Singh, Professor, Department of Physics and Astronomy

ROLE OF MULTIPLE REPRESENTATIONS IN PHYSICS PROBLEM SOLVING

Alexandru Maries, PhD

University of Pittsburgh 2013

ROLE OF MULTIPLE REPRESENTATIONS IN PHYSICS PROBLEM SOLVING

Alexandru Maries, PhD

University of Pittsburgh 2013

This thesis explores the role of multiple representations in introductory physics students' problem solving performance through several investigations. Representations can help students focus on the conceptual aspects of physics and play a major role in effective problem solving. Diagrammatic representations can play a particularly important role in the initial stages of conceptual analysis and planning of the problem solution. Findings suggest that students who draw productive diagrams are more successful problem solvers even if their approach is primarily mathematical. Furthermore, students provided with a diagram of the physical situation presented in a problem sometimes exhibited deteriorated performance. Think-aloud interviews suggest that this deteriorated performance is in part due to reduced conceptual planning time which caused students to jump to the implementation stage without fully understanding the problem and planning problem solution. Another study investigated two interventions aimed at improving introductory students' representational consistency between mathematical and graphical representations and revealed that excessive scaffolding can have a detrimental effect. The detrimental effect was partly due to increased cognitive load brought on by the additional

steps and instructions. Moreover, students who exhibited representational consistency also showed improved problem solving performance.

The final investigation is centered on a problem solving task designed to provide information about the pedagogical content knowledge (PCK) of graduate student teaching assistants (TAs). In particular, the TAs identified what they considered to be the most common difficulties of introductory physics students related to graphical representations of kinematics concepts as they occur in the Test of Understanding Graphs in Kinematics (TUG-K). As an extension, the Force Concept Inventory (FCI) was also used to assess this aspect of PCK related to knowledge of student difficulties of both physics instructors and TAs. We find that teaching an independent course and recent teaching experience do not correlate with improved PCK. In addition, the performance of American TAs, Chinese TAs and other foreign TAs in identifying common student difficulties both in the context of the TUG-K and in the context of the FCI is similar. Moreover, there were many common difficulties of introductory physics students that were not identified by many instructors and TAs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE AND ACKNOWLEDGEMENTS

First and foremost, I would like to thank my academic advisor, Dr. Chandralekha Singh for her continued support and guidance through my graduate career. I feel privileged to have had the opportunity to work closely with her and learned much about physics education research and cognitive science from her. I also learned from her much about teaching physics and how to interact with students in ways that can help them improve their attitudes towards physics.

I would like to thank Dr. Jeremy Levy for the numerous insightful comments and suggestions that aided some of my research, especially the TA and instructor study.

I would like to thank my committee members: Dr. Robert P. Devaty, Dr. Larry J. Shuman, Dr. Russell J. Clark and Dr. Arthur Kosowsky for their comments and critiques of my research. I would like to express additional gratitude to Dr. Robert P. Devaty for taking the time to read all my papers and provide innumerable fruitful comments which improved the quality of these papers.

I am also grateful to Emily Marshman and Seth DeVore for their feedback and helpful comments on my research.

I would also like to thank all the TAs, instructors and students who participated in my research. This thesis was made possible by their willingness to participate.

I would also like to thank my brother and all my friends for making my life easier during the drudgery of graduate school. I would like to extend additional thanks to my brother and also my good friend and fellow graduate student Louis Lello both of whom have repeatedly helped me practice my presentations and most likely found it annoying to hear the same presentation repeatedly.

1.0 INTRODUCTION

The goal of physics education has been described as transitioning students from an initial state to a desired final state (Reif 1995). Many instructional approaches (both traditional and based on Physics Education Research) either explicitly or implicitly attempt to improve the problem solving skills of introductory physics students [1-9]. This is because in order to learn physics concepts thoroughly, one must manipulate and work with these concepts in many different contexts and representations. Physics experts develop expertise through practice [10]; therefore, in a typical physics course, problem solving is the main modus through which students develop a functional understanding of physics principles.

1.1 PROBLEM AND PROBLEM SOLVING: DEFINITION

Before one can discuss problem solving it is necessary to have a definition of the word "problem". Some researchers have argued that the lack of clearly defining what one means by "problem" can lead to issues when interpreting physics education research on problem solving [11]. There are many definitions of what a problem is in the literature [12-16] which typically share features and do not necessarily provide discrepant descriptions. We adopt Newell's definition that "A person is confronted with a problem when he wants something and does not know immediately what series of actions he can perform to get it'' [14]. By this definition, a typical end-of-chapter problem from a typical introductory physics textbook [17] constitutes a problem for an introductory student, but does not for a physics expert, because experts have much compiled knowledge [18] of principles applicable in specific situations after long-term practice of solving problems, and therefore immediately know what steps must be taken in order to solve a typical end-of-chapter problem. However, when an end-of-chapter problem presented to an expert is non-intuitive, it can also be a problem for a physics expert [19]. Problem solving would then entail devising a strategy consisting of discrete steps which would provide the desired goal in a reasonable amount of time [20-22], which, in an introductory physics problem, is typically one or several physical quantities. As mentioned earlier, the understanding of how to solve problems is a central part of transitioning from an initial knowledge state to the desired final knowledge state.

1.2 INFLUENCES FROM COGNITIVE SCIENCE

Problem solving is a cognitive process; therefore, much research in cognitive science has been devoted to problem solving, and while the findings of cognitive science may not be directly related to physics problem solving or classroom instruction, these findings provide important instructional implications [23-24].

1.2.1 Memory

Problem solving is a process which takes place in Short Term Memory, also known as Working Memory, one of the two broad components of human memory, the other being Long Term

Memory (LTM) [25,26]. Working memory has a finite capacity to store information of roughly 7 "slots" [27] while the LTM does not appear to have any limits in the amount of information it can store. During the problem solving process, the working memory receives inputs from sensory buffers (eyes, ears, hands) and information from the LTM which needs to be distinguished from the vast amount of other information that is stored in LTM. Since the amount of information that can be processed at any given time in the working memory is finite, one must carefully process the particular information that makes the end goal easier to reach [14].

1.2.2 Chunking and Cognitive Load Theory.

Experts in a given domain can extend the limits of working memory by chunking more than one piece of information in one memory slot. A good example of chunking comes from chess [28]. Participants in a study were asked to reconstruct a chess board after viewing it briefly. The chess experts performed much better in this task than chess novices when the organization of the chess pieces on the board was from a good chess game; however, when the pieces were randomly placed, both the experts and novices exhibited similar performance in reconstructing the board. The reason argued by Chase et al. about why in the first situation the experts performed better than the novices is because their extended chess knowledge allowed them to "encode the position into larger perceptual chunks, each consisting of a familiar subconfiguration of pieces" [28]. For a chess expert, "pieces within a single chunk are bound by relations of mutual defense, proximity, attack over small distances and common color and type" [28]. Similarly, when engaged in problem solving, physics experts often group several pieces of information together into a single chunk which would take up one slot in working memory; however, for a physics novice (student) those pieces of information could seem disparate and require different slots in order to be processed. For example, while engaged in problem solving and processing information in working memory, an expert could group together information about a vector such as its magnitude, direction, *x* and *y* components into one single memory slot because of the relationships that connect them. In contrast, a novice could perceive these pieces of information as distinct and require one slot of working memory for processing each. Thus, the amount of information that a novice can process at any given time while engaged in problem solving is reduced compared to an expert, because experts, due to their compiled knowledge acquired through much problem solving experience, can chunk information into one single slot, whereas novices typically cannot. The amount of information that must be processed at any given time while engaged in problem solving in order to move forward with a solution is known as cognitive load, and due to their reduced information processing capabilities, introductory students can experience cognitive overload when solving problems (the amount of information that must be processed overloads the processing capacity of the working memory).

Sweller developed cognitive load theory in an effort to explain how people learn and extend their knowledge [29]. Cognitive load theory is based on a view that the knowledge structures stored in LTM are combinations of elements, otherwise known as schemas [30,31] which, although not known precisely, can be discerned through experimental research. According to Sweller [29] learning requires a change in the schemas stored in LTM because the main difference between experts and novices is that experts possess those schemas, while novices do not. As learners progress from novice to expert, their performance on problem solving tasks specific to the domain learned increases because the cognitive characteristics inherent in processing the material are altered so that the material can be processed more efficiently. One manner through which this occurs is chunking, which is supported by the research finding that as the expertise of an individual increases in a particular field, the cognitive load decreases [28-29]. Sweller therefore argues that, since information is first processed in working memory which has a finite processing capacity, in order for a learner to acquire the desired schemas, instructional strategies must be designed to reduce cognitive load. Many instructional strategies developed by physics education researchers, although not necessarily based on Sweller's cognitive load theory, are designed to reduce cognitive load [5,6,24,32-34].

1.3 CONNECTION WITH PHYSICS EDUCATION RESEARCH

1.3.1 Knowledge structure: Novices and Experts

The concept of schemas has been adopted by some physics education researchers [35] while others discuss an almost identical construct, knowledge structure [32,36], which describes how information about a particular domain is stored in LTM. The knowledge structure of physics experts is organized in terms of physics principles and is hierarchical with the most fundamental principles (which include applicability conditions) at the top (such as Newton's laws of motion, conservation of energy principles, etc.), and less fundamental principles further down the chain [20,37,38]. In addition, there are many connections that link related concepts together. In contrast, the knowledge structure of a physics novice (typical student) is comprised of facts and formulas that are only loosely connected. It is important to mention that "novice" and "expert" are two ends of a continuous spectrum, and that individual students in an introductory physics class can be somewhere in the middle [39].

1.3.2 Problem solving strategies: Novices and Experts

The manner in which experts and novices engage in problem solving is connected to the way their knowledge structure is organized [13]. Since the knowledge structure of experts is hierarchically organized in terms of physics principles, an expert's problem solving approach begins with a qualitative analysis of the problem (which can include drawing one or several diagrams to ensure the problem situation is well understood) and then a decision about which physics principles are applicable. Experts then make a plan and implement it, occasionally assessing their progress. After the goal is reached, they examine it to ensure it agrees with their physical expectation, and reflect on the problem to determine what can be learned from the problem solving process and how their knowledge structure can be extended. On the other hand, the problem solving strategies of novices differ markedly. Since the knowledge structure of novices consists of loosely connected facts and formulas, novices solve problems by focusing on pieces of information from the problem that look familiar. Then, they search for equations that match the disparate pieces and often do not ensure that the equations they found are applicable. Novices rarely spend time planning a solution and try to convert the verbal description of a problem to a mathematical description directly. Novices typically apply this "formula seeking" strategy in a superficial manner that does not require a thorough understanding of the physics principles involved.

1.4 THEORETICAL LEARNING FRAMEWORKS FROM COGNITIVE SCIENCE

Cognitive scientists have studied learning long before the establishment of the field of Physics Education Research and many of their findings provide important guidelines that the physics education researcher can use to design effective instructional strategies. It is therefore not a coincidence that many of the learning models developed by cognitive scientists are connected to the concept of knowledge structure outlined in the previous paragraphs. In particular, instructional strategies based on these learning models applied to the context of physics provide opportunities for learners to develop a good knowledge structure. The desired good knowledge structure is closer to that of an expert: organized hierarchically in terms of physics principles with the core physics principles at the top. In the following paragraphs, I will discuss four theoretical learning frameworks which informed much of my research presented in this thesis. Two of the frameworks were developed before the establishment of the field of physics education research, one was developed during its establishment and one was developed after the field of physics education research was well established. In particular, I will discuss how the application of these four learning frameworks in the context of physics improves the knowledge structure of the learner. The frameworks are Vygotsky's "zone of proximal development" model, Piaget's learning theory of assimilation, accommodation and optimal mismatch, the Cognitive Apprenticeship model and Schwartz, Bransford and Sears's framework of preparation for future learning.

1.4.1 Zone of proximal development

Vygotsky's theoretical framework of learning [40] is related to the concept of zone of proximal development (ZPD) which is defined as the difference between what a learner can achieve without support (initial knowledge state) and what they can achieve under the guidance of an expert or in collaboration with more capable peers. In Vygotsky's view, in order for meaningful learning to occur, one must stay within the ZPD of the students (which is itself dynamic). In the context of physics, this entails an understanding of the knowledge state of students, designing activities that students can actively engage in and, through scaffolding provided by an instructor or through collaboration with peers, the students can improve their knowledge state and by extension, their knowledge structure of physics principles. Through this repeated procedure, one can gradually move students from an initial knowledge state to a final desired knowledge state and, in the process, provide students with many activities designed to advance their knowledge structure to the level of (or close to) an expert's knowledge structure.

1.4.2 Assimilation, accommodation and optimal mismatch

Piaget's framework of learning involves the concepts of assimilation, accommodation and optimal mismatch [41]. This framework explains how new knowledge is internalized by learners as follows: if the new knowledge conforms to the pre-existing mental structures, it is assimilated in the learner's knowledge structure. If the new knowledge does not conform to the pre-existing mental structures, the pre-existing knowledge structure must be accommodated to incorporate the new knowledge. This latter mode of internalizing new knowledge is more common in the context of physics because students often begin the study of physics with conceptions that are not aligned with, and are frequently contrary to, the scientifically accepted way of reasoning about physics [42-44]. In Piaget's view, the second mode of internalizing new knowledge works best when the state of disequilibrium between the pre-existing knowledge structure and the new information to be assimilated is "optimal"; in other words, the gap between what is known and what must be learned is neither too great, nor too little, so that the learner is motivated to resolve the imbalance and can do it without finding the task too cognitively demanding (i.e., experience cognitive overload), which might lead to frustration and giving up and would result in little or no meaningful learning to occur. In the context of physics, instructional activities based on this notion of "optimal mismatch" can improve students' knowledge structure, because the state of disequilibrium between new knowledge to be learned and the existing knowledge structure can motivate students to modify their knowledge structure if the new knowledge is not outside of their learning capabilities (through expert scaffolding, guided activities, collaboration with peers, etc.). If the tasks are carefully chosen so as to promote conceptual thinking in students, their knowledge structure would be gradually improved and made to resemble the hierarchical knowledge structure of experts.

1.4.3 Cognitive apprenticeship model

The Cognitive Apprenticeship model [45] is based on the constructivist model of learning, in which knowledge is constructed rather than transmitted. In other words, an instructor cannot simply pour knowledge into students' brains by lecturing information; instead, students construct their own knowledge by making sense of the material for themselves. In the cognitive apprenticeship model, an expert first models a task for a student, then repeatedly coaches and guides the student while he/she attempts to follow the model and gradually reduces the support (also known as “fading”) until the student achieves independence. In the context of physics, if one desires to improve the knowledge structure of students so that it is closer to an expert’s (i.e., hierarchical and organized in terms of physics principles), during the coaching phase of problem solving for example, the attention of the student can be directed to processing information in the desired manner: start by conceptually analyzing the problem and decide which physics principles are applicable, devise and then implement a plan. This would have as a result the improvement of the knowledge structure of students because their approach to thinking about physics would be gradually shifted away from formula centered towards being centered on concepts and principles, along with their applicability conditions – similar to the approach of experts. As the students gradually achieve independence in this method, they can begin to practice expert-like problem-solving behavior on their own and require less and less support to improve their knowledge structure, thus gradually becoming experts. One example of the application of the Cognitive Apprenticeship model in physics and how it improves the knowledge structure of students is the Hierarchical Analysis Tool (HAT) of Mestre and Touger [24].

1.4.4 Preparation for future learning

While researching transfer, Schwartz et al. [46] introduced a two dimensional learning and performance space which they used to propose an optimal learning trajectory. The two dimensions they discuss are efficiency and innovation, which they argue play a significant role advancing students’ learning. Efficiency “includes a high degree of consistency that maximizes success and minimizes failure.” [46]. In other words, in the context of problem solving, efficiency-oriented practice does not require in-depth understanding, but rather rote memorization of procedures to solve problems. Focus on efficiency can yield “routine experts” who are good at solving only certain types of problems and do not know how to transfer this knowledge to new contexts [47]. Innovation on the other hand, requires rearranging one's thinking to handle new types of problems or information. In the context of physics this means dealing with complex problems (such as the context-rich problems of the University of Washington group) which require adaptive application of knowledge of physics (rearranging one's thinking) in new, unfamiliar contexts (unfamiliar in the sense of different from the typical abstract and non-specific contexts of typical textbook problems, e.g., an object sliding down an inclined plane). Tasks that focus on innovation are typically far beyond students' prior knowledge and can lead to frustration and little learning. Since Schwartz et al. consider both dimensions as important for preparation for future learning, they argue that instruction should go in a diagonal direction (a direction that includes both) in this 2D plane defined by perpendicular axes of efficiency and transfer. In the context of physics, teaching along both the efficiency and innovation directions advances students towards physics experts because being able to efficiently carry out procedures is required in order to free up memory slots when processing information. Increasing the information processing capability of students would in turn make it more likely that they adopt a more global view of problem solving which begins with a conceptual analysis and development of a plan. In addition, it would make it more likely that they can assess their progress while implementing a problem solving approach.

1.5 REPRESENTATIONS AND PROBLEM SOLVING

All the learning frameworks presented in section 1.4, when applied to the context of physics are connected in at least one aspect: problem solving tasks are an essential part of any instructional strategy whose purpose is to improve students' knowledge structure of physics and align it with the way physics is represented in the minds of experts. Representations play a major role both in knowledge structure and in problem solving. The concepts of physics, although abstract, are understood by experts in some form or representation [48], and therefore their knowledge structure consists of many representations that are directly connected to physics concepts. This implies that in order to improve the knowledge structure of students, they must be guided to represent physical concepts in different and complementary ways. Problem solving and use of multiple representations are connected as many studies of physics problem solving revealed that students who are consistent across different representations perform better on problem solving tasks [49-54]. When students are taught problem solving strategies that emphasize use of different representations of knowledge, they construct higher quality and more complete representations and exhibit better performance than students who are taught traditional problem solving approaches [8] akin to those of typical college textbooks, e.g. Halliday and Resnick [17]. Furthermore, teaching students to represent problems in different ways has a significant influence in deterring them from following novice-like formula centered problem solving approaches [24]. In addition, experts employ multiple representations in their initial conceptual analysis stage of problem solving, thus, if the instructional goal of an introductory physics course is the transition of students from novices to experts, the instructional design must place significant emphasis on multiple representations.

There are many reasons why using multiple representations in problem solving is conducive to an improved knowledge structure of physics concepts. Problem solving is the principal process through which students develop their understanding of physics and physics concepts must be represented in some form in LTM. It therefore stands to reason that using multiple representations in problem solving will lead to physics concepts being understood better by students because in the process of using multiple representations, students can learn meaningful and appropriate ways of representing these concepts. Furthermore, representations can reduce students' cognitive load by providing an external rather than an internal representation of physical information, a process known as "distributed cognition" [55]. For example, in one study presented in this thesis, students were given a problem which required addition of two vectors (non collinear). Students who explicitly drew the components of the vectors performed better than students who did not and interviews suggested that this may be a result of distributed cognition. Students who did not use an external representation, or ones who did not explicitly include the components of the vectors performed worse than students who did in part because students who did not represent the components externally had to keep more information in their working memory while engaged in problem solving and this may have increased their cognitive load. In addition, the process of drawing a diagram can provide a thorough understanding of the physical situation presented and greatly assist during the key stage of conceptual planning. This can help students focus attention on relevant concepts and increase the worth of their qualitative analyses. More attention devoted to the qualitative aspects of physics while engaged in problem solving can gradually modify the novice perspective of physics as a collection of facts and formulas, or what Hammer describes as "knowledge in pieces" [56] to a view of physics as a hierarchical construction of concepts that are interconnected, or what Hammer describes as viewing physics as "coherent". The knowledge structure of students would therefore become more closely aligned with that of experts, i.e., hierarchical and organized in terms of physics principles.

1.6 TA AND INSTRUCTOR KNOWLEDGE OF STUDENT DIFFICULTIES RELATED TO REPRESENTATIONS OF CONCEPTS

Since instruction which endeavors to improve the knowledge structure of students is greatly aided by emphasis on multiple representations of concepts, instructors and TAs should be aware of common difficulties that students encounter while learning to use various representations. Awareness of student difficulties is part of what Shulman defines as Pedagogical Content Knowledge (PCK) [57-58]. In addition, in Shulman's view, PCK includes "a veritable armamentarium of alternative forms of representation" [57]. Therefore, educators should not only be aware of student difficulties with interpreting various representations of concepts, but also possess many alternative ways to represent physics concepts in order to help students develop mental models of the concepts that span different representations.

TAs in particular can play a major role in teaching students multiple representations because, typically, they interact much more closely with students than instructors, especially at large universities with very high enrollments of undergraduate students in introductory physics courses. TA duties typically include teaching relatively small (compared to class sizes) recitations sections, grading assessments such as homework, quizzes and in some cases, exams, and holding regular office hours with students. If TAs are aware of student difficulties, they can guide the numerous interactions with students to address their difficulties and improve their knowledge structure of physics. In recitations, TAs can discuss different representations of concepts and pay particular attention to aspects found to be difficult by students; in grading homework, quizzes and exams, they can provide valuable feedback to students; and in interacting with students in office hours (typically with one or several students) they can pay individual attention to each student's difficulties.

In order to assess one aspect of the PCK of teaching assistants, in particular, their knowledge of student difficulties with different representations of concepts, a problem solving task for the teaching assistants was designed in the context of the Test of Understanding Graphs in Kinematics, or TUG-K [59]. The task (described in chapter five of this thesis) was for first-year teaching assistants to identify introductory students' most common incorrect answer choices for each item on the TUG-K. This study revealed that the teaching assistants were unaware of many representational difficulties of introductory students.

This research was extended in the context of the Force Concept Inventory, or FCI (the revised version, see [60] also printed in [61]) to investigate this aspect of the PCK of both teaching assistants and instructors because many of the items on the FCI assess student understanding of physics concepts posed in certain representations (e.g., Newton's second law in pictorial and verbal representations, concepts of velocity and acceleration in diagrammatic representation, etc.) This study is the last of this thesis and it discusses many results including but not limited to: experience teaching an independent course does not improve the ability to identify student difficulties, American teaching assistants perform no better than foreign teaching assistants and many student difficulties are not identified by both instructors and teaching assistants.

1.7 A STUDY OF THE ROLE OF REPRESENTATIONS IN PROBLEM SOLVING

Several studies in this thesis explore the role of diagrammatic representations in problem solving performance. The first study examined a problem which could be solved by employing a diagrammatic approach almost exclusively, and revealed that the diagrammatic approach is adopted by most physics experts and therefore it is the expert-like approach. Students who drew diagrams performed better than students who did not draw diagrams even if their chosen approach was primarily mathematical. In addition, mathematical difficulties of algebra-based students related to solving a system of equations with two unknowns were explored in detail and found to be in part due to a lack of transfer of mathematical knowledge to the context of physics which may be explicated by employing the framework of cognitive load theory.

In the studies discussed in chapters two and three, students were prompted to draw diagrams via explicit instructions, given a basic diagram (similar to an initial expert sketch) or neither (comparison group) in quiz problems. Analysis of the results revealed that in certain problems, providing a basic diagram, although intended as scaffolding support, resulted in deteriorated performance while in the rest of the problems, it did not result in improved performance. Interviews carried out using a think-aloud protocol [62,63] revealed that the cause of this deteriorated performance could partly be attributed to reduced conceptual planning time. In addition, in most problems, the students prompted to draw a diagram drew more productive diagrams (as defined from the point of view of an expert) and regardless of the instructions received, students with more detailed diagrams exhibited better problem solving performance. These results indicate that explicit instruction to draw a diagram can lead to improved problem solving performance and by extension a better knowledge structure of physics and therefore this instruction should be incorporated in problem solving tasks given to students.

The study discussed in chapter four investigated calculus-based students' ability to translate between graphical and mathematical representations a problem solution involving the electric field for spherical charge symmetry. Interventions were implemented to help scaffold students' representational consistency. Evaluation of the interventions revealed that a lot of scaffolding, which was considered by experts to likely improve the consistency of students, had the opposite effect. In addition, significant student difficulties related to translating between mathematical and graphical representations were encountered and explored in depth via think aloud interviews. A lack of transfer of mathematical knowledge in the context of physics was determined to partly account for some of the difficulties. Cognitive load theory was found to be useful in providing a learning framework that could account both for the detrimental effect of increased scaffolding and for the representational difficulties of students. Finally, representational consistency was found to correlate with performance, thus confirming earlier research [49].

The two studies discussed in the last two chapters of this thesis explored how knowledgeable TAs and instructors are about common student representational and conceptual difficulties as described in section 1.6.

1.8 CHAPTER REFERENCES

1. F. Reif (1995). Understanding basic mechanics. New York, Wiley.

2. J. Mestre, R. Dufresne, W. Gerace, P. Hardiman and J. Touger (1993). “Promoting skilled problem solving behavior among beginning physics students.” Journal of Research in Science Teaching 30, 303-317.

3. J. Larkin and F. Reif (1979). “Understanding and teaching problem solving in physics.” Eur. J. Sci. Ed. 1(2), 191-203.

4. F. Mateycik, D. Jonassen and N. S Rebello (2009). “Using similarity rating tasks to assess case reuse in problem solving.” AIP Conf. Proc. 1179, 201-204.

5. A. H. Shoenfeld (1980) “Teaching problem solving skills.” Am. Math. Mon. 87, 794-805.

6. A. Van Heuvelen (1991). “Overview, Case Study Physics.” Am. J. Phys. 59(10), 898-907.

7. A. Van Heuvelen (1991). “Learning to think like a physicist: A review of research-based instructional strategies.” Am. J. Phys. 59(10), 891-897.

8. D. Huffman (1997), “Effect of explicit problem solving strategies on high school students’ problem-solving performance and conceptual understanding of physics.” J. Res. Sci. Teach. 34(6), 551-570.

9. L. C. McDermott and P. S. Schaffer (1998). Tutorials in Introductory Physics, Upper Saddle River, NJ, Prentice-Hall.

10. J. Larkin (1981). “Cognition of learning physics.” Am. J. Phys. 49(6), 534-541.

11. D. P. Maloney (2011). “An Overview of Physics Education Research on Problem Solving.” *Getting Started in PER.* Reviews in PER vol. 2. College Park, MD: American Association of Physics Teachers.

12. J. R. Hayes (1981). The Complete Problem Solver, Philadelphia, PA, Franklin Institute Press.

13. F. Reif (1995). “Millikan lecture 1994: Understanding and teaching important scientific thought processes.” Am. J. Phys. 63(1), 17-32.

14. A. Newell and H. A. Simon (1972). Human Problem Solving, Englewood Cliffs, NJ, Prentice Hall.

15. G. Polya (1962). Mathematical Discovery. New York, Wiley.

16. J. R. Anderson (2000). Problem solving. Cognitive Psychology and its Implications. J. R. Anderson, New York, Worth: 239–278.

17. D. Halliday, R. Resnick and J. Walker (2007). Fundamentals of Physics (8^{th} ed.) Wiley.

18. E. F. Redish (2004). “A theoretical framework for physics education research: Modeling student thinking. Research on Physics Education.” Proceedings of the International School of Physics, “Enrico Fermi,” Course CLVI. E. F. Redish and M. Vicentini, Varenna, Italy, IOS Press: 1-65.

19. C. Singh (2002). “When physical intuition fails.” Am. J. Phys. 70, 1103-1109.

20. B. Eylon and F. Reif (1984). “Effects of knowledge organization on task performance.” Cognition Instruct. 1(1), 5-44.

21. J. Heller and F. Reif (1984). “Prescribing effective human problem-solving processes: Problem description in physics.” Cognition Instruct. 1(2), 177-216.

22. F. Reif and J. Larkin (1991). “Cognition in scientific and everyday domains: Comparison and learning implications.” J. Res. Sci. Teach. 28(9), 733-760.

23. E. F. Redish (1994). “The implications of cognitive studies for teaching physics.” Am. J. Phys. 69(2), 796-803.

24. J. Mestre and J. Touger (1989). “Cognitive research – What’s in it for physics teachers?” Phys. Teach. 27, 447-456.

25. J. R. Anderson (1995). Learning and Memory. New York, Wiley.

26. H. Simon (1974). “How big is a memory chunk?” Science 183(4124), 482-488.

27. G. Miller (1956). “The magical number seven, plus or minus two: Some limits on our capacity for processing information.” Psychol. Rev. 63(2), 81-97.

28. W. Chase W. and H. Simon (1973). “Perception in chess”, Cog. Psy. 4, 55-81.

29. J. Sweller (1988). “Cognitive load during problem solving: effects on learning.” Cog. Sci. 12(2), 257-285.

30. P. W. Cheng and K. J. Holyoak (1985). “Pragmatic reasoning schema.” Cog. Psych. 17, 391-416.

31. J. Sweller, J. V. Merriënboer and F. Paas (1998). “Cognitive architecture and instructional design.” Educ. Psych. Rev. 10(3), 251-296.

32. J. Heller and F. Reif (1982). “Knowledge structure and problem solving in physics.” Educ. Psych. 17(2), 102-127.

33. F. Reif and L. A. Scott (1999). “Teaching scientific thinking skills: Students and computers coaching each other.” Am. J. Phys. 67(9) 819-831.

34. W. R. Gerace, R. Dufresne, W. Leonard and J. P. Mestre (2000). “Minds-on Physics: Materials for Developing Concept-based Problem-solving Skills in Physics.” PERG 8. http://www.srri.umass.edu/publications/gerace-1999mdc

35. R. J. Dufresne, W. J. Gerace, P. T. Hardiman and J. P. Mestre (1992). “Constraining novices to perform expertlike problem analyses: Effects on schema acquisition.” J. Res. Learn. Sci. 2(3), 307-331.

36. I. D. Beatty and W. J. Gerace (2002). “Probing physics students’ conceptual knowledge structures through term association.” Am. J. Phys. 70(7), 750-758.

37. M. T. H. Chi, P. J. Feltovich and R. Glaser (1981). “Categorization and representation of physics knowledge by experts and novices.” Cog. Sci. 5, 121-151.

38. A. Schoenfeld and D. J. Herrmann (1982). “Problem perception and knowledge structure in expert novice mathematical problem solvers.” J. Exp. Psych.: Learning Memory and Cognition 8, 484-494.

39. A. Mason and C. Singh (2011). “Assessing expertise in introductory physics using categorization task.” Phys. Rev. ST. Phys. Educ. Res. 7, 020110.

40. L. S. Vygotsky (1978). Mind in society: The development of higher psychological processes. Cambridge, MA, Harvard University Press.

41. H. Ginsberg and S. Opper (1969). Piaget’s theory of intellectual development. Englewood Cliffs, NJ, Prentice Hall.

42. R. Hake (1998). “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses.” Am. J. Phys. 66, 64-74.

43. N. Lasry, E. Mazur and J. Watkins (2008). “Peer Instruction: From Harvard to the two-year college.” Am. J. Phys. 76, 1066-1069.

44. L. C. McDermott, M. L. Rosenquist and E. H. van Zee (1987). “Student Difficulties in Connecting Graphs and Physics: Examples from Kinematics” Am. J. of Phys. 55, 503-513.

45. A. Collins, J. S. Brown and S. E. Newman (1989). Cognitive Apprenticeship: Teaching the crafts of reading, writing and apprenticeship. Knowing, Learning and Instruction: Essays in Honor of Robert Glaser. R. Glaser and L. Resnick. Hillsdale, NJ, Lawrence Erlbaum Associates, 453-494.

46. D. Schwartz, J. Bransford and D. Sears (2005). Efficiency and innovation in transfer. Transfer of Learning: Research and Perspectives. J. Mestre. Greenwhich, CT, Information Age Publishing.

47. G. Hatano and Y. Oura (2003). “Commentary: Reconceptualizing school learning using insight from expertise research.” Educ. Res. 32(8), 26-29.

48. D. Meltzer (2005). “Relation between students’ problem solving performance and representational mode.” Am. J. Phys. 73(5), 463-478.

49. P. Nieminen, A. Savinainen and J. Viiri (2012). “Relations between representational consistency, conceptual understanding of the force concept, and scientific reasoning.” Phys. Rev. ST Physics Educ. Res. 8, 010123.

50. A. Van Heuvelen and X. Zou (2001). “Multiple representations of work-energy processes.” Am. J. Phys. 69(2), 184-194.

51. X. Zou (2001). “The role of work-energy bar charts as a physical representation in problem solving.” Proceedings of the 2001 Physics Education Research Conference. S. Franklin, J. Marx and K. Cummings, Rochester, NY, PERC Publishing: 135-138.

52. M. W. van Someren, P. Reimann, H. P. A. Boshuizen and T. de Jong (1998). Learning with Multiple Representations. New York, NY, Elsevier Science, Inc.

53. R. Plötzner (1994). The Integrative Use of Qualitative and Quantitative Knowledge in Physics Problem Solving. Frankfurt am Main, Peter Lang Publishing, Inc.

54. M. Reiner (1990). Conceptual change through problem solving. Paper presented at the AAPT meeting, Minneapolis 1990 (June).

55. J. Zhang (2006). Distributed cognition, representation and affordance, Prag. Cogn. 14(2), 333-341.

56. D. Hammer (1994). “Epistemological beliefs in introductory physics.” Cogn. Instruct. 12(2), 151-183.

57. L. S. Shulman (1986). “Those who understand: Knowledge growth in teaching.” Educ. Res. 15(2), 4-31.

58. L. S. Shulman (1987). "Knowledge and teaching: Foundations of the new reform." Harv. Educ. Rev. 57(1), 1-22.

59. R. Beichner (1994). "Testing student interpretation of kinematics graphs." Am. J. Phys. 62(8), 750-762.

60. I. Halloun, R.R. Hake, E.P. Mosca and D. Hestenes (1995). "Force Concept Inventory." (Revised): online (password protected) at http://modeling.la.asu.edu/R&E/Research.html.

61. E. Mazur (1997). Peer Instruction: A User's Manual. Engelwood Cliffs, Prentice-Hall.

62. K. Ericsson and H. Simon (1980). "Verbal reports as data." Psychol. Rev. 87(3), 215-251 (1980).

63. K. Ericsson and H. Simon (1993). Protocol Analysis: Verbal Reports as Data. Boston, MA, MIT Press.

2.0 A GOOD DIAGRAM IS VALUABLE DESPITE THE CHOICE OF A MATHEMATICAL APPROACH TO PROBLEM SOLVING

2.1 INTRODUCTION

Introductory physics is a challenging subject to learn partly because students rarely associate the abstract concepts they study in physics with more concrete representations that facilitate understanding without an explicit instructional strategy aimed to aid them. They often tend to have formula oriented problem solving strategies that do not require understanding of physical concepts. Unfortunately, these inferior strategies are rewarded in a traditional introductory physics course [1,2].

There are many reasons to believe that multiple representations of concepts along with the ability to construct, interpret and transform between different representations that correspond to the same physical system or process play a positive role in learning physics. First, physics experts often use multiple representations as a first step in a problem solving process [1,3-8]. Second, students who are taught explicit problem solving strategies emphasizing use of different representations of knowledge at various stages of problem solving construct higher quality and more complete representations and perform better than students who learn traditional problem solving strategies [9]. Third, multiple representations are very useful in translating the initial, mostly verbal description of a problem into a representation more suitable to mathematical manipulation [10-13] because the process of constructing a representation of a problem makes it easier to generate appropriate decisions about the solution process. Also, getting students to represent a problem in different ways helps shift their focus from merely manipulating equations toward understanding physics [14]. Some researchers have argued that in order to understand a physical concept thoroughly, one needs to be able to recognize and manipulate the concept in a variety of representations [11,15]. As Meltzer puts it [16], a range of diverse representations is required to "span" the *conceptual* space associated with an idea. Since traditional courses which don't emphasize multiple representations lead to low gains on the Force Concept Inventory [17,18] and on other assessments in the domain of electricity and magnetism [19], in order to improve students' understanding of physics concepts, many researchers have developed instructional strategies that place explicit emphasis on multiple representations [3,10,11,20-26] while other researchers developed strategies with implicit focus on multiple representations [27-35]. Van Heuvelen's approach, for example, [10,11] starts by ensuring that students explore the qualitative nature of concepts by using a variety of representations of a concept in a familiar setting before adding the complexities of mathematics. Many other researchers have stressed the importance of students becoming facile in translating between different representations of concepts [20,36-42] and that significant positive learning occurs when students develop facility in the use of multiple forms of representation [43,44]. However, careful attention must be paid to instructional use of diverse representational modes as specific learning difficulties may arise as a consequence [16] because students can approach the same problem posed in different representation differently without support [16,45].

One representation useful in the initial conceptual analysis and planning stage of a solution is a schematic diagram of the physical situation presented in the problem. Diagrammatic representations have been shown to be superior to exclusively employing verbal representations when solving problems [4,6-8]. It is therefore not surprising that physics experts automatically employ diagrams in attempting to solve problems [15,46,47]. However, introductory physics students need explicit help understanding that drawing a diagram is an important step in organizing and simplifying the given information into a representation which is more suitable to further analysis [48]. Therefore, many researchers who have developed strategies for teaching students effective problem solving skills attempt to make students realize how important the step of drawing a diagram is in solving a physics problem. In Newtonian mechanics, Reif [1,3] has suggested that several diagrams be drawn: one diagram of the problem description which includes all objects and one diagram for each system that needs to be considered separately. Also, he detailed concrete steps that students need to take in order to draw these diagrams: (a) describe both motions and interactions, (b) identify interacting objects before forces, (c) separate long range and contact interactions and (d) label contact points by the magnitude of the action-reaction pair of forces. Van Heuvelen's Active Learning Problem Sheets (ALPS) [10,11] adapted from Reif follow a very similar underlying approach. Other researchers who have emphasized, among other things, the importance of diagrams in their approach to teaching students problem solving skills have found significant improvements in students' problem solving methods [4,10,49]. In mathematics, Shoenfeld [50,51] advocates drawing a diagram (if possible) as the first step.

Previous research shows that students who draw diagrams even if they are not rewarded for it are more successful problem solvers [52]. An investigation into how spontaneous drawing of free body diagrams (FBDs) [53] affects problem solving [54,55] shows that only drawing correct FBDs improves a student's score and that students who draw incorrect FBDs do not perform better than students who draw no diagrams. Heckler [56] investigated the effects of prompting students to draw FBDs in introductory mechanics by explicitly asking students to draw clearly labeled FBDs. He found that students who were prompted to draw FBDs were more likely to follow formally taught problem solving methods rather than intuitive methods (i.e., thinking about the problem conceptually) which caused deteriorated performance.

The research presented in this study is closely related to student understanding of the concept of mechanical waves as it relates to harmonics of standing waves in cylindrical tubes. Conceptions of mechanical waves have been researched in young children [57,58], middle school and high-school students [59-61], introductory undergraduate students [62,63] and advanced undergraduate students [64]. Eshach and Schwartz [59] investigated whether Reiner's [65] earlier finding that the initial knowledge that students bring to the study of science is often "substance based" (which Reiner termed "substance schema") also holds for mechanical waves. They found that students do hold this view in some respects. However, sometimes students perceive the "substance" that they associate with sound waves differently from other "regular" substances. Wittmann [66] reported similar findings, namely that students often use reasoning that is focused on object-like properties when discussing waves which can be problematic to the goal of shifting student understanding of mechanical waves from a substance-based ontology to a sequence of events. In addition, Wittmann and Hrepic [67,68] were interested in identifying students' mental models of mechanical waves and how knowledge of these mental models can be used to improve students' understanding of mechanical waves. One tool for identifying these mental models is Wittmann's Wave Diagnostic Test [69], an open-ended questionnaire. Tonghchai et al., on the other hand, developed a multiple choice assessment tool for mechanical waves [70] and used it [71] to evaluate the consistency of students' conceptions. They found that students solve problems involving mechanical waves across different contexts inconsistently, much like what other researchers have found for other areas of physics [72-76]. Student understanding of other types of waves such as light waves [77] and electromagnetic waves [78] has also been researched. However, the use of multiple representations and its role to understanding mechanical waves has not been researched in much depth. Among the few who have investigated the role of non-verbal representations to understanding sound waves, Eshach and Schwartz [59] found that students have a variety of non-verbal representations that they employ while explaining their understanding of sound waves. During the interviews they conducted with high-school students, they allowed them to draw or gesticulate to explicate their reasoning. They concluded this research by suggesting that these non-verbal representations students use could be a good starting point to help them construct the correct visual representations needed to fully understand wave propagation phenomena.

The study presented here is concerned with the use of diagrammatic representations in the context of problem solving related to standing waves and the extent to which a diagram can improve student performance on problems related to standing waves in cylindrical tubes. More specifically, we investigated how algebra-based introductory physics students' performance on a problem (given in a quiz) related to standing waves in a tube is affected when students are given a diagram as opposed to when they are asked to draw a diagram (without being more specific than that). The performance of these students was also contrasted with that of a comparison group which was not given any instructions related to diagrams when solving the same problem related to standing waves. Moreover, a second similar problem was given in a midterm exam for which all introductory physics students received the same instructions regarding diagrams. We found that students who were explicitly asked to draw diagrams drew more productive diagrams than students in the other two groups and that both in the quiz problem and in the midterm problem students who used a mathematical approach, but also drew productive diagrams, performed better than students who used a mathematical approach without drawing a productive diagram. In addition, we found that many students employing the mathematical approach had difficulties manipulating two equations symbolically in the context of solving the problem involving standing waves in a tube. In order to investigate these findings in depth, we conducted think-aloud interviews with eight students enrolled in another algebra-based introductory physics course. The interviews were helpful in furnishing or corroborating possible interpretations of the quantitative results.

2.2 METHODOLOGY

A class of 118 introductory physics students in an algebra-based course was broken up into three different recitations. All recitations were taught in the traditional way in which the teaching assistant (TA) worked out problems similar to the homework problems and then gave a 15-20 minute quiz at the end of the recitation. Students in all recitations attended the same lectures, were assigned the same homework, and had the same exams and quizzes. In the recitation quizzes throughout the semester, the three groups were given the same problems but with the following interventions: in each quiz problem, the first intervention group, which we refer to as the “prompt only group” or “PO”, was given an explicit prompt or instruction to draw a diagram along with the problem statement; the second intervention group (referred to as the “diagram only group” or “DO”) was given a diagram drawn by the instructor that was meant to aid in solving the problem and the third group was the comparison group which was not given any diagram or an explicit instruction to draw a diagram with the problem statement ("no support group" or "NS").

The sizes of the different recitation groups varied from 22 to 55 students because the students were not assigned a particular recitation, they could go to whichever recitation they wanted. For the same reason, the sizes of each recitation group also varied from week to week, although not as drastically because most students (≈ 80%) would stick with a particular recitation. Furthermore, each intervention was not matched to a particular recitation. For example, in one week, NS was the Tuesday recitation while another week, NS was a different recitation section. This is important because it implies that individual students were subjected to different interventions from week to week and we do not expect cumulative effects due to the same group of students always being subjected to the same intervention.

In order to ensure homogeneity of grading, rubrics were developed for each problem analyzed and the rubrics were used to ensure that there was at least 90% inter-rater-reliability between two different raters. The development of the rubric for each problem went through an iterative process. During the development of the rubric, the two raters discussed students' scores separately from the ones obtained using the preliminary version of the rubric and adjusted the rubric if it was agreed that the version of the rubric was too stringent or too generous. After each adjustment to the rubric, all the students' scores were computed again using the improved rubric.

Here, we discuss two similar problems involving standing waves in tubes. One was given in a quiz where the interventions were implemented and the other was given in a midterm where the interventions were not implemented and all students received the same instructions. The two problems are the following:

Quiz problem (comparison group version):

"A tube with air is open at only one end and has a length of 1.5 m. This tube sustains a standing wave at its third harmonic. What is the distance between a node and the adjacent antinode?"

The diagram given to the students in intervention DO along with the above description contained an empty tube. The intervention PO students were explicitly asked to draw a diagram after the above problem statement.

Midterm problem:

The midterm problem was identical to the quiz problem except that the tube was open at both ends instead of just one end.

There are two approaches to solving the quiz problem (the midterm problem can also be solved by employing a very similar strategy for a tube that is open at both ends). One strategy is to draw the standing wave in question as shown in Figure 1.

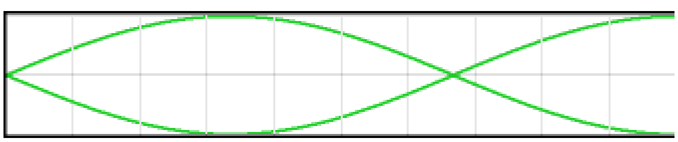


Figure 2.1. 3^{rd} harmonic for a standing wave in a tube open only at one end.

Then, for example, one can identify that three node to antinode distances fit in the tube with length L=1.5 m. Therefore, the distance between a node and the adjacent antinode is 1.5/3 = 0.5 m. This diagrammatic approach is the more expert-like approach because it requires understanding of a physics concept in its diagrammatic representation (third harmonic of a standing wave) and how it applies to a tube which is open at only one end (node at the closed end and antinode at the open end). The second approach to solving this problem (very similar to the second approach for the midterm problem) is to use the equation for the frequency of the n^{th} harmonic of a standing wave in a tube of length L open at only one end $\left(f_n = \frac{nv}{4L} \right)$ and the relation between the speed v, frequency f and wavelength of a wave, $v = f\lambda$**,** solve for wavelength λ given L and n and finally divide the wavelength obtained by 4 to get the distance between a node and the adjacent antinode. We refer to this latter approach as the "mathematical" approach because it does not necessarily require understanding the physics principles involved and the two equations can be used as mathematical algorithms if students have the mathematical skills required to manipulate them.

Students in DO were not given the diagram in Figure 2.1 because it would have greatly reduced the difficulty of the problem. Instead, they were given a partial diagram: an empty tube. It was intended that students would regard the partial diagram as a hint to complete it and be more likely to follow the expert like diagrammatic approach.

The quiz problem was also given to 26 first year physics graduate students (physics experts for this study) enrolled in a TA training course in order to assess how often physics experts use the diagrammatic approach, which was hypothesized to be a more expert-like approach. We also were interested in comparing the average score that graduate students obtained on this problem with that of introductory students. In order to make sure that the graduate students did not use a diagram simply because they did not remember the relevant equations for the frequency of a standing wave with different harmonics in a tube open at one end, they were given the relevant equations similar to the introductory physics students. The scores of the graduate students will be discussed and compared to the scores of the introductory students in the quantitative results section.

We investigated how the different interventions impacted the students in terms of how likely they were to draw productive diagrams. How much value one derives from drawing a particular type of diagram and how the person employs the diagram (and the process of drawing it) to solve a problem depends on the expertise of the individual. However, for the purposes of this research, a diagram was considered to be productive if it could have aided students in solving the problem based upon a cognitive task analysis of the problem. The productive diagrams were classified in two broad categories: diagrams of third harmonics (whether correct or not) and diagrams of one wavelength (whether drawn as standing or single sinusoidal waves). A diagram from a student attempting to draw a third harmonic was considered to be productive even if it did not represent a third harmonic, or the third harmonic of the correct situation (tube open at one end and closed at the other). This is because these diagrams can be used to solve the problem by use of the more expert-like approach. The second type of diagram (diagrams of one wavelength of a single sinusoidal/standing wave) was considered to be productive because it could be used to determine what fraction of a wavelength is the distance between a node and the adjacent antinode (the other type of productive diagram could be used to this end as well).

Furthermore, because there are two approaches to the solution of this standing wave problem, one primarily diagrammatic and another primarily based on mathematical manipulations, rubrics were developed to score the performance of students employing each approach. The summary of the rubric used to score students out of 10 points who chose the mathematical approach (both in the quiz and midterm problem) is shown in Table 2.1.

It is important to emphasize that a tube open only at one end can only sustain the odd harmonics, therefore n in the formula $f(n) = nv/4L$ takes only odd values. However, the third harmonic corresponds to $n = 3$ (and not the third possible value for n, namely 5) because of the convention that the frequency of the n^{th} harmonic of a standing wave must be n times the fundamental frequency. This is a common source of confusion for both experts and novices. We surveyed several physics experts (including graduate students and instructors) and found that almost all of them do not realize that the even harmonics do not arise for a standing wave in a tube open at only one end and associate the third harmonic with the third possible value for n, namely 5, which is incorrect based on the convention. Similarly, they incorrectly believed that the diagrammatic representation of the third harmonic corresponds to the third possible way of drawing a standing wave in a tube open at only one end, which instead corresponds to the fifth harmonic because of the same convention. Since the majority of experts confuse the fifth and third harmonic, the researchers considered that one should not penalize introductory students for this confusion. However, we are aware that not everyone would agree with this approach so we also performed data analysis for when one point (out of a maximum of 10) was taken off for mistaking the third harmonic with the fifth or using $n = 5$. All of the results are identical: every comparison which yielded a statistically significant difference in one instance (not taking off points for this mistake) also yielded a statistically significant difference in the other instance (taking off one point for this mistake).

Table 2.1 shows that there are two parts to the rubric: Correct Knowledge and Incorrect Ideas. Table 2.1 also shows that in the Correct Knowledge part, the problem was divided into different sections and points were assigned to each section (10 maximum points). Each student starts out with 10 points and in the Incorrect Ideas part, the common mistakes students made in each section and the number of points that were deducted for each of those mistakes are listed. It is important to note that each mistake is connected to a particular section (the mistakes labeled 1 and 2 are for the first and second sections respectively, the two mistakes labeled 3.1 and 3.2 are

Table 2.1. Summary of the rubric used to score the performance in the quiz of students employing the mathematical approach out of 10 points.

Correct Knowledge		
Section 1	1. Used given equation $f(n) = nv/4L$	1 p
Section 2	2. Chose $n = 3$ or $n = 5$	1 p
Section 3	3. Wrote down $v = f\lambda$	3 p
Section 4	4. Solved for λ correctly	2 p
Section 5	5. Found answer by dividing λ by 4	2 p
Section 6	6. Correct unit for answer	1 p
Incorrect Ideas		
Section 1	1. Used incorrect equation	-1 p
Section 2	2. Chose value for n other than 3 or 5	-1 p
Section 3	3.1 Did not write $v = f\lambda$	-3 p
	3.2 Tried to write down $v = f\lambda$, but made a mistake (i.e., wrote something like $v = f/\lambda$)	-2 p
Section 4	4.1 Did not solve for λ	-2 p
	4.2 Used a value for v other than that for sound wave	-1 p
	4.3 Made an error and obtained incorrect λ	-1 p
	4.4 Unclear how λ was found or other error	-1 p
Section 5	5. Did not divide λ by 4 to obtain the answer or did not obtain an answer	-2 p
Section 6	6. Incorrect units	-1 p

for the third section and so on) and that for each section, the rubric cannot be used to subtract more points than that section is worth. For example, the two mistakes in section 3 (3.1 and 3.2) are mutually exclusive. Similarly, mistake 4.1 is exclusive with all other mistakes in section 4 and mistakes 4.3 and 4.4 are mutually exclusive. Finally, if the mistake a student made was not common and not in the rubric, it would correspond to the mistake labeled as 4.4.

A rubric was also developed to score the performance of students employing the diagrammatic approach. The summary of the rubric is shown in Table 2.2.

Table 2.2. Summary of the rubric used to score the performance in the quiz of students employing the diagrammatic approach, out of 10 points.

Correct Knowledge		
Section 1	1. Drew a diagram of a wave	4 p
Section 2	2. Used diagram correctly to obtain the answer	5 p
Section 3	3. Correct units for answer	1 p
Incorrect Ideas		
Section 1	1.1 Diagram is a sinusoidal wave that does not clearly indicate locations of nodes and antinodes	-1 p
	1.2 Diagram has either two nodes or two antinodes at the endpoints	-2 p
	1.3 Diagram does not represent the third or fifth harmonic* (if endpoints are a node and an antinode)	-1 p
	1.4 Diagram does not represent the third harmonic** (if endpoints are both nodes or both antinodes)	-1 p
Section 2	2.1 Answer found is not the distance between a node and an antinode, nor the distance between two nodes (based on student's diagram)	-4 p
	2.2 Used diagram correctly, but found the distance between two nodes	-2 p
	2.3 Unclear how answer was obtained or other error	-1 p
Section 3	3. Incorrect units	-1 p

* Due to the confusion of experts of third and fifth harmonic for a standing wave in a tube open at only one end, both diagrams were considered correct.

** For the case when the tube is open or closed at both ends, experts do not have difficulties because all harmonics are possible, including even ones, therefore, it was considered that students should know in those instances how the third harmonic should be drawn.

The basic form of the summary of the rubric shown in Table 2.2 is the same as the one shown in Table 2.1; it has the same two main parts (Correct Knowledge and Incorrect Ideas), and again the problem is broken up into sections and the common mistakes students made in each section are listed. In section 1, mistakes 1.4 and 1.3 are mutually exclusive and in section 2, mistakes 2.1 and 2.2 are mutually exclusive. Similarly to the rubric used for the mathematical approach, we left ourselves a small window (labeled 2.3) if a mistake of a student was not explicitly in the rubric (a very rare occurrence, less than 5% of the cases).

We note that the rubrics are designed to be similar in terms of penalizing for mistakes that could be considered as analogous. For example, the rubric used for the mathematical approach treats the cases n = 3 and n = 5 as both correct because of the confusion of experts. Similarly, the rubric used for the diagrammatic approach does not penalize of student for drawing the fifth harmonic instead of the third of a standing wave in a tube open at only one end. Since this confusion was not penalized in one rubric, it was also not penalized in the other. Another example is provided by the analogy between the mistake of section 5 in the mathematical rubric and the mistake labeled 2.2 in the diagrammatic one. Students who use the mathematical approach and find the wavelength must have an understanding of what a node and an antinode are and divide the wavelength by four. This understanding of what a node and an antinode are is also required to use a diagrammatic representation of a standing wave to determine the distance between the two. This is why the mistakes are penalized equally (-2 points).

In addition to analyzing the quantitative data collected from the 118 students, interviews were conducted with eight students using a think-aloud protocol [79,80] in order to obtain an in-depth account of their difficulties while solving the quiz problem and in addition provide some insights that would account for the performance of these students. The results of the interviews will be discussed after the quantitative results section.

2.3 QUANTITATIVE RESULTS

2.3.1 Comparison between introductory students and graduate students

Before presenting the quantitative results it is important to mention that the data were analyzed using two grading approaches, one which penalized students for confusing the third with the fifth harmonic or selecting $n = 5$ instead of $n = 3$ and one which did not penalize for these mistakes. While this results in a slight change in averages and standard deviations, the statistical comparisons of performance of different groups of students yielded the same exact results. We present the data obtained with the latter grading approach.

As mentioned earlier, the quiz problem was also given to a group of 26 first year graduate students (physics experts for this study) enrolled in a TA training course. It is a straightforward mathematical exercise for a physics graduate student to solve for the wavelength, as previously described in the methodology section, using the mathematical approach. However, we found that 76% of them elected to draw a diagram to solve the problem (and ignored the equations provided to them completely), thus confirming our hypothesis that experts are more likely to follow the diagrammatic approach to solve this problem. The performance of both introductory physics students and graduate students on the quiz problem is listed in Table 2.3.

T-tests [81] reveal that graduate students who used the diagrammatic approach performed better than the introductory students who used the same approach. Also, the overall scores of graduate students were better than the scores of introductory students (both p values less than0.001). Interestingly, there does not appear to be a difference between the graduate students and the introductory physics students who used the mathematical approach. However, there were only 10 graduate students in this group and therefore a *t*-test is not appropriate.

Table 2.3. Number of students (N), averages (Avg.) and standard deviations (Std. dev.) for both the graduate students and the introductory students who used the diagrammatic approach and students who used the mathematical approach in the quiz problem.

Introductory physics students	**N**	**Avg.**	**Std. dev.**
1. Used diagrammatic approach	41	7.7	2.0
2. Used mathematical approach	77	7.8	2.0
4. Overall average	118	7.8	2.0
Graduate students	**N**	**Avg.**	**Std. dev.**
1. Used diagrammatic approach	16	9.4	1.0
2. Used mathematical approach	10	8.0	1.9
3. Overall average	26	8.8	1.6

2.3.2 Quantitative results pertaining to introductory student performance and drawing/use of diagrams

Students who primarily used a mathematical approach but drew productive diagrams performed better than students who used math without drawing productive diagrams.

We investigated how drawing a diagram and/or using the diagrammatic approach vs. the mathematical approach impacted students' scores both in the quiz problem and in the midterm problem. All the students were placed in groups based on whether they used the more expert-like diagrammatic approach ("Used diagram" in Table 2.4) or primarily used the mathematical approach. Among the students primarily using the mathematical approach (as discussed earlier), we classified students in two categories based upon whether they drew a productive diagram or not. We wanted to investigate whether a productive diagram helped improve scores or not (hence, the students who used the mathematical approach were divided into two groups in Table 2.4: “Used math, but also drew a productive diagram”, and “Used math without a productive diagram”).

Comparison of the performance of students in the quiz with that on the midterm yields no statistically significant differences between any of the groups of students shown in Table 2.4 (students who used a diagram, students who used math, but also drew a productive diagram etc.) However, both in the quiz and the midterm, students who primarily employed the mathematical approach but also drew a productive diagram performed better than students who chose the mathematical approach without drawing a productive diagram (p values for comparing these two groups of students are, 0.002 and 0.006 in the quiz and the midterm, respectively).

Table 2.4. Number of students (N), averages (Avg.) and standard deviations (Std. dev.) for students who used an expert-like diagrammatic approach (without math), used math and drew a productive diagram, and used only math without a productive diagram both in the quiz and in the midterm.

Quiz	**N**	**Avg.**	**Std. dev.**
1. Used diagram	41	7.7	2.0
2. Used math, but also drew a productive diagram	45	8.3	1.7
3. Used math without a productive diagram	24	6.7	2.1
Midterm	**N**	**Avg.**	**Std. dev.**
1. Used diagram	29	8.1	2.0
2. Used math, but also drew a productive diagram	68	8.8	2.0
3. Used math without a productive diagram	24	7.3	2.3

Students provided with a diagram of a tube (DO) performed worse than the students asked to draw a diagram (PO) and the students provided with no support regarding diagrams (NS).

The average scores on the quiz problem along with group sizes and standard deviations for the three different groups are shown in Table 2.5.

Table 2.5. Numbers of students (N), averages and standard deviations for the two intervention groups and the comparison group (NS) in the quiz problem.

Quiz	**N**	**Average**	**Standard deviation**
PO	50	8.1	1.7
DO	39	6.9	2.4
NS	29	8.6	1.0

It can be seen from Table 2.5 that students in group DO performed worse than students in the other groups. We performed *t*-tests to determine whether these differences were statistically significant. Table 2.6 shows the results, which indicate that students in DO performed statistically significantly worse than students in the other two groups.

Table 2.6. p values for comparisons between the scores of the different groups.

Quiz	PO-DO	DO-NS	PO-NS
	0.007	< 0.001	0.138

Students in PO drew more productive diagrams than students in the other groups.

Table 2.7 shows the percentages of students who drew productive diagrams. Almost all the students in PO drew productive diagrams compared to only 60% and 79% of students in groups DO and NS respectively. A chi-square test [81] shows the difference between PO and DO is statistically significant ($p < 0.001$). When comparing PO with NS, a chi-square test is not appropriate because some expected cell frequencies are less than 10 [82], so Fisher's exact test was performed [82], which yielded a statistically significant difference ($p = 0.046$).

Table 2.7. Percentages of students who drew productive diagrams in each group.

Quiz	PO	DO	NS
Percentage	96%	60%	79%

2.3.3 Quantitative data pertaining to introductory students' mathematical difficulties

Another important finding is that introductory physics students who primarily used the mathematical approach had great difficulties in solving for the wavelength without plugging in a value for the speed of the wave, v. They were given the equation for the frequency of the n^{th} harmonic of a wave in a tube open at one end $\left(f(n) = \frac{nv}{4L}\right)$, but they were not given the relationship between the speed, frequency and wavelength of a wave ($v = \lambda f$). Therefore, students had to remember the equation $v = \lambda f$ in order to solve for the wavelength. Table 2.8 lists how many students, among those who wrote down $v = \lambda f$, were able to solve for the wavelength correctly without plugging in a value for the speed, how many students were not able to do so, and how many students plugged in some numerical value for the speed (despite the fact that it was not explicitly given) in order to solve for the wavelength both in the quiz problem and in the midterm problem.

Table 2.8 shows that both in the quiz and the midterm, less than half of the students (48% in the quiz and 36% in the midterm) were able to eliminate the undesired quantities from the two equations and solve for the target variable without resorting to plugging in numerical information about speed that was not explicitly given. We note that the approach chosen by students in category 3 from Table 2.8 (plugging in a numerical value for the speed, *v*) is not necessarily an unproductive approach because it can help students reduce their cognitive load. However, the fact that so many students substituted a number for the speed of the wave in order to solve the quiz and midterm problems (when the speed would have canceled out between the two equations when solving for the wavelength) implies that a large fraction of the students in the algebra-based introductory physics course are uncomfortable manipulating two equations symbolically in order to eliminate the undesired quantities and determine the target variable.

As noted by others [83,84], students are not always facile in transferring mathematical knowledge to a physics context. We examined whether students in algebra based classes could solve an isomorphic, purely mathematical problem. The problem is as follows:

In the two equations underneath, *C* is a constant. Solve for *x* in terms of *C*. Show your work!

$$\begin{cases} y = \dfrac{C \cdot z}{4} \\ z = x \cdot y \end{cases}$$

Clearly, this is equivalent to the system of equations that students employing a purely mathematical approach must solve in the quiz and midterm problems if the following correspondences are made: $y \leftrightarrow f$, $C \leftrightarrow n$, $z \leftrightarrow v$, $x \leftrightarrow \lambda$ and $4 \leftrightarrow 4L$ (*L* was given as 1.5 so

Table 2.8. Numbers of students who were able to find the wavelength algebraically, who were not able to do so, and who plugged in a value for the speed of the wave (although not given) in order to solve for the wavelength (among the students who wrote down $v = \lambda f$)

Quiz	**N**
1. Solved correctly for λ (algebraically, i.e. without plugging in a value for v)	28
2. Did not solve correctly for λ or did not solve at all	6
3. Solved for λ by plugging in a numerical value for v	24
Midterm	**N**
1. Solved correctly for λ (algebraically, i.e. without plugging in a value for v)	17
2. Did not solve correctly for λ or did not solve at all	4
3. Solved for λ by plugging in a numerical value for v	26

students could plug it in to get $4L = 6$). We find that 64% of algebra-based students at the beginning of the first semester course are able to solve this system correctly and 89% of students at the beginning of their second semester course are able to solve this system correctly. It may not be appropriate to compare these percentages with the percentages of students who solved the quiz and midterm problems using the mathematical approach algebraically, without plugging in a value for v, because in those cases students had the option of plugging in a value for one of the unknowns (v) which greatly simplifies the task. It is of course possible that at least some of the students (if not the majority) who plugged in a value for v in the equation did so because otherwise they would have been unable to solve the problem. However, it does appear that students are more adept at solving for the desired variable from the system of two equations in the purely mathematical context than in the physics context, especially students in the second semester algebra-based class.

2.4 QUALITATIVE RESULTS FROM INTERVIEWS

As mentioned earlier, interviews were conducted with eight introductory physics students in order to get an in-depth account of their difficulties in solving the quiz problem. These students were at the time enrolled at the same university in an equivalent second semester algebra-based introductory physics course in which these concepts related to waves had been covered in the lectures and homework. They had also been tested (via a midterm) on concepts related to waves before the interviews were conducted. We found that some of their homework assignments involved very similar problems to the ones analyzed in this study and their first midterm contained a problem requiring students to draw different harmonics of standing waves in tubes. The interviews were conducted using a think-aloud protocol. Students were first asked to solve the problems to the best of their ability without interruption except they were asked to talk when they became quiet for a long time. After students were finished with the problem to the best of their ability, they were asked clarification questions if their reasoning at one point or another was unclear or questions related to other specific aspects of their problem solving approach. Also, they were asked to solve the tube problem using another approach (i.e. if a student solved it using the mathematical approach he/she was asked if he/she can solve it using a diagrammatic approach and vice-versa). If students found it difficult to solve the problem using either approach, they would often be asked questions intended to provide scaffolding and guide them (some examples will be provided below).

This qualitative results section is broken up into two subsections, the first reports qualitative findings obtained via interviews that are related to the quantitative results presented and the second reports student difficulties in employing the diagrammatic approach observed in the interviews.

2.4.1 Qualitative results via interview related to the quantitative results

1) **Qualitative results via interview related to the quantitative results**

One of the main findings from the quantitative investigation is that a good diagram is valuable for solving a problem related to a standing wave in a tube even when a student employs a primarily mathematical approach to problem solving. In particular, we found that students who used a mathematical approach but drew a productive diagram performed better than students who used the mathematical approach without drawing a productive diagram. As noted earlier, when defining “productive diagram” for these problems, it was considered that any diagram of a third harmonic (whether or not correct) could be productive because it gives students an opportunity to perform a conceptual analysis and planning related to the problem and students could use the insight derived from drawing this diagram to solve the problem. Moreover, even in the case when primarily a mathematical approach was chosen, the process of drawing a productive diagram can be helpful in conceptually analyzing the problem and such a diagram could be used to determine what fraction of a wavelength was represented by the distance between a node and the adjacent antinode. Another type of diagram that could be useful and was considered productive was a diagram of one wavelength of a standing or single sinusoidal wave. Interestingly, half of the students interviewed (four) chose the mathematical approach. Most of these students (three) drew a diagram of one wavelength of a single sinusoidal wave in order to determine that the distance between a node and the adjacent antinode is one quarter of the wavelength (and did so correctly). One of the interviewed students who drew a diagram, drew a diagram of the third harmonic but did not explicitly use the diagram she drew in solving the problem (and only focused on the equations). This latter student divided the wavelength by three (instead of four) to obtain the distance between the node and the antinode because she claimed that the number she needed to divide the wavelength by in order to determine the distance between a node and the adjacent antinode was related to the harmonic (i.e. she divided by *three*, because the problem involved the *third* harmonic of a standing wave).

Among the four students who used the diagrammatic approach, only one used it exactly in the way that an expert would most likely use it (and consistent with the approach of graduate students who were asked to solve it). He determined how many distances between a node and an antinode would fit in the length of the tube and then divided the length of the tube by that number. The other three students used the diagram of the third harmonic they had drawn to determine the wavelength. After finding the wavelength, they were at the same point as the students who used the mathematical approach to determine the wavelength, and just like those students, they then proceeded to determine the number by which to divide the wavelength in order to get the distance between a node and the adjacent antinode. To this end, one of these three students drew an additional diagram of one wavelength of a single sinusoidal wave (and used it incorrectly to obtain the distance between a node and an antinode) while the other two students explicitly used the diagrams of the third harmonic they had drawn (and used the diagrams correctly to find the distance between a node and an antinode). What the interviewed students did on their own while solving the problem and thinking aloud and what they said when asked for clarification of the points they had not made earlier suggests that drawing a diagram of a third harmonic for the problem or even one wavelength of a single sinusoidal wave can be helpful in finding the relationship between the distance between a node and the adjacent antinode and the wavelength of a standing wave because it helps students focus on relevant information in order to proceed with the problem solution. As noted earlier, the interviewed student who neither used her diagram of the third harmonic nor drew one wavelength of a single sinusoidal wave to determine the distance between the node and antinode with respect to the wavelength made a mistake (divided the wavelength by three because the problem involved the third harmonic). However, out of the other students who either used their diagrams of the third harmonic or diagrams of one wavelength of a single sinusoidal wave to determine the distance between the mode and antinode, only one made a mistake (divided the wavelength by 2 instead of 4). During the initial think aloud process while solving the problem without interruptions and later when asked for clarification, these students were able to articulate how the diagram was helpful in shaping the problem solving process. It appears from the interviews that students who drew productive diagrams performed better even if their chosen approach was primarily mathematical because the diagram helped them think about the problem solution conceptually.

Another finding discussed in the quantitative results section is that many algebra-based introductory physics students who selected the mathematical approach had difficulties in performing a substitution in order to eliminate the undesired quantities from the two equations, $v = f\lambda$ and $f_n = \frac{nv}{4L}$, and solve for the target variable, λ, without resorting to plugging in numerical information for speed that was not explicitly given in the two problems. In fact, among the students who knew the first equation, $v = f\lambda$, the percentage of them who were able to manipulate these two equations algebraically without plugging a numerical value for the speed and solve for the wavelength went down from 48% in the quiz to 36% in the midterm. These types of difficulties were also observed in the interviews. Some students approached the problem mathematically at first, but then changed their approach to diagrammatic when they had difficulty determining what mathematical steps to perform next. Dan, for example plugged in n=3 and L=1.5 m in the equation for frequency and solved for f/v to get 1/2 (he did not write down the units of 1/m). At this point he appeared to be stuck and after some thinking, he changed his approach to the diagrammatic one. After the think aloud part was over and Dan was probed further, he noted that he was aware of the other equation, $v = f\lambda$. He noted that at one point while solving the problem he thought about using this equation to find the wavelength. However, he did not explicitly write it down because he was not sure if it would help him to determine the wavelength. In particular, after he switched to solving the problem using the diagrammatic approach and attempted to solve it to the best of his abilities, the interviewer asked him if he was aware of the connection between speed of a wave, frequency and wavelength. At this point, Dan wrote this equation on the paper and identified correctly that the wavelength would equal 2 m. It was interesting that Dan noted that in his mind he had tried to think if he could solve for the wavelength using this equation $v = f\lambda$ along with $f/v = 1/2$ when he was solving the problem during the think aloud part of the interview without probing. However, he gave up on trying to solve the problem using these equations and did not realize that writing down the equation $v = f\lambda$ on paper may have reduced the cognitive load during problem solving and may have helped facilitate the problem solving process. Furthermore, while Dan himself noted that he contemplated using the equation $v = f\lambda$ along with $f/v = 1/2$ to solve for the wavelength; he was unable to do this and resorted to another approach. However, when he was given a system of two equations (without a physics context) with two unknowns of the form $\begin{cases} x + 2y = 3 \\ 3x - y = 2 \end{cases}$, he was able to readily solve this purely mathematical problem for variables x and y without much effort.

Another student, Karen, initially solved the problem using the diagrammatic approach during the think aloud part of the interview. During the second part of the interview, the interviewer asked her to solve the problem using the mathematical approach and gave her the equation $v = f\lambda$. At this point, Karen's first step was to substitute this equation into the other equation provided with the problem and plug in n=3 and L=1.5. She thus obtained $f = \frac{3(f\lambda)}{4(1.5)}$ which was a productive way to solve the problem. However, after this step, she was unsure about what to do and after some thinking, she gave up and noted that she did not know how to proceed (she did not realize that the frequency can be canceled from both sides of the equation). The interviewer then gave her a system of two equations with two unknowns (traditional x and y variables without the physics context) similar to the one Dan had to solve, and she was able to solve it correctly without much effort. In this situation, Karen was aware of what needed to be done next, but in the tube problem situation, after the substitution step, she was not able to determine what to do next. She did not realize that f was on both sides of the equation and she could cancel it or that she could multiply both sides of the equation by the denominator of the fraction on the right to get a simpler equation for the wavelength.

Another student, Tara, approached the problem mathematically from the beginning. She knew the two equations that needed to be used, $f = \frac{nv}{4L}$ and $v = f\lambda$, wrote them down, and then said the following:

Tara: If I knew v [speed of the wave] I could plug in this equation [$f = \frac{nv}{4L}$*], get f and then plug that in this equation [v=fλ] to get the wavelength.*

She then drew one wavelength of a travelling wave and said that she would divide the wavelength that she obtains by 4 to get the distance between the node and antinode. At this point she indicated that she was done to the best of her ability and her statement indicated that she could not solve this problem since the speed of the wave was not given. At this point, the interviewer then asked her:

Interviewer: Could you do this without knowing what v is?

Tara: Could I? Probably...

She then thought about how she could solve for the wavelength using the two equations for some time (a little less than a minute) and said:

Tara: I can't think of another way.

Interviewer: If you look at this equation [pointing with finger to $f = \frac{nv}{4L}$*] and this equation [pointing with finger to v=fλ]…*

Tara interrupted the interviewer before he could ask the question "could you solve for λ?":

Tara: I can plug it all in […] use substitution […] you would plug the frequency and the wavelength in for v [she meant plug in the frequency times the wavelength for v] in the equation given […] so you can solve for lambda that way.

She then correctly solved for the wavelength without plugging in a value for the speed of the wave. It appears that even though Tara had the information about how to perform substitution algebraically in her long term memory, she did not retrieve this information even after being explicitly asked if she could solve the problem without plugging in a value for the speed. Moreover, the fact that after the interviewer directed her attention to the two equations that had to be manipulated, Tara realized immediately what she needed to do, suggests that when she earlier paused (for about a minute) to think about whether she could solve the problem without knowing the speed, she may have not been focusing on the relevant information about the two equations.

These examples from interviews suggest that students were able to solve two simultaneous equations without a physics context without any difficulty but there was a lack of transfer of the mathematical knowledge to a physics context. This difficulty in transferring from the mathematical context to the physics context could be due to the fact that in the physics context there may be other information which can distract students from processing the relevant information, while such distractions are not present when engaged in a purely mathematical exercise. Moreover, mathematics is used differently in physics courses from mathematics courses [83] and this could also lead to difficulties in transfer from one context to another. For example, solving for the wavelength from the two equations, $f = \frac{nv}{4L}$, and $v = f\lambda$ may be more difficult for a student from solving a general 2x2 system of equations with x and y variables (as observed with some students in the interviews). In the first case, the symbols that go into those equations have physical meanings and it may be more difficult for students to focus on the relevant information (e.g. substitute one equation in the other, cross-multiply etc.) because the physical meanings add more information that is not present in the second case in which the variables x and y are devoid of physical meaning. This could partly account for the difficulties students exhibited in solving for the wavelength from the two equations without resorting to plugging in information that is not given, observed in the quantitative data and in the think-aloud interviews.

2.4.2 Student difficulties while using the diagrammatic approach

Interviews also revealed some difficulties algebra-based introductory physics students encountered while using the diagrammatic approach to standing waves in the tube while solving the problem. Karen for example, stated at the beginning after reading the problem that at the closed end of the tube there will be a node and that at the open end of the tube, there will be an antinode. She then tried to draw the third harmonic of this wave, and her attempts reflected this knowledge. However, she had difficulty drawing the third harmonic directly and she decided to start with a drawing of the first harmonic on the side (not in the tube) and work up to the third. However, the diagrams of the harmonics she drew on the side had nodes at both ends and therefore corresponded to a different situation (tube closed at both ends). Karen was unaware of this mistake in her drawing despite the fact that she explicitly stated at the beginning after reading the problem statement that at one end of the tube there should be a node and at the other end there should be an antinode. When solving the problem, she appeared to have forgotten about her initial correct statement (there should be a node at one end and antinode at the other end) and used the incorrect third harmonic she drew on the side (which corresponded to a standing wave in a tube closed at both ends) to solve the problem.

Another student, Sara, drew the 5^{th} harmonic for the wave in the tube open at only one end. However, the last section of the wave she drew (last 1/4 wavelength) looked on her diagram to be of the same length as the other two sections of the wave (1/2 wavelength) as shown in Figure 2.2.

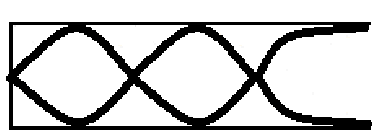


Figure 2.2. Diagram of the fifth harmonic as drawn by Sara (a student).

Consequently, she divided the length of the tube by three to get the distance between two nodes and then divided that distance by two to get the distance between a node and an antinode. After she was satisfied with her answer and was done with the problem to the best of her ability, the interviewer asked her why she divided the length of the tube by three. Here is a short excerpt:

Sara: Well, from what it said about third harmonic, I drew waves so you get one node here, another node here [the two middle nodes] and then the rest of the wave just opens up to the outside of the tube. Assuming the tube was closed, I think you would get another node right at the end of the tube [right side].

Interviewer: Yeah, but it's not closed.

Sara: Right, but I assumed that these nodes [the two middle ones] would automatically split the tube into three.

Interviewer: Okay, so you're thinking that here as well [at the right end] you would have a node?

Sara: If it was closed […] I know it's not closed so you don't get the node, it just kind of opens out, but I assumed, if it was closed you would get that node and these nodes [the two middle ones] would split the tube into three equal […] lengths.

Instead of correcting her diagram to fit the problem situation, Sara seemed to have modified the problem to fit her diagram and essentially ended up solving a different problem. She was also aware that it was a different problem ("I know it's not closed") but this didn't seem contradictory enough to her to change her diagram or interpretation. At this point the interviewer continued with further questioning to draw attention to her mistake:

Interviewer: Sure, but you're solving a different problem, because you're assuming it's closed and it's not. What would change if it's open?

Sara: So basically this [the last section on the right which should have been ¼ wavelength] is not the same as this and this [she pointed to the other two sections of the wave she drew].

Sara then stopped to think about her diagram and used the knowledge that the last section is half the length of the other two sections to correctly solve for the distance between a node the adjacent antinode. Similar to Karen, Sara also did not use knowledge she possessed (the last section of the wave was shorter than the other two sections) when she initially solved the problem without interruption. After the interviewer explicitly pointed her attention to the diagram she drew and pointed out that she had solved a different problem, Sara was able to retrieve the correct information and use it to solve the problem correctly based on her diagram of the fifth harmonic.

Another student, Brian, used information that was not applicable in the quiz problem. He thought that the distance between a node and an antinode decreases as you move away from the closed end. The diagram Brian drew is shown in Figure 2.3 and has a standing wave in which the distance between nodes decreases away from the closed end of the tube. It is unclear why he thought this to be true, but he explicitly stated that he remembered his instructor drawing a diagram where this was the case. It is unlikely that the instructor drew such a diagram because the book the students used had no such diagram or any discussion of a situation where the distance between nodes of a standing wave changes. It is possible that he misinterpreted a diagram drawn by the instructor.

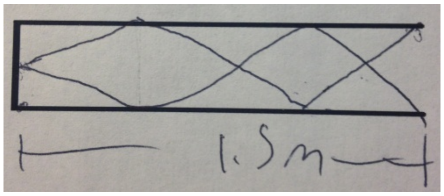


Figure 2.3. Diagram of the fifth harmonic drawn by Brian (a student).

Brian did not realize that the diagram he drew cannot be correct for the situation presented in the problem because if his reasoning was correct, then the problem should have specified *which* node to adjacent antinode distance students had to find (this distance would be different for his situation depending on which node/antinode you choose). He was therefore unable to solve the problem using this diagram and proceeded to try the mathematical approach.

2.5 DISCUSSION AND SUMMARY

We found that among the students who chose primarily the mathematical approach, those who drew productive diagrams performed better than those who did not. It is unclear whether the students who drew diagrams were the ones who generally had more expert-like approaches to problem solving which included a conceptual analysis stage that started with or involved drawing a diagram or whether the process of drawing a diagram helped students regardless of their general problem solving approaches. However, it is important to note that the interviews suggested that students who did draw diagrams were attempting to make sense of the problem conceptually and that the students who explicitly used the diagrams they drew were less likely to make mistakes than the students who did not. We therefore conclude that students should be explicitly taught good problem solving heuristics which include drawing a diagram in the conceptual planning stage and that instructors should emphasize and reward students for drawing diagrams.

We note that the quiz problem was also administered to a set of graduate students for two reasons: to confirm that the more expert-like approach is indeed the diagrammatic approach and also to obtain a benchmark for what would be the upper-limit of the performance of introductory physics students. We found that the majority of graduate students selected the diagrammatic approach even when the equation for the n^{th} harmonic frequency was provided, thus confirming that the diagrammatic approach is indeed a more expert-like approach. We also found that graduate students outperformed introductory physics students by an average of about 13%.

Furthermore, we found that the students who were given a diagram of an empty tube performed statistically worse than the students who were asked to draw a diagram and worse than students who were not given any instructions regarding diagrams. In a previous investigation [85] (also discussed in chapter 3 of this dissertation), we found the same result while examining introductory students' performance on two problems in electrostatics that involved considerations of initial and final situations. In addition, for the electrostatics problems, the differences in score between group DO and the other two intervention groups were even more pronounced (average of students given diagrams was more than 20% lower than the averages of the other two groups and the p values were smaller than 0.001). The research in Ref. [85] involved the same methodology as that described here. However, the diagrams given to students in group DO were very similar to what most instructors would initially draw in order to solve those two problems and were intended as scaffolding support. Instead of helping students solve those two problems, the given diagrams had the opposite effect, statistically worsening their performance as compared to students in the other two groups. Unlike the diagrams that were provided in the earlier study involving electrostatics problems, in the research presented here, students in group DO were given only a partial diagram (empty tube). Providing students the partial diagram was intended as a hint or prompt for them to complete it and attempt to solve the problem in an expert-like manner (drawing a diagram of a third harmonic of a standing wave and using it to solve the problem). However, similar to the study involving electrostatics problems, here we also we find that providing the diagram of the empty tube had the opposite effect from what was intended. In particular, the students who were given this diagram drew fewer productive diagrams than those who were not provided a diagram. This may be part of the reason why students in group DO performed worse than students in the other intervention groups.

We also found that in the context of solving the physics problem, students had great difficulty manipulating two equations symbolically. However, when students had to solve an isomorphic mathematical system of two equations devoid of physics context, the vast majority of them were able to perform the manipulations and solve for the target variable correctly in terms of other variables. This discrepancy between students' mathematical ability in a physics context and their mathematical ability in a mathematical context was also observed in the interviews. Many students expressed the need to substitute a numerical value for the speed of the wave before they solved for the wavelength using the two simultaneous equations even though the speed would have canceled out between the two equations. When they were asked to solve a 2x2 system of equations in a purely mathematical context, all the students were able to do so with little effort. One framework that can be used to partially interpret these findings is the cognitive load theory [86,87]. In this framework, problems are solved by a problem solver by processing relevant information in the working memory [88,89]. The relevant information for a problem includes both the information that comes from the problem itself and the possible matches that are found with the relevant knowledge in the long term memory of the problem solver. Research has shown that working memory is finite (5-9 “slots”) for any person regardless of intellectual capabilities [90,91]. Therefore, in order to solve a problem, one can only process 5-9 “chunks” of information at a given time to move forward with the solution. Experts have a hierarchically organized knowledge structure in their domain of expertise [92] and are also able to “chunk” knowledge and focus on important features of the problem which helps them retrieve appropriate information from their long term memory without experiencing cognitive overload [93-95]. In contrast, novices do not have a robust knowledge structure and while engaged in problem solving, they are typically unable to chunk more than one piece of information into one short term memory slot and therefore have reduced information processing capabilities (as compared to experts). Novices are also more likely to focus on unimportant features of the problem, and often retrieve information that is not necessarily useful or relevant [96,97]. These constraints are likely to overload the working memory while a novice is solving a physics problem.

The interviews suggest that cognitive load theory is one appropriate theoretical framework to reason about the mathematical difficulties exhibited by many students. In particular, in the initial problem solving phase of the interview, some students did not retrieve mathematical knowledge relevant to algebraically manipulate two simultaneous equations, even though this knowledge was present in their long term memory. It is possible that as a result of their expertise level, the physics context had too much information to be processed at a given time in their working memory and caused cognitive overload. Consequently, it was more difficult for them to focus on the relevant information that had to be processed at one time and make productive decisions in order to move forward with the solution. For example, as discussed earlier, one interviewed student was unable to determine how to solve the problem without plugging in information about the speed of the wave, which was not given, even after the interviewer explicitly asked her to do so. However, once the interviewer directed the student's attention to the two equations that had to be manipulated, that student immediately realized the next step (algebraic substitution) and solved the problem correctly. This interview suggests that while the student was thinking about how to solve the problem on her own she may have been allocating all of her cognitive resources at a given time to processing information related to the physical situation which distracted her from processing and retrieving the relevant mathematical information (e.g., perform a substitution). However, the fact that the interviewed students were able to solve two simultaneous equations not involving a physics context easily while they struggled to solve for the wavelength from two simultaneous equation in a physics context can also be interpreted as a lack of transfer of knowledge across disciplines [98,99]. In the context of the use of mathematics in physics it has been argued that the very fact that mathematics is used differently in physics courses than it is in mathematics courses may be adding to the difficulties encountered by students in using mathematics to solve physics problems [83].

In addition, in the interviews, difficulties were also observed when students were engaged in solving the problem using the diagrammatic approach. Some of these difficulties could also be interpreted using the framework of cognitive load theory. In particular, sometimes students did not make use of knowledge about waves that they possessed, which was explicitly mentioned (and, at times, even used briefly) by the same student at another stage of problem solving. Interviews suggest that at various points in problem solving, students had cognitive overload while they focused on certain aspects of the problem, and they completely lost track of other important information that they had in their long term memory which led to deteriorated performance.

One instructional implication of this research is that students should be encouraged to draw productive diagrams by rewarding them for drawing them. One of the many frameworks that may be useful for helping students learn to draw productive diagrams and other effective approaches to solving physics problems is the field tested cognitive apprenticeship model [100]. Within this cognitive apprenticeship model, the instructor can model productive diagrams while exemplifying effective approaches to problem solving, then coach students and provide feedback while they practice these skills and then gradually remove the support as they develop self-reliance. Another instructional implication is that it is important for instructors to keep in mind that algebra-based introductory physics students can have cognitive overload while solving physics problems as they must manage both the mathematical manipulations and how to use the underlying physical principles simultaneously to proceed successfully in the vast problem space. Trying to juggle both these tasks at the same time can be cognitively demanding particularly for introductory physics students in algebra-based courses who are not facile in algebra. A major fraction of their working memory may be used either in comprehending the mathematical procedure or in processing the related physics concepts. For example, students whose significant cognitive resources are allocated to parsing the mathematics involved rather than in sense making of the underlying physics principles and *why* certain concepts were used, may find it difficult to build a good mental picture of the concepts involved and may have difficulty in solving physics problems successfully. Since mathematical difficulties can make it challenging for students to build a good knowledge structure [101] of physics, suitable scaffolding should be provided to students which takes into account their physics and mathematics competencies to take them gradually from their initial knowledge state to the final knowledge state based upon the goals of the course [102].

2.6 CHAPTER REFERENCES

1. J. Heller and F. Reif (1984). "Prescribing effective human problem solving processes: problem description in physics." Cogn. Instruct. 1(2), 177-216.

2. C. Henderson, E. Yerushalmi, V. H. Kuo, P. Heller, and K. Heller (2004). "Grading student problem solutions: The challenge of sending a consistent message." Am. J. Phys. 72, 164-169 (2004).

3. F. Reif (1994). "Millikan lecture 1994: Understanding and teaching important scientific thought processes." Am. J. Phys. 63(1), 17-32.

4. J. Larkin, The role of problem representation in physics, in *Mental Models* edited by D. Gentner & A. Stevens (Hillsdale, NJ: Erlbaum, 1983).

5. J. Larkin and H. Simon, Why a Diagram is (Sometimes) Worth Ten Thousand Words, Cog. Sci. **11**(1), 65-99 (1987).

6. Y. Qin and H. Simon (1992). "Imagery and mental models in problem solving." AAAI Technical Report, SS-92-02.

7. J. Zhang and D. Norman. (1994) "Representations in Distributed Cognitive Tasks." Cog. Sci. 18(1), 87-122.

8. J. Zhang (1997). "The nature of external representations in problem solving." Cog. Sci. 21, 179-217.

9. D. Huffman (1997). "Effect of explicit problem solving strategies on high school students' problem-solving performance and conceptual understanding of physics." J. Res. Sci. Teach. 34(6), 551–570.

10. A. Van Heuvelen (1991). "Learning to think like a physicist: A review of research-based instructional strategies." Am. J. Phys. 59(10), 891-897.

11. A. Van Heuvelen (1991). "Overview, Case Study Physics." Am. J. Phys. 59(10), 898-907.

12. L. C. McDermott (1990). “A view from physics.” in Toward a Scientific Practice of Science Education, edited by M. Gardner, J. G. Greeno, F. Reif, A. H. Schoenfeld, A. diSessa and E. Stage (Lawrence Erlbaum, Hillsdale, New Jersey, 1990) pp. 3-30.

13. D. R. Jones and D. A. Schkade (1995). “Choosing and translating between problem representations.” J. Organ. Behav. Hum. Dec. Proc. 61(2), 214-223.

14. R. J. Dufresne, W. J. Gerace and W. J. Leonard (1997). “Solving physics problems with multiple representations.” Phys. Teach. 35(5), 270-275.

15. D. Hestenes (1997). “Modeling methodology for physics teachers” in The Changing Role of Physics Departments in Modern Universities: Proceedings of the International Conference on Undergraduate Physics Education, edited by E. F. Redish and J. S. Rigden [AIP Conf. Proc. 399] (American Institute of Physics, Woodbury, New York, 1997), Part Two, pp. 935-957.

16. D. Meltzer (2005). “Relation between students’ problem solving performance and representational mode.” Am. J. Phys. 73(5), 463-478.

17. D. Hestenes, M. Wells and G. Swackhammer (1992). “Force Concept Inventory.” Phys. Teach. 30(3), 141-158.

18. R. R. Hake (1998). “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses.” Am. J. Phys. 66(1), 64-74 (1998).

19. L. Ding, R. Chabay, B. Sherwood and R. Beichner (2006). “Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment.” Phys. Rev. ST Phys. Educ. Res. 2, 010105.

20. W. Maarten, van Someren, Peter Reimann, Henry P. A. Boshuizen and Ton de Jong (1998). editors, Learning with Multiple Representations (Pergamon, Amsterdam, 1998).

21. J. D. H. Gaffney, E. Richards, M. B. Kustusch, L. Ding and R. Beichner (2008). “Scaling up educational reform.” J. Coll. Sci. Teach. 37(5), 48-53.

22. W. Gerace, R. Dufresne, W. Leonard, J. P. Mestre (2000). “Minds-on physics: Materials for developing concept-based problem-solving skills in physics.” In *PERG*, 8. http://www.srri.umass.edu/publications/gerace-1999mdc

23. M. S. Sabella and S. A. Barr (2008). “Implementing research–based instructional materials to promote coherence in physics knowledge for the urban STEM student.” in Proceedings of the American Society for Engineering Education, pp. 395-409.

24. D.-H. Nguyen, E. Gire and N. S. Rebello (2010). "Facilitating students' problem solving across multiple representations in introductory mechanics." AIP Conf. Proc. 1289, 45-48.

25. A. Savinainen, A. Mäkynen, P. Nieminen, and J. Viiri (2013). "Does using a visual-representation tool foster students' ability to identify forces and construct free-body diagrams?" Phys. Rev. ST Phys. Educ. Res. 9, 010104

26. E.F. Redish, J. M. Saul, and R. N. Steinberg (1997). "On the effectiveness of active-engagement microcomputer-based laboratories." Am. J. Phys. 65 (1), 45-54.

27. Z. Hrepic, N. S. Rebello and D. A. Zollman (2009). "Remedying shortcomings of lecture-based physics instruction through pen-based, wireless computing and DyKnow software." in Reading: Assessment, Comprehension, and Teaching, edited by N. H. Salas and D. D. Peyton (Nova Science Publishers, 2009), pp. 97-129.

28. M. S. Sabella (2010). "What we learned by moving beyond content knowledge and diversifying our research agenda." AIP Conf. Proc. 1279, 53-56.

29. H. R. Sadaghiani (2012). "Controlled study on the effectiveness of multimedia learning modules for teaching mechanics." Phys. Rev. ST Phys. Educ. Res. 8, 010103.

30. V. P. Coletta and J. A. Phillips (2010). "Developing thinking & problem solving skills in introductory mechanics." AIP Conf. Proc. 1289, 13-16.

31. S. Aalie and D. Demaree (2010), "Toward meaning and scientific thinking in the traditional freshman laboratory: Opening the 'Idea Space' ", AIP Conf. Proc. 1289, 1-4.

32. L. Ding, N. W. Reay, A. Heckler and L. Bao (2010). "Sustained effects of solving conceptually scaffolded synthesis problems." AIP Conf. Proc. 1289, 133-136.

33. R. Teodorescu, C. Bennhold, and J. Feldman (2007). "Pedagogical coherence and consistency in an introductory physics course." AAPT National Meeting, Greensboro NC.

34. M. A. Kohlmyer, M. D. Caballero, R. Catrambone, R. W. Chabay, L. Ding, M. P. Haugan, M. J. Marr, B. A. Sherwood, and M. F. Schatz (2009). "Tale of two curricula: The performance of 2000 students in introductory electromagnetism." Phys. Rev. ST Phys. Educ. Res. 5, 020105.

35. R. J. Duchovic, D. P. Maloney, A. Majumdar, and R. S. Manalis (1998). "Teaching science to the non-science major – An interdisciplinary approach." J. Coll. Sci. Teach. 27, 258-262.

36. R. Beichner (1994). "Testing student interpretation of kinematics graphs." Am. J. Phys. 62(8), 750-762.

37. J. Clement (1998). "Observed methods for generating analogies in scientific problem solving." Cog. Sci. 12(4), 563-586.

38. R. Plötzner (1994) The Integrative Use of Qualitative and Quantitative Knowledge in Physics Problem Solving (Peter Lang, Frankfurt am Main, 1994), pp. 33-46.

39. R. K. Thornton and D. R. Sokoloff (1998). "Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and lecture curricula." Am. J. Phys. 66(4), 338-352.

40. A. Van Heuvelen and X. Zou (2001). Multiple representations of work-energy processes, Am. J. Phys. 69(2), 184-194.

41. X. Zou (2001). "The role of work-energy bar charts as a physical representation in problem solving." Proceedings of the 2001 Physics Education Research Conference, edited by S. Franklin, J. Marx and K. Cummings (PERC Publishing, Rochester, NY 2001), pp. 135-138.

42. R. Beichner, R. Chabay, and R. Sherwood (2010). "Labs for the Matter & Interactions curriculum." Am. J. Phys. 78(5) 456-460.

43. R. Lesh, T. Post and M. Behr (1987). "Representations and translating between representations in mathematics learning and problem solving." in Problems of Representations in the Teaching and Learning of Mathematics, edited by C. Janvier (Lawrence Erlbaum Hillsdale, New Jersey, 1987) pp. 33-40.

44. P. White and M. Mitchelmore (1996). "Conceptual knowledge in introductory calculus." J. Res. Math. Educ. 27(1), 79-95.

45. B. Ibrahim and N. S. Rebello (2012). "Representational Task Formats and problem solving strategies in kinematics and work." Phys. Rev. ST Phys. Educ. Res. 8, 010126.

46. J. Larkin (1980). "Skilled problem solving in physics: A hierarchical planning model." J. Struct. Learn. 6, 271-297.

47. J. Larkin (1980). "Skilled problem solving in physics: A hierarchical planning approach." J. Struct. Learn. 6, 121-130.

48. L. C. McDermott, M. L. Rosenquist and E. H. van Zee (1987). "Student difficulties in connecting graphs and physics: Examples from kinematics." Am. J. Phys. 55, 503-513.

49. M. Ward, and J. Sweller (1990). "Structuring effective worked examples." Cog. Instruct. 7(1), 1–39.

50. A. H. Schoenfeld (1987). "What's all the fuss about metacognition." in Cognitive Science and Mathematics Education, edited by A. H. Schoenfeld (Lawrence Erlbaum Associates, Hillsdale, NJ, 1987), pp. 189-215.

51. A. H. Shoenfeld (1980). "Teaching problem solving skills." Amer. Math. Monthly 87, 794-805.

52. A. Mason and C. Singh (2010). "Helping students learn effective problem solving strategies by reflecting with peers." Am. J. Phys. 78(7), 748-754.

53. J. E. Court (1993). "Free-body diagrams." Phys. Teach. 31, 104–108.

54. D. Rosengrant (2007). Ph.D. Dissertation, Rutgers University.

55. D. Rosengrant, A. Van Heuvelen and E. Etkina (2005). "Free-body diagrams: Necessary or sufficient?" AIP Conf. Proc. 790, 177–180.

56. A. F. Heckler (2010). "Some consequences of prompting novice physics students to construct force diagrams." Int. J. Sci. Educ. 32(14), 1829-1851.

57. K. Mazens (1997). "Conceptual change in physics: naïve representations of sounds in 6- to 10-year old children." paper presented at the EARLI conference, Athens, pp. 1–9.

58. K. Mazens and J. Lautrey (2003). "Conceptual change in physics: Children's naïve representations of sound." Cogn. Dev. 18, 159–176.

59. H. Eshach and J. Schwartz (2007). "Sound stuff? Naïve materialism in middle-school students' conceptions of sound." Int. J. Sci. Educ. 28(7), 733-764.

60. E. Boyes and M. Stanisstreet (1991). "Development of pupils' ideas about seeing and hearing – The path of light and sound." Res. Sci. Tech. Educ. 9, 223-251.

61. M. E. Houle and G. M. Barnett (2008). "Students' Conceptions of Sound Waves Resulting from the Enactment of a New Technology-Enhanced Inquiry-Based Curriculum on Urban Bird Communication." J. Sci. Educ. Tech. 17(3), 242-251.

62. C. J. Linder and G. L. Erickson (1989). "A study of tertiary physics students' conceptualizations of sound." Int. J. Sci. Educ. 11, 491–501.

63. C. J. Linder (1993). "University physics students' conceptualizations of factors affecting the speed of sound propagation." Int. J. Sci. Educ. 15, 655–662.

64. T. R. Rhoads and R. J. Roedel (1999). "The wave concept inventory – A cognitive instrument based on Bloom's taxonomy." paper presented at the 28th Annual Frontiers in Education Conference, Tempe Mission Palms Hotel, Tempe, AZ.

65. M. Reiner, J. D. Slotta, M. T. H. Chi and L. B. Resnick (2000). "Naïve physics reasoning: A commitment to substance-based conceptions." Cog. Instruct. 18(1), 1-34.

66. M. Wittmann, R. N. Steinberg and E. F. Redish (2003). “Understanding and affecting student reasoning about sound waves.” Int. J. Sci. Educ. 25, 991-1013.

67. M. Wittmann, R. Steinberg, and E. Redish (1999). “Making sense of how students make sense of mechanical waves.” Phys. Teach. 37(1), 15.

68. Z. Hrepic, D. Zollman, and N. S. Rebello (2010). “Identifying students’ mental models of sound propagation: The role of conceptual blending in understanding conceptual change.” Phys. Rev. ST Phys. Educ. Res. 6, 020114.

69. M. C. Wittman (1998). “Making sense of how students come to an understanding of physics: An example from mechanical waves” Ph.D. thesis, University of Maryland.

70. A. Tongchai, M. D. Sharma, I. D. Johnston, K. Arayathanitkul and C. Soankwam (2009). “Developing, evaluating and demonstrating the use of a conceptual survey of mechanical waves.” Int. J. Sci. Educ. 31, 2437.

71. A. Tongchai, M. D. Sharma, I. D. Johnston, K. Arayathanitkul and C. Soankwam (2011). “Consistency of students’ conceptions of wave propagation: Findings from a conceptual survey in mechanical waves.” Phys. Rev. ST Phy. Educ. Res. 7, 020101.

72. B. W. Frank, S. E. Kanim, and L. S. Gomez (2008). “Accounting for variability in student responses to motion questions.” Phys. Rev. ST Phys. Educ. Res. 4, 020102.

73. E. E. Clough and R. Driver (1986). “A study of consistency in the use of students’ conceptual frameworks across different task contexts.” Sci. Educ. 70, 473.

74. D. Palmer (1993). “How consistently do students use their alternative conceptions?” Res. Sci. Educ. 23, 228.

75. M. Finegold and P. Gortsky (1991). “Students’ concepts of force as related to physical systems: A search for consistency.” Int. J. Sci. Educ. 13, 97.

76. J. R. Watson, T. Prieto and J. S. Dillon (1997). “Consistency of students’ explanations about combustion.” Sci. Educ. 81, 425.

77. Mila Kryjevskaia, MacKenzie R. Stetzer, and Paula R. L. Heron. “Student difficulties measuring distances in terms of wavelength: Lack of basic skills or failure to transfer?” Phys. Rev. ST Phys. Educ. Res. 9, 010106.

78. N. S. Podolefsky and N. D. Finkelstein (2007). “Analogical scaffolding and the learning of abstract ideas in physics: Empirical studies.” Phys. Rev. ST Phys. Educ. Res. 3, 020104.

79. K. Ericsson and H. Simon (1980). “Verbal reports as data.” Psychol. Rev. 87, 215.

80. K. Ericsson and H. Simon (1993). Protocol Analysis: Verbal Reports as Data, (MIT Press, Boston, MA 1993).

81. G. V. Glass and K. D. Hopkins (1996). Statistical Methods in Education & Psychology, (3rd ed.), Boston: Allyn & Bacon (1996).

82. R. A. Fisher (1992). "On the interpretation of χ^2 from contingency tables, and the calculation of P." J. Roy. Stat. Soc. 85, 87.

83. E. Redish (2005). "Problem solving and the use of math in physics courses." paper presented at the World View on Physics Education in 2005, Delhi.

84. D. J. Ozimek, P. V. Engelhardt, A. G. Bennett, and N. S. Rebello (2005). "Retention and Transfer from Trigonometry to Physics." AIP Conf. Proc. 790, 173-176.

85. A. Maries and C. Singh (2011). "Should students be provided diagrams or asked to draw them while solving introductory physics problems?" AIP Conf. Proc. 1413, 263-266.

86. J. Sweller, R. Mawer, and M. Ward (1983). "Development of expertise in mathematical problem solving." J. Exp. Psychol. Gen. 112, 639.

87. J. Sweller (1988). "Cognitive load during problem solving: Effects on learning", Cog. Sci. 12, 257.

88. H. Simon (1979), Models of Thought, (Yale University Press, New Haven, CT 1979) Vols. 1 and 2.

89. J. R. Anderson (1995), Learning and Memory, (Wiley, New York 1995).

90. G. Miller (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information." Psychol. Rev. 63, 81.

91. A. Miyake, M. A. Just, and P. Carpenter (1994). "Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts." J. Mem. Lang. 33, 175.

92. B. Eylon and F. Reif (1984). Effect of knowledge organization on task performance, Cogn. Instruct. 1, 5.

93. M. T. H. Chi, P. J. Feltovich, and R. Glaser (1981). "Categorization and representation of physics knowledge by experts and novices." Cog. Sci. 5, 121-152.

94. K. Johnson and C. Mervis (1997). "Effects of varying the levels of expertise on the basic level of categorization." J. Exp. Psych. Gen. 126, 248.

95. J. L. Docktor, J. P. Mestre, and B. H. Ross (2012). “Impact of a short intervention on novices’ categorization criteria.” Phy. Rev. ST Phys. Educ. Res. 8, 020102.

96. P. W. Cheng and K. J. Holyoak (1985). “Pragmatic reasoning schema.” Cogn. Psychol. 17, 391.

97. M. S. Sabella and E. F. Redish (2007). “Knowledge organization and activation in physics problem solving.” Am. J. Phys. 75, 1017.

98. N. S. Rebello (2009). “Can we assess efficiency and innovation in transfer?” AIP Conf. Proc. 1179, 241-245.

99. J. F. Wagner (2010). “A transfer-in-pieces consideration of the perception of structure in the transfer of learning.” J. Learn. Sci. 19, 443.

100. A. Collins, J. S. Brown and S. E. Newman (1989), “Cognitive Apprenticeship: Teaching the crafts of reading, writing and apprenticeship.” in Knowing, Learning and Instruction: Essays in Honor of Robert Glaser, R. Glaser and L. Resnick (eds.) Hillsdale, NJ, Lawrence Erlbaum Associates, 453-494.

101. E. Bagno and B. Eylon (1997). “From problem solving to a knowledge structure: An example from the domain of electromagnetism.” Am. J. Phys., 65(8), 726-736.

102. L. S. Vygotsky (1978). Mind in Society: The Development of Higher Psychological Processes, Cambridge, MA, Harvard University Press.

3.0 SHOULD STUDENTS BE PROVIDED DIAGRAMS OR ASKED TO DRAW THEM WHILE SOLVING INTRODUCTORY PHYSICS PROBLEMS?

3.1 INTRODUCTION

For a literature review of previous research related to the role of multiple representations in problem solving, refer to the introduction in the study presented in Chapter Two.

In this research study we investigate how the student performance will be affected when students are given a diagram instead of being asked to draw it and compare their performance to the performance of students who are asked to draw a diagram (without being any more specific than that) and to the performance of a comparison group which is neither asked to draw diagrams nor provided a diagram. We found that students who were provided diagrams performed worse than the other students on two problems in electricity discussed here which involved considerations of initial and final conditions. One possible interpretation that we provide for this result is that students who were provided with a diagram were more likely to spend less time on the conceptual planning stage and sometimes jumped into the implementation stage without understanding the problem situations fully. This interpretation was evaluated by conducting interviews with fourteen students, six of them being conducted using a think-aloud protocol, while in the others, students were observed by a researcher while solving the problems. These interviews provided evidence to support our interpretation.

3.2 METHODOLOGY

A class of 111 algebra-based introductory physics students was broken up into three different recitations. All recitations were taught in the traditional way in which the TA worked out problems similar to the homework problems and then gave a 15 minute quiz at the end of class. Students in all recitations attended the same lectures, were assigned the same homework, and had the same exams and quizzes. In the recitation quizzes throughout the semester, the three groups were given the same problems but with the following interventions: in each quiz problem, the first intervention group, which we refer to as "prompt only group" or "PO", was given explicit instructions to draw a diagram with the problem statement; the second intervention group (referred to as "diagram only group" or "DO") was given a diagram drawn by the instructor that was meant to aid in solving the problem and the third group was the comparison group and was not given any diagram or explicit instruction to draw a diagram with the problem statement ("no support group" or "NS").

The sizes of the different recitation groups varied from 22 to 55 students because the students were not assigned a particular recitation; they could go to whichever recitation they wanted. For the same reason, the sizes of each recitation group also varied from week to week, although not as drastically because most students (≈ 80%) would stick with a particular recitation. Furthermore, each intervention was not matched to a particular recitation. For example, in one week, students in the Tuesday recitation comprised the comparison group, while another week the comparison group was a different recitation section. This is important because it implies that individual students were subjected to different interventions from week to week and we do not expect cumulative effects due to the same group of students always being subjected to the same intervention.

In order to ensure homogeneity of scoring, we developed rubrics for each problem we analyzed and made sure that there was at least 90% inter-rater reliability between two different raters. The development of the rubric for each problem went through an iterative process. During the development of the rubric, the two raters also discussed a student's score separately from the one obtained using the rubric and adjusted the rubric if it was agreed that the version of the rubric was too stringent or too generous. After each adjustment, all students were scored again on the improved rubric.

In this study, we analyze two problems from electrostatics which involve consideration of initial and final states. The goal was to investigate if there were any statistical differences in the scores of the groups of students subjected to different interventions.

The two problems discussed and the diagrams given to students in DO are the following:

Problem 1

Two identical point charges are initially fixed to diagonally opposite corners of a square that is 1 m on a side. Each of the two charges q is 3 C. How much work is done by the electric force if one of the charges is moved from its initial position to an empty corner of the square?

Problem 2

A particle with a mass 10^{-5} kg and a positive charge q of 3 C is released from rest from point A in a uniform electric field. When the particle arrives at point B, its electrical potential is

25 V lower than the potential at A. Assuming the only force acting on the particle is the electrostatic force, find the speed of the particle when it arrives at point B.

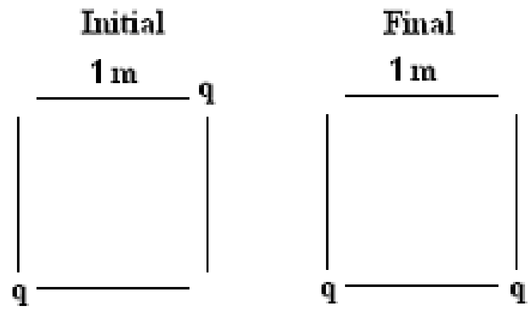


Figure 3.1. Diagram for problem 1 given to students in DO.

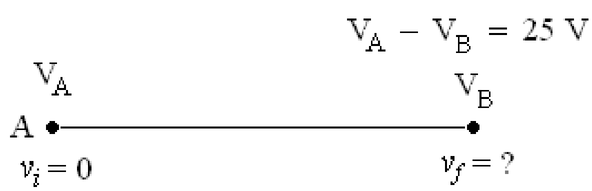


Figure 3.2. Diagram for problem 2 given to students in DO

These diagrams were drawn by the instructor and they are very similar to what most physics experts would generally draw in order to solve the problems. Furthermore, the second diagram also includes an important piece of information from the problem statement that would normally be included in a known quantities/target quantities section of a solution. Neither diagram was meant to trick the students, but rather they were provided as a scaffolding support for them.

As mentioned earlier, we developed rubrics for each problem. In Table 3.1, we provide the summary of the rubric for the first problem. The rubric for the other problem is similar.

Table 3.1. Summary of the rubric used for Problem 1.

Correct Ideas		
Section 1	1. W = -qΔV or W = -ΔEPE	2 p
Section 2	2. Obtain V_f, V_i and find $\Delta V = V_f - V_i$ **or** obtain EPE_f, EPE_i and find $\Delta EPE = EPE_f - EPE_i$	7 p
Section 3	3. Correct units	1 p
Incorrect Ideas		
Used the electrostatic force incorrectly: if provided correct units (-8 p), if no units (-10 p)		
Section 1	1. Used incorrect equation	-2 p
Section 2	2.1 Obtained one potential or one EPE incorrectly	-2 p
	2.1 Obtained both potentials or EPEs incorrectly	-4 p
	2.2 Did not subtract the electric potentials/EPEs (-2 p), and/or other mistake (-1 p)	-3/-1 p
	2.3. Incorrect sign	-1 p
Section 3	3. Incorrect or no units	-1 p

Table 3.1 shows that there are two parts to the rubric: Correct and Incorrect Ideas. Table 3.1 also shows that in the Correct Ideas part, the problem was divided into different sections and points were assigned to each section (10 maximum points). Each student starts out with 10 points and in the Incorrect Ideas part we list the common mistakes students made and how many points we deducted for each of those mistakes. Using the electrostatic force for this problem is not an effective strategy for algebra based students (this approach involves calculus), so students who attempted to use the electrostatic force had to be graded separately because their approach is not productive. The rest of the rubric in the Incorrect Ideas part was used for grading the students who chose a productive approach. For each mistake, we deducted a certain number of points. We note that it is not possible to deduct more points than a section has (e.g., the two mistakes that are

both labeled 2.1 in Table 3.1 are mutually exclusive). We also left ourselves a small window (labeled 2.2) to account for possible mistakes not listed in the rubric.

In order to explore further how students' problem solving approach and reasoning can be affected by being provided diagrams along with the physics problems, interviews were conducted with fourteen students who were at the time enrolled in an equivalent algebra-based second semester introductory physics. It was not clear a priori how the interview protocol would affect students' reasoning and sense making during problem solving. Therefore, we decided to use one type of interview protocol for some of the students followed by another type of protocol for another set of students. In particular, six of these interviews were conducted using a think-aloud protocol, while in the other eight interviews, the students solved the problems while being observed by one of the researchers. In order to compare how a student approaches problems when diagrams are not given as opposed to when the diagrams are given, the students were asked to solve an additional problem which required use of the same concepts (conservation of energy/work, electric potential, electric potential energy, etc.) as the two problems discussed in this paper. However, in this additional problem, a diagram was not provided.

Additional problem

A particle of mass 10^{-4} kg and charge q_1 = 1μC is shot at a speed of 10 m/s directly towards another particle with charge q_2 = 1μC that is held fixed. If the initial distance between the two particles is 1m, how close does the particle with charge q_1 get to q_2?

3.3 QUANTITATIVE RESULTS

Before discussing the findings for the two problems outlined, we note that the two problems analyzed in this paper were part of the same three problem recitation quiz. In the third problem of that quiz, we did not find any statistical differences between the different groups. Furthermore, students in different groups exhibited almost identical performance in midterm and final examinations and we therefore believe that the groups are comparable in terms of students' physics problem solving abilities and any differences in student performance on these problems are due to the interventions.

It is evident from Table 3.2 that students who were given the diagram (DO) performed significantly worse than all the others (PO and NS). In particular, their averages are lower by roughly 20% compared to the other intervention groups. We also performed *t*-tests [1] to investigate if the differences are statistically significant. The p-values from the *t*-tests are shown in Table 3.3.

Table 3.2. Group sizes (N), averages and standard deviations for the scores of the two intervention groups and the comparison group on the two problems.

Problem 1	**N**	**Avg.**	**Std. dev.**
PO	26	8.5	1.88
DO	34	6.9	2.82
NS	51	9.0	1.39
Problem 2	**N**	**Avg.**	**Std. dev.**
PO	26	9.0	1.44
DO	34	6.4	3.06
NS	51	8.6	1.34

Table 3.3. p values for t-test comparisons between the different groups.

	DO-PO	DO-NS	PO-NS
Problem 1	0.015	< 0.001	0.193
Problem 2	< 0.001	< 0.001	0.343

Table 3.3 shows that students in DO (who were given the diagram) performed significantly worse than students in the other two groups. More noteworthy is how small the p values are (three of them being less than 0.001). Table 3.3 also shows that the scores of PO and NS are comparable on both problems. We note that, for Problem 1, virtually all students drew a diagram even if they were not specifically asked to do so. However, for Problem 2, only 57% of the students in NS drew a diagram. But within NS, there are no statistical differences between the performance of the students who drew a diagram and those who did not draw a diagram. We performed a *t*-test to compare the performance of students in NS who did not draw a diagram and all students in DO. We found that students in DO performed significantly worse ($p = 0.004$) than those in NS who did not draw a diagram. Thus, on Problem 2, students who did not draw a diagram performed better than those who were given a diagram (drawn by the instructor) with the problem statement. Some possible reasons for this surprising counter-intuitive result will be discussed.

Table 3.4 shows that the percentage of students who performed poorly on this problem (obtained a score less than 5) from DO is significantly larger than those in PO and NS but percentages with an intermediate score are comparable.

Table 3.4. Percentages (and numbers) of students in each group who earned below 5 or (5, 6 and 7) or above 8 (out of 10).

Problem 1	**score ≤4**	**5≤ score <8**	**score ≥8**
PO	4% (1)	23% (6)	73% (19)
DO	26% (9)	21% (7)	53% (18)
NS	2% (1)	16% (8)	82% (42)
Problem 2	**score ≤4**	**5≤ score <8**	**score ≥8**
PO	4% (1)	15% (4)	81% (21)
DO	38% (13)	21% (7)	41% (14)
NS	2% (1)	22% (11)	76% (39)

3.4 QUALITATIVE RESULTS

This section is broken up into two subsections. The first subsection discusses qualitative results from interviews with students which suggested that the hypothesis we developed to account for the quantitative results may be befitting. The second subsection discusses qualitative results from discussions with faculty. The instructors discussed what they expect would be the consequence of providing diagrams for the two problems discussed in this study and their general viewpoint regarding diagrams and problem solving.

3.4.1 Qualitative results from student interviews

As mentioned earlier, interviews were conducted with students who were at the time enrolled in an equivalent second semester algebra-based introductory physics course. All these interviews occurred after students learned and were tested in their course on the relevant concepts required for successfully solving these problems. The students participating in these interviews were asked to solve three problems. Two of the problems were the ones investigated in this research study, for which diagrams were provided. The students interviewed were specifically asked to comment on the diagrams and on being provided diagrams. None of them mentioned anything negative about the diagrams and in general they thought that the diagrams were helpful. A few students said that they didn't necessarily gain anything from being provided diagrams because if they had not been provided diagrams, they would have drawn something similar anyway.

As discussed earlier, in order to compare the problem solving approaches of students to problems which provide diagrams with their approaches to problems which do not provide diagrams, interviewed students were asked to solve an additional problem (described in the Methodology section earlier) which required use of the same concepts. The additional problem was carefully chosen by the researchers because it was considered that it should fulfill the following three criteria:

1) In order to solve it one must make use of the same concepts as the other two problems investigated in this study.
2) The additional problem had to be comparable in difficulty with the other two problems investigated in this study.
3) The physical description of the additional problem should be such that a student could potentially solve it without having to draw a diagram.

These three criteria were chosen because the goal of the interviews was to gain a better understanding of the reasoning behind the quantitative results discussed earlier. In particular, the interviews were designed to evaluate our hypothesis which we believed could partly account for the deteriorated performance of students who were provided a diagram with the problem statement (intervention DO) compared to students who were not provided a diagram. We hypothesized that the deterioration may partly be due to students being more likely to spend less time on (or completely skipping) the important step of conceptually analyzing the problems when a diagram is provided compared to when it is not provided. It is possible that the diagram provided prompted students to jump into the implementation of problem solution early without adequate conceptual analysis, planning and decision making related to the problem solving. This may in turn make it more likely for students to follow formula centered approaches and perhaps use equations that are not appropriate for the given problem, which would cause deteriorated performance. We note that the step of drawing a diagram can be very helpful in conceptually analyzing a problem. In order to compare how much time students spend conceptually analyzing the additional problem in which a diagram was not provided with how much time they spend conceptually analyzing the other two problems in which diagrams were provided, the researchers considered that the additional problem posed during interviews had to deal with the same concepts since students may find some concepts more challenging than others. The additional problem also had to be of comparable difficulty because if the difficulty level was different from those of the other two problems, students will not spend the same time conceptually analyzing it compared to the other two problems. Finally, the additional problem was chosen such that an interviewed student was not necessarily compelled to draw a diagram in order to solve it. In particular, if the additional problem posed during the interviews would necessarily require all students to draw a diagram due to the physical situation presented in the problem (for example, a two dimensional problem, like Problem 1 investigated in this research study, for which all the students drew a diagram), that would be counterproductive to the goals of the interviews. Furthermore, the additional problem was given first to the interviewed students (before the other two problems) due to a concern that had it been given last, the students' approach to solving it might be somewhat influenced by the other two problems which provided diagrams with the problem statement. In particular, we did not want students to be prompted to draw a diagram because in the other two problems the diagrams were provided. Thus, the order in which the problems were solved by the students who participated in the interviews was: 1) Additional problem, 2) Problem 1 and 3) Problem 2 (as described in the Methodology section).

Furthermore, in order to make the interview situation similar to the quiz situation, students were given an equation sheet which was photocopied from the textbook's [2] end of chapter summary (chapter 19, which discusses electrostatic potential and electrostatic potential energy). This was because in the quiz, the students were given equations from this chapter by their teaching assistant who wrote them on the board.

During the first six interviews, students were asked to solve the three problems (Additional Problem, Problem 1, Problem 2) while thinking aloud. The amount of time students spent conceptually analyzing a problem was estimated by timing students from when they first started reading the problem until they wrote down an equation from the equation sheet provided. These interviews revealed that in the think-aloud setting, students spent about the same time conceptually analyzing each of the three problems and also spent about the same time solving each problem. It is possible that because they were asked to verbalize their though process, each student approached the three problems in very similar ways and was not influenced by having been given diagrams in the last two problems. Half of the interviewed students drew a diagram for the additional problem and made some effort to connect the diagrams given in the other two problems with the verbal description of those problems (i.e. they looked at the diagrams as they read the problem statements, occasionally adding information). The other half of the interviewed students made use of more formula-centered approaches for all three problems: they did not draw a diagram for the additional problem and did not seem to pay too much attention to the diagrams provided in the other two problems. While the think-aloud setting does not reproduce the quiz setting very well, these six think-aloud interviews provided valuable information since they offered evidence that the additional problem was well chosen. In particular, students spent about the same time conceptually analyzing this additional problem as they did the other two problems and they spent about the same amount of time solving this additional problem as they did the other two problems. This indicated that the additional problem was of comparable difficulty. Also, the students who had more formula centered approaches to solving problems did not draw a diagram for the additional problem indicating that our third criterion for selecting the additional problem was met (these students did not consider that drawing a diagram was necessary to solve the problem).

Since it appears that asking the students to think aloud resulted in them spending about the same time conceptually analyzing each problem whether or not a diagram was provided, more interviews were conducted which were designed to provide an environment more similar to the written quiz setting than the think-aloud interviews. In these interviews, students solved the three problems, but were not asked to talk during this time, rather, a researcher observed and took detailed notes about what the students were doing, what they were writing down and at what times. Similar to the think-aloud interviews, the amount of time students spent conceptually analyzing each problem was estimated by timing them from when they first started reading the problem until they wrote down an equation from the equation sheet provided. During these interviews, most students (five out of the eight interviewed) wrote down an equation noticeably quicker when solving the second and third problems in which the diagrams were provided than when solving the first problem in which a diagram was not provided and in a few of those cases, in one problem or the other, this quicker focus on manipulation of equations appeared to impact their performance. One of these students, for example, while solving Problem 2 (which was given as the last problem, after the other two problems), wrote down two electric potential energies: $EPE_A = 25V$ and $EPE_B = 0V$, even though the diagram provided contained an equation relating electric potentials ($V_A - V_B = 25V$). In the first problem, however, she was aware that electric potential and electric potential energy are different because she used the equation which relates electric potential to electric potential energy, $V = EPE/q_0$. In addition, she explicitly solved for the electric potential energy using this equation ($V = EPE/q_0$) and solved for the electric potential due to a point charge q at a distance r from that charge using $V = kq/r$. In this first problem, which did not provide a diagram, she correctly obtained a different equation for the electric potential energy ($EPE = kqq_0/r$) than that for electric potential. It is possible that this student proceeded to manipulate equations earlier in Problem 2 (which provided a diagram) because while solving Problem 2, this student did not spend sufficient time conceptually analyzing this problem. In particular, similar to four other interviewed students, almost immediately after reading this problem which included a diagram, she looked at the equation sheet and copied a formula on her paper and proceeded to solve the problem. Despite the fact that she had previously realized, while solving the first problem, that electric potential energy and electric potential are different, in Problem 2 she confused one with the other which resulted in an incorrect solution.

As mentioned earlier, the fact that most students looked at the equation sheet and copied an equation from it to their paper noticeably more quickly while solving the second and third problems (Problem 1 and 2 from this study) for which diagrams were provided than while solving the first problem (Additional problem) for which a diagram was not provided might be taken as an indication that these students were spending less time conceptually analyzing the problems when diagrams were provided. It is also possible that students were spending more time conceptually analyzing the first problem because it took them longer to recall the concepts which needed to be used for the first problem. We note, however, that in the six think-aloud interviews there was no noticeable difference and students spent about the same amount of time thinking about each of the three problems before writing down any equations. It is therefore possible that the longer time to recall the concepts in the first problem in which a diagram was not provided compared to the later problems in which diagrams were provided was due to a difference in the time for conceptual analysis and planning.

3.4.2 Qualitative results from discussions with faculty

To evaluate the opinions of instructors who had taught introductory physics frequently, we presented the three interventions for the two problems discussed here to seven physics faculty members and asked them to predict which group is likely to perform the best. Interestingly, some faculty members automatically assumed that the diagram would help and tried to answer the question "why would the diagrams help students" despite the fact that we asked them a neutral question about the group which is likely to perform the best. Also, similar to our original hypothesis, all seven faculty members incorrectly predicted that students in DO would perform the best because they were given explicit diagrams clarifying the situation. Some of them also mentioned that the second problem discussed here is more difficult than the first and that the given diagram should help more with the first problem than the second one because the first problem involves a situation with charges situated in two dimensions.

When the faculty members were told how the students actually performed, two of them recalled that they had observed in the past that providing a diagram had sometimes worsened student performance. Some of them mentioned that when they themselves solve a physics problem, they perform an initial conceptual analysis and often draw a diagram to make the situation clearer. Similar to our hypothesis, they noted that the absence of this important stage of problem solving when a diagram is provided to students can derail the entire problem solving process. Others noted that when a diagram is given, students may not read the problem statement carefully. Some claimed that for the first problem, students in DO were more likely to resort to a solution method involving force instead of energy because students are more likely to encounter diagrams with charges at the corner of a square or rectangle in problems involving the electrostatic force in books and homework problems. Furthermore, when the faculty members were explicitly asked whether their students would find any aspect of the diagrams confusing, their responses were negative. The disconnect between the faculty members' initial predictions about the usefulness of providing diagram and students' actual performance further suggests that the manner in which the cognitive processes of the novices was negatively affected by the given diagrams is quite complex.

3.5 DISCUSSION AND SUMMARY

Prior research has shown that students in classes which promote conceptual understanding through active-learning methods outperform students from traditional classes even on quantitative tests [3]. This finding suggests that students who perform poorly on physics problem solving may do so not because they have poor mathematical skills, but rather because they do not effectively analyze the problem conceptually. In particular, they may not employ effective problem solving heuristics and transform the problem into a representation which makes further decision making and consideration of relevant physics principles easier. For example, converting a physics problem from the verbal to the diagrammatic representation by drawing a diagram is a heuristic that can facilitate better understanding of the problem and aid in solving it.

One hypothesis for why students in DO who were given a diagram performed significantly worse than the other two groups is that, due to being provided diagrams, students in DO were more likely to skip or spend less time on the important step of conceptual analysis of the problem before implementing the plan for how to solve the problems. Therefore, they had difficulty in conceptualizing the problem and formulating a correct solution. The data in Table 3.4 suggest that students in DO on average performed significantly worse and more students in that group than in the other groups performed very poorly. The fact that many students who were given the diagrams failed to understand the problem conceptually is also evident from observing their individual solution strategies. For example, more students in DO than in the other groups explicitly employed formula-based approaches and it was unclear by observing their written work how they arrived at the decision to use those formulas (which were sometimes not productive for the problems). Interviews conducted with students who solved these problems (and an additional one, which did not include a diagram) while being observed by a researcher provided evidence that is consistent with this interpretation since most students (five out of eight) spent less time thinking about the problem conceptually when a diagram was provided compared to when it was not. Some of these students looked at the equation sheet almost immediately after reading the problem which provided a diagram, but in the problems which did not provide a diagram, they spent more time thinking about the problem first (presumably performing some sort of conceptual analysis or trying to understand the physical situation presented) before looking for a relevant equation to use. Therefore, it appears that both the qualitative and quantitative data presented suggest that providing a diagram can be detrimental to students' problem solving performance in these types of introductory physics problems involving considerations of initial and final situations.

As mentioned earlier, in Problem 2, even the students who did not draw a diagram from the comparison group (NS) performed better than the students who were given a diagram (DO). One possible reason may be that Problem 2 (actually, both problems discussed here) is not a difficult or multi-part problem requiring the use of many physics principles. Therefore, cognitive load theory [4] suggests that the cognitive load while solving the problem may not be high even if an explicit diagram is not drawn and algebra-based introductory physics students may be able to process all the relevant information in their working memory while engaged in solving the problem. Students' written work from the three groups also suggests that a higher percentage of students who were not provided the diagram went through an explicit process of making sense of the problem than the students provided with the diagram.

The interventions from this study were implemented in all the quizzes throughout the semester and a total of ten problems were analyzed. In only one other problem (the quiz problem discussed in section 2) did we find that students provided with a diagram performed worse than students in other groups. However, this particular problem involved the third harmonic of a standing wave in a tube open at only one end and students in DO were only provided with a partial diagram (empty tube) which was intended as a hint for them to complete it (draw the harmonic in question) and use the diagram to solve the problem. The two problems discussed here were the only ones in which providing students with a diagram similar to what most experts would draw to solve the problems resulted in deteriorated performance. What makes these problems special in this respect is unclear. However, we note that what these two problems have in common other than the fact that they require use concepts from electricity is that they both involve considerations of initial and final conditions. This latter characteristic was not present in any of the other problems we analyzed. It is also important to note that in none of the other problems analyzed did it happen that students provided with diagrams performed better than students in the other intervention groups. In fact, most of the time, they performed slightly worse. Furthermore, students who were asked to draw diagrams were almost always statistically more likely to draw productive diagrams (as defined from an expert's point of view) than students in the other intervention groups and usually performed slightly better. Therefore it appears that introductory physics students should be explicitly asked to draw diagrams while solving problems because this makes it more likely that they draw useful diagrams which could in turn prove to be a helpful step in getting students accustomed to using productive problem solving heuristics.

We also found that when physics faculty members were asked which intervention group is more likely to perform the best, some instructors automatically assumed that providing diagrams would help and attempted to answer the question of how and why they would help despite the fact that they were asked a neutral question. In addition, all of the instructors incorrectly predicted that students in DO would exhibit the best performance. This discrepancy between instructor predictions and student outcome suggests that the manner in which providing diagrams for these two problems which involve considerations of initial and final conditions affects students' performance is not at all intuitive and in fact quite complex.

3.6 CHAPTER REFERENCES

1. G. V. Glass and K. D. Hopkins (1996). Statistical Methods in Education & Psychology, (3rd ed.), Boston: Allyn & Bacon.

2. J. D. Cutnell and K. W. Johnson (2009). Physics (8th ed.), Wiley.

3. R. R. Hake (1998). “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses.” Am. J. Phys. 66(1), 64-74.

4. J. Sweller (1988). “Cognitive load during problem solving: Effects on learning.” Cog. Sci. 12, 257.

4.0 TO USE OR NOT TO USE DIAGRAMS: THE EFFECT OF DRAWING A DIAGRAM IN SOLVING INTRODUCTORY PHYSICS PROBLEMS

4.1 INTRODUCTION

For a literature review of previous research related to the role of multiple representations in problem solving, refer to the introduction in the study presented in Chapter Two.

In this research we investigate how prompting students to draw diagrams affects their performance in two electrostatics problems and how the performance is affected when students are provided with a diagrammatic representation of the physical situation described in the problems. We also investigate how the quality of a diagram affects performance and compare performance on identical problems dealing with electric force and electric field both immediately after instruction (quiz) and a few weeks after instruction (midterm) as well as performance on one-dimensional (1D) and two-dimensional (2D) electric force problems. Finally, think-aloud interviews were conducted with nine students who were taking an equivalent introductory algebra-based physics course at the time. The interviews provided support for some of the interpretations discussed and were helpful in identifying some difficulties students still exhibited after having learned the concepts of electric field and electric force and after having been tested on them in a midterm exam.

4.2 METHODOLOGY

For the quantitative part of the research, a class of 111 algebra-based introductory physics students was broken up into three different recitations. All recitations were taught in the traditional way in which the TA worked out problems similar to the homework problems and then gave a 15 minute quiz at the end of class. Students in all recitations attended the same lectures, were assigned the same homework, and had the same exams and quizzes. In the recitation quizzes throughout the semester, the three groups were given the same problems but with the following interventions: in each quiz problem, the first intervention group, which we refer to as "prompt only group" or "PO", was given explicit instructions to draw a diagram with the problem statement; the second intervention group (referred to as "diagram only group" or "DO") was given a diagram drawn by the instructor that was meant to aid in solving the problem and the third group, the comparison group, was not given any diagram or explicit instruction to draw a diagram with the problem statement ("no support group" or "NS").

The sizes of the different recitation groups varied from 22 to 55 students because the students were not assigned a particular recitation; they could go to whichever recitation they wanted. For the same reason, the sizes of each recitation group also varied from week to week, although not as drastically because most students ($\approx$ 80%) would stick with a particular recitation. Furthermore, each intervention was not matched to a particular recitation. For example, in one week, students in the Tuesday recitation comprised the comparison group, while another week the comparison group was a different recitation section. This is important because it implies that individual students were subjected to different interventions from week to week and we do not expect cumulative effects due to the same group of students always being subjected to the same intervention.

In order to ensure homogeneity of grading, we developed rubrics for each problem we analyzed and made sure that there was at least 90% inter-rater-reliability between two different raters. The development of the rubric for each problem went through an iterative process. During the development of the rubric, the two graders also discussed a student's score separately from the one obtained using the rubric and adjusted the rubric if it was agreed that the version of the rubric was too stringent or too generous. After each adjustment, all students were graded again on the improved rubric.

We analyzed two problems: the first problem is one dimensional and has two almost identical parts, one on electric field and the other on electric force. This problem was given both in a quiz (a week after learning about these concepts) and in a midterm exam (several weeks after learning the concepts). The second problem is a two dimensional problem on electric force which was given in a quiz only. The two problems and the diagrams (given only to students in DO) are the following:

<u>Problem 1</u>

Two equal and opposite charges with magnitude 10^{-7} C are held 15 cm apart.

(a) What are the magnitude and direction of the electric field at the point midway between the charges?

(b) What are the magnitude and direction of the force that would act on a 10^{-6} C charge if it is placed at that midpoint?

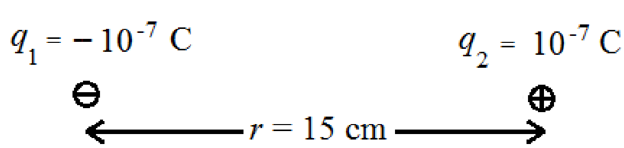


Figure 4.1. Diagram for Problem 1 given only to students in DO.

Problem 2

Three charges are located at the vertices of an equilateral triangle that is 1 m on a side. Two of the charges are 2 C each and the third charge is 1 C. Find the magnitude and direction of the net electrostatic force on the 1 C charge.

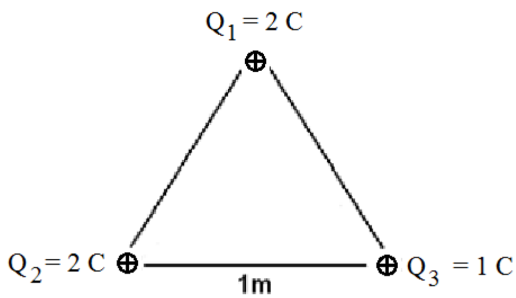


Figure 4.2. Diagram for Problem 2 given only to students in DO.

These diagrams were drawn by the instructor and they are very similar to what most physics experts would initially draw in order to solve the problems. Of course, subsequently they would most likely draw arrows to indicate the directions of electric field/force vectors. Neither diagram was meant to trick the students, but rather they were provided as a scaffolding support.

As mentioned earlier, we developed rubrics for each problem. For Problem 1, one research objective was to compare student performance on electric field with electric force. Therefore, parts (a) and (b) were scored separately. In Table 4.1, we provide the summary of the rubric for part (a) (electric field) of the first problem. The rubric for part (b) (electric force) is similar.

Table 4.1. Summary of the rubric for part (a) of Problem 1 ("E" stands for electric field).

Correct Ideas		
Section 1	Used correct equation for E	1 p
Section 2	Added the two fields due to individual charges correctly	7 p
Section 3	Indicated correct direction for net electric field	1 p
Section 4	Correct units	1 p
Incorrect Ideas		
Section 1	Used incorrect equation for E	-1 p
Section 2	2.1 Did nothing in this section	-7 p
	2.2 Did not find electric fields due to both charges	-6 p
	2.3 Used Pythagorean theorem (not relevant here) or obtained zero for electric field	-4 p
	2.4 Did not use *r*/2 to find E	-2 p
	2.5 Minor mistake(s) in finding E	-1 p
Section 3	Incorrect or no mention of direction of net electric field	-1 p
Section 4	Incorrect or no units	-1 p

Table 4.1 shows that there are two parts to the rubric: Correct and Incorrect Ideas. Table 4.1 also shows that in the Correct Ideas part, the problem was divided into different sections and points were assigned to each section. Each student starts out with 10 points and in the Incorrect Ideas part we list the common mistakes students made in each section and how many points were deducted for each of those mistakes. We note that it is not possible to deduct more points than a section has (the mistakes labeled 2.1 and 2.2 are exclusive with respect to all other mistakes in section 2 and with each other). We also left ourselves a small window (labeled 2.5) if the mistake a student made was not explicitly in the rubric.

In addition to the quantitative data collected, individual interviews were conducted with nine students who were at the time enrolled in a second semester algebra-based introductory physics course. During the interviews, students were asked to solve the problems while thinking aloud and, after they were finished working on the problems, they were asked short follow-up questions related to the physics concepts required for successfully solving the problems. The interviews provided qualitative data which supported some of our quantitative findings and helped us identify some student difficulties. These will be presented in the qualitative results section.

4.3 QUANTITATIVE RESULTS

4.3.1 Problem 1

The overall averages on the electric field and electric force questions in the quiz are comparable.

Table 4.2 lists the average score for each group in the two different parts when the problem was given in a quiz (one week after learning about electric field and electric force).

Table 4.2. Number of students (N) and averages on the two parts of the quiz for the two intervention groups and the comparison group out of 10 points.

Quiz	N	Field average	Force average	Problem average
PO	29	6.9	8.6	7.8
DO	40	7.5	6.6	7.0
NS	51	8.0	6.7	7.3
All students	120	7.4	7.2	7.3

We performed *t*-tests [1] on the data in Table 4.2 to determine whether there were any differences between the scores of the different groups on each part. PO performed better on the electric force part than both of the other groups (p values are 0.017 and 0.011 for comparison with DO and NS, respectively). However, the scores on the electric field part and the overall scores on the problem were not statistically different between the different groups. Moreover, the overall averages on the electric field and electric force parts for a given group were not statistically different. It appears that a week after learning the concepts of electric field and electric force, students show comparable performance on the problem (Problem 1) dealing with these two related concepts although the concept of field is more abstract than the concept of force.

Students in PO were more likely to draw productive diagrams.

We investigated differences between the groups resulting from the differences in instructions regarding diagrams (draw a diagram in PO, diagram given in DO, or no instructions in NS). We found that although all the students had a diagram drawn for this problem (some drew it themselves while others had it drawn for them) regardless of the instructions they received, those asked to draw a diagram (PO) were more likely to draw productive diagrams where we defined a productive diagram as follows. We considered that a productive diagram should have, in addition to the two charges, either two electric field or two electric force vectors or all four explicitly drawn at the midpoint. Any other diagram was considered unproductive (for example, a diagram containing just the two charges or diagrams containing the two charges and arrows drawn somewhere other than at the midpoint). It is worthwhile to note that students in DO were given a diagram containing the two charges (unproductive). We hypothesized that some students might modify it by adding vectors at the midpoint that indicate the directions of electric fields or electric forces in order to make it productive. Therefore, in addition to investigating the number of students who drew productive diagrams in each group, we also investigated the number of students in each group who had diagrams of only the two charges. The results are shown in Table 4.3.

Table 4.3. Percentages (and numbers) of students who drew productive diagrams ("Prod. diag.") and those who only drew two charges ("Only 2 charges") in each group in the quiz.

Quiz	Prod. diag.	Only 2 charges
PO	66% (19)	14% (4)
DO	45% (18)	48% (19)
NS	41% (21)	33% (17)

We performed Chi-squared tests [1] to investigate if the differences in Table 4.3 are statistically significant. The results are shown in Table 4.4 (Table 4.4 lists p values for comparison between groups; e.g., the value 0.036 under "PO-NS" for "prod. diag." means the p value for comparison of the percentage of students who used productive diagrams in PO and NS is 0.036).

Table 4.4. p values for comparison of percentage of students who drew productive diagrams ("Prod. diag.) with those who drew only the two charges ("Only 2 ch.") in the different groups in the quiz.

Quiz	PO-DO	PO-NS	DO-NS
Prod. diag.	0.092	0.036	0.190
Only 2 ch.	0.001	0.056	0.170

Table 4.4 shows that students in PO are statistically more likely to draw productive diagrams than students in NS. It also shows that students in DO are statistically more likely than students in PO to only use a diagram of the two charges. Students in DO were given this diagram so if they were using this type of diagram, they had not modified it into a productive diagram (Table 4.3 shows that almost half of them did not modify it to make the diagram provided productive). On the other hand, Table 4.3 shows that students in PO who were asked to draw a diagram were more likely to draw and use productive diagrams. Below, we provide evidence that for Problem 1, drawing a productive diagram improves students' scores. It appears that students in PO who were asked to draw a diagram performed significantly better (in the force part of the problem at least) perhaps because they were more likely to draw productive diagrams.

In the midterm exam, students performed better on the electric force part than on the electric field part.

Problem 1 was also given again in a midterm exam (several weeks after students learned about electric field and electric force). The three interventions implemented in the quiz were not implemented in the midterm exam and all students received the same instructions corresponding to NS in the quiz. The performances of students in different groups (defined earlier for the quiz intervention) were comparable in the midterm exam. Therefore, we only provide the overall midterm averages including all students in Table 4.5.

Table 4.5. Number of students and averages on the midterm exam on the two parts of Problem 1 out of 10 points.

Midterm	N	Field average	Force average	Problem average
All students	120	7.2	8.8	8.0

Comparison of the quiz and midterm exam performances (shown in Tables 4.2 and 4.5) shows that the average on the electric field part of the problem did not improve. In the quiz, the overall average on electric field was 7.4 (see Table 4.2) and in the midterm it was 7.2 (see Table 4.5). However, the average on the electric force part of the problem improved significantly from the quiz (7.2 – see Table 4.2) to the midterm exam (8.8 – see Table 4.5). A *t*-test reveals that the score differences between the quiz and the midterm exam on the electric force problem are statistically significant ($p < 0.001$). Thus, the performance on the electric field part of the problem remained stagnant from the quiz to the midterm exam while there was a significant increase in the performance on electric force. Furthermore, fewer students in the midterm exam than in the quiz used the connection between electric field and force, namely $\vec{F} = q\vec{E}$, which is an efficient method for calculating the force on a point charge at a point once the electric field at that point due to all the other charges has been calculated. The percentage of students who used this connection went down from 58% for the quiz to 41% for the midterm exam, a difference that is statistically significant ($p = 0.008$). In contrast, all introductory physics instructors who have been asked to solve or comment on this problem have noted that $\vec{F} = q\vec{E}$ should be used to find the force on the charge after the field at the point has been calculated.

Both in the midterm exam and the quiz, students who drew productive diagrams performed better on Problem 1 than those who did not.

We stratified all the students into three categories based on the quality of their diagrams and analyzed their scores. A lower category corresponds to a lower quality diagram. The results are shown in Table 4.6. The different levels of diagram quality in Table 4.6 are: Diagram Quality 1 (DQ1 in Table 4.6) is an unproductive diagram, Diagram Quality 2 (DQ2) is a diagram which includes either two electric field or two electric force vectors at the midpoint, but not both and Diagram Quality 3 (DQ3) is a diagram which includes all four vectors (corresponding to both field and force) at the midpoint.

Table 4.6. Numbers of students (N), averages and standard deviations for groups of students with different quality diagrams for problem 1.

Quiz	**N**	**Average**	**Standard deviation**
DQ1	62	6.4	2.6
DQ2	49	8.3	2.2
DQ3	9	8.9	1.4
Midterm	**N**	**Average**	**Standard deviation**
DQ1	45	7.0	2.6
DQ2	51	8.4	2.0
DQ3	25	9.0	4

We also performed *t*-tests on the data in Table 4.6 to compare the performance of students who had different categories of diagram quality. The results are shown in Table 4.7, which lists the p values obtained when comparing the performance of students from different categories (which are defined above). For example, the first value in Table 4.7 ($p<0.001$) is the p value comparing the performance of the DQ1 group of students (students who drew an unproductive diagram) with the performance of the DQ2 group of students (students who drew either two electric field or two electric force vectors at the midpoint, but not both). Table 4.7 shows that students who drew productive diagrams (DQ2 and DQ3) performed better than those who did not (DQ1).

Table 4.7. p values for comparison of the performance of students with different quality diagrams (the categories are defined in the text right before Table 4.6) for Problem 1.

	DQ1-DQ2	DQ1-DQ3	DQ2-DQ3
Quiz	< 0.001	< 0.001	0.284
Midterm	0.003	< 0.001	0.133

4.3.2 Problem 2

A higher level of detail in a student's diagram corresponds to a better performance.

For Problem 2, a two dimensional (2D) problem given in the quiz only, there were no statistically significant differences between the different intervention groups (PO, DO and NS), both in terms of scores and in terms of percentages of students drawing productive diagrams. One possible explanation for this result is that Problem 2 is two dimensional and it is very difficult (one might say even impossible) for a novice to solve correctly without the use of a productive diagram which would at least include the directions of the individual electric forces acting on the 1C charge due to each of the other charges (before finding the net electric force). Therefore, students in all groups were more likely to draw productive diagrams in order to help them solve this problem regardless of the instructions they received involving diagrams. On the other hand, we found that there was a correlation between the level of detail in students' diagrams and their performance.

We stratified the students based on three categories of diagram quality and analyzed their scores. Diagram Quality 1 (DQ1 in Table 4.8) corresponds to diagrams with just the three charges, Diagram Quality 2 (DQ2 in Table 4.8) corresponds to diagrams with the three charges and the two forces acting on the 1C charge and Diagram Quality 3 (DQ3 in Table 4.8) corresponds to diagrams with the three charges, the two forces acting on the 1C charge, and the *x* and *y* components of those forces. Since students were explicitly asked to indicate the direction of the net force acting on the 1C charge, whether or not a student drew a vector for the net force was not taken into consideration when determining the different levels of diagram quality. We only took into consideration levels of detail that the students themselves thought would help them solve the problem, not what they were explicitly asked to draw. The results are shown in Table 4.8.

Table 4.8. Number of students (N), averages and standard deviations for students in different categories (by diagram detail) for Problem 2 which was given in a quiz.

	N	Average	Standard deviation
DQ1	27	4.1	2.5
DQ2	58	5.7	2.9
DQ3	33	8.0	2.2

Table 4.8 shows that there is a correlation between the level of detail in the diagrams drawn and the score: a higher level of detail corresponds to a better score. We performed *t*-tests on the data in Table 4.8 and found that students who, in addition to the three charges, drew two force vectors (DQ2), outperformed the students who only drew the three charges (DQ1) (p =

0.008). Similarly, students who drew the two forces due to individual charges and their *x* and *y* components (DQ3), outperformed students who drew only the two forces (DQ 2) ($p < 0.001$). The p values for these comparisons are quite small and the differences between the averages of the groups are quite noticeable. Students with the highest level of detail performed better than students with the lowest level of detail by almost 100%!

4.4 QUALITATIVE RESULTS FROM INDIVIDUAL STUDENT INTERVIEWS

In order to investigate related student difficulties in more depth, individual interviews with nine students who were at the time taking an equivalent second semester of an introductory algebra-based physics course were carried out using a think aloud protocol [2]. Five of these interviews were conducted one week after the second exam, which covered the material required for the two problems. The other four were carried out after the third exam, which covered material from Magnetism. As mentioned before, during the interviews students were asked to solve the problems while thinking aloud and, after they were finished working on the problems to the best of their ability, they were asked for clarifications and short follow-up questions related to the physics concepts which needed to be used in order to successfully solve the problems. Several related student difficulties were understood in more depth and sometimes uncovered during these interviews. The results from the interviews related to Problem 1 will be discussed first.

4.4.1 Qualitative results related to Problem 1

1) Students encountered more difficulties with the concept of electric field than with the concept of electric force.

This finding supports the quantitative results presented earlier which indicated that students performed worse on the concept of electric field than on the concept of electric force a few weeks after learning about these concepts on the midterm exam. John's interview provided a very prominent instance of the discrepancy between facility with electric force compared to electric field. In the electric field part of the problem, John only included the contribution to the net electric field from one charge and was unable to determine the direction of the electric field even due to that charge. However, in the electric force part of the problem, he readily recognized that two interactions would affect the net force on the charge placed at the midpoint, and then explicitly reasoned that these interactions would cause equal forces on the charge in the same direction (left). After this correct reasoning, he found the magnitude and direction of the net force. He did make one minor mistake, however, in that he included a negative sign in the magnitude of the net force, which may be because he was trying to indicate direction in his numerical answer (i.e., this student was thinking that a negative force points to the left, in the negative x direction and a positive force points to the right, in the positive x direction).

Another student, Karen, in the electric field part of the problem, identified two contributions to the net electric field (using the equation $E = kQ/r^2$, plugged in the distance between the charges and the midpoint for r and added them together, but was visibly unsure about the reason because when she added the two contributions she said "you … plus them" in a questioning tone of voice (the "…" indicates a short pause). After she finished the problems, she was asked why she decided to add the two contributions she found and she said:

> *Karen: Cause I thought they were moving towards each other [she meant that the positive and negative charges are attracted to one another], so then I thought that the E should be added […] if this one [the negative charge] was a positive then they'd be moving away, and then it would be subtracting.*

Her reasoning is related to electric forces, not electric fields; she also did not mention the fact that she was calculating the electric field at the midpoint. In a nutshell, what she said was that the contributions to the net electric field that she found should be added because the charges attract one another. If she had to find the electric field on the extended straight line joining the two point charges at a point not between the charges but where both charges were on the same side of the point, her reasoning would yield an incorrect answer for the net electric field (because the contributions to the electric field due to the two charges with opposite signs at that point would subtract). Discussions with her and other interviewed students suggest that this type of difficulty is also related to the fact that students often do not clearly differentiate between the concepts of electric field and electric force; they use these words interchangeably (i.e., say "electric field" when they mean "electric force" and vice versa) and, in some cases, completely mistake one concept for the other. This difficulty is discussed in more detail later in this section. In contrast, in the electric force part of the problem, she encountered no difficulties and solved that part of the problem correctly along with providing sound reasoning. She also indicated the correct direction for the net force.

As mentioned earlier, during the think aloud interviews, after students were finished with the problems to the best of their ability, they were asked for clarifications and short follow-up questions related to the concepts required to solve the problems. Some later questions from the interviewer asked about simpler situations than the ones presented in the problems or presented modified versions of the problems and asked students if and how their solution/answer would change. For example, if a student only included one contribution to the electric field (due to one of the two charges only) in solving part (a) of the problem, he/she was asked if his/her answer would change if one of the charges was removed. Other questions probed for reasoning because, often, the reasoning presented by students during the think aloud process was either unclear or incorrect. Karen, for example, was asked why she added the two magnitudes she found, as noted earlier, because it was evident that she was unsure and she did not provide reasoning when she added the two contributions. After students finished solving the two problems to the best of their ability (and struggled to solve them), some students were asked to look back at one of the problems and the interviewer would ask directed questions which were intended to provide scaffolding support (examples will be provided below) and, in addition, help the interviewer understand their problem solving strategies better. Through directed questioning, we found that the students interviewed either found it less challenging to remember (or be reminded of) electric force concepts than electric field concepts or, while unable to grasp the electric field concept despite the hints provided, managed to reason correctly about the concept of electric force (as exhibited by their understanding of how these two problems can be solved). For example, after directed questioning, some students still did not realize that two contributions to the net electric field must be considered due the two charges for Problem 1, but almost immediately recognized this for the electric force when given a modest hint. Other times, if directed questioning was successful in assisting them to recognize that two charges would contribute to the field and force in the two parts of Problem 1, it would take more directed (scaffolded) questioning in the electric field part than in the electric force part.

One good example comes from an interview with Alex. Both in the electric field and in the electric force parts, Alex had only considered one contribution during the initial phase of the problem solving process while thinking aloud. The discussion that followed probed why he only considered one of the two contributions to each (after he was finished with both problems). Below, we give an example from the discussion between the interviewer and Alex in which the interviewer asks Alex about the electric field part of the problem, then moves on (when indicated) to the electric force part of the problem.

Interviewer: If I remove this charge [the negative charge] would anything change in your answer?

Alex answers "No", but then provides an explanation which seems to contradict this answer:

Alex: It's still emitting the same amount of electric field from the one [positive charge] which decreases the farther you get away, but also increases with the other [charge], so it would still be equal.

While his reasoning is not entirely clear, it appears that his explanation includes contributions from both charges. Therefore, the interviewer probed further:

Interviewer: So even if I remove this charge [the negative one], the magnitude of the electric field at the middle would be the same?

Alex: Yes.

Interviewer: What if I change the sign of this [the negative charge] from negative to positive?"

Alex: Then it would double […] cause you would add the two together.

Interviewer: So there's two electric fields and you would add them together?

Alex: Yes.

Interviewer: And in this situation where it's negative [the initial problem situation] you have just one electric field?

Alex: Well it [the electric field] is only going in one direction […] if two positive charges were there, they [the two electric fields] would both be facing away from each positive charge so there would be two different fields.

He then added further that if he was asked to find the electric field in this new situation he would simply multiply the contribution he found for one charge by two.

Interviewer: Why would you multiply by two?

Alex: Cause if it's in the middle then it's gonna get exactly the same electric field from both.

Thus, Alex did not include the direction in determining the net electric field (if one considers the direction of the electric field due to each charge correctly, if both charges were positive and equal, the electric field midway between the charges will be zero because the contributions due to the two charges are in opposite directions). Interestingly though, it can be inferred from Alex's answers that, at least to some degree, he realized that the electric field has a direction. Alex was then asked scaffolding questions intended to help him understand that both the positive and negative charges affect the net force on the middle charge.

Interviewer: So again, suppose I remove this [the negative charge]…

Alex: Yea, in that case, then it would decrease because it's only getting pushing [force from the positive charge] instead of pushing and pulling force.

Interviewer: Is the fact that it's both being pushed and pulled reflected in your answer here?

Alex: [After a short pause] No.

Interviewer: Why not?

Alex: Cause the equation only accounts for two charges and I did not know how to incorporate a third.

Interviewer: Would you know how to do that now?

Alex: Yes, I guess I would just calculate the force from one on that charge and then also add the force from the other.

Interviewer: Why add?

Alex: Cause the forces are in the same direction.

Interviewer: And what would happen if I changed the negative charge to be positive?

Alex; Then the forces would be opposing and if it was in the center, they'd be equal and they would cancel.

It is evident from this short discussion with Alex (and from the rest of the interview with him), that he had a much more difficult time understanding electric field than electric force. The guiding questions are almost identical (“If I remove this negative charge, would anything change in your answer?”, “What if I changed the negative charge to be positive?” etc.), and while these questions were not sufficient in helping him correctly interpret electric field in this context, they were quite successful in guiding him to grasp the concept of electric force while expressing correct reasoning.

Another example that points to the difficulties students have with the concept of electric field comes from Sam’s interview. Sam was the only student who solved Problem 2, which asked for the electric force by finding the net electric field at the right corner of the triangle (position of 1C charge) and then multiplying this field by the 1C charge at that vertex of the triangle. Sam found the net electric field without considering the fact that electric field is a vector. She did not consider adding the two electric field vectors due to each of the 2C charges by choosing a coordinate axis, using the *x* and *y* components of the electric field and then adding them vectorially. The Interviewer then asked her a few directed questions intended to help her realize that the electric field has a vectorial nature which must be taken into account when adding or subtracting two or more electric fields. After acknowledging the vector nature of the electric field, in response to why she had not used the vector nature of the field initially, she noted:

Sam: I didn’t even think about vectors, I just saw field and I was like, oh, field, another field.

As she said “field” and “another field” she drew two ‘fields’ (reproduced in Figure 4.3) that looked like sinusoidal travelling waves emanating from each 2C charge reaching the vertex where the 1C charge was located.

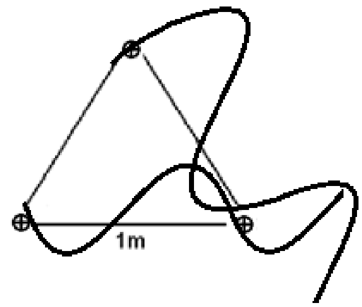


Figure 4.3. Two 'fields' emanating from the two 2C charges as drawn by a student, Sam.

It can be inferred from Sam's comments and the picture she drew that she had a mental representation of electric field but her mental model was not adequate to prompt her that electric field is a vector while solving the problem involving electric field. Several other interviewed students also had mental models of field which were ineffectual in guiding them to solve the problems involving electric field appropriately and relate the concept of electric field with the concept of electric force.

2) Students had difficulty differentiating between the concepts of electric field and electric force

This difficulty was most evident when students had to determine the direction of the net electric field, or determine whether the two contributions to the net electric field at the midpoint due to the individual charges should be added. The example mentioned earlier from Karen's interview suggests this difficulty. She struggled to differentiate between the concept of electric field and electric force and answered the question "Why did you add them?" (the two contributions to the net field that she found due to the two charges) with reasoning directly related to the electrostatic attraction between the two charges. Not only did her reasoning about the problem related to field not mention “electric field”, but it also did not mention the midpoint where she had to calculate the field.

Tara had a similar difficulty in distinguishing between the electric field and the electric force and sometimes used them almost interchangeably in her explanations. Initially, she had not indicated a direction for the net electric field. Therefore, when she was explicitly asked about the direction, she noted:

Tara: This [the midpoint] would end up being a positive charge, so it’s gonna want to go that way, so to the left […] cause opposites attract and like repel.

The interviewer wanted to ensure that she was not mentally placing a positive test charge at the midpoint and using that to determine the direction of the electric field via its definition ($\vec{E} = \frac{\vec{F}}{q_0}$) so he asked her: “and this would give you the direction of electric *field*?” (stressing the word “field”). In response to this explicit question about the direction of electric field, Tara acknowledged that her reasoning was related to electric forces, not electric fields and that she did not know how to find the direction of the electric field.

Further, the interviewer asked her if she could indicate the direction of the electric field produced by a positive charge to the right of that charge (which is a simpler problem with only one charge). The interviewer drew a positive charge and indicated a specific point to the right of that charge. Tara could not answer this question, but she remembered that a charge should be placed at that point in order to find the electric field at that point. She got confused because when she placed a positive (test) charge at that point, she concluded that the electric field should point to the right but when she placed a negative (test) charge at the same point, she concluded that the electric field should point to the left. She gave up because she obtained different directions for the electric field depending upon whether the charge placed at that point was positive or negative. What she did for the negative (test) charge would give her the direction of the electric force on the negative charge, and not the electric field because the electric force and electric field point in the opposite directions for a negative point charge. Unfortunately, she was unable to disentangle the concepts of field and force.

Another illustrative example of the confusion between the concepts of field and force comes from Megan's interview. When solving for the electric field at the midpoint between the charges in Problem 1 she drew two vectors on the diagram that are reproduced in Figure 4.4.

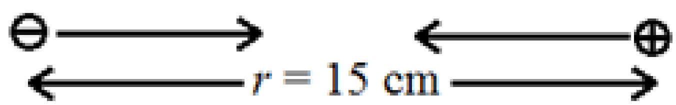


Figure 4.4. Diagram drawn by a student, Megan, when solving part (a) of Problem 1.

Then she said *"I'm thinking that the electric field on q_1 due to q_2 would be kq/r^2"* after which she plugged in r = 15 cm instead of 7.5 cm. Not only did she draw vectors that indicated electrostatic forces between the two charges, she also used the distance between the charges rather than the distance from one charge to the midpoint. Also, the language she used throughout including "the electric field <u>on</u> q_1 <u>due to</u> q_2" indicated that she was thinking about electric forces and not realizing the difference between electric force and electric field.

As mentioned earlier, even after some questioning (e.g., would your answer change if a charge was removed?, what if the negative charge is replaced with a positive charge? etc.), Alex was unable to realize that two individual contributions to the net electric field due to the two charges must be considered. The interviewer then explicitly asked Alex if he was aware of a connection between electric field and electric force. He noted that he knew there was a connection, but did not remember it. The interviewer provided the equation ($\vec{E} = \vec{F}/q_0$) and explained that the electric field at a point is the force that would act on a small positive test charge placed at that point divided by that charge. The following is an excerpt from a discussion after this:

Interviewer: If you look at this equation [$\vec{E} = \vec{F}/q_0$], does this cause you to change your answer here [pointing to part (a) of the problem] cause you included one electric field?

Alex: Yes. (At this point, Alex was still thinking about the case when both charges are positive because he was previously asked what would change if the negative charge was replaced by a positive charge.)

Interviewer: Why?

Alex: Because the two electric fields would cause opposing forces meaning they would cancel instead of add.

Interviewer: In this situation where it's – and +?

Alex: Oh, no, not in the – and + …

Interviewer: In the + and +?"

Alex: Yes.

Interviewer: What about in the – and +?

Alex: It would be double because they'd both exert a force so the field would be twice as much as that.

Further discussions also suggest that after reminding Alex about the connection between electric force and electric field, with minimal questioning, he was able to use it correctly to reason about electric field. However, Alex was a student who was quite concerned with conceptual understanding (a fact that was identified after a later discussion), and in general, an above average student (obtained an A in the first semester algebra-based physics class). Other interviewed students had a more difficult time using this connection correctly despite being reminded of it and required more scaffolded questioning.

For example, when asked about it, John said that the direction of the electric field at some point away from a negative charge should be towards the charge because "charges flow from positive to negative". Below, a part of the conversation between the interviewer and John during the interview is reproduced.

Interviewer: The definition for the electric field is $\vec{E} = \vec{F}/q_0$*. Do you remember anything about that? Would that cause you to change that direction [towards the positive charge]?*

John: Ok, so, E equals F over q, so then the electric field would be in the same direction as the force, so the field would go away from this charge.

He then drew electric field lines emanating away from the positive charge.

Interviewer: Why is that?

John: If F over q is the electric field and this [the q in the formula] is a positive charge, so then the electric field would move in the direction of the force and this [the positive charge drawn on the paper] would apply a positive force outward.

Interviewer: Apply a force on what?

John: If there was another charge, it would apply a force outward.

He was then asked if he could answer the question about the direction of the electric field. He got quite confused and was unable to answer the question. The interviewer attempted to help him again by pointing to his picture of the electric field emanating from a positive charge and asking him how it would look like for a negative charge. He correctly answered that question and drew electric field lines pointing inward towards the negative charge.

Interviewer: So if the direction of electric field due to negative charges is towards those charges and I was looking at this point here [middle]…

John: Oh, so then the whole field would be towards the negative charge.

Interviewer: Yes and why is that?

John: […] since it's going away from a positive and towards a negative, the whole field would be going that way [left].

These discussions with John and those with other students suggest that in order to help students, the questions must be framed to take advantage of their knowledge resources and help them focus on important information in order to apply their existing resources and additional information provided appropriately (e.g., in John's case the statement "If the direction of electric field due to negative charges is towards those charges and *I was looking at this point here* [middle]" was revealing to him). Without appropriate scaffolding tailored to take advantage of students' knowledge resources, they may have a difficult time realizing what is relevant and should be considered, and therefore focus on information that is neither relevant nor helpful.

4.4.2 Qualitative findings related to Problem 2

1) Cognitive load theory may possibly explain why students who explicitly draw the components of the two forces perform better.

Two of the students interviewed were almost identical in terms of their majors and grades (both in the current physics course and the previous one). Karen and Dan were both Biology majors; in the first semester of physics they both obtained similar grades (B+ and A-, respectively). In the second semester physics class, in the first exam (class average 75/100), they both obtained 81/100 and in the second exam (class average 65/100) they also both obtained 81/100.

When solving the second problem, Karen recognized that she needed to find the *x* and *y* components of both forces due to each of the 2C charges and, before she proceeded to find them, she drew all the components on the diagram provided as shown in Figure 4.5.

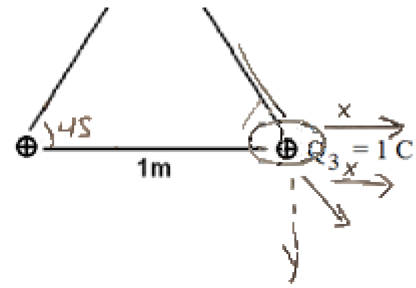


Figure 4.5. Forces due to the two individual charges on the 1C charge and their components as drawn by Karen (student).

She then figured out all the components and combined them correctly to determine both the magnitude of the net force and its direction (angle below the *x* axis). While working on this problem, it was evident that Karen was focusing on only a few things at a time and was being systematic about the way in which she found the net force. For example, when finding the components of the oblique (not horizontal) force, she redrew a triangle in which this force was the hypotenuse and identified the angles. Karen's only mistake was using an angle of 45° instead of 60° to find these components.

Dan also immediately recognized that components should be considered and proceeded to find them after redrawing the 1C charge (see Figure 4.6) and the two forces acting on it due to the 2C charges. He worked more slowly than Karen on this problem, but after some time, he correctly determined the *x* and *y* components of the oblique force and wrote them down (trigonometric functions were still included, i.e., he wrote down the *y* component as $18 \times 10^9 \cos 30$). However, unlike Karen, he did not draw these components on his diagram; his diagram of the forces (shown in Figure 4.6) only included the two forces and their magnitudes.

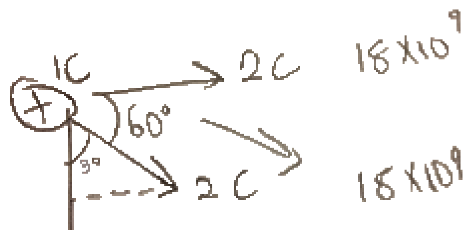


Figure 4.6. Forces acting on the 1C charge due to the two 2C charges as drawn by Dan (student)

When Dan combined the components, he made two mistakes: 1) his net *y* component did not include the trigonometric function which he had previously written down (when he found the *y* component of the oblique force). As he was figuring out the net *y* component he said: "this one [horizontal force] doesn't have a *y* component, so it [the *y* component of the net force] is just 18 times 10^9" and 2) he subtracted the *x* components instead of adding them (he subtracted the horizontal force from the *x* component of the oblique force). In particular, he wrote the following on the paper for the net *x* component: Net x = $18 \times 10^9 \sin 30 - 18 \times 10^9$. It is possible that part of the reason why he subtracted the components is because he didn't explicitly draw the *x* component of the oblique force and perhaps, due to the fact that the oblique force is in the fourth quadrant (which should be dealt with carefully), he implicitly assumed that one of its components must be negative, or that something must be subtracted. He subtracted the horizontal force from the *x* component of the oblique force even though the picture he drew clearly indicated that the horizontal force is in the positive *x* direction. After he finished working on all problems to the best of his ability, in the second phase of the interview, he was asked for clarifications of points he had not made clear earlier and some additional questions. For example, Dan was asked a simpler question. He was asked to add two forces: one in the positive *y* direction, the other in the first quadrant, making an angle of 30° with the horizontal. Here too, he didn't draw the components explicitly in the diagram and ended up subtracting the *y* components of the two forces in exactly the same way in which he subtracted the *x* components in Problem 2 (the triangle problem) i.e., he subtracted the vertical force from the *y* component of the oblique force. When asked why he subtracted these components he looked at the diagram for a few seconds and said:

Dan: Actually, you're adding […] sorry, I don't know why [he was going to say 'I don't know why I did that'] […], you're adding because there's a positive y component here [vertical force] and a positive y component here [of the oblique force].

The approaches of these two students differed mainly in that Karen explicitly drew all forces and components, whereas Dan only drew the forces. Dan subtracted the *x* components without providing a reason, and when he was asked to add two forces in a simpler mathematical context, he made the same exact mistake for the two components that were supposed to be added. When questioned about why he subtracted them, he realized this mistake on his own almost immediately, which suggested that when he solved both problems (Problem 2 and the simpler mathematical problem which had similar addition of vectors) he wasn't focusing on the appropriate information. Once his attention was drawn to the issue of whether the vectors should be added or subtracted in the simpler mathematical problem, he clearly knew that the *y* components must be added. Without being questioned, he did not draw the components of the oblique force and appeared to be subtracting the components automatically, without a clear reason. When asked why he subtracted the components, he did not start by trying to justify this (for example by beginning a sentence with "I subtracted them because…"), which suggested that there was no clear reason for why he subtracted the *y* components. Further discussions with Dan suggest that since he had not explicitly drawn the components of the oblique force along the x and y directions in his diagram, he was keeping the information about the components in his head. However, when it was time to utilize this information about the components of the oblique force to find the x component of the net force, he forgot to correctly account for the x component. On the other hand, Karen had the components explicitly drawn on the paper as opposed to keeping this information in her head and she was able to look back at her components and account for the sign of the x component of the oblique force correctly. Cognitive load theory [3], which incorporates the notion of distributed cognition [4], provides one possible explanation for Dan's unsuccessful and Karen's successful addition of vectors in this context: lack of information about components on Dan's diagram required him to keep this information in his working memory, while Karen did not need to keep this information in her working memory since she included the components explicitly in her diagram. As Dan's working memory was processing a variety of information during problem solving, he may have had cognitive overload and the information about the components that he planned to use at the opportune time to find the components of the net force was not invoked appropriately.

2) Most students (six of the nine interviewed) possessed correct mathematical knowledge about adding non-collinear forces that they did not use while solving Problem 2.

Five out of the nine students interviewed solved Problem 2 without considering the vectorial nature of forces. They found the magnitudes of the two forces and added them like scalars. One student arbitrarily multiplied both forces by sin45° and later, when asked about why she did that, said "it's on an angle, there's a triangle". When pressed a little further about this issue it was clear that there was no well thought out reason for doing this except because "it's on an angle". All these six students were asked a follow-up question that required them to add two forces that were non-collinear. They were given a diagram like the one shown in Figure 4.7 (not all diagrams were identical; sometimes there was a vertical force instead of horizontal, but there was always an oblique force at a given angle from the horizontal).

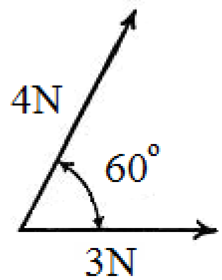


Figure 4.7. Example of a follow-up question used to assess whether students had mathematical knowledge about adding vectors that they did not use in Problem 2.

All the six students who had not considered adding forces by components while solving Problem 2, in this context similar to that shown in Figure 4.7, invoked and proceeded to find horizontal and vertical components of the oblique force. Some were successful, some were not, but all of them understood that these two forces must be added by finding their *x* and *y* components. In Problem 2 on the other hand, none of these six students used this systematic approach. The interviews suggest some possible reasons for why students used components in the "mathematical" context but not in the physics context (quotations are used because forces which are inherently related to physics are being added, but the question itself is not necessarily a physics question per se since no physics specific knowledge is required in order to solve it). The interviews suggest that in Problem 2 (physics context) some students added the magnitudes of the two forces instead of considering vector addition because both forces "push" the 1C charge away from the configuration. They did not consider directions because they did not apply mathematical knowledge to find the magnitude of the net force; instead, they used intuition or gut feeling. John's interview provides an interesting example of how, for some interviewed students, knowledge of components was retrieved correctly even in the physics context for predicting the direction of the net force but not applied for finding the magnitude of the net force. When trying to determine the direction of the net force he said:

Since these [the two forces acting on the 1C charge] are both coming in at the same angle, I believe that one of the components will cancel out […] so it will be […] splitting it in half.

What he meant by "splitting in half" is that the 60° angle will be bisected by the line along which the net force will be pointing. He then drew a vector that indicated the direction of the net force which bisected the 60° angle at the 1C charge's corner as reproduced in Figure 13.

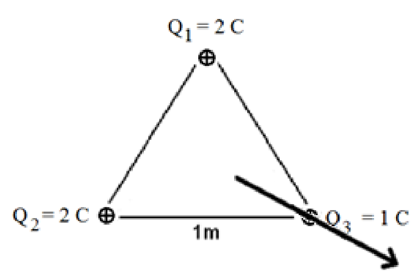


Figure 4.8. Vector indicating the direction of the net force on the 1C charge drawn by John (student).

What John did in order to figure out the direction of the net force on the 1C charge (which is indeed correct) is he used symmetry. For example, if the triangle is placed with the 1C charge at the top, the *x* components (horizontal) cancel out and the net force points in the +*y* direction (vertical) and the 60° angle at the top will be "spilt in half" in John's language by the line along which the force will point. It is very interesting that although John was able to use symmetry to determine the direction of the net force which involved reasoning related to components ("one of the components will cancel out"), he did not use information about the components of the forces due to each of the 2C charges to determine the magnitude of the net force (he added them like scalars). Therefore, one can argue that knowledge about how to add components of forces to find the net force was present in John's memory ("components cancel"), and was used correctly when John determined the direction of the net force, but was not used at all when John determined the magnitude of the net force (added the two forces like scalars).

Another possible reason suggested by the interviews for why students used components correctly to find the net force in the mathematical context but not in the physics context (Problem 2) is that students may be prompted to use components because the angle is drawn explicitly on the figure for the mathematical problem but not for Problem 2. This triggering of proper protocol for adding vectors using components when an angle is explicitly given in the diagram might be due to the fact that when learning addition of vectors, many of the problems used (both in mathematics and physics) explicitly contain angles. For example, in the interview conducted with Megan, when asked why she considered components to find the net force for the mathematical problem she said:

Megan: I just remember, we had to find components of forces before, so, I saw angle and assumed, maybe we'd have to do that.

4.5 DISCUSSION AND SUMMARY

We found that for Problem 1, students who were explicitly asked to draw a diagram were more likely to draw a productive diagram. We also found that students who drew productive diagrams performed better than those who drew unproductive diagrams. Among the students in DO who were provided with a diagram (which was unproductive unless modified by the student by adding force and/or field arrows) less than half attempted to draw the arrows, which is statistically significantly lower than the fraction of students in PO who were not provided any diagram and explicitly asked to draw one. This finding suggests that in an algebra-based introductory physics course the intervention for PO is likely to provide better scaffolding for solving problems than that for DO and should be incorporated in helping students learn effective problem solving strategies.

We also found that more detailed diagrams (in general a more detailed diagram is also a higher quality diagram) corresponded to better performance. In a previous investigation [5] related to free body diagrams and their impact on student performance, Rosengrant [5] found that only drawing correct FBDs improves a student's score and that students who draw incorrect FBDs do not perform better than students who draw no diagrams. In the study presented here, the correctness of the diagrams (correctness of the vector arrows representing electric fields or forces) did not impact students' scores significantly. It is possible that the reason for this difference between the two studies is that for both problems in this study, students who drew incorrect vector arrows in the diagrams had the incorrect direction of electric field or electric force vectors due to both charges. These students differ from the students with the correct direction in that they obtained a direction for the net electric field or force which was 180º from the correct direction. Referring back to Table 4.1, an incorrect direction would only cost a student 1 out of 10 points because partial credit was given. On the other hand, Rosengrant's study included multiple choice problems only and involved some problems in which students were often completely mistaken about the direction of the force (e.g., in an inclined plane problem, some students claimed that the normal force on the block is collinear with the gravitational force as opposed to perpendicular to the incline). Since some answer choices to the multiple choice problems in their study were based on common student errors, one would expect that the correctness of the diagrams would make a larger difference in Rosengrant's study than in ours.

As noted earlier, one theoretical framework that can provide a possible explanation for why students with more detailed diagrams performed better is the cognitive load theory [3,6-11], which incorporates the notion of distributed cognition [4]. In Problem 2, students had to add forces by using components, so students who did not draw the force vectors or their components they had to add vectorially would have to keep too much information in their working memory while engaged in problem solving (individual components of the two forces, angles required to get those components, what trigonometric function needs to be used for each component, etc.). This can lead to cognitive overload and deteriorated performance. Explicitly drawing the forces and their components can reduce the amount of information that must be kept in the working memory while engaged in problem solving and may therefore make the student more able to go through all the steps necessary without making mistakes. As noted earlier, individual think-aloud interviews conducted with two students who were nearly identical in terms of performance on class examinations also suggested that this interpretation may be appropriate. When solving Problem 2, both of these students were able to determine the x and y components of the two forces correctly. However, the student who did not explicitly draw the components in the diagram made two mistakes when combining these components. Furthermore, the interview suggested that his incorrect choice of subtracting the *x* components (instead of adding them) was not done for any particular reason because, when explicitly asked to explain this choice and why he subtracted those components, he immediately realized that the two components must be added. This can be interpreted as an indication that this student, although aware of how components must be combined, had cognitive overload and did not retrieve the information about components from his working memory appropriately while engaged in problem solving.

It is also important to note that these problems were given in the second semester of a one year introductory physics course for algebra based students. These students had done problems for which they had to find the net force in Newtonian mechanics, and still less than 30% of the students realized that they should draw the components of the electric force in Problem 2 presented here. Also, only 42% of all students indicated a direction for the net force. This can partly be an indication of a lack of transfer from one context to another [12-18]. Students' performance also suggests that many algebra-based introductory students do not have a robust knowledge structure of physics nor do they employ good problem solving heuristics and their familiarity with addition of vectors may also require an explicit review. Earlier surveys have found that only about 1/3 of the students in an introductory physics class have enough knowledge about vectors to begin the study of Newtonian mechanics [19]. Here we find that even after a semester of instruction in physics that involves quite a fair amount of vector addition, the fraction remains about the same and students have great difficulty dealing with vector addition in component form.

The interviews also revealed that many students may not use mathematical knowledge about adding vectors that they possess while solving a problem in a physics context, but they can apply this knowledge correctly to a vector addition problem in a mathematical context as was posed during interviews after students had answered all problems initially given to them to the best of their ability. This dichotomy between students' facility with problems devoid of context and difficulty with problems with a physics context can be interpreted as a lack of knowledge transfer from one situation to another. This context-specific knowledge and lack of transfer can partly account for the finding that a small percentage of students (30%) attempted to draw the components of the two forces they added in Problem 2. Six out of the nine students interviewed did not add the two forces in Problem 2 by determining *x* and *y* components, but when asked a very similar force addition problem *all of these six students* tried to find their components in order to add them. Some were successful and some were not, but nonetheless, the knowledge that forces must be added by components was not used when solving the physics problem, even though students clearly had this knowledge. Interviews also suggested two possible reasons for why students would use this knowledge in the mathematical context, but not in the physics context. One reason is that several students used intuition and gut feeling when adding the forces rather than their mathematical knowledge about vector addition. For example, they would look at the situation as the two 2C charges pushing the 1C charge "away" and not take into consideration that the two forces due to the interaction of the 1C charge with the other charges were not along the same direction. When probed about why they added the magnitudes of the forces as scalars, some students specifically mentioned their gut feeling. A second reason is that students may be prompted to find components while solving the mathematical problem because an angle was drawn on the diagram provided; however, no angles were drawn in the physics context (Problem 2).

For Problem 1, we also found that several weeks after instruction, students' performance on electric force improved while their performance on electric field remained stagnant. Interviews with students (which were all conducted after an exam which covered these topics) also reveals that several weeks after instruction students exhibited more difficulties on the concept of electric field than on the concept of electric force. In particular, some students who were not able to recognize that contributions coming from each of the two charges must be considered when evaluating the net electric field at the midpoint, readily recognized this in the electric force part of the problem. Furthermore, students who exhibited difficulties in both parts when solving Problem 1, after finishing both problems, were asked directed (scaffolding) questions by the interviewer which were intended to improve students understanding of these concepts (at least as it pertains to these problems). Some of the students were not able to grasp the concept of electric field even after scaffolding support and had a difficult time determining the correct method for solving for the electric field part of the problem while being able to use the scaffolding (sometimes very little) to solve the force part of the problem. Other students managed to take advantage of scaffolding during the interviews to solve both parts of the problem; however, all of them took more time and significantly more directed questioning in the electric field part than in the electric force part.

The lack of a robust knowledge structure and the abstract nature of electric field compared to electric force may contribute to this finding. Experts extend their knowledge by connecting new information with prior knowledge already stored in their long term memory. Moreover, after years of sense making, even abstract concepts do not appear very abstract to the experts. Introductory students' knowledge about physics is fragmented. They have information about forces from Newtonian mechanics and it is easier for them to connect the new concept of electric force with what they already know. However, prior to being introduced to the electric field, they have little or no knowledge of the abstract concept of fields especially in an algebra-based course. In particular, electric fields are generally the first fields to be introduced in an algebra-based introductory physics course because the concept of gravitational field is skipped. Therefore, they have difficulty connecting this new abstract concept of electric field with their prior knowledge and whatever short term gain there is while practicing homework problems immediately before a quiz appears to be lost later in the midterm performance. It is also important to mention that our research suggests that the percentage of students who used the essential relationship ($\vec{F} = q\vec{E}$) between electric field and electric force decreases as the semester progresses. If instructional design stresses this relationship that connects the two concepts within a coherent curriculum that focuses on helping students build a robust knowledge structure and also stresses the vectorial nature of both field and force, students may make a better connection between electric field and electric force and improve their performance on both while practicing problems.

The fact that the relationship between electric force and electric field ($\vec{F} = q\vec{E}$) can be used effectively to help students develop a better understanding of the abstract concept of electric field by connecting it with a more familiar concept of force was also suggested by the interviews. Students who had a difficult time solving the electric field part of the problem were asked directed questions intended to help them use this connection in order to reason about electric field. Some students were able to use it correctly after only several questions, whereas others required more involved directed questioning. However, in almost all cases, starting with a simple situation (such as using this relationship to determine the direction of the electric field due to a positive charge) where the interviewer guided the student to use the equation $\vec{F} = q\vec{E}$ through scaffolding questions (such as "What does this equation say about the direction of the electric field relative to the direction of the electric force?", "Can you use this equation to figure out the direction of the electric field at some distance from a negative charge?" etc.) and working up to the more complex situation present in the original problem, students were able to reason about the problem correctly and understand how the concept of electric field is used to solve it.

Finally, the interviews suggest some difficulties students have when dealing with the concepts of electric field and electric force. One difficulty is in developing a good mental representation of the concept of electric field which can be used to solve problems such as those discussed in this research. The abstract notion of "field" is difficult for students to conceptualize and they may not even have an idea what it represents. The comment (and drawing) made by one of the students was very representative. This student attempted to solve Problem 2 by finding the electric field at the corner of the 1C charge and then multiplying it by the charge. However, when finding the net electric field she did not consider the vectorial nature of the field. When asked why she did not add the electric fields due to each of the 2C charge vectorially, she said she wasn't thinking of vectors and drew pictures of 'fields' emanating from each of the 2C charge towards the 1C charge that looked like travelling waves. Within this model, the student felt that the electric field is a scalar quantity. A second issue with which students struggled only adds to this first difficulty: students often did not differentiate between the concepts of electric field and electric force. This was evident when students would use the words "field" and "force" interchangeably (i.e. use "field" when they meant "force" and vice versa) or when trying to answer questions about why they were adding two contributions to the net electric field in the first part of problem 1 (because they would often use reasoning directly related to electric forces). Furthermore, when asked what is the direction of the electric field at a point due to a point charge, students often had difficulty realizing that a non-zero electric field could exist at a point in empty space; many claimed that a charge must be placed at that point in order to answer the question. Moreover, when negative charges were placed at the point to find the electric field due to a positive point charge, students often got confused between the direction of the electric force and the electric field.

4.6 CHAPTER REFERENCES

1. G. V. Glass and K. D. Hopkins (1996). Statistical Methods in Education & Psychology, (3rd ed.), Boston: Allyn & Bacon.

2. K. Ericsson and H. Simon (1993). Protocol Analysis: Verbal Reports as Data, (MIT Press, Boston, MA 1993).

3. J. Sweller (1998). “Cognitive load during problem solving: Effects on learning.” Cog. Sci. 12(2), 257-285.

4. J. Zhang (2006). “Distributed cognition, representation and affordance.” Prag. Cogn. 14(2), 333-341.

5. D. Rosengrant (2007). Ph.D. Dissertation, Rutgers University.

6. H. Simon (1974). “How big is a chunk?” Science 183 (4124), 482-488.

7. P. Kyllonen and R. Christal (1990). “Reasoning ability is (little more than) working memory capacity?!” Intelligence 14, 389-433.

8. A. Fry and S. Hale (1996). “Processing speed, working memory and fluid intelligence: Evidence for a developmental cascade.” Psychol. Sci. 7(4), 237-241.

9. R. Kail and T. Salthouse (1994). “Processing speed as a mental capacity.” Acta Psychologica 86, 199-225.

10. G. Miller (1956). “The magical number seven, plus or minus two: Some limits on our capacity for processing information.” Psychol. Rev. 63, 81-97.

11. A. Miyake, M. A. Just and P. Carpenter (1994). "Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts." J. Mem. Lang. 33(2), 175-202.

12. L. Novick (1988). "Analogical transfer, problem similarity and expertise." J. Exp. Psychol. Learn. 14(3), 510-520.

13. M. Gick and H. Holyoak (1983). "Schema induction and analogical transfer." Cognitive Psychol. 15, 1-38.

14. M. Bassok and H. Holyoak (1989). "Interdomain transfer between isomorphic topics in algebra and physics." J. Exp. Psychol. Learn. 15(1), 153-166.

15. F. Mateycik, D. Jonassen, and N. S. Rebello (2009). "Using similarity rating tasks to assess case reuse in problem solving." AIP Conf. Proc. 1179, 201-204.

16. P. Adey and M. Shayer (1993). "An exploration of long-term far-transfer effects following an extended intervention program in the high school science curricula." Cognition Instruct. 11, 1-29.

17. A. Brown (1989). "Analogical learning and transfer: What develops?" Similarity and Analogical Reasoning, edited by S. Vosniadu and A. Ortony (Cambridge University Press, NY, 1989), pp. 369-412.

18. D. K. Detterman and R. J. Sternberg (1993). Transfer on trial: Intelligence, Cognition and Instruction (Ablex Pub. Corp., Norwood, NJ, 1993).

19. R. Knight (1995). "The Vector Knowledge of Beginning Physics Students." Phys. Teach. 33, 74-80.

5.0 STUDENT DIFFICULTIES IN TRANSLATING BETWEEN MATHEMATICAL AND GRAPHICAL REPRESENTATIONS IN INTRODUCTORY PHYSICS

5.1 INTRODUCTION

For a literature review of previous research related to the role of multiple representations in problem solving, refer to the introduction in the study presented in Chapter Two.

This study investigates students' ability to transform between mathematical and graphical representations and how this relates to their problem solving performance. Student difficulties in interpreting graphical representations have been extensively researched in kinematics [1-6]. Instructional strategies have also been developed to remedy student difficulties [7-13]. Other researchers have investigated student understanding of *P-V* (pressure vs. volume) diagrams both in upper level-thermodynamics courses [14,15] as well as in introductory physics calculus-based courses [16]. Pollock et al. [14] also looked at student performance on similar questions devoid of physical context and found that some of the difficulties students exhibited in a physical context could be attributed to mathematical difficulties related to the concept of an integral. In a later study, Christensen and Thompson [17] investigated student difficulties with the concept of slope and derivative in a mathematical (graphical) context.

This investigation focused on students' ability to translate from the mathematical description of an electric field to the corresponding graphical representation, which is directly related to the concept of function and graphing of a function. Student difficulties with the concept of function have been researched by mathematics education researchers [18,19]. Hitt [20] identified five levels in the construction of the particular concept of a function which vary between imprecise ideas about a concept (Level 1) to coherent articulation of different systems of representation in the solution of a problem (Level 5). Hitt also found that, sometimes, even secondary mathematics teachers cannot always articulate between the various systems of representation involved in the concept of a function. Vinner and Dreyfus [21] distinguished between a concept image and a concept definition because they saw students repeatedly misuse and misapply terms like function, limit, tangent and derivative. For many students, the image evoked by the term "function" is of two expressions separated by an equal sign [22,23]. Thompson found [23] that many students who had successfully passed a Calculus and a Modern Algebra course still saw no problem with a definition like $f(x) = \frac{n(n+1)(2n+1)}{6}$ because it fits their concept image of a function. Also, students in algebra courses often have an *action* conception of a function because a function is seen as a command to calculate and therefore they must actually apply it to a number before the recipe will produce anything. The way many introductory physics students manage equations in solving physics problems is often very predictable: they plug numbers into an equation and figure out an unknown which can in turn be plugged into another equation. This process is continued until the target variable is found. When numbers are not given or when students run into a situation with two equations and two unknowns, they have a much more difficult time solving the problem. As evidenced by these examples and others in [23], students' concept images are often not consistent with concept definitions. However, for mathematics "experts" the concept images become tuned over time so that they are consonant with the conventionally accepted concept definitions. One proposed instructional method of overcoming some of these difficulties involves real-world investigations that use realistic data and scenarios [24-27]. Mathematics education researchers have also investigated student difficulties in connecting various representations of functions, in particular graphical and algebraic [28,29] and some have stressed that this process of translating between the graphical and algebraic representations of functions presents one of the central difficulties for students to construct an appropriate mental image of a function [30]. Other mathematics researchers have investigated the intertwining between the flexibility of moving from one representation of a function to another and other aspects of knowledge and understanding [31] as well as students' abilities to extract meaningful information from graphs [32].

In physics, there is the added difficulty of understanding the relevance of certain mathematical knowledge and procedures to the solution of physical problems. Students may have the requisite mathematical knowledge that needs to be applied to a physical situation but they may fail to invoke it at the appropriate time because they are unaware of its usefulness. This is supported by Hammer's observation that high-school students take little out of an initial mathematical review of procedures divorced from physics [33] and by research on difficulties of transferring mathematical knowledge across disciplines [34-36]. Also, the physics context typically requires additional information processing, which can lead to an increased cognitive load and deteriorated performance [37]. In this investigation we explore the facility of students in a calculus-based introductory physics course in transforming a problem solution involving the electric field for spherical charge symmetry from mathematical to graphical representation. Our previous research suggests that students have great difficulty in transforming the electric field from one representation to another for the problem discussed in this paper and the research presented here provides several insights that could partly account for this difficulty. Our major finding is that more scaffolding, which experts might consider should help students, can instead hinder students' performance. Therefore, it is important to optimize the level of scaffolding via research (for students with a given prior knowledge) to ensure that they benefit from the support provided in the intended manner and learn to translate appropriately from a mathematical to a graphical representation.

5.2 METHODOLOGY

A class of 95 calculus-based introductory physics students was enrolled in three different recitations. The three recitations formed the comparison group and two intervention groups for this investigation. In addition, ten students in different but equivalent calculus-based introductory physics classes were interviewed individually in paid interviews using a think-aloud protocol [38,39] to understand their thought processes better while they solved the problem. Below, the two interventions used in two of the recitations are described first. All recitations were taught in a traditional manner in which the TA worked out problems similar to the homework problems and then students were given a 15-20 minute quiz at the end of class. Students in all recitations attended the same lectures, were assigned the same homework, and had the same exams and quizzes. Students' difficulties in transforming the solution to the following problem (which was given in a quiz) from the mathematical to the graphical representation are investigated. The version of the problem below is the one with no scaffolding support and was given to the comparison group (referred to as the "scaffolding level zero group" or SL0).

A solid conductor of radius *a* is inside a solid conducting spherical shell of inner radius *b* and outer radius *c*. The net charge on the solid conductor is +*Q* and the net charge on the concentric spherical shell is –*Q* (see figure).

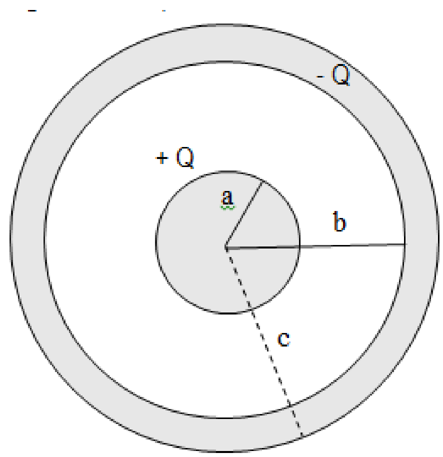


Figure 5.1. Problem diagram provided to all students.

(a) Write an expression for the electric field in each region.

(i) $r < a$

(ii) $a < r < b$

(iii) $b < r < c$

(iv) $r > c$

(b) On the figure below, plot $E(r)$ (which is the electric field at a distance r from the center of the sphere) in all regions for the problem in (a).

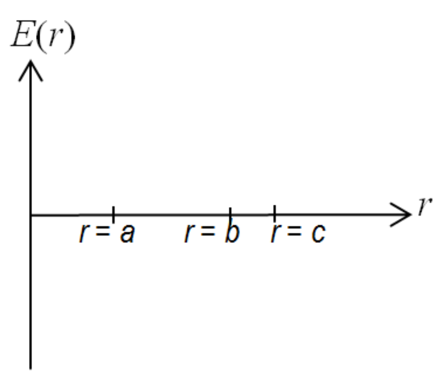


Figure 5.2. Coordinate axes provided to all students for sketching the electric field in part (b).

Previous preliminary research in a different introductory calculus-based physics class suggested that students have great difficulty in graphing the electric field after writing an expression for the electric field in each region. In particular, a majority of students (~70% – 80%) drew graphs that were not consistent with their mathematical expressions in one or more regions. Motivated by these preliminary findings, two scaffolding interventions were implemented in two of the recitations by giving students some scaffolding support in order to assess if it helps them make a better connection between the two representations. Theoretical task analysis from an expert perspective [40-42] of the process of transforming from mathematical to graphical representation was used to design the two interventions. The students that received the first level of scaffolding (which will be referred to as "SL1" – Scaffolding Level 1) were asked to draw the electric field in each region before graphing it in part (b) shown above. Their instructions were as follows:

> *(a) Write an expression for the electric field in each region and sketch the electric field in that region on the coordinate axes shown (in the shaded regions, please do not draw).*

For each region ($r<a$, $a<r<b$, etc.) right after calculating the expression for the electric field in that region, they were given coordinate axes with the irrelevant parts shaded out. For example, for region $r<a$, they were given the coordinate axes shown in Figure 5.3 for drawing the graph of the electric field in that region.

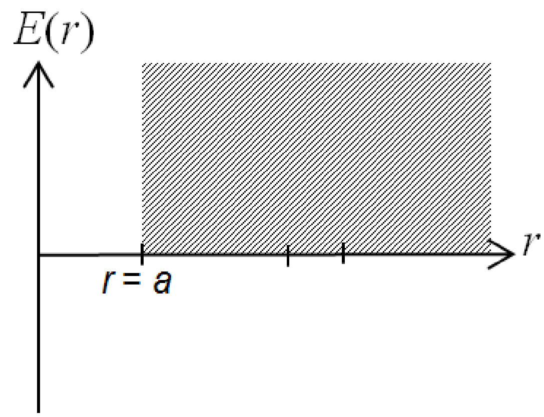


Figure 5.3. Coordinate axes provided to students in the SL1 and SL2 groups for sketching the electric field in region r < a.

Students who received the second level of scaffolding ("SL2" – Scaffolding Level 2) were given all the support of SL1 described above and, in addition, they were asked to evaluate the electric field at the beginning, mid and end points of each region before graphing it in that region. For example, for region $r<a$, they were also asked to fill in the following blanks after writing an expression for the electric field for that region, but before graphing it:

When $r = 0$, $E(r = 0) =$ ___________

When $r = a/2$, $E\left(r = \frac{a}{2}\right) =$ ___________

When $r \rightarrow$ a, $E(r \rightarrow 0) =$ ___________

For convenience, a brief description of the three scaffolding levels is provided in Table 5.1.

Both the SL1 and SL2 interventions were designed to help students perform better on graphing the electric field. It was hypothesized that asking students to graph the field in each region first, after writing an expression for the field in that region, but before constructing the graph for the field everywhere, may help them make a connection between the graphical and mathematical representations better. In particular, it was anticipated that some students would

Table 5.1. Brief description of the three scaffolding levels

Scaffolding Level 0 (SL0)	Asked to draw the electric field at the end, after finding an expression for it in all regions.
Scaffolding Level 1 (SL1)	In each region, right after finding the electric field, they are asked to draw it and they are provided with coordinate axes with the irrelevant regions shaded out. They are also asked to draw it again at the end.
Scaffolding Level 2 (SL2)	**Everything given to SL1 and in addition**, in each region, they were asked to evaluate the electric field at the beginning, mid, and endpoint of that region. They are also asked to draw it again at the end.

realize that in this problem the electric field takes the form of a piece-wise defined function (with discontinuity in the electric field where one crosses a surface charge distribution) and in order to graph it, they must individually plot each forms of this function in the corresponding region. The additional support in the SL2 intervention, namely, the instructions to find the electric field at the beginning, mid, and end point of each region before graphing in that region were intended to, on the one hand, give another hint that the electric field has different forms in different regions, and on the other hand, help students realize that the electric field has discontinuities at charged interfaces and thus help them perform better on graphing it.

The researchers jointly determined the grading rubric iteratively. After extensive discussions among two researchers, the way the problem was finally scored is summarized in Table 5.2.

Table 5.2 shows that the region $a<r<b$ was assigned three times as many points as each of the other regions. This consideration was made because region $a<r<b$ was the only one with a non-zero electric field. In finding the expression for the electric field in parts (a)(i) through (a)(iv), students were given 80% for the correct expression and 20% for the correct reasoning

Table 5.2. Summary of the scores assigned to each part of the problem.

(a) Find an expression for the electric field			
(i) $r<a$	(ii) $a<r<b$	(iii) $b<r<c$	(iv) $r>c$
10 points	30 points	10 points	10 points
(b) Plot $E(r)$ in all regions			
$r<a$	$a<r<b$	$b<r<c$	$r>c$
5 points	15 points	5 points	5 points

that led to that expression. For example, if a student wrote $E = \frac{kQ}{r^2}$ for the expression without any explanation in region $a < r < b$, he/she would obtain 24/30 points. Table 5.2 also shows that plotting the electric field in part (b) was worth 30 points, which is half of the points assigned to finding the expressions for the electric field in part (a). Part (b) was broken up into individual regions and in each region we investigated whether the student's graph was consistent with the expression found for the electric field in that region. Full credit was given if the form of the graph matched the expression; students were not expected to label endpoints, or even have correct endpoints in order to receive full credit (for graphing). For example, if a student found $E(r) = kr/3$ in region $b < r < c$, and drew a graph similar to the one shown in Figure 5.4 (an increasing linear graph that starts from the r-axis), this student would be considered to be consistent (and obtain the 5 points assigned to this part) because he/she selected the correct type of graph (linear) consistent with the expression in that region, even though the left endpoint is clearly incorrect (based on the expression, $E(r = b) = kb/3$, but in the graph $E(r = b) = 0$). We note that the maximum score that can be obtained on this problem is 90 points rather than 100. This is because it was considered that matching (in terms of scoring) between regions in the parts that required finding an expression and plotting the electric field (i.e., plotting an expression correctly is worth half the points assigned to finding the correct expression) was more important than making the maximum 100 points.

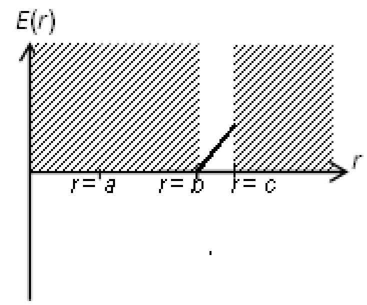


Figure 5.4. Example of a graph drawn by a student in region $b<r<c$.

5.3 QUANTITATIVE RESULTS

Before presenting the results it is worthwhile mentioning that students' scores on the final exam were analyzed in order to ensure that the three groups exhibited similar performance. There were no differences in the performance of students in different groups on the final exam (the difference between the lowest and highest average performance of students in the different groups on the final exam was 5.4 points out of 100).

5.3.1 Primary Finding

The additional scaffolding given to students in the SL2 group (as compared to SL1) had the opposite effect to the one intended as evidenced by three factors:

- **1) Students in the SL2 group performed worse at finding the correct expressions for the electric field than those in the SL1 group;**

- **2) Students in the SL2 group performed worse at graphing the electric field than students in the SL1 group;**
- **3) Students in the SL2 group were also less consistent between the expressions found and the graphs drawn in each part.**

Each of the three components of the primary finding is discussed below.

5.3.1.1 Students in the SL2 group performed worse than students in the SL1 group in finding the correct expressions for the electric field

Table 5.3 shows the averages and standard deviations of the SL1 and SL2 groups on the first four parts of the quiz combined (the parts which required finding expressions for the electric field). Although the TA gave all students sufficient time to finish the quiz, some did not complete it because it was a low-stakes quiz (students received credit for one recitation quiz which counted for less than one percent of the course grade). Due to the fact that students in the two scaffolding groups had more instructions to take care of due to the scaffolding provided (e.g., draw the electric field in each region, find $E(r \rightarrow a)$, $E(r \rightarrow b)$ etc.), a few of those students did not work on more than two of those four parts. The numbers in Table 5.3 are based only on the students who had done work on at least two out of the four parts. A *t*-test [83] on the data in Table 5.3 reveals that students in the SL2 group performed worse than students in the SL1 group ($p=0.040$). A calculation of Cohen's d [44,45] (which yields 0.622) suggests that this difference corresponds to an effect size between medium and large. Cohen's d refers to the standardized mean difference [44,45]. As defined by Cohen, large, quite noticeable effects correspond to a value for Cohen's d around (or larger than) 0.8, medium effects correspond to 0.5 and small correspond to 0.3 or less. To put this into context, an effect size of 0.8 corresponds to the height difference between 13 and 18 year old girls.

Table 5.3. Numbers of students (N) in the SL1 and SL2 groups, averages (Avg.) and standard deviations (Std. dev.), renormalized for 100 maximum points, for the scores of students in the SL1 and SL2 groups on the first four parts combined. Only students who did work in at least two out of the four parts (a majority of students) are included in these statistics.

	N	Avg.	Std. dev.
SL1	27	58	33
SL2	30	39	32

The performance of the SL1 group was also compared with the performance of the SL2 group in each individual part of the quiz. This had the benefit of eliminating a few students who had not done work in a part from the total pool of students and obtaining a more accurate picture of their performance.

Table 5.4. Numbers of students (N) in the SL1 and SL2 groups, averages (Avg.) and standard deviations (St.d.) for the scores in parts (a)(i) through (a)(iv) of the SL1 and SL2 groups out of 10 points (part (a)(ii) was renormalized to 10 maximum points).

	(a)(i) ($r<a$)			(a)(ii) ($a<r<b$)			(a)(iii) ($b<r<c$)			(a)(iv) ($r>c$)		
	N	Avg	St.d.	N	Avg	St.d.	N	Avg	St.d.	N	Avg	St.d.
SL1	30	5.7	4.2	27	5.7	4.0	26	5.4	4.4	26	5.9	4.6
SL2	32	2.8	3.9	30	4.7	4.1	22	2.0	3.7	19	3.4	4.6

Table 5.4 shows the averages and standard deviations for the scores of students from the SL1 and SL2 groups in the parts of the quiz that required finding an expression for the electric

field (for each of the parts in Table 5.4, only those students who worked on those particular parts are included in the statistics, thus, the numbers of students sometimes differ in different parts). Comparison of these groups yields statistically significant differences (SL2 group performing worse than the SL1 group) in part (a)(i) (p = 0.006, Cohen's d = 0.816) and in part (a) (iii) (p = 0.005, Cohen's d = 0.843). It is worthwhile to note that the SL1 group outperformed the SL2 group by at least 20% in each part, but due to the large standard deviations, those differences are not statistically significant in two of the cases.

Another measure of student performance can be obtained by investigating the number of students who determined that the electric field is zero inside the conductors (parts (a)(i) and (a)(iii)) and in region $r > c$ (part (a)(iv)). Results are shown in Table 5.5. Chi-squared tests [83-85] on these data reveal that students in the SL2 group underperformed students in the SL1 group in all three parts: (a)(i): p=0.002, (a)(iii): p=0.007, (a)(iv):p=0.047. The difference between the SL1 and SL2 groups in terms of percentage of correct answers ($E = 0$) is 30% or higher in every part.

Table 5.5. Percentages (and numbers) of students who found that the electric field is zero and non-zero in the regions where it is supposed to be zero.

	Part (a)(i)		Part (a)(iii)		Part (a)(iv)	
	E=0	E≠0	E=0	E≠0	E=0	E≠0
SL1	60% (18)	40% (12)	58% (15)	42% (11)	62% (16)	38% (10)
SL2	22% (7)	78% (25)	17% (4)	83% (19)	32% (6)	68% (13)

Table 5.6 shows the performance on the last part, which asked students to graph the electric field everywhere. A *t*-test on the data in Table 5.6 shows that students in the SL2 group performed worse than students in the SL1 group (p = 0.012, Cohen's d = 0.732).

Table 5.6. Numbers of students (N) in the SL1 and SL2 groups, averages (Avg.) and standard deviations (Std. dev.) for the scores on graphing the electric field (renormalized to 10 maximum points).

	N	Average	Std. dev.
SL1	27	6.2	3.9
SL2	24	3.5	3.3

The score on graphing the electric field is based on how consistent the students were between the expressions they found and the graphs they drew in part (b), thus the results shown in Table 5.6 provide the first indication that students in the SL2 group were less consistent than students in the SL1 group. However, these scores were based on the final graph. We also investigated if the students were consistently plotting their expressions immediately after finding them.

5.3.1.2 Students in the SL2 group were less consistent than students in the SL1 group between the expressions they found and the graphs they drew in three out of the four parts.

Students in the two scaffolding interventions were asked to sketch the electric field in each region immediately after finding it, and in addition, they were provided with coordinate axes with the irrelevant parts shaded out. Table 5.7 shows, in each of the four parts, how many students plotted their expressions correctly and incorrectly by showing whether students were consistent between the expressions they found and the graphs they drew. Chi-squared tests on the data in Table 5.7 reveal that students in the SL2 group were less consistent than students in the SL1 group in all but the last part (p values for comparison are 0.010, 0.024 and 0.002 for parts (a)(i), (a)(ii) and (a)(iii) respectively; the difference in part (a)(iv) is not statistically significant).

Table 5.7. Percentages (and numbers) of students from the SL1 and SL2 groups who were consistent between the graphs they drew and the expressions they found in each of the first four parts.

	(a)(i) consistent		(a)(ii) consistent		(a)(iii) consistent		(a)(iv) consistent	
	Yes	No	Yes	No	Yes	No	Yes	No
SL1	86% (24)	14% (4)	67% (18)	33% (9)	77% (20)	23% (6)	69% (18)	31% (8)
SL2	54% (17)	46% (14)	37% (11)	63% (19)	32% (6)	68% (13)	58% (11)	42% (8)

The data in Table 5.7 suggest that sometimes students were consistent in one or more parts, but not all. We also investigated how many students were always consistent between the expressions they found and the graphs they drew. This result is shown in Table 5.8. Once again, a chi-squared test on the data in Table 5.8, reveals that students in the SL2 group were significantly less consistent than students in the SL1 group (p=0.008).

Table 5.8. Percentages (and numbers) of students from the SL1 and SL2 groups who were consistent in all parts.

	Consistent in all parts	
	Yes	No
SL1	59% (16)	41% (11)
SL2	24% (7)	76% (22)

It is important to keep in mind that the electric field is zero in three out of four parts in this problem. Therefore, a better measure of how adept students are at translating between a mathematical and graphical representation of electric field (consistency) can be obtained by investigating how adept students are at graphing non-zero electric fields, because if a student finds that the electric field is zero in a region, it should be relatively straightforward for this student to graph that electric field and be consistent. In order to investigate how well students could graph non-zero electric fields, one excludes students who found E = 0 in a particular region and plotted it accordingly. Therefore, for each student, only the regions where the student found a non-zero electric field were considered (regardless of whether the electric field is supposed to be zero in that region or not). Then, all the times when students were (and were not) consistent were added up to obtain two numbers. These numbers represent the number of times students in each of the two scaffolding intervention groups were able (and were not able) to graph a non-zero electric field in a particular region. These numbers are referred to as "consistencies yes" and "consistencies no" in Table 5.9. A chi-squared test on the data in Table 5.9 shows that students in the SL2 group are statistically less consistent in graphing non-zero electric fields than students in the SL1 group ($p = 0.025$).

Table 5.9. Percentages (and numbers) of yes and no consistencies for the SL1 and SL2 groups.

	Consistencies yes	Consistencies no
SL1	57% (31)	43% (23)
SL2	38% (31)	62% (51)

Students in the two scaffolding intervention groups were essentially asked to graph the electric field twice – *immediately* after finding it in each region *and at the end*. An expert would simply put together the graphs found in the individual regions to obtain the final graph. However, some students drew final graphs that were not consistent with the graphs they drew in one or more regions. For example, a student drew a graph that looks like $1/r$ for region $a<r<b$ right after finding an expression for the electric field in that region, but in the final graph, in the same region, the student drew a constant non-zero electric field. Table 5.10 shows the number of times students from the scaffolding interventions were (and were not) consistent in this respect. A chi squared test is not appropriate for data in Table 5.10 because the numbers are too small

and not all the expected cell frequencies are larger than 10 [43]. Therefore, Fisher's exact test [46] was performed, which revealed that students in the SL2 group were less consistent than students in the SL1 group between the graphs drawn in each part and the final graph.

Table 5.10. Percentages (and numbers) of students from the SL1 and SL2 groups who were (and were not) consistent between graphs drawn in each part and the final graph.

	Consistent	Not consistent
SL1	86% (19)	14% (3)
SL2	57% (12)	43% (9)

5.3.2 Secondary Findings

Students in the SL1 group performed better at finding the correct expressions than the comparison group.

Table 5.11 shows the average scores on the first four parts which required finding an expression for the electric field. Similar to the previous comparison in this respect between SL1 and SL2 students, only the students who worked on at least two of the four parts are included. A *t*-test reveals that the difference is statistically significant (p=0.004). Also, Cohen's d (0.843) shows a large effect size.

Table 5.11. Sizes (N), averages and standard deviations (renormalized for 100 maximum points) for the scores of students in the SL0 and SL1 groups on the first four parts combined (only students who did work in at least two out of the four parts are included in these numbers).

	N	Avg.	Std. dev.
SL0	32	34	29
SL1	27	58	33

The scores in each of the four parts were also compared. Table 5.12 shows the averages and standard deviations for the scores of students in the SL0 and SL1 groups. *T*-tests on data in Table 5.12 reveal that students in the SL1 group outperformed students in the SL0 group in part (a) (ii) ($p = 0.022$, Cohen's $d = 0.628$) and in part (a) (iii) ($p = 0.019$, Cohen's $d = 0.654$).

Table 5.12. Numbers of students (N) in the SL0 and SL1 groups, averages and standard deviations for the scores in parts (a)(i) through (a)(iv) of the SL0 and SL1 groups out of 10 points (part (a)(ii) was renormalized to 10 maximum points).

	(a)(i) ($r<a$)			(a)(ii) ($a<r<b$)			(a)(iii) ($b<r<c$)			(a)(iv) ($r>c$)		
	N	Avg	St.d.	N	Avg	St.d.	N	Avg	St.d.	N	Avg	St.d.
SL0	32	4.3	4.0	30	3.4	3.3	30	2.8	3.7	31	3.4	4.4
SL1	30	5.7	4.1	27	5.7	4.0	26	5.4	4.4	26	5.9	4.6

One last measure of performance in terms of accuracy in finding the correct expressions was investigated by comparing the number of correct responses (E=0) in regions (a)(i), (a)(iii) and (a)(iv). Table 5.13 shows the findings. A chi-square test reveals that the difference is statistically significant (p=0.002).

Table 5.13. Percentages (and numbers) of correct responses from students in the SL0 and SL1 groups in the regions where the electric field was zero: regions (a)(i), (a)(iii) and (a)(iv).

	$E = 0$	$E \neq 0$
SL0	37% (34)	63% (59)
SL1	60% (49)	40% (33)

More students in the SL1 group were always consistent than students in the SL0 group.

It was also investigated whether there were more students in the SL1 group who were always consistent than students in the SL0 group. Results are shown in Table 5.14. A chi-squared test on these data shows that students in the SL1 group were performing better in this respect than students in the SL0 group (p=0.002).

Table 5.14. Percentages (and numbers) of students from the SL0 and SL1 groups who were always consistent.

	Consistent in <u>all</u> parts	
	Yes	No
SL0	29% (9)	71% (22)
SL1	59% (16)	41% (11)

Performance of students in the SL2 group was not statistically significantly better or worse in any respect than the performance of students in the SL0 group.

We also performed *t*-tests and chi-square tests to compare the various aspects of performance mentioned so far (performance in finding expressions, performance in figuring out that the electric field is zero in the three regions where it should be zero, performance in graphing the field in part (b) and performance in consistency between expressions found and graphs drawn). There were no statistically significant differences in any of these aspects.

Regardless of the intervention, students who were always consistent between the expressions they found and the graphs they drew performed better than the other students.

It was also investigated whether students who were consistent between the expressions they found and the graphs they drew in <u>all</u> parts also performed better regardless of the intervention, i.e., is consistency correlated with performance for all students? Since in the graphing part of the problem, part (b), the scores were given based on how consistent students were, one would expect a correlation to exist between scores and consistency on this part. Therefore, one should look at each one of the other parts, where the scores were based solely on the expressions students found. Table 5.15 shows the averages and standard deviations in each of the four parts, (a)(i) through (a)(iv), which were graded based on the expressions students found. The averages and standard deviations were computed for students who were consistent in all parts ("Consistent" in Table 5.15) and students who were not consistent in one or more parts ("Not cons." in Table 5.15). We performed *t*-tests to compare the performance of students who were consistent with the performance of students who were not consistent and found that the students who were consistent outperformed the other students in *every part*. The p values for comparing these groups are also very small; three of them are less than 0.001 (for the last three parts) and the other is 0.002 (for the first part). The effect sizes (Cohen's d) also show significant effects; three of them are above 1.0 (for the last three parts) and the other is 0.73. (Important note:

Cohen's d is defined as the difference in means of the two groups one compares divided by the standard deviation of the population from which the samples were taken. In practice, the standard deviation of the population is almost never known and is most commonly estimated by the standard deviation of the control/comparison group. For the two groups we compare here, neither is a control/comparison group. In this case one can estimate the population standard deviation by using a pooled standard deviation based on the two standard deviations of the samples being compared. This pooled standard deviation is defined as $\sigma_{pooled} = \sqrt{(\sigma_1{}^2 + \sigma_2{}^2)/2}$ and can be used as an estimation of the population standard deviation [83]. This is what was used in this case).

Table 5.15. Numbers of students (N), averages and standard deviations in each part where the scores were based on expressions of students who were consistent in all parts ("Consistent") and of those who were not consistent in one or more parts ("Not cons.").

	Part (a)(i)			Part (a)(ii			Part (a)(iii)			Part (a)(iv)		
	N	Avg.	St. d.	N	Avg.	St. d.	N	Avg.	St. d.	N	Avg.	St. d.
Consistent	32	6.3	4.0	32	7.1	3.6	32	5.8	4.2	32	6.4	4.6
Not cons.	55	3.5	3.9	55	2.9	3.3	55	1.5	3.0	55	2.0	3.7

Comparison of the performance of students from different quiz intervention groups on the same problem given in the final exam in multiple choice form

As noted earlier, there is no statistically significant difference on the final exam overall between the three groups. However, the problem discussed was also given in the final exam in a multiple choice format (also, the scaffolding interventions were not implemented). For each region, students were asked to choose an expression for the electric field from a list of four choices or provide an expression if what they found for the expression was different from the choices given. After finding all four expressions they were asked to select the graph that represented the electric field in all regions from a choice of four different graphs. They were also given a fifth choice: blank coordinate axes similar to Figure 5.2 onto which they could draw their own graph if none of the four graphs given matched how they would plot the electric field. The incorrect graphs were based on incorrect expressions the students could choose in the previous multiple choice question. The sizes, averages and standard deviations are shown in Table 5.16. (The group sizes (N), are less than they were in the quiz because a few students from each group dropped out by final exam time). Table 5.16 shows that students in the SL1 group performed better than students in the other groups by more than 20% (although the differences are not statistically significant).

Table 5.16. Numbers of students (N), averages and standard deviations (Std. dev.) for the scores on the final exam multiple choice problem for the students in the different groups.

	N	Average	Std. dev.
SL0	30	5.2	2.7
SL1	29	6.3	3.2
SL2	29	5.0	2.9

Finally, for the multiple choice problem given in the final exam, we investigated the percentages of students who were consistent in all the parts in each group. Results are in Table 5.17.

None of the differences in Table 5.17 are statistically significant. However, it is interesting to note that the percentages of students who were consistent in all the parts in the no scaffolding group (SL0) and Scaffolding Level 2 group (SL2) did not change by much (24% and

Table 5.17. Numbers of students (N), percentages (and numbers) of students who were consistent in all the parts in the final exam multiple choice problem.

Final	N	% (number)
SL0	30	27% (8)
SL1	29	45% (13)
SL2	29	31% (9)

29%, respectively, in the quiz as shown in Table 5.8 and Table 5.14 and 31% and 27%, respectively, in the final exam as shown in Table 5.17), whereas the percentage of students from the Scaffolding Level 1 group who were consistent in all the parts went from 59% in the quiz (see Table 5.8) to 45% in the final exam multiple choice problem (see Table 5.17).

5.3.2.1 Graduate students' performance

One would not expect differences between graduate students who take this quiz that include the two different scaffolding levels because graduate students are not very likely to require any support in order to perform well on this problem. In order to ensure that graduate students are not affected negatively by the second scaffolding intervention as opposed to the first (as we found for introductory physics students discussed earlier), SL1 and SL2 versions of the problem were also given to a group of 26 first-year graduate physics students enrolled in a TA training class. Roughly half (14) of them (GrSL1 – Graduate Scaffolding Level 1) were randomly assigned to solve the version of the quiz that included Scaffolding Level 1 and the other half (GrSL2 – Graduate Scaffolding Level 2) solved the version of the quiz that included Scaffolding Level 2 (as described in Table 5.1). The graduate students were graded in the same manner as the introductory students. No statistically significant differences between the scores of the two groups of graduate students were found, both in each individual part of the problem and overall. Comparison (*t*-tests) with the corresponding introductory physics student groups reveals that graduate students performed statistically significantly better (both p values are less than 0.001). Table 5.18 lists the overall scores of the two graduate student groups on the problem.

Table 5.18. Numbers of graduate students (N), averages (Avg.) out of 10 points, and standard deviations (Std. dev.) for the scores of different groups of graduate students on the problem.

	N	Avg.	Std. dev.
GrSL1	14	8.5	2.3
GrSL2	12	8.9	0.8

5.4 QUALITIATIVE RESULTS FROM INDIVIDUAL STUDENT INTERVIEWS

5.4.1 Qualitative results relevant to the main quantitative finding

In the quantitative section we discussed that the fact that students in the SL2 group were less consistent than students in the SL1 group is surprising because those interventions differ by something that was intended to make students in group SL2 *better* at graphing, not *worse*. We also found it puzzling that students in the SL2 group obtained lower scores than students in the SL1 group. It was difficult to come up with a reasonable hypothesis that would explain these unexpected findings. Therefore, we conducted in-depth individual interviews using a think-aloud protocol with six students in order to obtain a better grasp of what may be hindering their (students in the SL2 group) reasoning and to figure out what may be causing their poor performance.

Interviews provided a possible explanation for students' poor performance in the SL2 group. One cognitive framework that can explain this poor performance is cognitive load theory [47,48] (short term memory or STM). In this framework problems are solved by processing relevant information in the working memory or STM [49-52]. However, working memory has been shown to be finite (5-9 "slots") for any person regardless of their intellectual capabilities [53,54]. In order to solve a problem one has to figure out the relevant information that must be processed at a given time in order to move forward with a solution. Some of the relevant information to solve a problem must be retrieved from long term memory (for example, relevant principles, e.g., Newton's second law, conservation of energy, physics concepts, mathematical information, etc.) Experts generally solve problems by focusing on important features of the problem and by retrieving the appropriate information from their long-term memory [55-59], which has a well-organized knowledge hierarchy in their domain of expertise. Novices do not have a robust knowledge structure and they are more likely to focus on unimportant features of the problem and retrieve information that is not necessarily useful [60-63]. Since their knowledge chunks are smaller, novices are also more likely to have cognitive overload while solving problems if there is too much information to keep track of during problem solving.

The Scaffolding Level 2 (SL2) group task included the extra instructions (as compared to SL1) to find the electric field at the beginning, mid and endpoint of each interval. A cognitive task analysis from an expert point of view suggests that these are good things to calculate before graphing a function because they give you explicit information about the function which is helpful for graphing it. Discussions with the graduate students in the TA training class after they had solved the two different versions indicated that they thought these instructions (although they did not need them) would definitely be helpful for introductory physics students. However, the interviews suggested that the introductory physics students in the SL2 group for whom the additional instructions were included did not discern the relevance of these instructions to graphing the function in the next part, and to them, evaluating the function at various points in a given interval was just another chore. For example, asking them to calculate $E\ (r \rightarrow a)$ in region $a < r < b$ implies they had to "find the limit of E as r approaches a from the right" (we will henceforth refer to these instructions, i.e. $E(r \rightarrow a)$, $E(r \rightarrow b)$, etc. as "limits"). While asking introductory students these additional questions before graphing was meant to provide scaffolding for graphing the function, interviews suggested that these additional questions may have caused cognitive overload since they required additional information processing. Interviews suggested that these students were more likely to lose track of important, relevant information and sometimes even omitted reading instructions carefully. *Every single student interviewed* who had to evaluate the electric field at three points in each interval before graphing it did not read the instructions carefully at one point or another. Some forgot to graph the electric field in a particular region, some went straight to evaluating the limits even before finding an expression for the electric field in that region. An interesting example of losing track of important information comes from an interview with John. In finding the limits of the function in regions $r < a$ and $a < r < b$, John did not plug in the corresponding values for r. For example, he wrote $E(r \rightarrow a) = kQ/r^2$ without plugging $r = a$ into the expression. But then when he got to the first limit in region $b < r < c$ ($E(r \rightarrow b)$), after writing down an initial expression in which he did not plug in $r = b$, he suddenly realized, without the interviewer saying anything, that he should plug in $r = b$.

John: Oh, should I plug in […] 'cause it's r approaching b?

Researcher: I can't tell you that. […] What do you think?

John: I'll just write it to be safe.

He then went back and changed all the previous limits where he had not plugged in the corresponding values for *r*. Thus, it appears that the piece of information "when you find a limit of a function, you have to plug in the value for the variable in that function" was present in his long term memory but he did not retrieve it until a particular point. He appeared to be focusing on and processing other information in the problem that was not helpful for figuring out the limits correctly. As noted earlier, *every single student interviewed* overlooked something in a somewhat similar manner while solving the different parts of the problem and the intended scaffolding involving explicit evaluation of the function at three points in each region did not help them in transforming the equation for the electric field in a particular region to the graphical representation correctly.

All of the interviewed introductory physics students made some mistakes in finding some limits (the most common one was copying down a limit from a previous region without thinking about what the expression of the electric field is in the region they were working in – for example copying down $E(r \to b)$ from region $a < r < b$ for $E(r \to b)$ in region $b < r < c$ even though the expressions for the electric field in those two regions do not match). After they were done with the problem, the students were asked to answer a follow-up question related to limits before discussing their solutions to the problem. They were given a piece-wise defined function in three different regions and were asked for three limits as follows:

$$f(x) = \begin{cases} 2, & 0 < x < a \quad (region\_1) \\ x + 2, & a < x < b \quad (region\_2) \\ 5, & x > b \quad (region\_3) \end{cases}$$

Region 1: $\lim_{x \to a} f(x)$

Region 2: $\lim_{x \to a} f(x)$

Region 3: $\lim_{x \to b} f(x)$

Even though students had to apply the same reasoning to determine the limits in the electric field problem and *none* of them managed to find them correctly there for all parts as discussed earlier, nearly all students solved this problem correctly without much trouble. One student solved it correctly after reasoning about this task for a long time; he did not readily figure it out. This is an indication of the difficulty students sometimes have with connecting physics and mathematics. The difference in performance on these two problems (one in the physics context and one without a physics context) suggests that while students were working on the limits in the electric field problem, they may have had difficulty processing the appropriate information systematically. This difficulty may partly be due to cognitive overload because they were focused on various details of the problem that were not relevant for computing the limits in a given region and they did not comprehend that evaluating the function at various points should be useful for graphing it.

5.4.2 Qualitative results from interviews relevant to the secondary quantitative results

Another surprising finding was that students provided with Scaffolding Level 1 exhibited improved performance in determining the correct expressions for the electric field in the different regions (they were more successful because they were typically outperforming students in the other groups in the first four parts of the problem, and the scores on those parts were based solely on the accuracy of the expressions). This was surprising because the intervention given to students in this group was intended to help them graph better and not at all aimed towards helping them find these expressions. It was also found that students in the SL1 group were more consistent, and although this was not surprising because the intervention was intended to make them more consistent, we wanted to identify how the intervention was causing their improved performance in terms of consistency. Therefore, we conducted another four interviews using a think-aloud protocol with students taking a similar second semester calculus-based introductory physics class in which we asked them to solve the SL1 version of the quiz. Unfortunately, we were unable to identify through these interviews how the intervention impacted them and helped them be more successful in finding the correct expressions and be more consistent. However, these interviews did provide some valuable insights into some possible reasons for the poor performance of students, both in finding expressions and in being consistent.

One of our findings from these interviews (observed also in some of the six earlier interviews where the problem given contained Scaffolding Level 2) was that some students were reluctant to think that $E = 0$ is not an acceptable mathematical expression. In these interviews, some students applied Gauss's law qualitatively correctly (most commonly in region $b < r < c$) which implied that the electric field vanishes, but instead of writing down $E = 0$ and moving on to the next region, they attempted to find a mathematical expression with variables in it (or constants from the problem, i.e., a, b, c, Q). Sara's interview provided one of the clearest examples of this because, when she got to region $b < r < c$, she said:

Sara: In there it should be zero because it's within a conductor.

Then, after a short pause:

Sara: Now, if only I could find an expression for that.

Instead of writing down $E = 0$ she tried to use Gauss's law mathematically, did so incorrectly and obtained $E = -4\pi c^2 + 4\pi b^2$. She then explicitly said:

Sara: The electric field will be equal to negative four pi c squared minus four pi b squared, and <u>it will be equal to zero</u>, I just know that.

It was very interesting to observe how some students did not observe the inconsistency of trying to find an expression other than $E = 0$ for a vanishing electric field. Even more interesting was the fact that Sara *was aware* that her expression was not consistent with what she was expecting ($E = 0$) because she said:

Sara: Hmm… that's not always gonna work out, that four pi c squared and four pi b squared will cancel out in the equation to give zero […] but I don't have anything better in my head right now.

In Sara's case, this reluctance to take $E = 0$ as an acceptable expression was partly influenced by her reluctance to believe her qualitative (conceptual) reasoning using Gauss's law. She was more inclined to trust a result after it followed from a mathematical procedure. This is why she wrote $E = -4\pi c^2 + 4\pi b^2$ instead of $E = 0$ in region $b < r < c$. She also did something similar in the first region, $r < a$, where she initially used Gauss's law qualitatively to obtain that the electric field was zero. But instead of trusting this result, she tried to solve this part in a different, more mathematical way. She incorrectly remembered that $E = qF$ (she was trying to recall the connection between electric field and electric force, namely $F = qE$) and then noted that because the charge is zero, E will be zero. This second approach, although incorrect, was more mathematical and Sara trusted the result more now than when she used qualitative reasoning. She even made a comment that indicated she was not very sure that the equation she remembered, $E = qF$, was correct, but still trusted this more. She then noted "either way, you get zero", which indicates that she was aware that she solved this part with two different approaches, both of which yielded the same result. When these two approaches resulted in different answers, she trusted the mathematical result more (as she did in region $b < r < c$). When asked about this (after the interview) she noted the following:

Sara: Sometimes, I need the conceptual to pull me into the math, but when they don't line up, […] you just have to go with the math.

These types of reasoning can partly account for the poor performance exhibited by students in regions $r < a$, $b < r < c$ and $r > c$. At most 60% of students wrote that $E = 0$ in any of these regions (for students in the SL1 group); however, in the SL0 and SL2 groups, the percentages were much lower (sometimes as low as 17% for SL2 as shown in Table 5.5; and 37% for the SL0 group as shown in Table 5.13).

Interviews also suggested a reason which can partly account for the lack of consistency. In particular, when graphing the electric field students sometimes did not trust the results coming from the mathematical procedure/approach (i.e. graphing their expressions), and graphed the behavior they expected from qualitative reasoning. Sara, for example, in region $b < r < c$, even though her expression was $E = -4\pi c^2 + 4\pi b^2$, graphed a zero electric field because she knew that it was supposed to vanish. Similarly in region $a < r < b$, her expression was $E = -4\pi b^2 + 4\pi a^2$, but instead of graphing this she said:

Sara: For r between distances a and b […] we dropped off with E being proportional to 1/r²

She then graphed a function that decreases in this way instead of graphing her expression (a constant negative function).

Another interviewed student, Joe, had very similar approaches. In region $b<r<c$, he found a non-zero mathematical expression, $k|Q||\rho\ |/r^2$, (in this formula, ρ refers to volume charge density) but when he had to graph it in this region, he said the following:

Joe: There's gonna be no electric field inside this region because the charges [–Q] are all on this [inner] surface.

Even though he was aware that the electric field should vanish ("no electric field inside this region") he trusted the mathematical expression he found and did not modify his expression to $E = 0$. Similarly to Sara, when he graphed the electric field in this region, he also graphed a zero electric field instead the expression he found ($\sim 1/r^2$).

Thus, some students were aware that the electric field vanished in a region ($r < a$, $b < r <$ c or $r > c$), but they did not believe that $E = 0$ was an acceptable expression and attempted to find "the real" expression by using a mathematical procedure (either using Gauss's law mathematically, or trying to remember a formula that may be applicable). If the attempt resulted in the same answer ($E = 0$) students would write it down. However, more often (especially for regions $b < r < c$ and $r > c$) the attempts resulted in non-zero mathematical expressions which most of the interviewed students trusted despite the fact that they were *aware* at least at some point in the problem solving process that the electric field *should vanish* in those regions. This may be a factor contributing to the poor performance exhibited by students in these regions. A few students encountered a similar difficulty in the region which had a non-zero electric field ($a < r < b$) for which the mathematical procedure did not result in the expected behavior of the electric field ($\sim 1/r^2$, or constant in the case in which the student considered the situation as a spherical capacitor and used contributions both from the inside and the outside and thought that as one contribution gets stronger, the other gets weaker, but the sum is constant). Being unable to reconcile the two approaches, qualitative and mathematical, students trusted the mathematical approach more. As mentioned before, when some students graphed the electric field, instead of graphing the corresponding functions obtained from their mathematical procedures, they graphed the expected behaviors obtained from qualitative reasoning. This can account for the lack of consistency observed. In other words, some students are not inconsistent between the expressions they find and the graphs they draw not because they do not know *how* various functions ($\sim r$, $\sim r^2$, $\sim 1/r$, constant etc.) *are supposed to be graphed*, but because they *are not graphing those functions*. Instead, they are graphing *other functions* which they did not write down, that were obtained through qualitative reasoning.

5.5 DISCUSSION AND SUMMARY

We found that providing calculus-based introductory physics students with additional scaffolding (by asking them to evaluate the electric field at the beginning, mid, and endpoint of each interval), although intended to help students be more consistent in graphing the electric field, had an adverse effect on their performance (both in terms of score and consistency). On the other hand, the graduate students in a TA training course who were randomly given the SL1 and SL2 versions of the problem did not exhibit any statistically significant differences in their performance on the two versions. Discussions with graduate students after 14 of them solved the SL2 version of the problem indicated that, similar to the researchers, they also thought that these additional instructions were good instructions to include when one is asked to sketch the electric field. During the discussions, some of the graduate students who were given the SL1 version noted that they implicitly calculated the electric field at various points using the functional form in order to graph it in the relevant region. On the other hand, conducted think-aloud interviews with introductory physics students suggested that they did not discern the relevance of these additional instructions in the SL2 version and showed signs of having cognitive overload due to the additional instructions while engaged in solving this problem.

We also found that asking introductory students to graph the electric field in each region immediately after finding an expression for it in that region (the SL1 intervention) impacted students positively, resulting in better performance in determining the correct expressions for the electric field and improved likelihood of being consistent in graphing the expressions they found. We hypothesized that giving introductory students coordinate axes with the irrelevant regions shaded out may have focused their attention on the relevant information in the problem that must be taken into account while finding the expressions for the electric field. Often, introductory students may focus on pieces of information that are not necessarily helpful for solving a problem, which can in turn cause cognitive overload due to finite capacity of working memory and less chunking of knowledge related to physics for beginning students. Therefore, if students' attention is drawn mostly to the relevant information (as interviews suggested was the case in intervention SL1) and they do not have a cognitive overload they would be more effective problem solvers. Thus, cognitive load theory is one framework which can partly account for their improved performance in terms of finding the correct expressions for the electric field for students in intervention SL1. In terms of consistency, asking students to draw the electric field in each region immediately after writing an expression for the electric field in that region, and shading out the irrelevant regions, was an attempt to help them keep track of information related to graphing in a more expert-like manner, which would entail making sense of the different functions in the different regions and drawing each of those functions carefully on the graph. It appears that more students in the SL1 group, due to their much higher rates of consistency, followed an expert-like way of graphing a piece-wise defined function. This may account for why these students were more consistent than students in the comparison group. In addition, we found that students who were always consistent in plotting the expressions they found were outperforming the other students. Both the p values and the effect sizes showed large significant differences.

Another finding was that the percentages of students from the SL2 and comparison groups who always drew graphs consistent with the expressions they found did not change from the quiz to the final exam (see Tables 5.8, 5.14 and 5.17). This finding may partly be due to the fact that in the class, graphing was not emphasized and this type of exercise was not common in the quizzes or class examples. Therefore, the performance of these two groups in this respect is a baseline for what students can achieve without instruction. It is interesting that for students in the SL1 group, the story is very different. In the quiz, their performance on consistency (see Tables 5.8 and 5.14) is significantly better than the other two groups. Not only that, but even though the quiz was virtually the only time students received some help (via the intervention) with graphing the function, in the final exam, although their performance decreased, it did not drop back down to the baseline (see Table 5.17). Therefore, it appears that the effects of the intervention have stuck somewhat with some of the students for quite some time. It is possible that, while solving the quiz problem, these students were learning more about graphing and what they learned was more likely to be remembered later because they learned it *on their own*. One could even go one step further and hypothesize that if there had been more problems both in class and recitations that dealt with graphing, and more quiz problems with the same intervention, by the end of the semester, the performance of students in the SL1 group may have gotten better in the final exam compared to the quiz instead of worse. These hypotheses will be investigated in future research.

Also, think-aloud interviews conducted with students asked to solve the version of the quiz containing SL1 (which asked them to graph the electric field in each region immediately after finding an expression) provided some reasons for the poor performance of students in terms of consistency and finding an expression for the electric field. In particular, we found that some students were aware that the electric field is supposed to vanish in a region (either from using Gauss's law qualitatively or from remembering that the electric field is always zero inside a conductor) but were reluctant to think that $E = 0$ is an acceptable expression and tried to find a mathematical expression with variables and constants from the problem. While writing down an expression, they preferred the mathematical, non-zero expression (despite concluding intuitively at the beginning that the electric field should vanish), but while graphing the electric field, they preferred the expression agreeing with the behavior that they expected (E=0). This would lead to a lower score because the scores in the first four parts are based on the expressions and lower rates of consistency (as well as lower scores on the graphing part), since we investigated how consistent students were between the expressions they found and the graphs they drew.

Another finding from the interviews which corroborates previous research done on students' understanding of electricity and magnetism concepts [64,65] is the inability or reluctance of students, even high achieving ones, to use Gauss's law mathematically. Only one out of the ten interviewed students applied Gauss's law correctly. Students either tried to remember a formula derived in class for spherical symmetry or, when they tried to apply Gauss's law for a spherically symmetric charge distribution, they made mistakes, mostly because they ended up evaluating an integral that they were not sure how to simplify by symmetry arguments (i.e., their integral did not reduce to $EA_{surface}$ as it should in this case with a spherically symmetric charge distribution). Also, sometimes they applied the boundaries of the region they were working in as the lower and upper limits of the integral (i.e., in region $a<r<b$ they used $r=a$ as the lower limit and $r=b$ as the upper limit) and evaluated a definite integral by making use of those limits instead of choosing a Gaussian sphere with a radius equal to the distance from the center of the sphere where the electric field was to be calculated. One instructional implication is that Gauss's law is challenging to apply correctly to calculate the magnitude of the electric field for a highly symmetric charge distribution, e.g., spherically symmetric distribution, even for high-achieving students. Students should be guided to understand how Gauss's law simplifies in highly symmetric cases if a Gaussian surface is chosen appropriately and symmetry arguments are used to infer the direction of the electric field and how this simplification is essential in order for it to be useful to find the magnitude of the electric field (in most, if not all cases discussed in

a typical introductory physics class). It may be helpful to stress that Gauss's law is always true for electric flux through any closed surface, but rarely helpful to find the magnitude of the electric field, and ask students to consider physical situations in which the law is not useful for finding the magnitude of electric field because there may be a lack of symmetry, or one cannot find a Gaussian surface such that the flux through each part of the surface is either zero or can easily be written in terms of the magnitude of the electric field and the area of that part of the Gaussian surface. One example of this would be to ask students if Gauss's law can easily help us find the magnitude of electric field inside and outside a configuration of charges equal in magnitude and sign placed at the corners of a cube. Despite the apparent symmetry, finding a Gaussian surface that can simplify the problem is virtually impossible and if one attempts to use the obvious choice (the surface of a cube with a larger side having the same center) the requirements of the Gaussian surface in order for it to be useful to find the magnitude of the electric field due to the charge distribution are not satisfied. Other approaches to teaching Gauss's law are discussed in [65].

5.6 CHAPTER REFERENCES

1. R. Beichner (1994). "Testing student interpretation of kinematics graphs." Am. J. Phys. 62, 750.

2. L. C. McDermott, M. L. Rosenquist, and E. H. van Zee (1987). "Student difficulties in connecting graphs and physics: Examples from kinematics." *Am. J. Phys.* 55, 503.

3. F. M. Goldberg and J. H. Anderson (1989). "Student difficulties with graphical representation of negative velocity." Phys. Teach. 27, 254.

4. C. A. Berg and P. Smith (1994). “Assessing students’ abilities to construct and interpret line graphs: Disparities between multiple-choice and free-response instruments.” Sci. Educ. 78, 527.

5. I. Testa, G. Muoroy, and E. Sassi (2002). “Students’ reading images in kinematics: the case of real-time graphs.” Int. J. Sci. Educ. 24, 235.

6. J. Clement (1985). “Misconceptions in graphing.” Proceedings, Ninth Conference of the International Group for the Psychology of Mathematics Education, edited by L. Streefland (Noordwijkerhout, The Netherlands 1985), pp. 369-375.

7. M. L. Rosenquist and L. C. McDermott (1985). Kinematics (ASUW Publishing, University of Washington, Seattle, 1985).

8. M. L. Rosenquist and L. C. McDermott (1987). “A conceptual approach to teaching kinematics.” Am. J. Phys. 55, 407.

9. W. L. Barclay III (1986). “Graphing misconceptions and possible remedies using microcomputer-based labs.” Technical Report No. TERC-TR-85-5, Cambridge, MA.

10. R. K. Thornton (1987). “Tools for scientific thinking: Microcomputer-based laboratories for physics teachers.” Phys. Educ. 22, 230.

11. J. R. Makros and R. F. Tinker (1987). “The impact of microcomputer-based labs on children’s ability to interpret graphs.” J. Res. Sci. Teach. 24, 369.

12. L. K. Wilkinson, J. Risley, J. Gastineau, P. V. Engelhardt, and S. F. Schultz (1994). “Graphs and tracks impresses as a kinematics teaching tool.” Computers in Physics 8, 696.

13. E. Gire, D.-H. Nguyen, and N. S. Rebello (2011). “Characterizing students' use of graphs in introductory physics with a graphical analysis.” Proceedings of the 2011 National Association for Research in Science Teaching Annual Meeting, April 3-6, 2011, Orlando, FL.

14. E. B. Pollock, J. R. Thompson, and D. B. Mountcastle (2007). “Student understanding of the physics and mathematics of process variables in P-V diagrams.” Proceedings of the 2007 Physics Education Research Conference, edited by L. Hsu, C. Henderson and L. McCullough, AIP Conf. Proc. No. 951 (AIP, Melville, NY, 2007), pp. 168-171.

15. M. E. Loverude, C. H. Kautz, and P. R. L. Heron (2002). “Student understanding of the first law of thermodynamics: Relating work to the adiabatic compression of an ideal gas.” Am. J. Phys. 70, 137.

16. D. E. Meltzer (2004). “Investigation of students’ reasoning regarding heat, work, and the first law of thermodynamics in an introductory calculus-based general physics course.” Am. J. Phys. 72, 1432.

17. W. M. Christensen and J. R. Thompson (2012). "Investigating graphical representations of slope and derivative without a physics context." Phys. Rev. ST Phys. Educ. Res. 8, 023101.

18. T. A. Romberg, E. Fennema, and T. P. Carpenter (1993). Integrating research on the graphical representation of functions (Erlbaum, Hillsdale, NJ 1993).

19. G. Harel and E. Dubinsky (1991). The Concept of Function: Aspects of Epistemology and Pedagogy (MAA Notes, Vol. 25. Mathematical Association of America, Washington D. C. 1991).

20. F. Hitt (1998). "Difficulties in the articulation of different representations linked to the concept of function." J. Math. Behav. 17, 123.

21. S. Vinner and T. Dreyfus (1989). "Images and definitions for the concept of function." J. Res. Math. Educ. 20, 356.

22. R. Thompson (1993). "Students, functions and the undergraduate curriculum." Paper presented at the Annual Joint Meeting of the American Mathematical Society and the Mathematical Association of American, San Antonio, 12-16 January 1993.

23. S. Vinner (1991). "The role of definitions in the teaching and learning of mathematics." Advanced Mathematical Thinking, edited by D. Tall, (Kluwer, Dordrecht, The Netherlands, 1991), pp. 65-81.

24. S. Arnold (2006). "Investigating functions using real-world data." Aust. Sen. Math. Jour. 20, 44.

25. M. Borenstein (1997). "Mathematics in the real world." Learn. Lead. Tech. 24, 34.

26. R. Hershkowitz and B. B. Schwartz (1997). "Unifying cognitive and sociocultural aspects in research on learning the function concept." Proceedings of the conference of the international group for the psychology of mathematics education, vol. 1, edited by E. Pehkkonnen, Lathi, Finland.

27. R. Y. Shorr (2003). "Motion, speed, and other ideas that 'should be put in books'." J. Math. Behav. 22, 465.

28. G. Leinhardt, O. Zaslavsky, and M. K. Stein (1990). "Functions, graphs, and graphing: Tasks, learning, and teaching." Rev. Educ. Res. 60, 1.

29. H. Brasell and M. Rowe (1989). "Graphing skills among high-school physics students." Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

30. T. Dreyfus and T. Halevi (1991). "QuadFun: A case study of pupil computer interaction." J. Comp. Math. Sci. Teach. 10, 43.

31. R. Even (1998). "Factors involved in linking representations of functions." J. Math. Behav. 17, 105.

32. A. Bell and C. Janvier (1981). "The interpretations of graphs representing situations." For the Learning of Mathematics 2, 34.

33. D. Hammer (1995). "Epistemological considerations in teaching introductory physics." Sci. Educ. 79, 393.

34. N. S. Rebello (2009). "Can we assess efficiency and innovation in transfer?" Proceedings of the 2009 Physics Education Research Conference, edited by M. Sabella, C. Henderson and C. Singh, AIP Conf. Proc. No. 1179 (AIP, Melville, NY, 2010), 241-245.

35. See Transfer of Learning from a Modern Multidisciplinary Perspective, edited by J. Mestre (Information Age, Greenwich, CT, 2005).

36. J. F. Wagner (2010). "A transfer-in-pieces consideration of the perception of structure in the transfer of learning." J. Learn. Sci. 19, 443.

37. J. Larkin (1980). "Skilled problem solving in physics: A hierarchical planning approach." J. Struct. Learn. 6, 121.

38. K. Ericsson and H. Simon (1980). "Verbal reports as data." Psychol. Rev. 87, 215.

39. K. Ericsson and H. Simon (1993). Protocol Analysis: Verbal Reports as Data (MIT Press, Boston, MA 1993).

40. S. F. Chipman, J. M. Schraagen, and V. L. Shalin (1996). "Introduction to cognitive task analysis." Cognitive Task Analysis, edited by J. M. Schraagen, S. F. Chipman & V. J. Shute (Lawrence Erlbaum Associates Mahwah, NJ 2000), pp. 3-23.

41. R. E. Clark and F. Estes (1996). "Cognitive task analysis." Int. J. Educ. Res. 25, 403.

42. D. H. Jonassen, M. Tessmer, and W. H. Hannum (1999). Task analysis methods for instructional design (Lawrence Erlbaum Associates, Mahwah, NJ 1999).

43. G. V. Glass and K. D. Hopkins (1996). Statistical Methods in Education & Psychology (Allyn & Bacon, Boston, MA 1996).

44. J. Cohen (1988). Statistical Power Analysis for the Behavioral Sciences (Erldbaum, Hillsdale, NJ 1988).

45. R. L. Rosnow and R. Rosenthal (1996). “Computing contrasts, effect sizes, and counternulls on other people’s published data: General procedures for research consumers.” Psychol. Methods 1, 331.

46. R. A. Fisher (1922). “On the interpretation of χ^2 from contingency tables, and the calculation of P.” J. Roy. Stat. Soc. 85, 87.

47. J. Sweller (1988). “Cognitive load during problem solving: Effects on learning.” Cog. Sci. 12, 257.

48. F. Paas, A. Renkel, and J. Sweller (2004). “Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture.” Instruct. Sci. 32: 1–8.

49. J. R. Anderson (1995). Learning and Memory, (Wiley, New York 1995).

50. P. Kyllonen and R. Christal (1990). “Reasoning ability is (little more than) working memory capacity?!” Intelligence 14, 389.

51. A. Fry and S. Hale (1996). “Processing speed, working memory and fluid intelligence: Evidence for a developmental cascade.” Psychol. Sci. 7, 237.

52. R. Kail and T. Salthouse. “Processing speed as a mental capacity.” Acta Psychol. 86, 199.

53. G. Miller (1956). “The magical number seven, plus or minus two: Some limits on our capacity for processing information.” Psychol. Rev. 63, 81.

54. A. Miyake, M. A. Just, and P. Carpenter (1994). “Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts.” J. Mem. Lang. 33, 175.

55. M. T. H. Chi, P. J. Feltovich, and R. Glaser (1981). “Categorization and representation of physics knowledge by experts and novices.” Cogn. Sci. 5, 121-152.

56. K. Johnson and C. Mervis (1997). “Effects of varying the levels of expertise on the basic level of categorization.” J. Exp. Psychol. Gen. 126, 248.

57. B. Eylon and F. Reif (1984). “Effect of knowledge organization on task performance.” Cogn. Instruct. 1, 5.

58. F. Reif and J. I. Heller (1982). “Knowledge structure and problem solving in physics.” Educ. Psychol. 17, 102.

59. J. L. Docktor, J. P. Mestre, and B. H. Ross (2012). “Impact of a short intervention on novices’ categorization criteria.” Phy. Rev. ST Phys. Educ. Res. 8, 020102.

60. A. Schoenfeld and D. J. Herrmann (1982). "Problem perception and knowledge structure in expert novice mathematical problem solvers." J. Exp. Psychol. Learn. 8, 484.

61. P. W. Cheng and K. J. Holyoak (1985). "Pragmatic reasoning schema." Cogn. Psychol. 17, 391.

62. J. Larkin (1980). "Skilled problem solving in physics: A hierarchical planning model." J. Struct Learn. 6, 271.

63. M. S. Sabella and E. F. Redish (2007). "Knowledge organization and activation in physics problem solving." Am. J. Phys. 75, 1017.

64. C. Singh (2006). "Student understanding of symmetry and Gauss's law of electricity." Am. J. Phys. 74, 903.

65. J. Li (2009). "Improving Students' Understanding of Electricity and Magnetism." Ph.D. Thesis, University of Pittsburgh.

6.0 EXPLORING ONE ASPECT OF PEDAGOGICAL CONTENT KNOWLEDGE OF TEACHING ASSISTANTS USING THE TEST OF UNDERSTANDING GRAPHS IN KINEMATICS

6.1 INTRODUCTION

The Test of Understanding Graphs in Kinematics (TUG-K) [1] is one of many multiple-choice tests designed to assess conceptual understanding in introductory physics [2-11]. Some of these tests, e.g., the Force Concept Inventory [3], have been widely used by instructors and education researchers for various purposes, for example, to identify student difficulties [2,12], to compare the effectiveness of curricula and pedagogies [13], and to investigate gender differences [14,15]. The TUG-K was developed by Beichner to assess students' understanding of kinematics graphs after early physics education research which revealed that introductory physics students have many difficulties with constructing and interpreting graphs in kinematics [1,16-24]. Helping introductory physics students become facile with different representations of concepts is a critical component of the development of expertise in physics. Facility with graphical representations is particularly important and thus, graphical representation have been emphasized extensively in research-based instructional tools, e.g., in multimedia learning modules [25-27].

The TUG-K was developed by taking the common difficulties of introductory students in interpreting graphs, revealed by research, into consideration and many items on the test include strong distractor choices which uncover that some difficulties are very common. Beichner subjected the test to much statistical analysis (including calculation of KR-20, point biserial coefficients, Ferguson's delta and others) to ensure that it is a reliable instrument for assessing understanding of kinematics graphs. In addition, in the construction phase of the test, he asked many educators at different institutions for feedback on the items on the test in order to ensure content validity.

There are several theoretical frameworks that inspire our research and focus on the importance of instructors familiarizing themselves with students' prior knowledge (including what students learn from traditional instruction) in order to scaffold their learning with appropriate curricula and pedagogies. In the context of this study, they point to the importance of being knowledgeable about student difficulties in order to help students learn better. For example, Piaget [28] emphasized "optimal mismatch" between what the student knows and where the instruction should be targeted in order for desired assimilation and accommodation of knowledge to occur. A related framework is Posner et al.'s theory of conceptual change [29]. In this model, they suggest that conceptual changes or "accommodations" can occur when the existing concepts students have are not sufficient for or inconsistent with new phenomena. They also suggest that these accommodations can be very difficult for students, particularly when students are firmly committed to their prior understanding. This model suggests that it is important for instructors to be knowledgeable about student ideas, which, when applied to particular physics contexts can lead to difficulties. Within this model, if students are motivated by an anomaly which provides a cognitive conflict that illustrates how their conceptions are inadequate for explaining a newly encountered physical situation, they can become dissatisfied with their current concepts and improve their understanding. But instructors must be aware of what conceptions students have, and what difficulties these conceptions can lead to in order to design a task that produces the desired cognitive conflict.

The research presented here uses the TUG-K (along with the original student data in Ref. [1]) to explore one aspect of the pedagogical content knowledge of first-year graduate students, namely, knowledge of common introductory student difficulties. The graduate students were enrolled in a semester long TA training course at the University of Pittsburgh (Pitt). Towards the end of the semester, the graduate students performed a task which used the TUG-K survey to investigate how knowledgeable the graduate students are about common student difficulties related to graphical representations of motion. For each item on the TUG-K, the graduate students were asked to identify which one of the four incorrect answer choices was, in their view, the most common incorrect answer choice of introductory physics students if they did not know the correct answer after instruction in relevant content. The graduate students first carried out this task individually followed by repeating the task in groups of two or three. A class discussion related to their responses followed these exercises.

Pedagogical content knowledge (PCK) is a term coined by Shulman [30,31] to mean the subject matter knowledge *for teaching* and many researchers in K-16 education have used this construct [32-36]. Shulman defines PCK as “a form of practical knowledge which guides the pedagogical practices of educators in highly contextualized settings” [30]. According to Shulman, PCK is comprised of the most useful forms of representations of the topics and concepts, powerful analogies, illustrations and examples, and “understanding of what makes the learning of specific topics easy or difficult” [30]. Therefore, knowledge of student difficulties is an important aspect of PCK and the research presented here was designed to explore this aspect of the PCK of graduate students: knowledge of common introductory student difficulties with kinematics graphs identified by the TUG-K. We refer to this as the "TUG-K related PCK" of graduate students. The graduate students who teach recitations for introductory physics courses typically have a closer association with introductory students than the course instructors because they hold regular office hours and interact with introductory students in the physics resource room at Pitt where they help introductory students. In addition, recitation sizes are usually much smaller than the sizes of lecture classes taught by instructors. Therefore, TAs who are knowledgeable about introductory student difficulties in interpreting kinematics graphs can play a significant role in improving introductory student understanding of kinematics and they can address these difficulties directly in their interactions with students. Of course, it is also important for instructors to be knowledgeable of student difficulties in order to design instruction to effectively address and remedy these difficulties.

Research questions: Performance of graduate students at identifying introductory physics students' difficulties related to kinematics graphs on the TUG-K

The following research questions were developed for the purpose of investigating the TUG-K related PCK of graduate students:

I. To what extent are American physics graduate students, who have been exposed to undergraduate teaching in the United States, better at identifying introductory student difficulties than foreign physics graduate students?

Graduate programs across the United States are populated by many foreign graduate students. According to recent AIP statistics, almost half of the first-year physics graduate students in US universities are non-US citizens [37], and more than half of the awarded physics PhDs are to non-US citizens [38]. A majority of physics departments in the United States require that graduate students become TAs for undergraduate courses at least for one or two semesters. Since the influence of foreign graduate students in physics undergraduate education is becoming commensurate (at least in terms of numbers of TAs) with that of American graduate students, it is worthwhile comparing the knowledge that these two different groups of graduate students have regarding introductory student difficulties with physics. The educational backgrounds of these two groups of graduate students are very different and it is unclear whether these backgrounds have a significant effect on developing an understanding of the difficulties of introductory physics students with physics content, in particular, with kinematics graphs for our research presented here.

II. To what extent do graduate students identify introductory students' difficulties more often when working in groups than when working individually (i.e., do discussions improve graduate students' understanding of introductory students' difficulties with kinematic graphs?)

Peer discussions have been found to be productive in the context of learning physics [12,39]. It is useful to investigate if discussions with peers are also productive in the context of learning about student difficulties related to kinematics concepts.

III. To what extent do graduate students identify 'major' introductory student difficulties compared to 'moderate' ones? (Major and moderate difficulties are defined later.)

Research in physics education has shown that introductory students encounter many common difficulties in learning physics that must be taken into account in the design of curricula and pedagogies to help students build good mental models. These difficulties are of varying degrees, and while one may assume that the more common difficulties are easier to identify, this may not be true. In particular, in a particular content area, cognitive task analysis of the underlying knowledge from the expert perspective can fail to identify common difficulties that are actually found via research. Therefore, in the context of difficulties with kinematics graphs, we investigated to what extent the major difficulties of introductory students were identified by graduate students compared to the moderate ones.

IV. To what extent do graduate students identify specific introductory student difficulties with kinematic graphs? Is their ability to identify these difficulties context dependent? (A particular graphical concept is probed in different contexts in different questions on TUG-K)

The TUG-K reveals several different types of student difficulties with kinematics graphs which are identified by student responses to several questions. We investigated the extent to which graduate students are able to identify specific difficulties of introductory students. Physics education research has shown that introductory student performance is context dependent, i.e., correct application of physics concepts depends on the contexts of the questions posed. Here, we investigate whether the ability of graduate students to identify common introductory student difficulties is also context dependent.

For multiple choice questions, the context is comprised of both the physical situation presented in the problem and the answer choices because different answer choices can change the difficulty of a question. For example, a multiple-choice question is easier for introductory students if the incorrect answer choices are not chosen to reflect common student difficulties, and are challenging for students when they are chosen to reflect common difficulties [2-3]. For the TUG-K, our use of the term context refers to the type of graph presented (position, velocity, acceleration), the type of task (conceptual vs. quantitative) and the answer choices. A conceptual and a quantitative question posed with the same type of graph provide different contexts (for example, items 2 and 6 on the TUG-K). Similarly, two quantitative questions with the same type of graph provide different contexts if their answer choices do not reflect the same type of student difficulties (for example, items 6 and 7 on the TUG-K: item 6 provides an answer choice which corresponds to the student difficulty related to computing slopes by calculating y/x instead of $\frac{\Delta y}{\Delta x}$, but item 7 does not use this type of answer choice).

6.2 METHODOLOGY

6.2.1 Materials and Participants

The materials used for this study were the TUG-K survey developed by Beichner along with the data in Beichner's original paper [1], which was collected from more than 500 college and high-school students.

The participants of this study were twenty-five first-year physics graduate students enrolled in a TA training class in their first semester in graduate school. The TA training class is a pedagogy oriented semester long course which is required of all first-year graduate students at Pitt. The course meets once a week for two hours and is designed to help graduate students be more effective teachers. During the course, students learn about cognitive research and physics education research (PER) and discuss their instructional implications. Students are also introduced to curricula and pedagogies based on physics education research which stress the importance of being knowledgeable about introductory students' difficulties in order to help them transition toward expertise. Each graduate student also discusses the solution of a physics problem in the class in the manner in which they would discuss it if they were teaching introductory students and they receive feedback from the other graduate students and the instructor. Also, during the course, students complete various reflective exercises aimed at helping them perform their TA duties in a student-centered manner.

All but three of the graduate students who participated in this study were teaching introductory physics recitations or labs for the first time. Two of the three who were not teaching had physics teaching experience as undergraduates, either as a teaching assistant or as a tutor for introductory physics courses. Only one student did not have teaching experience with physics, but this student tutored mathematics as an undergraduate. Also, in the TA training course introductory student difficulties were discussed, however, not in the specific context of interpreting kinematics graphs (until after students completed all tasks related to the TUG-K as described below).

6.2.2 Methods

Toward the end of the TA training class (so that a majority of graduate students had almost a semester worth of teaching experience), the graduate students were asked to complete three different tasks related to the TUG-K: (1) while working individually, they were asked to identify the correct answers for each question; (2) while working individually, for each question on the TUG-K, they were asked to identify which one of the four *incorrect* answer choices, in their view, would be most commonly selected by introductory physics students after instruction in relevant concepts if the introductory students did not know the correct answers and (3) they repeated the second task, except working in groups of two or three. The graduate students performed task (1) first, then task (2) and finally task (3) followed by a class discussion during a two hour TA training class. We refer to tasks (2) and (3) as individual and group TUG-K related PCK tasks. The graduate students were allowed as much time as they needed for each task. All graduate students finished the first task within the first 30 minutes and the second task within the first hour. The third task (group work) was completed by all groups within 40 minutes followed by a full class discussion about the PCK task.

In order to investigate the TUG-K related PCK of graduate students, scores were assigned to each graduate student as follows: a graduate student who selected a particular answer choice in a particular question received a score which was the fraction of introductory students who selected that particular answer choice. If a graduate student selected the correct answer choice, they would be assigned a score of zero because they were explicitly asked to indicate which *incorrect* answer choice is most commonly selected by introductory students. For example, on question 1, the percentages of introductory students who selected A, B, C, D and E are 40%, 16%, 4%, 22% and 17% respectively (see Table A1 in appendix A). Answer choice B is correct, thus, the score assigned for each answer choice on question 1 was 0.4, 0, 0.04, 0.22 and 0.17 (A, B, C, D and E). The score a graduate student would obtain on this PCK task for the whole test can be obtained by summing over all of the questions. A mathematical description of how this calculation was performed is included in the appendix.

In order to determine whether the graduate students performed better than random guessing on the TUG-K related PCK task, a population of random guessers was generated. The population was generated by choosing $N = 24$ 'random guessers' in order to have a reasonable group size when performing *t*-tests [40]. Random guessing on this task would correspond to selecting one of the four incorrect answer choices for each question with equal probability (25%). Therefore, one quarter of the random guessers always selected the first incorrect answer choice, one quarter selected the second incorrect answer choice, etc. Since the graduate students were not told the correct answers before they performed the TUG-K related PCK task, random guessing would not perfectly correspond to selecting one of the four incorrect answer choices with equal probability. For a particular question, there is a small probability that a graduate student does not know the correct answer. However, our data indicate that this probability is very small because in all but two questions, at least 24 out of 25 graduate students knew the correct answers. In the other two questions, 23 out of 25 and 22 out of 25 of the graduate students knew the correct answers (see table A1 in appendix A). Moreover, since for a given question, one quarter of the random guessers selected each of the four incorrect answer choices, one can calculate a mean and a standard deviation that can be used to perform comparison with the graduate student scores. Furthermore, our choice of random guessers maximizes the standard deviation.

We note that our approach to determine the TUG-K related PCK score of graduate students weighs the responses of graduate students by the percentage of introductory students who selected a particular incorrect response. This weighing scheme was chosen because the more prevalent an introductory student difficulty is, the more important it is for the graduate students to be aware of it and take it into account in their instruction.

6.2.3 Approach for answering the research questions

Performance of graduate students at identifying introductory physics students' difficulties related to kinematics graphs on the TUG-K

The researchers analyzed whether graduate students performed better at identifying introductory students' difficulties on the TUG-K than random guessing by performing statistical analysis.

I. To what extent are American physics graduate students, who have been exposed to undergraduate teaching in the United States, better at identifying introductory student difficulties than foreign physics graduate students?

Out of the twenty-five first year graduate students who participated in this study, nine were American, nine were Chinese and seven were from other foreign countries (Asia and Europe). The PCK scores of three groups of graduate students were compared (American, Chinese and other foreign students). The reason we divided the graduate students into three groups is because the American graduate students were exposed to teaching in the United States as opposed to the foreign students, who were not exposed to US teaching practices before graduate school and many were taught physics in their own native languages. The nine Chinese graduate students were placed in a separate group because, although they fit the category of foreign graduate students, it is possible that their backgrounds are different from the backgrounds of most of the other foreign graduate students.

II. To what extent do graduate students identify introductory students' difficulties more often when working in groups than when working individually? (i.e., do discussions improve graduate students' understanding of introductory students' difficulties with kinematic graphs?)

Previous studies have found that introductory students exhibit improved performance and conceptual understanding after engaging in discussions with one another [12,39]. We investigated whether discussions among graduate students related to introductory student difficulties improve their PCK performance related to kinematics graphs. Since the graduate students first performed the TUG-K related PCK task individually and then in groups, we investigated if their PCK performance increased in the group exercise compared to the individual exercise. In addition, we investigated whether the discussions shifted graduate students' selections towards more common introductory student incorrect answer choices. In particular, we identified how often two or three graduate students who worked together in the group TUG-K related PCK task, when completing the individual task, did not select the same answer as the most common difficulty with that question and when completing the group task, selected an answer choice which was connected to a more common (by 5% or more) introductory student difficulty.

III. To what extent do graduate students identify 'major' introductory student difficulties compared to 'moderate' ones?

Most of the questions on the TUG-K have strong distractor choices that are selected by many introductory students even after instruction. The researchers selected a heuristic such that an incorrect answer choice was connected to a 'major' student difficulty if more than 33% (or 1/3) of introductory students selected that answer choice. An incorrect answer choice was considered to be connected to a 'moderate' difficulty if between 20% and 33% of the introductory students selected that answer choice. In order to answer this research question, the average TUG-K related PCK scores of graduate students on questions that had major difficulties were compared to the average scores on questions that had moderate difficulties. However, for each question, the minimum and maximum possible scores are different because they correspond to the smallest and largest fraction of introductory students who select a particular incorrect answer choice. Therefore, for each question, the average score of graduate students was normalized to be on a scale from zero to a maximum possible score of 100 in order to make a comparison between different questions (see Table A2). This was done for each question in the following manner: grad student normalized score = 100 * (grad student average PCK score – minimum possible score) / (maximum possible score – minimum possible score). The normalized graduate student score on a particular question on the TUG-K is then zero if they obtained the minimum possible score and 100 if they obtained the maximum possible score.

IV. To what extent do graduate students identify specific introductory student difficulties with kinematic graphs? Is their ability to identify these difficulties context dependent?

This question was answered by identifying common introductory student difficulties on different questions and analyzing graduate students' PCK performance in identifying these common difficulties in different contexts.

6.3 RESULTS

Analysis of the PCK performance of the graduate students was performed on each of the questions on the TUG-K which revealed a moderate or major introductory student difficulty and it is shown in Tables A1 and A2 (included in Appendix A). Table A1 shows the percentages of introductory physics students and graduate students who selected each answer choice in each question on the TUG-K. The introductory students were asked to identify the correct answers, and the graduate students were asked to identify the incorrect answers which, in their view, were most common among introductory students for each question after instruction in relevant concepts. In Table A1, correct answers are indicated by the green shading, major introductory student difficulties (incorrect answer choices selected by more than 33% of the introductory students) are indicated by red shading and moderate difficulties are shown in red font. In addition, the second column (>RG) indicates whether the graduate students performed better than random guessing on each question (Yes/No).

Table A2 shows the normalized average TUG-K related PCK score (on a scale from 0 to 100) for the graduate students on each question that had moderate or major difficulties. The TUG-K related PCK performance of the graduate students on a given question was considered 'good' (and shaded green) if their normalized average PCK score is 67% or more of the maximum possible score, 'moderate' (and shaded yellow) if their normalized average PCK score is between 50% and 67% of the maximum possible score and 'poor' (shaded red) if their normalized average PCK score is less than 50% of the maximum possible score. These cutoffs were selected based on the normalized scores of the graduate students. The scores were put in order from smallest to largest and the bottom 1/3 of the scores correspond to poor performance, the middle 1/3 correspond to moderate performance and the top 1/3 of the scores correspond to good performance. Moreover, in Table A2, for questions that had moderate difficulties, the question numbers are in red font and for questions that had major difficulties, the question numbers are shaded red.

I. To what extent are American physics graduate students, who have been exposed to undergraduate teaching in the United States, better at identifying introductory student difficulties than foreign physics graduate students?

In order to answer this question, we compared the average PCK scores of different subgroups of graduate students. As noted earlier, the maximum PCK score on this task for any given question that a graduate student could obtain is the largest percentage of introductory students who selected a particular incorrect answer choice. The maximum PCK score on this task for the whole test is the sum of all these percentages which turns out to be 6.70.

Table 6.1. Numbers of American/Chinese/Other foreign graduate students, their averages (and percentage of those averages out of the maximum PCK score) and standard deviations (Std. dev.) for the PCK scores obtained for determining introductory student difficulties on the TUG-K out of a maximum PCK score of 6.70.

	N	Average	Std. dev.
American	9	4.00 (60%)	0.54
Chinese	9	4.24 (63%)	0.55
Other foreign	7	4.46 (66%)	0.59

Table 6.1 shows the averages and standard deviations of the PCK scores of the three different groups of graduate students. The group sizes are too small for meaningful statistics to be extracted from the data. However, it appears that the averages of the American, Chinese and Other foreign graduate students (60%, 63% and 66% of the maximum PCK score, 6.70,

respectively) are comparable. Therefore, it appears that American graduate students do not perform better at identifying introductory student difficulties (in fact, their average performance was somewhat lower than the performance of the foreign graduate students).

II. To what extent do graduate students identify introductory students' difficulties more often when working in groups than when working individually? (i.e., do discussions improve graduate students' understanding of introductory student difficulties with kinematics graphs?)

1) Graduate student TUG-K related PCK performance is significantly better when they worked in groups compared to when they worked individually.

Table 6.2 shows that the performance of graduate students when they worked in groups was better than when they worked individually. A *t*-test indicates that the difference in performance is statistically significant (p=0.033). In addition, calculation of Cohen's *d* [40] gives a reasonable effect size of 0.78.

Table 6.2. Number of graduate students/groups, averages (and percentage of those averages out of the maximum PCK score) and standard deviations for the PCK scores obtained for identifying the most common introductory student difficulties on the TUG-K out of a maximum PCK score of 6.70.

Individual	**N**	**Average**	**Std. dev**
	25	4.21 (63%)	0.57
Group	**N**	**Avg.**	**Std. dev**
	12	4.67 (70%)	0.59

2) Discussions among graduate students tend to converge on a more common introductory student difficulty.

We investigated how often graduate students who selected different answers in the individual TUG-K related PCK task, while working in groups, selected a 'better' answer (i.e., an incorrect answer choice which was connected to a more common, by 5% or more, introductory student difficulty). There were 74 instances in which two or three graduate students who did not all select the same answer in the individual TUG-K related PCK task (while identifying common introductory student difficulties) converged to one answer. In 45 of those instances (61%), they selected an incorrect answer which was more common (by 5% or more) among introductory students who did not know the correct answer. It therefore appears that discussions among graduate students were productive and led to a better understanding of introductory student difficulties related to kinematics graphs.

III. To what extent do graduate students identify 'major' student difficulties compared to 'moderate' ones?

As mentioned earlier, 'moderate' difficulties were considered to be connected to incorrect answer choices selected by between 20% and 33% of introductory students, while 'major' difficulties were those had by more than 33% of introductory students. There are 17 questions on the TUG-K which fit at least one of these two criteria (see Table A1 or A2 in Appendix A), eight of which have major introductory student difficulties and nine of which have moderate difficulties. Table A2 shows that the four questions on the TUG-K with the lowest graduate student PCK performance (questions 6, 8, 9 and 17) all contain a major introductory student difficulty. Moreover, the average PCK score of graduate students on the questions that had major difficulties is 48% compared to 61% on the questions that had moderate difficulties. It appears that the average graduate student TUG-K related PCK performance is better by 13% on questions with moderate introductory student difficulties than on questions with major ones. In other words, overall, graduate students identified moderate difficulties better than major ones.

IV. To what extent do graduate students identify specific introductory student difficulties? Is their ability to identify these difficulties context dependent?

These questions was answered by identifying common introductory student difficulties along with the questions in which these difficulties occurred and analyzing the graduate student TUG-K related PCK performance on those questions. Whenever a particular difficulty occurred in more than one question, it was investigated whether the PCK performance of graduate students was context dependent in that it was significantly different on different questions which had different contexts. We note that any interpretation of student difficulties presented here is taken from the original TUG-K paper. The focus of this research is not to discuss these difficulties, but to discuss the performance of the graduate students in identifying them.

Very few graduate students identified the common introductory student difficulty that graphs of time dependence of different kinematics variables that correspond to the same motion should look the same.

Table 6.3. Introductory student difficulty that graphs of time dependence of different kinematics variables that correspond to the same motion should look the same, items on the TUG-K which uncover this difficulty (TUG-K item #), percentage of introductory students who answer the items incorrectly (% overall incorrect), incorrect answer choices which uncover this difficulty, percentage of introductory students who have this difficulty based on their selection of these answer choices (% intro. stud. diff.) and percentage of graduate students who select these answer choices as the most common incorrect answer choices of introductory students (GS %). For convenience, short descriptions of the questions are given underneath.

<table>
<tr><th>Introductory student difficulty</th><th>TUG-K item #</th><th>% overall incorrect</th><th>Incorrect answer choices</th><th>% intro stud. diff</th><th>GS %</th></tr>
<tr><td rowspan="3">Graphs of time dependence of different kinematics variables that correspond to the same motion should look the same</td><td>11</td><td>64%</td><td>A</td><td>28%</td><td>8%</td></tr>
<tr><td>14</td><td>52%</td><td>A</td><td>25%</td><td>16%</td></tr>
<tr><td>15</td><td>71%</td><td>B</td><td>24%</td><td>8%</td></tr>
<tr><td colspan="6">11. Given a displacement-time graph, identify the velocity vs. time graph that represents the same motion.</td></tr>
<tr><td colspan="6">14. Given a velocity-time graph, identify the acceleration vs. time graph that represents the same motion.</td></tr>
<tr><td colspan="6">15. Given an acceleration-time graph, identify the velocity vs. time graph that represents the same motion.</td></tr>
</table>

As mentioned by Beichner in Ref. [1], the common difficulty of students in distinguishing between different kinematics variables is evidenced by the fact that some students claimed that the time dependence of different kinematics variables that correspond to the same motion should look the same. Table 6.3 shows that this difficulty was identified by very few graduate students on each of the three questions in which it occurs. The answer choices which uncover this difficulty (choice A for questions 11 and 14, and choice B for question 15) were selected by roughly 25% of introductory students; however, these answer choices were rarely selected by graduate students in the PCK task (see Table 6.3). The highest percentage of graduate students who selected any of these three incorrect answer choices was 16% on question 14. Beichner noted in Ref. [1] that these three questions are the ones with the highest discrimination indices (introductory physics students who answered these questions correctly performed well on the whole test), and he argued that this could be interpreted to mean that this difficulty is the one most critical to address to improve introductory students' understanding of kinematic graphs. However, our analysis suggests that graduate students are largely unaware that this difficulty exists and they are therefore unlikely to address it directly while performing their teaching duties as TAs. Many graduate students expressed astonishment in the discussions that followed the task that introductory physics students would have these difficulties.

The introductory students' difficulty that determining slopes does not require examining initial conditions was identified by very few graduate students, while other difficulties related to determining slopes were identified by more graduate students.

Table 6.4 shows that both questions 6 and 17 had incorrect answer choices selected by 46% of introductory students but identified by few graduate students. Again, discussions with the graduate students after they carried out the TUG-K related PCK task suggest that many of them were very surprised that introductory students would often not examine initial conditions when determining slopes (i.e., they computed the slope as y/x instead of $\Delta y/\Delta x$). The graduate students were more likely to think that the most common introductory student difficulty is to ignore the kinematics variables (axes) and read-off the corresponding ordinate value for a given abscissa value rather than compute the slope, i.e., slope-height confusion (incorrect answer choices E in both questions 6 and 17, selected by 36% and 44% of graduate students in this TUG-K related PCK task, but only 16% and 19% of introductory physics students as shown in Table A1). The performance of graduate students on

Table 6.4. Introductory student difficulties related to determining slopes, items on the TUG-K which uncover these difficulties (TUG-K item #), percentage of introductory students who answer the items incorrectly (% overall incorrect), incorrect answer choices which uncover these difficulties, percentage of introductory students who have these difficulties based on their selection of these answer choices (Intro stud. diff.) and percentage of graduate students who select these answer choices as the most common incorrect answer choices of introductory students (GS %). For convenience, short descriptions of the questions are given underneath.

Introductory student difficulty	TUG-K item #	% overall incorrect	Incorrect answer choices	% intro stud. diff.	GS %
Determining slopes does not require examining initial conditions	6	74%	A	46%	20%
	17	79%	B	46%	16%
Slope-height confusion in Ref. [1] (i.e., reading off the value from the vertical axis instead of computing the slope appropriately)	2	37%	C	24%	52%
	7	69%	D	28%	36%
Not taking into account the scales of the *x* and *y* axes when determining slope (i.e. slope = 2 units/1unit = 2m/s rather than 2*5m/1*10s = 1m/s) on question 7	7	69%	B	20%	28%

2. Given velocity-time graph, identify at which point/interval the acceleration is most negative.
6. Given a velocity-time graph, identify the acceleration at a particular time (must determine the slope of a straight line which does not go through the origin).
7. Given a velocity-time graph, identify the acceleration at a particular time (must estimate the slope of a straight line which does not pass through the origin).
17. Given displacement-time graph, identify the velocity at a particular time (must determine the slope of a straight line which does not go through the origin).

the other two questions related to slopes in which there were common introductory student difficulties is better; however, there is room for improvement even in those contexts. On question 2, 52% of graduate students identified the common difficulty of 37% of introductory students of confusing slope with height (see Table 6.4). On question 7, there were two common difficulties: the slope-height confusion (difficulty of 28% of introductory students, identified by 36% of graduate students as shown in Table 6.4) and not taking into account the scale of the *x* and *y* axes when determining the slope (difficulty of 20% of introductory students, identified by 28% of graduate students as shown in Table 6.4).

The performance of graduate students in identifying common introductory student difficulties related to determining areas under curves (including area-slope and area-height confusion in Ref. [1]) is context dependent.

Table 6.5. Introductory student difficulties related to determining areas under curves, items on the TUG-K which uncover these difficulties (TUG-K item #), percentage of introductory students who answer the items incorrectly (% overall incorrect), incorrect answer choices which uncover these difficulties, percentage of introductory students who have these difficulties based on their selection of these answer choices (% intro. stud. diff.) and percentage of graduate students who select these answer choices as the most common incorrect answer choices of introductory students (GS %). For convenience, short descriptions of the questions are given underneath.

Introductory student difficulty	TUG-K item #	% overall incorrect	Incorrect answer choices	% intro stud. diff.	GS %
Area-slope and/or area-height confusion	1	84%	A, D	63%	96%
	4	72%	C	23%	40%
	10	70%	C	62%	56%
	16	78%	B, C	70%	84%
	18	54%	C	32%	58%
Finding area by multiplying $y*x$ (i.e. distance traveled by an object until point (3m/s, 2s) is 6m	4	72%	E	32%	44%
1. Given 5 acceleration vs. time graphs, identify the graph in which the object has the greatest change in velocity during the time interval.					
4. Given a linearly increasing velocity vs. time graph, identify the distance covered in the first few seconds.					
10. Given 5 acceleration vs. time graphs, identify the graph in which the object has the smallest change in velocity during the time interval.					
16. Given a linearly increasing acceleration vs. time graph, identify the object's change in velocity in the first few seconds.					
18. Given a linearly increasing velocity vs. time graph, describe how you would find the distance covered in the first few seconds (read off *y* value, find the area under line segment, find the slope etc.)					

There are five questions on the TUG-K (items 1, 4, 10, 16 and 18) which require students to determine the area under a particular graph and which reveal major or moderate introductory student difficulties. Table 6.5 shows that the performance of graduate students in identifying these difficulties is context dependent. On questions 1, 4 and 16 the vast majority of graduate students identified these difficulties (96%, 84% and 84% in questions 1, 4 and 16 respectively as shown in Table 6.5), however, on questions 10 and 18, fewer graduate students identified the area-slope confusion of introductory students. This is interesting because questions 1 and 10 are posed in similar contexts: the five graphs of acceleration vs. time are almost identical; the most salient difference is that question 1 asks for the greatest change in velocity, whereas question 16 asks for the smallest change in velocity. Although on question 1, graduate students overwhelmingly selected answer choices A and D which correspond to graphs which have the highest slopes, on question 10, only 52% of them identified the most common introductory student difficulty and 28% of them selected an answer choice (D) which was selected by only 3% of introductory students (see Table A1). On question 18, 58% of graduate students identified the common area-slope confusion of 32% of introductory students (see Table 6.5). Based upon these variations, it appears that the PCK performance of graduate students in identifying the area-slope and area-height confusion of introductory students is context dependent.

Many introductory students match the verbal description of a motion with a graph superficially, without regard for the axes: this difficulty was identified by graduate students in the context of straight-line graphs, but not in the context of more complex graphs.

Table 6.6. Introductory student difficulty related to interpreting straight-line and more complex graphs, items on the TUG-K which uncover this difficulty (TUG-K item #), percentage of introductory students who answer the items incorrectly (% overall incorrect), incorrect answer choices which uncover this difficulty, percentage of introductory students who have this difficulty based on their selection of these answer choices (% intro. stud. diff.) and percentage of graduate students who select these answer choices as the most common incorrect answer choices of introductory students (GS %). For convenience, short descriptions of the questions are given underneath.

Introductory student difficulty	TUG-K item #	% overall incorrect	Incorrect answer choices	% intro. stud. diff.	GS %
Matching verbal description superficially with graph without regard for the axes in straight-line graphs	3	38%	C	20%	72%
	21	82%	B	73%	79%
Matching verbal description superficially with graph without regard for the axes in more complex graphs	8	63%	C	37%	8%
	9	76%	B	57%	28%

3. Given linearly increasing distance-time graph, select correct verbal description.
8. Given multi-part distance-time graph, select correct verbal description.
9. Given multi-part verbal description of motion (constant positive acceleration for some time, constant velocity after), select correct graph of position vs. time.
21. Given linearly decreasing velocity-time graph, select correct verbal description.

Questions 3 and 21 both ask students to interpret a straight-line graph. In question 3, the graph is of position vs. time (positive slope), and in question 21 the graph is of velocity vs. time (negative slope). On both of these questions, the most common introductory student selection essentially ignores the kinematic variable on the vertical axis and these students are matching the verbal description of a motion with a graph superficially, without regard for the vertical axis. On question 3, 20% of introductory students claimed that the graph represents an object moving with uniformly increasing velocity (which would be true if the vertical axis represented velocity instead of position) and on question 21, 73% of introductory students claimed that the graph represents an object moving with a uniformly decreasing acceleration (which would be true if the vertical axis represented acceleration instead of velocity). On both of these questions, the majority of graduate students identified this difficulty (72% and 79% in questions 3 and 21, respectively, as shown in Table 6.6). It is interesting that the performance of introductory students in interpreting graphs is vastly superior in the context of a position vs. time graph than in the context of a velocity vs. time graph (38% incorrect in question 3, compared to 82% incorrect in question 21 as shown in Table 6.6). This implies that introductory students find the concept of acceleration more difficult than the concept of velocity.

The fact that introductory students have greater difficulty in the context of acceleration than velocity is also supported by an examination of questions 12 and 19. The five graphs displayed in both of these questions are identical; however, question 12 asks them to identify the graphs that represent constant velocity and question 19 asks them to identify the graphs that represent constant acceleration. The introductory student performance on the acceleration question is much worse than the performance on the velocity question (37% compared to 63% correct). On question 19, almost 3/4 of the TAs performed well and identified the two most common incorrect answer choices (choices A and E). On question 12, there were no moderate or major introductory student difficulties.

Question 8 displays a more complex displacement vs. time graph and asks for the verbal description of this motion, and question 9 provides a verbal description of a motion and asks for the correct graph. As shown in Table 6.6, on both of these questions, the most common difficulty of introductory students is to match the verbal description of a motion with its graphical representation superficially without regard for the graph axes (identical to the difficulty in questions 3 and 21 which provide straight-line graphs). On question 8, 37% of introductory students select a description (choice C) which would be correct if the graph was of velocity vs. time rather than displacement vs. time; and on question 9, 57% of introductory students select a graph (choice B) that would be correct if it was of velocity vs. time rather than position vs. time (see Table 6.6). Few graduate students (8% and 28%, respectively) identify these answer choices as the most common incorrect choices of introductory students. Also, the PCK performance of graduate students on these two questions was the lowest among all TUG-K questions. During the whole class discussion after the task, many graduate students noted that they did not expect that introductory students would have this difficulty.

6.4 SUMMARY

In this research study, we explore one aspect of the pedagogical content knowledge of first year graduate students enrolled in a TA training course at the end of the course as it relates to knowledge of student difficulties with kinematics graphs revealed by the TUG-K. Most of the graduate students were teaching recitations or labs for introductory physics courses, and out of the three that were not, two had experience as teaching assistants or tutors for introductory physics courses and one had tutored mathematics in her undergraduate career. For each question on the TUG-K, the graduate students were asked to identify the most common incorrect answer choice selected by introductory students who did not know the correct answer after instruction in relevant concepts. The graduate students first performed this task while working individually and then while working in groups of two or three after which there was a class discussion about the task and specific introductory student difficulties.

The ability to identify introductory student difficulties on the TUG-K does not appear to be dependent on familiarity with US teaching practices.

We find that American graduate students who have been exposed to undergraduate teaching in the US and had been taught physics in English do not perform better at identifying the most common introductory student difficulties than foreign graduate students. The discussions in the TA training class related to this TUG-K related PCK task suggest that the foreign graduate students were similar to American graduate students in this regard. However, it is difficult to explain why these groups exhibit comparable PCK performance when identifying common student difficulties with kinematic graphs as revealed by the TUG-K despite their different backgrounds.

Discussions among graduate students improved their PCK performance in identifying common introductory student difficulties on the TUG-K.

The performance of graduate students in identifying introductory student difficulties with kinematics graphs as revealed by the TUG-K was significantly better when they worked in small groups compared to when they worked individually. In addition, when the individual answers of graduate students working in a group disagreed, discussions more often shifted towards the more common introductory student difficulty than the less common one. Furthermore, the class discussion with the graduate students after they performed the TUG-K related PCK tasks suggested that they found the tasks challenging but worthwhile. Many graduate students noted that they were surprised by the frequency of incorrect responses of introductory students in some of the questions and that they had not expected that introductory students would have certain difficulties with kinematics graphs. These findings suggest that performing individual and group activities about introductory student difficulties in the contexts of conceptual assessments like the TUG-K could prove to be beneficial in improving the pedagogical content knowledge related to common student difficulties of the participants and should be incorporated in professional development activities for TAs and instructors. In addition, this type of research should be carried out with other conceptual assessments to further explore the pedagogical content knowledge of instructors and/or teaching assistants related to understanding of common student difficulties in other areas.

Identifying some common introductory student difficulties related to kinematics graphs was very challenging for graduate students.

The three questions on the TUG-K with the highest discrimination indices (questions 11, 14 and 15) revealed a common introductory student difficulty that graphs of time dependence of different kinematics variables that correspond to the same motion should look the same. This difficulty was identified by very few graduate students. These questions have the highest discrimination indices according to Ref. [1] and introductory physics students who answered these questions correctly performed well on the whole test. Since these questions have the highest discrimination indexes, Beichner [1] noted that this difficulty might be the most critical to address to improve introductory students' understanding of graphs in the context of kinematics. However, we find that many graduate students are unaware that introductory students have this difficulty, and are therefore very unlikely to address this difficulty during instruction.

Another common difficulty of introductory students that determining slopes does not require examining initial conditions uncovered in question 6 and 17 was identified by few graduate students. Graduate students were more likely to think that on these questions, introductory students would read-off the corresponding ordinate value for a given abscissa value instead of trying to compute the slope, which was a difficulty much less common among introductory students.

Another common difficulty in interpreting more complex graphs than straight-line graphs of introductory students in questions 8 and 9 is to match the verbal description of the motion superficially with a graph without regard for what the axes represent. For example, on question 8, which provided a displacement vs. time graph, introductory students selected the verbal description which treated the graph as though it was of velocity vs. time. Very few graduate students were aware of this difficulty and their average PCK performance on these questions was the lowest of all questions.

For the common introductory student difficulties which were uncovered in more than one question, the ability of graduate students to identify them was context dependent.

When examining the PCK performance of graduate students in identifying introductory student difficulties in particular contexts (such as determining areas under curves, determining slopes, interpreting graphs, etc.) we find that the ability of graduate students to identify the most common difficulties is almost always context dependent. For example, difficulties of introductory students related to determining areas under curves or difficulties related to determining slopes were identified by very few graduate students on some questions, but by more graduate students on other contexts.

Graduate students, on the average, exhibited lower PCK performance when identifying major introductory student difficulties on the TUG-K than when identifying moderate ones.

There are 17 questions on the TUG-K which uncover moderate (nine questions) or major (eight questions) introductory student difficulties, and the graduate students performed better than random guessing on eight of these 17 questions. Moreover, graduate students had more difficulty in identifying major difficulties compared to moderate difficulties of introductory students. Furthermore, the analysis of the PCK score of the graduate students (as a percentage of the maximum possible score) on each question shows that on all four questions on which the average PCK score of graduate students was the lowest, there were major introductory student difficulties.

This result can be interpreted to mean that it is challenging to identify what introductory students would find difficult in a particular context. In other words, it is challenging for instructors to understand their students' perspective on what specific aspects of physics are difficult unless they have explicitly focused on these issues of student difficulties in their own classes or are familiar with physics education research which discusses student difficulties. The graduate students took introductory physics at least three or four years prior to this study and they may have lost track of what they found confusing during the learning process. It is even possible that most graduate students are not typical introductory physics students and did not have the same difficulties that many introductory students have. Therefore, activities like the one presented here, especially if they are designed to promote discussions about student difficulties, can prove valuable in preparatory courses for prospective physics instructors.

6.5 CHAPTER REFERENCES

1. R. Beichner (1994). "Testing student interpretation of kinematics graphs." Am. J. Phys. 62, 750.

2. D. Hestenes, M. Wells, and G. Swackhammer (1992). "Force Concept Inventory." Phys. Teach. 30, 141.

3. I. Halloun, R.R. Hake, E.P. Mosca, and D. Hestenes (1995). "Force Concept Inventory." (Revised, 1995); online (password protected) at http://modeling.la.asu.edu/R&E/Research.html and also printed in E. Mazur, *Peer Instruction: A User's Manual*, (Prentice-Hall, Englewood Cliffs, 1997).

4. R. Thornton and D. Sokoloff (1998). "Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation." Am. J. Phys. 66, 228.

5. P. Nieminen, A. Savinainen and J. Viiri (2010). "Force Concept Inventory-based multiple-choice test for investigating students' representational consistency." Phys. Rev. ST Phys. Educ. Res. 6, 020109.

6. D. Hestenes and M. Wells (1992). “A Mechanics Baseline Test.” Phys. Teach. 30, 159.

7. G. L. Gray, D. Evans, P. J. Cornwell, B. Self, and F. Constanzo (2005). “The Dynamics Concept Inventory Assessment Test: A progress report.” Proceedings of the 2005 American Society for Engineering Education Annual Conference, Portland, OR.

8. J. Mitchell, J. Martin, and T. Newell (2003). “Development of a Concept Inventory for Fluid Mechanics.” Proceedings, Frontiers in Education Conference, Boulder, CO, USA, T3D 23-28, DOI: 10.1109/FIE.2003.1263340.

9. M. C. Wittman (1998). “Making sense of how students come to an understanding of physics: An example from mechanical waves.” Ph.D. thesis, University of Maryland.

10. A. Tongchai, M. D. Sharma, I. D. Johnston, K. Arayathanitkul, and C. Soankwam (2009). “Developing, evaluating and demonstrating the use of a conceptual survey of mechanical waves.” Int. J. Sci. Educ. 31, 2437.

11. L. Ding, R. Chabay, B. Sherwood, and R. Beichner (2006). “Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment.” Phys. Rev. ST Phys. Educ. Res. 2, 010105.

12. E. Mazur (1997). Peer Instruction: A User’s Manual (Prentice-Hall, Engelwood Cliffs, 1997).

13. R. R. Hake (1998). “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses.” Am. J. Phys. 66, 64.

14. M. H. Dancy (2000). “Investigating animations for assessment with an animated version of the Force Concept Inventory.” Ph.D. dissertation, North Carolina State University.

15. J. Docktor and K. Heller (2008). “Gender differences in both Force Concept Inventory and introductory physics performance.” AIP Conf. Proc. 1064, 15.

16. Fred M. Goldberg and John H. Anderson (1989). “Student difficulties with graphical representation of negative velocity.” Phys. Teac. 27, 254.

17. J. R. Mokros and R. F. Tinker (1987). “The impact of microcomputer-based labs on children’s ability to interpret graphs.” J. Res. Sci. Teach., 24, 369.

18. L. C. McDermott, M. L. Rosenquist, and E. H. van Zee (1987). “Student difficulties in connecting graphs and physics: Examples from kinematics.” Am. J. Phys. 55, 503.

19. E. H. van Zee and L. C. McDermott (1987). “Investigation of student difficulties with graphical representations in physics.” Misconceptions and Educational Strategies in

Science and Mathematics. Proceedings of the International Seminar (2nd, Ithaca, NY, July 26-29, 1987), available at http://eric.ed.gov/?id=ED293686, pp. 531-539.

20. W. L. Barclay (1986). "Graphing misconceptions and possible remedies using microcomputer-based labs." Technical Report Number TERC-TR-85-5 (Cambridge, MA: Technical Education Research Center).

21. R. Thornton and D. Sokoloff (1990). "Learning motion concepts using real-time microcomputer-based laboratory tools." Phys. Teach. 30, 141, DOI: 10.1119/1.16350.

22. J. Larkin (1981). "Understanding and problem-solving in physics." Research in Science Education: New Questions, New Directions, edited by J. Robinson (Center for Educational Research and Evaluation, Louisville, CO, 1981), pp. 115-130.

23. A. B. Arons (1984). "Student patterns of thinking and reasoning, part three" Phys. Teach. 22: 88.

24. D. E. Trowbridge and L. C. McDermott (1981). "Investigation of student understanding of the concept of acceleration in one dimension", Am. J. Phys. 49, 242.

25. Z. Chen, T. Stelzer, and G. Gladding (2010). "Using multimedia modules to better prepare students for introductory physics lecture." Phys. Rev. ST Phys. Educ. Res. 6, 010108.

26. T. Stelzer, D. R. Brookes, G. Gladding, and J. P. Mestre (2010). "Impact of multimedia learning modules on an introductory course on electricity and magnetism." Am. J. Phys. 78, 755.

27. H. R. Sadaghiani (2011). "Using multimedia learning modules in a hybrid-online course in electricity and magnetism." Phys. Rev. ST Phys. Educ. Res. 6, 010102 (2011).

28. Ginsberg, H. and S. Opper (1969). Piaget's theory of intellectual development. Englewood Cliffs, NJ, Prentice Hall.

29. G. J. Posner, K. A. Strike, P. W. Hewson, and W. A. Gertzog (1982). "Accomodation of a scientific conception: Toward a theory of conceptual change." Sci. Educ. 66, 211-227.

30. L. S. Shulman (1986). "Those who understand: Knowledge growth in teaching." Educ. Res. 15(2), 4.

31. L. S. Shulman (1987). "Knowledge and teaching: Foundations of the new reform." Harv. Educ. Rev. 57, 1.

32. J. H. van Driel, N. Verloop, and W. de Vos (1998). "Developing science teachers' pedagogical content knowledge." J. Res. Sci. Teach. 35, 673.

33. P. L. Grossman (1991). “What are we talking about anyhow: Subject matter knowledge for secondary English teachers.” Advances in Research on Teaching, Vol. 2: Subject Matter Knowledge, edited by J. Brophy (JAI Press, Greenwich, CT), pp. 245–264.

34. J. Gess-Newsome and N. G. Lederman (2001) Examining Pedagogical Content Knowledge, (Kluwer Academic Publishers, Boston).

35. J. Loughran, P. Mulhall, and A. Berry (2004). “In search of Pedagogical Content Knowledge in science: Developing ways of articulating and documenting professional practice.” J. Res. Sci. Teach. 41, 370.

36. E. Etkina (2010). “Pedagogical content knowledge and preparation of high school physics teachers.” Phys. Rev. ST. Phys. Educ. Res. 6, 020110.

37. Available at: http://www.aip.org/statistics/trends/highlite/edphysgrad/table1b.htm.

38. Available at: http://www.aip.org/statistics/trends/highlite/edphysgrad/figure7.htm.

39. C. Singh (2002). “Effectiveness of group interaction on conceptual standardized test performance.” Proceedings of the Phys. Ed. Res. Conference, Boise, ID, 2002, Edited by S. Franklin, J. Marx, and K. Cummings (AIP, Melville, NY).

40. G. V. Glass and K. D. Hopkins (1996). Statistical Methods in Education & Psychology, (Allyn & Bacon, Boston, MA).

7.0 EXPLORING ONE ASPECT OF PEDAGOGICAL CONTENT KNOWLEDGE OF PHYSICS INSTRUCTORS AND TEACHING ASSISTANTS USING THE FORCE CONCEPT INVENTORY

7.1 INTRODUCTION

7.1.1 Background on previous research involving the Force Concept Inventory

The Force Concept Inventory (FCI) is a multiple choice survey developed in 1992 by Hestenes, Wells and Swackhammer [1] and later revised [2] after many early observations made by Halloun and Hestenes [3] and other physics education researchers [4-7] that many students enter and leave physics classes with conceptions that are not consistent with the scientifically accepted concepts taught in the physics classes. The FCI was designed to assess student understanding of the fundamental mechanics concepts related to force and motion and has been widely used for this purpose by many educators and physics education researchers. Similar assessments in mechanics have been designed for the same purpose by other physics education researchers [8-12]. Although the conclusion that the FCI consistently measures Newtonian thinking was subject to some debate [13-16], the general consensus is that the FCI score is a good indicator of Newtonian thinking [17-20]. Some researchers have investigated the validity of the items on the FCI using Item Response Analysis [17-19]. Morris et al. [18] have argued that Item Response

Analysis can be used to identify answer choices which do not discriminate between students in different ability groups (ability was mainly defined by them using FCI scores). Their analysis was also used to investigate student performance in more detail and gain some insights into student difficulties for some of the items on the FCI. Other researchers studied the FCI using the Rasch model [20] and concluded that the FCI "has succeeded in defining a sufficiently uni-dimensional construct for each population" (non-Newtonian and predominantly Newtonian). The analysis by Planinic et al. suggested that "the items in the test all work together and there are no grossly misfitting items which would degrade measurement" [20].

The FCI has played a key role in convincing many educators that traditional teaching methods which are primarily lecture oriented and do not actively engage students in the learning process do not promote conceptual and functional understanding [21-23]. Several studies have demonstrated that many students enter and leave introductory physics courses with the same alternate conceptions that are inconsistent with the accepted scientific ways of reasoning. Indeed, the use of the FCI in traditionally taught classes (even those taught by popular instructors) gave an impetus to the field of physics education research (PER) as educators increasingly realized that traditional methods were not working as intended, and consequently began to develop and evaluate instructional strategies designed to promote functional understanding of physical phenomena [21-24]. The FCI has often been used to assess whether a particular instructional strategy is effective in promoting conceptual and functional understanding. Hake [21] used the FCI for this purpose and found that courses that make use of research based instructional approaches such as collaborative peer instruction [25-27], modeling [28-30], concept tests [24], microcomputer-based labs [31-33], active-learning problem sheets (ALPS) [34,35] and others [36,37] result in higher normalized gains on the FCI than courses which employ traditional methods such as standard lectures. The average normalized gain is defined as the ratio of the change in the average post-test score (after instruction of Newtonian concepts) with respect to the average pre-test score (before instruction of Newtonian concepts) to the average maximum possible change from the average pre-test score, i.e., average normalized gain <g> = (<post percent>–<pre percent>)/(100%–<pre percent>). Hake's study included more than six thousand students from both college and high-school classes.

The FCI has also been used to explore gender differences in understanding of Newtonian concepts related to force and motion [38-41]. Typically, in a particular introductory course, males outperform females on the FCI. However, in other course assessment measures such as the final exam, the males and females exhibit comparable performance. The gender gap observed on the FCI can be effectively reduced [38,40], although not necessarily removed [41], through PER based teaching strategies including but not limited to peer instruction, cooperative problem solving or using tutorials such as *Tutorials in Introductory Physics* by the University of Washington group. Other researchers have argued that the worse performance of females on the FCI can be partly attributed to the context of the questions which is mostly masculine and/or abstract [42]. Previous research indicated that females are more successful when questions are phrased using real-life contexts [43]. Therefore McCullough developed a "gender" version of the FCI [44] in which the items were rephrased from formal or male-oriented contexts to daily-life and female-oriented contexts. McCullough showed [42,44] that there was significant context dependence in the performance of both males and females for some questions. However, on individual questions different trends were observed (e.g., male performance improved and female performance declined on some revised questions, performance of both stayed the same on some, and females improved and males declined on other revised questions) Overall, there was a decline in the overall performance of males, but the performance of females stayed about the same. Other researchers [45] have used differential item functioning to investigate whether some questions favor one gender over another (i.e., if a male student is statistically more likely than a female student of the same ability to answer a question correctly) and concluded that five questions may have a gender bias. Context dependent performance on FCI questions was also investigated by Dancy [46], who developed an animated version of the FCI and found differences in performance on seven questions.

Other researchers have used the FCI to investigate correlations between FCI scores and various other indicators of student performance: normalized gain on the FCI [47], problem solving ability [48], scientific reasoning ability [47,49], mathematics preparation [50], SAT scores [51], representational consistency [52], etc. In almost all these instances, significant positive correlations were found.

The FCI has often been administered by physics education researchers and curriculum developers as a pre-test to determine what initial knowledge students bring to the learning of physics. Knowing the initial knowledge state of students is important because instructional tools and pedagogies can be designed to take advantage of the knowledge resources students have and to effectively address the alternate conceptions which are not consistent with the accepted scientific way of reasoning about physical phenomena. In addition, the FCI has been administered as a post-test, e.g., to determine what concepts are difficult for students even after instruction and how effective instruction was at addressing student difficulties.

7.1.2 Focus of this study: Pedagogical Content Knowledge related to student difficulties revealed by the FCI

The study presented here used the FCI to explore one aspect of the pedagogical content knowledge of instructors and graduate teaching assistants (TAs), namely, knowledge of student difficulties related to mechanics concepts as revealed by the FCI. The instructors who participated in the study had varying degrees of experience teaching introductory physics courses. For each item on the FCI, the instructors and TAs at the University of Pittsburgh (Pitt) were asked to identify the most common incorrect answer choice of introductory physics students. We also discussed the responses individually with a few instructors and had a discussion with the TAs, who at the time of the study were enrolled in a TA training class at Pitt.

Pedagogical content knowledge (PCK) was defined by Shulman [53,54] as the subject matter knowledge *for teaching* and many researchers in K-16 education have adapted this construct [55-58]. According to Shulman, PCK is a form of practical knowledge used by experts to guide their pedagogical practices in highly contextualized settings. In addition to the knowledge of the most useful forms of representation of the topics, use of powerful analogies, illustrations and examples, etc., Shulman included in pedagogical content knowledge, "understanding of the conceptions and preconceptions that students bring with them to the learning of those most frequently taught topics and lessons" [54]. Our research presented here explores the aspect of PCK of the physics instructors and graduate TAs related to their knowledge of introductory physics students' alternate conceptions related to force and motion as revealed by the FCI. We refer to this as the "FCI related PCK" of instructors and TAs. In particular, we investigate whether instructors and graduate students are able to identify the common alternate conceptions of students on individual items in the FCI. Knowledge of the alternate conceptions which are inconsistent with the scientifically accepted way of reasoning about the concepts can be helpful in devising the curricula and pedagogical strategies to improve student understanding. Much physics education research has been devoted to devising and assessing such strategies.

We note that, in order conduct the research related to the FCI related PCK of instructors and TAs, we needed introductory student data for each answer choice on individual items on the latest version of the FCI from large populations of students. It is most appropriate to analyze instructor and graduate student PCK data at Pitt by comparing it to introductory physics students' FCI data at the same institution, which is a large, typical state related university of about 18,000 undergraduate students. Therefore, data were collected over a few years both in pre-tests (before instruction) and post-tests (after instruction) from about 900 algebra-based students and over 300 calculus-based students. The courses were all taught using traditional instructional methods at Pitt. These data were used to determine the common student alternate conceptions related to each item on the FCI and thus to assess the FCI related PCK of physics instructors and TAs.

7.2 RESEARCH QUESTIONS

7.2.1 Primary research questions – FCI related PCK of instructors and TAs

P.1. To what extent does teaching experience influence (if at all) the ability to identify introductory students' alternate conceptions?

P.2. To what extent are American physics graduate students, who have been exposed to undergraduate teaching in the United States, better at identifying introductory students' alternate conceptions than foreign physics graduate students?

P.3. To what extent do instructors and/or graduate students identify 'strong' student alternate conceptions compared to 'medium' level ones?

P.4. To what extent do graduate students identify introductory students' difficulties more often when working in groups than when working individually (i.e., do discussions improve graduate students understanding of introductory students' alternate conceptions related to force and motion as revealed by the FCI)?

P.5. To what extent do instructors/graduate students identify specific alternate conceptions of introductory physics students? Is their ability to identify these alternate conceptions context dependent?

7.2.2 Secondary research questions – Introductory student FCI performance

In order to answer the primary research questions, we needed data on the performance of students on individual items on the FCI (the revised version of the test from 1995). Therefore, we collected FCI data from about 900 students in algebra-based and more than 300 students in calculus-based introductory physics courses at Pitt. Subsequently, the following secondary research questions emerged, which were related to analysis of introductory student performance on individual questions on the FCI and comparison of the pre-test and post-test data for both algebra-based students and calculus-based students.

S.1. Which questions on the FCI pose significant challenges for students?

S.2. Are there any questions on the FCI for which there is little improvement (small normalized gain) from pre-test to post-test?

S.3. Are there any shifts in the most common alternate conceptions from the pre-test to the post-test?

S.4. On which questions do calculus-based students perform better than algebra-based students by 20% or more? Are there any questions in which the most common alternate conceptions of algebra-based students are different from the most common alternate conceptions of calculus-based students?

7.3 METHODOLOGY

7.3.1 Materials and Participants

The materials used in this study are the FCI and the pre-post introductory student data collected from 900 algebra based and more than 300 calculus based introductory physics courses at Pitt (see Tables B3 and B4 included in Appendix B). We compare the algebra-based and calculus-based classes in the results section. All classes from which these data were collected were taught in a traditional manner and the average unmatched (all students who took the pre-test and post-test were included regardless of whether they took both the pre-test and post-test) normalized gain was 0.26 for the algebra-based classes and 0.36 for calculus-based classes (almost identical to the matched normalized gains). These gains are close to gains for courses that do not employ PER based instructional strategies as reported by Hake [21].

The participants of this study were thirty physics instructors and twenty five first year graduate students. The instructors varied widely in terms of introductory physics teaching experience. In particular, some instructors were relatively new and had only taught introductory courses a few times, while others were emeritus professors who had not taught for many years (but had taught a long time ago) and yet others were instructors who taught introductory physics courses on a regular basis.

The graduate students were enrolled in a semester long pedagogy oriented TA training class. This course is required of all first year graduate students. The course meets once a week for two hours and is designed to help graduate students be more effective teachers. During the course, students learn about cognitive research and physics education research (PER) and discuss their instructional implications. Students are also introduced to curricula and pedagogies based on physics education research which stress the importance of being knowledgeable about introductory students' difficulties in order to help them transition toward greater expertise in physics. Each graduate student also discusses the solution of a physics problem in the class in the manner in which he/she would discuss it if he/she were teaching introductory students and they receive feedback from the other graduate students and the instructor. Also, during the course, students complete various reflective exercises aimed at helping them perform their TA duties in a student-centered manner.

All but three of the graduate students who participated in this study were teaching introductory physics recitations or labs for the first time. Two of the three who were not teaching had physics teaching experience as undergraduates, either as a teaching assistant or as a tutor for introductory physics courses. Only one student did not have teaching experience in physics, but this student tutored mathematics as an undergraduate. Also, the TA training course included discussions of introductory student difficulties, however, not in the specific context of the FCI (until after students completed all tasks related to the FCI as described below).

7.3.2 Methods

The physics instructors were given the FCI survey and for each question, they were asked to identify which one of the four *incorrect* answer choices, in their view, would be most commonly selected by introductory physics students after instruction in relevant concepts if the students did not know the correct answers (we refer to this as the "FCI related PCK task"). The instructors were asked to complete the task at their convenience. Also, the task was originally given to 33 physics instructors at Pitt but three of them did not complete the task in a reasonable amount of time even after multiple reminders. After the instructors had completed the task, we discussed the reasoning for their responses individually with some of them, especially for the questions in which the reasoning was not explicitly provided (and subsequently with the graduate students in a class discussion).

Towards the end of the TA training class (so that a majority of graduate students had almost a semester worth of teaching experience), the graduate students were asked to complete three different tasks related to the FCI: (1) while working individually, they were asked to identify the correct answers for each question; (2) while working individually, they were asked to complete the FCI related PCK task; and (3) they repeated the FCI related PCK task, except working in groups of two or three. The graduate students performed task (1) first, then task (2) and finally task (3), followed by a class discussion during a two hour TA training class. The graduate students were allowed as much time as they needed for each task. All graduate students finished the first task within the first 30 minutes and the second task within the first hour. The third task (group work) was completed by all groups within 40 minutes followed by a full class discussion about the FCI related PCK task and why knowledge of student difficulties is critical for teaching and learning to be effective in general. The graduate student population at Pitt is consistent with that of a typical research focused state university and the nationality of the graduate students varied: nine graduate students were from the United States, nine were from China and the other seven were from other countries (Asian and European).

We note that the task given to the instructors and graduate students was framed such that they had to identify the most common incorrect option for each multiple choice question that introductory physics students would select after instruction if they did not know the correct answer (rather than before instruction), because individual discussions with some faculty members who had taught introductory physics before giving them the task indicated that they felt that they had no way of knowing the “pre-conceptions” of introductory physics students. Their reluctance to contemplate introductory physics students’ preconceptions about force and motion before instruction motivated us to ask them to identify the most common incorrect answer choice for each question if the student did not know the correct answer after instruction in relevant concepts. Although asking them to identify the most common alternate conception in a post-test made the task easier to complete, some faculty members who participated in the study were concerned about their ability to identify students' difficulties and explicitly noted that they have no way of knowing the most common difficulty of introductory students for each question.

We also note that it does not make a significant difference whether the question is phrased to the instructors and graduate students about introductory physics students' difficulties with each question in the post-test or pre-test because the common alternate conceptions of introductory students rarely changed after traditional instruction. Instead, typically, fewer students held the same common alternate conceptions (this was found to be true when we compared the pre-test and post-test data of introductory students). Therefore, the performance of experts (instructors and graduate TAs) at identifying these alternate conceptions provides an indication of their knowledge of the initial knowledge state of introductory students.

In order to compare the FCI related PCK performance of the physics instructors with that of the graduate students (and also to compare the FCI related PCK performance of different subgroups of instructors/graduate students), scores were assigned to each instructor/graduate student. An instructor/graduate student who selected a particular incorrect answer choice as the most common incorrect choice in a particular question received a PCK score which was equal to the fraction of introductory students who selected that particular incorrect answer choice. If an instructor/graduate student selected the correct answer choice as the most common incorrect answer (a rare occurrence), he/she was assigned a score of zero because he/she was explicitly asked to indicate the *incorrect* answer choice which is most commonly selected by introductory students if they did not know the correct answer. For example, in question 2, the fractions of algebra-based students who selected A, B, C, D and E are 0.44, 0.25, 0.06, 0.21 and 0.04, respectively (see Table B1 included in Appendix B). Answer choice A is correct, thus, the score assigned to instructors or graduate students for each answer choice if they selected it as the most common incorrect answer would be 0, 0.25, 0.06, 0.21 and 0.04 (A, B, C, D and E). The total score an instructor/graduate student would obtain on the task for the entire FCI can be obtained by summing over all of the questions. A mathematical description of how this calculation was performed is included in Appendix B.

In order to determine whether the instructors/graduate students performed better than random guessing on the FCI related PCK task, a population of random guessers was generated. The population was generated by choosing N = 24 'random guessers' in order to have a reasonable group size when performing *t*-tests [59]. Random guessing on this task would correspond to choosing one of the four incorrect answer choices for each question with equal probability (25%). Therefore, one quarter of the random guessers always selected the first incorrect answer choice, one quarter selected the second incorrect answer choice, etc. Since the instructors and graduate students were not provided with the correct answers before they performed the FCI related PCK task, random guessing would not perfectly correspond to selecting one of the four incorrect answer choices with equal probability. For a particular question, there is a small probability that an instructor/graduate student does not know the correct answer. However, our data indicate that this probability is very small (see table B1 in Appendix B). Moreover, since for a given question, one quarter of the random guessers selected each of the four incorrect answer choices, one can calculate a mean and a standard deviation for their scores which can be used to perform comparison with the graduate student scores. Furthermore, our choice of random guessers maximizes the standard deviation.

We note that our approach used to determine the PCK score related to FCI appropriately weighs the responses of instructors/graduate students by the fraction of introductory students who selected a particular incorrect response. This weighing scheme was chosen because the more prevalent an introductory student difficulty is, the more important it is for an instructor/graduate student to be aware of it and take it into account in his/her instruction.

7.3.3 Approach for answering the primary research questions

The researchers analyzed whether instructors and/or graduate students performed better at identifying introductory students' alternate conceptions than random guessers by performing statistical analysis. The analysis of the FCI related PCK performance was carried out with both the algebra based and calculus based student data yielding nearly identical results. We present the analysis with the algebra based student data.

P.1. To what extent does teaching experience influence (if at all) the ability to identify introductory students' alternate conceptions?

In order to answer this question, we compared the average FCI related PCK score of all instructors with the average FCI related PCK score of all graduate students and also compared the FCI related PCK scores of instructors who had recently taught introductory mechanics (either algebra-based or calculus-based) with those who had not taught introductory mechanics recently.

The PCK scores of instructors (all of whom had taught some introductory physics course in the near or distant past and several had taught them many times) were compared with the PCK scores of graduate students enrolled in the TA training course (at the end of the course) who had never taught an introductory physics course as lecturers before. All of the graduate students were at the time in their first semester in physics graduate school and most were doing a teaching assistantship for the first time. Since the teaching experience as lecturer of the graduate students was very limited compared to the teaching experience of most of the instructors who had taught some introductory courses, this comparison may provide some indication for whether teaching experience as lecturer influences the ability to identify student alternate conceptions. We note however, that the first year physics graduate students in the TA training course had taken introductory physics only a few years prior to the study as undergraduates and a majority of them were TAs for introductory recitations and laboratories, graded homework, quizzes and exams and held regular office hours in addition to spending time weekly in the resource room at Pitt to help introductory students throughout the semester. These experiences may help the graduate students understand the difficulties of introductory students and therefore increase their ability to identify introductory students' alternate conceptions. As a result, it is difficult a priori to predict how they will perform compared to the instructors (most of whom did minimal grading and had minimal direct contact with students in the large introductory classes) regardless of the fact that instructors had significantly more independent classroom teaching experience.

Therefore, we also compared the FCI related PCK scores of instructors who had taught introductory mechanics recently with those who had not taught it or had not taught it in the last seven years. Half of the instructors who participated in this study had taught introductory algebra-based or calculus-based mechanics courses at least a few times in the past seven years, while the other half had not taught these courses or taught them more than seven years prior to the study. This analysis was designed to investigate if recent teaching experience in introductory algebra-based or calculus-based mechanics courses played a role in the instructors' ability to identify introductory students' alternate conceptions about force and motion.

P.2. To what extent are American physics graduate students, who have been exposed to undergraduate teaching in the United States, better at identifying introductory students' alternate conceptions than foreign physics graduate students?

Out of the twenty-five first year graduate students who participated in this study, nine were American, nine were Chinese and seven were from other foreign countries (Asia and Europe). The FCI related PCK scores of three groups of graduate students were compared (American, Chinese and other foreign students). The reason we divided the graduate students into three groups is because the American graduate students were exposed to teaching in the United States as opposed to the foreign students, who were not exposed to US teaching practices before graduate school and most were taught physics in their own native languages. The nine Chinese graduate students were placed in a separate group because, although they fit the category of foreign graduate students, it is possible that their backgrounds are different from the backgrounds of most of the other foreign graduate students, and it is unclear whether these differences in backgrounds translate to differences in performance on the FCI related PCK task.

P.3. To what extent do instructors and/or graduate students identify 'strong' student alternate conceptions compared to 'medium' level ones?

Many of the questions on the FCI contain strong distractor choices that are selected by large numbers of introductory students even in a post-test. The researchers determined that an incorrect answer choice can be attributed to a 'strong' student alternate conception if more than 1/3 of introductory algebra-based students selected that answer choice. An incorrect answer choice was considered connected to a 'medium' level alternate conception if between 19% and 34% of the students selected that answer choice (initially, the lower cutoff was chosen to be

20%, but there were three questions on the FCI in which 19% of introductory students selected an incorrect answer choice, and the researchers considered that two of them were worth discussing, thus 19% was selected to be the lower cutoff).

In order to answer whether physics instructors or graduate students are better at identifying strong alternate conceptions than medium level ones, we compared how often instructors or graduate students performed better than random guessing on questions which contained strong alternate conceptions with how often they performed better than random guessing on questions which contained medium level alternate conceptions.

P.4. To what extent do graduate students identify introductory students' difficulties more often when working in groups than when working individually (i.e., do discussions improve graduate students understanding of introductory students' alternate conceptions related to force and motion as revealed by the FCI)?

Previous studies have found that student discussions improve performance on conceptual examinations [24, 60]. Mazur's Peer Instruction [24] approach has been developed because student discussions tend to converge to the correct answers rather than the incorrect answers. In particular, research suggests that if two students individually select different answers and one of them is correct, the student with the correct answer is more likely to convince the student with the incorrect answer through a discussion than otherwise. In addition, in the context of introductory calculus-based electricity and magnetism, Singh found that even if both students initially select an incorrect answer choice, in 29% of the cases, discussions among introductory students lead them to the correct answer [60]. We investigated whether discussions among graduate students that are centered on introductory students' alternate conceptions helped them identify the more common alternate conceptions.

The graduate students completed three tasks related to the FCI in a two hour long TA training class toward the end of the semester: first they were asked to provide the correct answers to the FCI, second, they individually performed the FCI related PCK task (identified the incorrect answers most commonly selected by introductory students), and third, they repeated the FCI related PCK task in groups of two or three. It was investigated whether discussions among graduate students improved their knowledge of introductory student alternate conceptions. Two factors would indicate that discussions improve graduate students' understanding of introductory students' alternate conceptions:

- 1) better FCI related PCK performance and
- 2) convergence to a more common introductory student alternate conception.

The second factor warrants further explanation: if in the individual PCK task, two graduate students selected two different incorrect answer choices (that they thought would be most common among introductory students who did not know the correct answer), and at least one of the incorrect answer choices is connected to a common student alternate conception, we investigated how often the two graduate students agreed on the incorrect answer choice which is selected by more introductory students. In order to answer this question, we identified all the instances in which two (or three) graduate students who selected different incorrect choices in the individual PCK task, while working in a group, agreed on one of the incorrect answers. Then, we determined how often the incorrect answer selected in the group PCK task was more common (by 5% or more) among introductory students than the other answers selected by the graduate students in the individual PCK task.

P.5. To what extent do instructors/graduate students identify specific alternate conceptions of introductory physics students? Is their ability to identify these alternate conceptions context dependent?

These questions were answered by identifying particular alternate conceptions (e.g., constant force implies constant velocity) in different questions and analyzing instructor/graduate student PCK performance in identifying these common alternate conceptions in different questions.

7.3.4 Approach for answering the secondary research questions

S.1. Which questions on the FCI pose significant challenges for students?

This question was answered while analyzing the PCK performance of instructors and graduate students at identifying students' alternate conceptions because this analysis was restricted to the alternate conceptions held by at least 19% of introductory students. For each alternate conception, the question in which it appears and the percentage of introductory students who hold the particular alternate conception was identified.

S.2. Are there any questions on the FCI in which there is very little improvement from pre- to post-test?

Introductory student performance in a post-test is not the sole indicator of how difficult a question is. If the percentage of introductory students who answer a question correctly does not improve significantly after instruction in the relevant concepts, it is an indicator of the difficulty of the question regardless of the percentage correct in the post-test. The determination of questions with "little" improvement from pre-test to post-test was done based on two criteria: the average normalized gain (see Table B3) and improvement in the percentage of students who harbor an alternate conception. For normalized gain, the questions were ordered from lowest to highest and the researchers determined that "little" improvement occurred in the bottom 1/3 of the questions. For the second criteria, it was considered that "little" improvement occurred in questions in which the improvement in the percentage of students who hold the most common alternate conception is less than 5%.

S.3. Are there any shifts in alternate conceptions from the pre-test to the post-test?

Previous research has found that many students enter introductory mechanics classes with naïve interpretations of real world phenomena that are inconsistent with physics principles [1-10]. One may expect that after instruction, the performance of introductory students on individual items would improve and the incorrect answers which were selected most commonly in the post-test would largely remain the same as the ones that would be selected most commonly in the pre-test, except by smaller percentages of students in the post-test. However, students might shift from one incorrect answer choice in the pre-test to another incorrect answer choice in the post-test. For example, in question 5 (identify all the forces that act on a ball while it is moving in a frictionless, circular channel), before instruction, many students do not know that the channel exerts a force on the ball, but know about the force of gravity and hold the alternate conception that in order for an object to be moving in a certain direction, a distinct force must be acting on it in the direction of motion. It is possible that after instruction, most of these students learn that the channel must exert a force on the ball, but do not abandon the idea that a distinct force must exist in the direction of motion. The post-test response based upon these notions would still be incorrect; however, the alternate conception will now be different than on the pre-test.

In order to determine whether algebra-based introductory students hold different alternate conceptions after instruction compared to before instruction we analyzed questions which had two or more common alternate conceptions either in the pre-test or the post-test. In these questions, it was considered that a shift occurred if either the following changes transpired from the pre-test to the post-test:

- 1) the percentage of introductory students who selected one of these incorrect answer choices decreased (by 10% or more) while the percentage of introductory students who selected the other incorrect answer choice(s) remained the same or increased or
- 2) the percentage of introductory students who selected one of the incorrect answer choices remained the same, while the percentage of introductory students who selected the other incorrect answer choice(s) increased.

S.4. On which questions do calculus-based students perform better than algebra-based students? Are there any questions in which the alternate conceptions of algebra-based students are different from the alternate conceptions of calculus-based students?

Previous research has found that students in calculus-based classes perform better than students in algebra-based classes on the FCI [21,61] and other conceptual assessments [8,58]. However, it is possible that on some FCI questions the differences are less pronounced than on others. We investigated on which questions on the FCI the calculus based students performed better than the algebra based students and on which questions the differences were small. In addition, we investigated whether there were any questions for which the most common alternate conceptions of algebra-based students were different from the common alternate conceptions of calculus-based students.

7.4 RESULTS

Many instructors and graduate students noted that the task of thinking from a student's point of view was challenging; some even confessed that they did not feel confident about their performance in identifying the most common incorrect answers. Also, the task was posed as the identification of the most common incorrect answer of introductory physics students for each FCI question *after* instruction if students did not know the correct answer. Thus, the primary data analysis in this section involves comparison of the instructors' and graduate students' responses with introductory physics responses on each FCI question *after* instruction. However, our analysis revealed that the introductory students' alternate conceptions are generally the same, except more pronounced before instruction compared to after instruction.

This section is broken up into two subsections. In the first, we discuss the primary research questions which focused on investigating one aspect of the pedagogical content knowledge of instructors and graduate students, namely, knowledge of common student difficulties related to force and motion as revealed by the FCI. In the second, we discuss the secondary research questions which focused on the performance of introductory students on the FCI.

7.4.1 Results: Primary research questions

There are 24 questions on the FCI which reveal strong and/or medium alternate conceptions: items 2, 4, 5, 9 and 11-30. Analysis of the FCI related PCK score of both instructors and graduate students was conducted on each of these questions and the results are displayed in Tables B1 and B2 (included in Appendix B).

Table B1 shows the percentages of introductory physics students who selected each answer choice when asked to select the correct choice for each question, and instructors and graduate students who selected each answer choice for what would be the most common incorrect choice of introductory physics students if they did not know the correct answer on each of the 24 questions in which strong or medium level alternate conceptions were identified. Correct answers are indicated by the green shading in Table B1, strong student alternate conceptions (incorrect answer choices selected by more than 1/3 of the introductory students) are indicated by red shading and medium alternate conceptions are written in red. In addition, the second column (titled >RG) in Table B1 indicates whether instructors and/or graduate students performed better than random guessing. For each question, "I" in the second column of Table B1 indicates that instructors performed better than random guessing, "GS" indicates that graduate students performed better than random guessing and "I, GS" indicates that both instructors and graduate students performed better than random guessing in identifying introductory physics students' most common incorrect answer for a particular question. If the field in the second column (titled >RG) of Table B1 is blank then neither instructors nor graduate students performed better than random guessing.

Table B2 shows, for each question, the normalized average FCI related PCK scores of the instructors and graduate students. Their FCI related PCK scores were normalized on a scale from zero to 100 because for each question on the FCI there is a minimum and a maximum possible score, which correspond to the smallest and largest fractions of introductory students who selected a particular incorrect answer choice among the four incorrect answer choices. The normalization was done in the following manner: normalized FCI related PCK score = 100 * (average FCI related PCK score – minimum possible score) / (maximum possible score –

minimum possible score). The normalized FCI related PCK score is then zero if the instructors/graduate students obtained the minimum possible score and 100 if they obtained the maximum possible score. This also provides a way to compare the FCI related PCK performance in different questions which have different minimum and maximum possible FCI related PCK scores. Table B2 also shows the difficulty of each of these questions via the percentage of introductory algebra-based students who answered each question correctly in a post-test, normalized gain and strength of the alternate conception(s), i.e., medium level or strong. The questions which contained a strong alternate conception are indicated by the red shading and those which contained a medium level alternate conception are written in red. Also, the performance of instructors and graduate students is considered 'good' (and shaded green) if their normalized FCI related PCK score is more than 2/3 of the maximum possible score, 'medium' level (and shaded yellow) if their normalized score is between 1/2 and 2/3 of the maximum possible score and 'poor' (horizontal stripes) if their performance is less than 1/2 of the maximum possible score. Examination of Table B2 indicates that the strength of an alternate conception is not correlated with the FCI related PCK performance of instructors and/or graduate students. Table B2 shows that there are questions with strong alternate conceptions in which both instructors' and graduate students' FCI related PCK performance is poor, other questions with strong alternate conceptions in which their FCI related PCK performance is medium level and yet others in which their performance is good. A similar observation can be made for questions in which there is a medium alternate conception. These results are discussed in more detail below, where we provide the results which helped answer research question P.3.

P.1. To what extent does teaching experience influence (if at all) the ability to identify introductory students' alternate conceptions?

In order to answer this question, one analysis involved comparison of the overall FCI related PCK scores of instructors, who on the average had significant experience teaching introductory physics courses as lecturers, with the FCI related PCK scores of graduate students enrolled in the TA training course, who had limited or no experience teaching introductory physics courses as lecturers. The maximum possible FCI related PCK score of instructors or graduate students on each question would be equal to the maximum fraction of introductory students who selected an incorrect answer choice. The maximum possible FCI related PCK score on the whole survey is 9.21, which is the sum of these fractions for all the questions. Table 7.1 shows that the average of instructors (68% of the maximum possible FCI related PCK score) and the average of graduate students (65% of the maximum possible FCI related PCK score) are very close. Also, *t*-tests revealed no significant difference between instructors and graduate students in terms of their FCI related PCK scores. Although, their overall PCK performance is the same, there were many differences observed in the performance of identifying specific student alternate conceptions. However, both the instructors and graduate students performed significantly better on the FCI related PCK task than random guessing (both p values when comparing instructors to random guessing and graduate students to random guessing were less than 0.001). We note that since the graduate students had taken introductory physics only four years prior to this study as undergraduates and the vast majority of them were TAs in an introductory recitation or laboratory class, did weekly grading of quizzes, homework and exams and held office hours in which they helped introductory students individually, they may identify with introductory physics students' difficulties related to FCI concepts.

Table 7.1. Numbers of instructors/graduate students/random guessers, averages and standard deviations (Std. dev.) for the FCI related PCK scores obtained (in determining student alternate conceptions on the FCI) out of a maximum of 9.21.

	N	Average	Std. dev.
Instructors	30	6.25	0.90
Graduate students	25	6.01	0.78
Random guessing	24	3.71	0.93

We also investigated whether recent teaching experience in algebra-based or calculus-based introductory mechanics course was related to the ability of instructors to identify students' alternate conceptions that emerge in the FCI. The average of the instructors who had taught introductory mechanics courses in the past seven years was nearly identical to the average of instructors who had not taught those introductory physics courses recently (see Table 7.2). It appears that recent teaching experience of these instructors in introductory mechanics is not related to their ability to identify introductory students' alternate conceptions.

Table 7.2. Numbers of instructors who had taught and who had not taught introductory mechanics in the past seven years, their averages and standard deviations (Std. dev.) for the scores obtained for determining students' alternate conceptions on the FCI out of a maximum of 9.21.

	N	Average	Std. dev.
Have taught in the past 7 years	15	6.33	0.77
Have not taught in the past 7 years	15	6.17	1.03

P.2. To what extent are American physics graduate students, who have been exposed to undergraduate teaching in the United States, better at identifying introductory students' alternate conceptions than foreign physics graduate students?

Our analysis suggests that it was not the case that American graduate students performed better than the others. In particular, the averages of these three groups of graduate students (American, Chinese, other foreign) were very similar as shown in Table 7.3. Statistical analyses using *t*-tests are not appropriate here because the group sizes are small, but it does appear that the averages are not very different. The Chinese students were placed in a separate group because they comprised more than half of the foreign graduate students and we did not want the performance of foreign graduate students to be skewed because of this.

Table 7.3. Numbers of American/Chinese/Other foreign graduate students, their averages and standard deviations (Std. dev.) for the scores obtained in determining student alternate conceptions on the FCI out of a maximum of 9.21.

	N	Average	Std. dev.
American	9	6.20	0.70
Chinese	9	6.04	0.76
Other foreign	7	5.71	0.91

P.3. To what extent do instructors and/or graduate students identify 'strong' student alternate conceptions compared to 'medium' level ones?

There are 11 questions on the FCI in which at least 1/3 of the introductory physics students selected a particular incorrect answer choice (see Table A1). The instructors' PCK score was better than random guessing on eight of these questions (73%) while the graduate students' PCK score was better than random guessing on five (45%) of these questions. In the other 13 questions (which contained 'medium level' alternate conceptions, i.e. conceptions held by 19%-33% of introductory students in a post-test), both instructors' and graduate students' PCK scores were better than random guessing on seven of them (54%). These numbers are too small to perform meaningful statistics, but it appears that instructors identified 'strong' misconceptions somewhat better than 'medium level' ones (73% compared to 54%), whereas for graduate students, the difference is minor (45% as compared to 54%).

P.4. To what extent do graduate students identify introductory students' difficulties more often when working in groups than when working individually (i.e., do discussions improve graduate students understanding of introductory students' alternate conceptions related to force and motion as revealed by the FCI)?

1. Graduate student FCI related PCK performance is better when they work in groups compared to when they work individually.

Table 7.4 shows the graduate students' FCI related PCK performance when they worked individually and in groups of two or three. A *t*-test shows that the group performance is better than the individual performance ($p = 0.040$).

Table 7.4. FCI related PCK performance of graduate students in the individual and in the group PCK tasks: number of graduate students/groups (N), averages (Avg.) and standard deviations (Std. dev.)

Graduate students' FCI related PCK performance			
Individual	**N**	**Avg.**	**Std. dev.**
	25	6.01	0.78
Group	**N**	**Avg.**	**Std. dev.**
	12	6.59	0.79

2. Discussions among graduate students often tend to lead them to agree on a more common introductory student alternate conception

There were 98 instances in which two or three graduate students who did not all select the same incorrect answer choice in the individual PCK task, when working in groups, converged to one of their original answers pertaining to introductory students' common difficulties. In 73 of those instances (74%) the graduate students converged to the 'better' option (i.e., the more common incorrect answer choice of introductory students by 5% or more) and in 25 of those instances (26%), they did not converge to the 'better' answer choice. It therefore appears that discussions among graduate students tend to lead them to agree on a more common introductory student alternate conception.

P.5. To what extent do instructors/graduate students identify specific alternate conceptions of introductory physics students? Is their ability to identify these alternate conceptions context dependent?

These questions were answered by identifying student alternate conceptions, the questions in which these alternate conceptions are connected to incorrect answer choices and analyzing the FCI related PCK performance of instructors and graduate students in those questions. Similar alternate conceptions were grouped whenever it was deemed appropriate by the researchers (e.g., alternate conceptions related to Newton's third law, alternate conceptions related to particular tasks, such as identifying all the distinct forces that act on an object, etc.) and, if a particular alternate conception appeared in more than one context, it was investigated whether instructors and/or graduate students performed better at identifying it in some contexts than in other contexts. For multiple choice questions, the context is comprised of both the physical situation presented in the problem and the answer choices, because different answer choices can modify the difficulty of a question. For example, a multiple-choice question is easier for introductory students if the incorrect answer choices are not chosen to reflect common student difficulties, and is challenging for students when they are chosen to reflect common difficulties [2-3].

We now turn to discussing the performance of instructors and teaching assistants in identifying specific introductory student alternate conceptions which arise in more than one context.

1) **Newton's third law: The alternate conceptions of students related to Newton's third law and the performance of both instructors and graduate students in identifying the most common alternate conceptions are both context dependent.**

Table 7.5 shows that in some contexts, introductory students hold alternate conceptions related to Newton's third law more strongly (questions 4, 15 and 28 for which at least 32% of introductory students hold an alternate conception) than in other contexts (question 16, in which only 19% of introductory students hold an alternate conception). Thus, these alternate conceptions are context dependent and they arise more often in certain contexts than in others. This is similar to the finding by Redish [61] that students can answer paired questions about the same concept differently in different contexts: in one context most students answer it correctly, while in another context the most common incorrect answer involves a common student alternate conception.

Table 7.5. Introductory students' alternate conceptions related to Newton's 3rd law, questions in which these alternate conceptions arise (FCI item #), percentage of introductory students who answer the questions incorrectly in the pre-test (% overall incorrect pre) and in the post-test (% overall incorrect post), incorrect answer choices on each question which uncover these alternate conceptions (incorrect answer choices), percentage of introductory students who hold the alternate conceptions based on their selection of these answer choices in the pre-test (Intro stud. alt. pre) and in the post-test (Intro stud. alt. post.) and percentage of instructors (Ins.) and graduate students (GS) who identify them as the most common incorrect answer choices. For convenience, brief descriptions of the problems are given underneath.

Introductory student alternate conceptions	FCI item #	% overall incorrect pre	% overall incorrect post	Incorrect answer choices	Intro stud. alt. pre.	Intro stud. alt. post	Ins.	GS
Newton's 3rd: while both objects exert forces on one another, if both objects are active (i.e., collision), the larger object exerts the larger force; if only one is active (i.e., car pushing truck), the active object exerts a larger force on passive object than vice versa	4	74%	40%	A	73%	39%	97%	84%
	15	75%	56%	C	61%	48%	60%	40%
	16	45%	27%	C	37%	19%	37%	16%
	28	76%	41%	D	61%	32%	38%	52%
Questions								
4. Truck colliding with car.								
15. Car pushing truck and speeding up.								
16. Car pushing truck and moving at constant speed.								
28. Student "a" puts his feet on student "b" and pushes against student "b".								

In addition, it appears that the FCI related PCK performance of both instructors and graduate students is also context dependent. For example, the vast majority of both instructors and graduate students identified the alternate conception related to Newton's third law in a typical context (question 4 – truck colliding with car – see Table 7.5), but they did not identify it as often in the other three contexts (question 15 – car pushing truck and speeding up, question 16 – car pushing truck at constant speed and question 28 – student "a" pushing student "b"). Also, in question 15, 10% of instructors and 12% of graduate students selected the correct answer choice as the most common incorrect answer choice selected by introductory students (see Table

B1 in Appendix B). It is likely that, in this context, they have the same alternate conception as introductory students, namely, that while the car is speeding up, it exerts a larger force on the truck than vice versa. In addition, as noted earlier, the graduate students were first asked to identify the correct answers on the FCI before performing the FCI related PCK task and 24% of them incorrectly selected this answer choice as the correct one (this was one of the two questions with the lowest graduate student performance when asked to select the correct answers for the FCI questions). The fact that even some experts hold this alternate conception after many years of practicing physics points out how strong this alternate conception is and how difficult it is to overcome it in this particular context. Question 16, although relatively easy for introductory students (73% of them answered it correctly in a post-test), revealed a medium level alternate conception, namely that the force the car exerts on the truck is larger than the force the truck exerts on the car. On the other hand, both instructors and graduate students performed very poorly on this question on the FCI related PCK task. In particular, a majority of them (60% of instructors and 76% of graduate students) selected answer choices B, D and E which were selected by only 8% of introductory students (see Table B1 in Appendix B). Similarly, in question 28, many instructors and graduate students performed poorly on the FCI related PCK task of identifying the most common alternate conception and selected answer choice B (45% instructors and 36% graduate students – see Table B1) which stated that student “a” exerts a force on student “b”, but student “b” does not exert a force on student “a”. However, very few introductory students selected this answer choice (2%) and the vast majority of them knew that both students exert forces on one another (91% who selected answer choice C, D or E).

2) **Identification of distinct forces: In the following questions, which ask introductory students to identify all the distinct forces acting on an object, neither instructors nor graduate students identified the most common student alternate conceptions and many graduate students, and even more instructors, selected answer choices which either ignored contact forces or all forces altogether, inconsistent with introductory students' most common incorrect answer choices.**

Table 7.6. Student alternate conceptions related to identifying forces, questions in which these alternate conceptions arise (FCI item #), percentage of introductory students who answer the questions incorrectly in the pre-test (% overall incorrect pre) and in the post-test (% overall incorrect post), incorrect answer choices on each question which uncover these alternate conceptions (incorrect answer choices), percentage of introductory students who hold the alternate conceptions based on their selection of these answer choices in the pre-test (Intro stud. alt. pre) and in the post-test (Intro stud. alt. post.) and percentage of instructors (Ins.) and graduate students (GS) who identify them as the most common incorrect answer choices (Ins.). For convenience, brief descriptions of the problems are given underneath.

Introductory student alternate conceptions	FCI item #	% overall incorrect pre	% overall incorrect post	Incorrect answer choices	Intro stud. alt. pre	Intro stud. alt. post	Ins.	GS
Do not know about any forces (including the force of gravity)	11	86%	65%	E	3%	4%	20%	0%
	29	58%	29%	E	4%	1%	45%	44%
Do not know about contact forces (normal force, tension)	5	90%	76%	A, C, E	64%	32%	70%	60%
	11	86%	65%	A, B, E	41%	17%	60%	40%
	18	88%	72%	A, C, E	62%	30%	70%	48%
	29	58%	29%	A, E	19%	3%	69%	64%
Moving objects are acted on by a distinct force in the direction of motion	5	90%	76%	C, D, E	86%	73%	80%	100%
	11	86%	65%	B, C	76%	56%	63%	80%
	18	88%	72%	C, D, E	86%	71%	84%	96%

Questions
5. Identify the forces acting on a ball while moving in a frictionless, circular channel.
11. Identify the forces acting on a puck while moving on a frictionless surface.
18. Identify the forces acting on a boy while swinging on a rope.
29. Identify the forces acting on a chair at rest on a floor.

Table 7.6 shows that the majority of both instructors and graduate students are aware that introductory students have the alternate conception that moving objects are acted on by a distinct force in the direction of motion. However, in all these questions, many instructors and graduate students claimed that introductory students will not identify contact forces (normal and tension forces), and to a lesser extent they will not identify any forces (including the force of gravity) even in the post-test. However, contrary to what instructors and graduate students claimed, introductory students rarely selected answer choices which correspond to these alternate conceptions. For example, in question 5, 70% of instructors and 60% of graduate students selected answer choices A, C and E which do not include the force that the channel exerts on the ball; however none of these choices was selected by 19% or more introductory students (see Table B1). Similarly, in question 11, 60% of instructors and 40% of graduate students selected choices A, B and E which do not include the normal force; however, these answer choices combined were only selected by 17% of introductory students (see Table 7.6). Moreover, in question 29, it is very interesting that almost half of both instructors (45%) and graduate students (44%) claimed that the most common incorrect answer choice selected by introductory students in the post-test is choice E, which states that no forces act on the ball because it is at rest (see Table B1). On the other hand, this answer choice was selected by only 1% of introductory students. Furthermore, 24% of instructors and 20% of graduate students selected choice A, which only included the force of gravity, an answer choice selected by only 2% of introductory students (see Table B1). Thus, instructors and graduate students did not identify introductory students' alternate conceptions related to identification of distinct forces in different contexts very well.

3) **Constant force implies constant velocity: This alternate conception of introductory students and the performance of both instructors and graduate students in identifying it are both context dependent.**

Table 7.7. Alternate conception that constant net force implies constant velocity, questions in which this alternate conception arises (FCI item #), percentage of introductory students who answer the questions incorrectly in the pre-test (% overall incorrect pre) and in the post-test (% overall incorrect post), incorrect answer choices on each question which uncovers this alternate conception (incorrect answer choices), percentage of introductory students who hold the alternate conception based on their selection of these answer choices in the pre-test (Intro stud. alt. pre) and in the post-test (Intro stud. alt. post.) and percentage of instructors (Ins.) and graduate students (GS) who identify them as the most common incorrect answer choices (Ins.). For convenience, brief descriptions of the problems are given underneath.

Introductory student alternate conception	FCI item #	% overall incorrect pre	% overall incorrect post	Incorrect answer choices	Intro stud. alt. pre	Intro stud. alt. post	Ins.	GS
Constant net force implies constant velocity (also: zero net force implies decreasing velocity)	17	92%	76%	A, D	82%	72%	90%	88%
	21	65%	67%	C	23%	38%	43%	44%
	22	70%	55%	A	37%	33%	67%	28%
	24	37%	30%	C	25%	22%	70%	68%
	25	88%	77%	D	58%	53%	57%	44%
	26	97%	86%	A, B	83%	73%	87%	74%
	27	46%	42%	A	31%	26%	63%	68%

Questions
17. Elevator being pulled up by a cable at constant speed.
21. Rocket drifting horizontally, constant thrust applied vertically, find path followed by the rocket.
22. What is the speed of the rocket during this time (constant, increasing, etc.)?
24. What is the speed of the rocket after thrust drops to zero (constant, increasing, etc.)?
25. Constant horizontal force exerted on a box which causes it to move at constant speed.
26. Force in question 25 is doubled, what happens to speed of box?
27. Force is removed. The box will (A) immediately come to stop, (B) continue moving at constant speed for a while and then slow to a stop, etc.

Examination of Table 7.7 reveals that this alternate conception in introductory students' responses to different questions arises more or less often depending on the context. Table 7.7 shows that, on FCI questions 17 and 26, the vast majority of students (72% and 73%) select answer choices which imply that a constant net force would cause a constant velocity. In question 25, about half answered that the force exerted by the woman has to be greater than the total force which resists the motion of the box in order for the box to move at a constant velocity. In questions 21 and 22, the fraction of students who selected answer choices corresponding to this alternate conception was about one third and in question 24, the fraction was about one fifth. Thus, this alternate conception is observed in introductory students' responses more or less frequently depending on the context.

Table 7.7 suggests that the performance of both instructors and graduate students in identifying this alternate conception, constant force implies constant velocity, is also context dependent and their performance varies significantly depending on the question. For example, the contexts of problems 17 and 25 are similar and in both cases an object is acted upon by two forces, one of which is applied in the direction of motion, and the other opposite to it. In question 17, they are the force exerted by the cable and the weight of the elevator and in question 25 they are the force exerted by the woman and the total force which resists the motion of the box. However, the performance of instructors and graduate students at identifying the alternate conception in these two questions is very different. In particular, in question 17 nearly all of them identified it (90% of instructors and 88% of graduate students – see Table 7.7) whereas in question 25, 57% of instructors and 44% of graduate students identified it. The rest of their choices regarding the most common incorrect answer of introductory students were spread over answer choices A, B and E, none of which was selected by more than 12% of introductory students (see Table B1 in Appendix B).

In addition, the performance of instructors and graduate students related to the "constant force implies constant velocity" alternate conception is not only context-dependent, but it is also not well correlated with the strength of the alternate conception. While instructors and graduate students performed well in the two questions in which more than 70% of introductory students selected answer choices which revealed this alternate conception, they performed better in the question with a medium level alternate conception (question 24) than in other questions in which this alternate conception was strong. In particular, in question 21 and question 25, many instructors and graduate students, and in question 22, many graduate students had difficulty in identifying this alternate conception as shown in Table 7.7).

We note that there is a large discrepancy between the performance of instructors and graduate students in question 22. While the majority of instructors (67%) correctly identified the alternate conception that constant net force implies constant velocity on this question, fewer graduate students identified it (28%) and a large percentage of them (40%) thought that the most common alternate conception is that the speed of the rocket would increase for a while and be constant thereafter, an answer choice (choice D) selected by fewer introductory students (see Table B1 in Appendix B). In addition, 24% of the graduate students mistakenly selected the correct answer choice as the most common incorrect answer chosen by the introductory students for this question.

4) **Confusion between position and velocity and velocity and acceleration: Graduate students are better than instructors at identifying that some introductory students confuse position with velocity, while instructors are somewhat better than graduate students at identifying that some introductory students confuse velocity with acceleration.**

Table 7.8. Student difficulties with interpreting strobe diagrams of motion, questions in which these difficulties arise (FCI item #), percentage of introductory students who answer the questions incorrectly in the pre-test (% overall incorrect pre) and in the post-test (% overall incorrect post), incorrect answer choices on each question which uncover the difficulties (incorrect answer choices), percentage of introductory students who have these difficulties based on their selection of these answer choices in the pre-test (Intro stud. alt. pre) and in the post-test (Intro stud. alt. post) and percentage of instructors (Ins.) and graduate students (GS) who identify them as the most common incorrect answer choices (Ins.). For convenience, brief descriptions of the problems are given underneath.

Introductory student difficulties	FCI item #	% overall incorrect pre	% overall incorrect post	Incorrect answer choices	Intro stud. alt. pre	Intro stud. alt. post	Ins.	GS
Confusing position with velocity	19	46%	49%	D	26%	29%	38%	76%
Confusing velocity with acceleration	20	68%	51%	C	36%	27%	72%	56%
Questions								
19. Diagrams of positions of two blocks at regular, successive time intervals. One block is accelerating, the other has constant velocity. Do they ever have the same speed?								
20. Diagrams of positions of two blocks at regular, successive time intervals. Both blocks move at constant velocities, one smaller than the other. Compare the accelerations.								

Table 7.8 shows that in question 19, graduate students performed better than instructors at identifying that the most common difficulty of introductory students is confusion between position and velocity. In particular, they selected answer choice D much more frequently than instructors (76% compared to 38%). Answer choice D states that the instances when the two blocks have the same speed are when the two blocks have identical positions. The answers of the instructors were spread over other answer choices which were selected by 12% or fewer introductory students (see Table B1 in appendix B).

In question 20, instructors performed better (although not significantly so) at identifying that the most common difficulty of introductory students is confusion between velocity and acceleration. In particular, more instructors selected answer choice D compared to the graduate students (72% instructors compared to 56% graduate students as shown in Table 7.8). Answer choice D states that the acceleration of block “b” is greater than the acceleration of block “a”, while the strobe diagram implies that the velocity of block “b” is greater than the velocity of block “a” (both velocities are constant).

We note that for question 19 there is virtually no change in the performance of algebra-based students from the pre-test to the post-test (54% in the pre-test, 51% in the post-test). There was an improvement in the performance of introductory students in question 20 (17% improvement from 32% correct to 49% correct). One reason why it is more difficult for students to improve in performance in question 19 compared to question 20 is due to the fact that in question 19, one motion is accelerated whereas for question 20, both blocks move at a constant velocity [1-3].

5) Instructors’ and graduate students’ difficulties in identifying other common alternate conceptions of introductory students

The student alternate conception related to an impetus view of motion identified in question 13 is that after a boy throws a ball in the air vertically, on the way up, in addition to the force of gravity, a steadily decreasing force also acts on the ball. On the way down, only the force of gravity acts on the ball. This alternate conception (which is held by 50% of introductory students) was identified by about half of the instructors (47%), but very few graduate students

Table 7.9. Three other common alternate conceptions/difficulties, questions in which these difficulties arise (FCI item #), percentage of introductory students who answer the questions incorrectly in the pre-test (% overall incorrect pre) and in the post-test (% overall incorrect post), incorrect answer choices on each question which uncover the difficulties (incorrect answer choices), percentage of introductory students who have these difficulties based on their selection of these answer choices in the pre-test (Intro stud. alt. pre) and in the post-test (Intro stud. alt. post.) and percentage of instructors (Ins.) and graduate students (GS) who identify them as the most common incorrect answer choices (Ins.). For convenience, brief descriptions of the problems are given underneath.

Introductory student alternate conceptions/difficulties	FCI item #	% overall incorrect pre	% overall incorrect post	Incorrect answer choices	Intro stud. alt pre	Intro stud. alt. post	Ins.	GS
Ball thrown vertically in the air: on the way up - steadily decreasing upward force and gravity, on way down, only gravity	13	88%	65%	C	64%	50%	47%	16%
Relative velocity and reference frame difficulties	14	64%	39%	A	35%	19%	17%	20%
If a constant force acts on an object for some time and then it is removed, the object will eventually go back to the direction in which it was originally moving	23	71%	61%	D	28%	23%	23%	24%

Questions
13. Ball thrown vertically in the air, no air resistance. Find the forces acting on the ball while in the air.
14. Bowling ball rolls off a plane while plane is travelling horizontally. Find the path of the ball.
23. Rocket moving horizontally. Constant thrust applied vertically for some time, then removed. Find the path of the rocket after thrust drops to zero.

(16%) as shown in Table 7.9. A sizeable percentage of both instructors (30%) and graduate students (44%) thought that the most common incorrect answer choice of introductory students for question 13 is choice B (see Table B1 in Appendix B) in which, on the way down, the force of gravity steadily increases. Only 11% of introductory students selected this answer choice. In addition, 20% of instructors and 36% of graduate students selected answer choice A as the most common alternate conception, which does not make a distinction between the forces acting on the object on the way up and on the way down (downward force of gravity along with a steadily decreasing upward force), and which was selected by only 4% of introductory students (see Table B1). Thus, the responses of instructors and graduate students suggest that they do not have a good understanding of introductory physics students' difficulty in this situation.

Question 14 reveals an interesting introductory student difficulty. Although the question was somewhat easy for introductory students in the post-test (61% correct), 19% of introductory students selected the trajectory which arches backwards even in the post-test. This question is one for which the reasons for explicitly selecting the answers are not provided, and it would be worthwhile knowing students' reasoning. We therefore added reasons for the each answer choice and administered the question as part of a final exam in a large algebra-based introductory physics class with 400+ students. The reasons for the path that arches backwards were (A) "because by the time it strikes the ground, the plane will cover some horizontal distance" and (B) "due to air resistance". Choice (C) in this multiple-choice question provided a justification for path (2) which goes straight down: "the force of gravity is the only force acting on the ball after the ball falls from the plane and it causes the ball to fall vertically downwards". These justifications increased the percentage of students who selected these answer choices by 5% each. The percentages of students who selected each incorrect answer choice (A), (B) and (C) are: 17%, 7% and 14%. It appears that the main reason students select the path that arches backwards is because they are having difficulty viewing the motion of the ball from the perspective of a person on the ground. They are implicitly in the airplane thinking that it keeps travelling after the bowling ball falls out and therefore covers more horizontal distance. In the FCI version of the question, few instructors (17%) and graduate students (20%) identified that the most common incorrect choice would be (A), the path which arches backwards. Both groups selected choices B (straight down) and C (straight oblique line) much more often (see Table B1 in Appendix B), and these answer choices were selected by fewer introductory students (10% and 9% – see Table B1).

Question 23 reveals another interesting student alternate conception. For this question, the most commonly selected answer choice (by 23% of introductory students) is choice D. The path described by choice D is one in which the direction of the velocity of the rocket gradually returns to its original orientation (to the right). This implies that students who selected this choice thought that forces which act for a finite time do not change the direction of motion indefinitely and the rocket eventually returns to its original orientation. Only about one quarter of both instructors and graduate students (23% and 24% – see Table 7.9) identified this as the most common incorrect answer choice. The answers of graduate students appear to suggest that they are random guessing (percentages between 20% and 28% for each incorrect answer choice – see Table B1), while 47% of instructors selected answer choice C (vertical path), which was selected by only 18% of introductory students. Thus, in this context also, the instructors and graduate students struggled to identify the most common difficulties of introductory physics students.

7.4.2 Results: Secondary research questions

S.1. Which questions on the FCI pose significant challenges for students even after instruction (poor performance)?

We note that the original paper describing the development of the FCI has a discussion of students' difficulties on the original version of the FCI even after instruction. For the later version of the FCI that we employed in our research, this question about introductory physics students' difficulties in post-test was answered earlier while discussing which introductory student alternate conceptions were correctly identified by instructors and graduate students. One can also refer back to Table B1 (in Appendix B) which provides the performance of introductory students on the post-test on the questions in the FCI in which at least one incorrect answer choice was selected by 19% or more introductory students.

S.2. Are there any questions on the FCI in which there is little improvement (less than 10%) for algebra-based students from pre- to post-test?

The researchers decided that "little" (not noteworthy) improvement occurred from pre- to post-test in questions in which the normalized gain was less than 0.173 (i.e. normalized gain is in the lower 1/3 based on the average normalized gain of 0.26), and questions in which the percentage of introductory students who hold a particular alternate conception decreased by 5% or less. There were twelve questions on the FCI (shown in Table 7.10) which fit at least one of these two criteria. Table 7.10 also shows the percentage of introductory students who answered each question incorrectly both in the pre-test and in the post-test, the normalized gain, the incorrect answer choices corresponding to the most common alternate conceptions and the percentage of introductory students who hold those alternate conceptions both in the pre-test and in the post-test.

Constant net force implies constant velocity and zero net force implies decreasing velocity

The most prevalent difficulty observed was that introductory students have a very difficult time abandoning the notion that a constant net force implies a constant velocity. In the questions which can be used to test for this alternate conception (questions 17, 21, 22, 24, 25, 26 and 27), there was either less than 0.173 normalized gain, and/or the percentage of students who hold this alternate conception did not decrease by more than 5% (see Table 7.10). In question 21, there was a slight shift in alternate conceptions because in the pre-test, students selected answer choice

Table 7.10. The 12 questions on the FCI on which there was little improvement (less than 0.173 normalized gain and/or difference of 5% or less in the percentages of introductory physics students harboring a particular alternate conception), student alternate conceptions/difficulties associated with these questions, percentage of introductory students who answered them incorrectly in the pre-test (% overall incorrect pre) and in the post-test (% overall incorrect post), normalized gain (Norm. gain), most common incorrect answer choices which uncovered these alternate conceptions/difficulties (incorrect answer choices), percentage of students who have these alternate conceptions/difficulties based on their selection of those incorrect answer choices in the pre-test (Intro stud. alt. pre) and in the post-test (Intro stud. alt. post). For convenience, short descriptions of the questions are given underneath.

FCI item #	**Introductory student alternate conceptions/difficulties**	% overall incorrect pre	% overall incorrect post	Norm. gain	Incorrect answer choices	Intro stud. alt. pre	Intro stud. alt. post
5	Moving objects have a distinct force in the direction of motion	90%	76%	0.16	C, D, E	86%	73%
30		88%	74%	0.16	B, D, E	87%	71%
9	Difficulties with addition of perpendicular velocities	57%	47%	0.17	C	20%	19%
19	Confusing position with velocity	46%	49%	-0.06	D	26%	29%
23	If a constant force acts on an object for some time and then it is removed, the object will eventually go back to the direction in which it was originally moving	71%	61%	0.14	D	28%	23%
17		92%	76%	0.17	A	60%	62%
21		65%	67%	-0.02	C	23%	38%
22	Constant net force implies constant	70%	55%	0.22	A	37%	33%
24	velocity (also: zero net force implies	37%	30%	0.20	C	25%	22%
25	decreasing velocity)	88%	77%	0.13	D	58%	53%
26		97%	86%	0.11	A	41%	41%
27		46%	42%	0.09	A	31%	26%

Questions
5. Identify the forces acting on a ball while moving in a frictionless, circular channel.
9. Puck sliding horizontally with speed "v_0". Kicked vertically (if at rest kick would give the puck speed "v_k". What is the speed of the puck just after the kick?
17. Elevator being pulled up by a cable at constant speed.
19. Diagrams of positions of two blocks at regular, successive time intervals. One block is accelerating, the other has constant velocity. Do they ever have the same speed?
21. Rocket drifting horizontally, constant thrust applied vertically. Find path followed by the rocket.
22. What is the speed of the rocket during this time (constant, increasing, etc.)?
23. Path of the rocket after thrust drops to zero.
24. What is the speed of the rocket after thrust drops to zero (constant, increasing, etc.)?
25. Constant horizontal force exerted on a box which causes it to move at constant speed.
26. Force in question 25 is doubled, what happens to speed of box?
27. Woman stops applying horizontal force (from question 26). The box will: immediately come to a stop, immediately start slowing to a stop, etc.

30. Tennis player hits a tennis ball against strong wind. Identify the forces acting on the tennis ball while in the air.

i.e., implying that constant vertical force results in constant vertical speed) equally (21% and 23%). However, in the post-test, more students selected choice C than choice B (38% compared to 13%). It appears that some students have learned that the initial motion of the rocket must be taken into account when determining the path after the engine of the rocket was turned on, but they still harbor the alternate conception that a constant force implies a constant velocity.

Moving objects are acted upon by a distinct force in the direction of motion

When it comes to this alternate conception, it appears that introductory students improved in some contexts after instruction, but not in others. Out of the four questions in which this alternate conception can be identified, the normalized gain was less than 0.173 in two (questions 5 and 30 – see Table 7.10). However, even in the other two questions in which the normalized gain was not in the lower one third, it was not very large (question 11 – puck sliding across a frictionless surface: normalized gain = 0.24 and question 18 – boy swinging on a rope – normalized gain = 0.19 – see Table B3 in Appendix B).

Difficulties with addition of perpendicular velocities

In question 9, it appears that the same number of students (20% in the pre-test and 19% in the post-test) noted that the final speed of the puck will be the arithmetic sum of "v_0" (initial speed of the puck) and "v_k" (the speed the kick would have imparted, had the puck been stationary). These students had difficulty realizing that the Pythagorean Theorem must be applied to add the two perpendicular velocities. Even after instruction, only about half of the introductory students correctly reasoned that the final speed will be greater than either of the speeds "v_0" and "v_k", but less than their sum.

Confusing position with velocity

In question 19, which assesses students' ability to extract information about speed from strobe diagrams of motion, the percentage of correct answers decreased from 54% before instruction to 51% after instruction. In addition, the percentage of introductory students who confused position with velocity (choice D: the two objects have the same speed at points 2 and 5 on the strobe diagram, at which points they have the same position) remains approximately the same (26% in the pre-test and 29% in the post-test). Interestingly, in question 20, which assesses students' ability to extract information about acceleration from strobe diagrams of motion, the normalized gain was 0.25 (see Table B3 in Appendix B).

If a constant force acts on an object for some time and then it is removed, the object will eventually return to the direction in which it was originally moving before the constant force was applied.

Many introductory physics students incorrectly believe this**.** In addition to low normalized gain on question 23, which assessed understanding of this concept (normalized gain = 0.14 – see Table B3), the percentage of students who hold this alternate conception decreased by only 5% from the pre-test to the post-test.

S.3. Are there any shifts in the most common alternate conceptions from the pre-test to the post-test?

1) Identify all of the distinct forces that act on an object

For questions 5, 11 and 18, the answer choices which include the force of gravity and a force in the direction of motion are choices C, B and C, respectively, while the answer choices which include the force of gravity, the contact force and a force in the direction of motion are choices D, C and D, respectively. Table 7.11 shows that the percentage of introductory students who selected these incorrect answer choices in each question are comparable in the pre-test (question 5: C – 31%, D – 25%; question 11: B – 31%, C – 45% and question 18: C – 14% and D – 27%). In the post-test, the incorrect answer choices shift significantly towards the answer choices which include the force of gravity, the contact force and a force in the direction of motion (question 5: 44% compared to 12%, question 11: 48% compared to 8%, question 18: 42% compared to 4%). It appears that, before instruction, some students are not aware of contact forces (normal force, tension force), and after instruction they are aware of them. However, introductory students often do not abandon the alternate conception that if an object is moving in a certain direction, a distinct force must be acting on it in that direction.

The only major shift in alternate conceptions of introductory algebra-based students which occurred for more than one question was observed on questions which asked students to identify all the distinct forces that act on an object. Before instruction, algebra-based students selected incorrect answer choices which corresponded to the force of gravity and force in the direction of motion with similar frequency compared to the incorrect answer choices which

Table 7.11. Introductory students' alternate conceptions related to identifying all the distinct forces that act on an object and alternate conceptions related to question 2, the questions in which these alternate conceptions occurred, the percentage of introductory students who answered the questions incorrectly in the pre-test (% incorrect pre) and in the post-test (% incorrect post), the most common incorrect answer choices which uncovered these alternate conceptions and the percentage of students who hold these alternate conceptions based on their selection of those incorrect answer choices in the pre-test (Intro stud. alt. pre) and in the post-test (Intro stud. alt. post). For convenience, short descriptions of the questions are given underneath.

Introductory student alternate conceptions	FCI item #	% overall incorrect pre	% overall incorrect post	Incorrect answer choices	Intro stud. alt. pre	Intro stud. alt. post
Force of gravity and force in the direction of motion	5	90%	76%	C	31%	12%
	11	86%	65%	B	31%	8%
	18	88%	72%	C	14%	4%
Force of gravity, contact force and force in the direction of motion	5	90%	76%	D	25%	44%
	11	86%	65%	C	45%	48%
	18	88%	72%	D	27%	42%
Ball twice as heavy that rolls off horizontal table travels half as far	2	73%	56%	B	21%	25%
Ball twice as heavy that rolls off horizontal table travels considerably less, but not half	2	73%	56%	D	37%	21%
Questions						
2. Two metal balls, same size, one twice as heavy as the other, roll off a horizontal table: (A) both balls hit the floor at the same distance, (B) heavier ball hits the floor at half the distance of the lighter ball, etc.						
5. Identify the forces acting on a ball while moving in a frictionless, circular channel.						
11. Identify the forces acting on a puck while moving on a frictionless surface.						
18. Identify the forces acting on a boy while swinging on a rope.						

corresponded to the force of gravity, contact forces and force in the direction of motion. After instruction, they overwhelmingly selected the latter compared to the former (this is also true for calculus-based students). The only other shift occurred on question 2, for which, before instruction, more algebra-based students thought that the ball twice as heavy will strike the floor considerably closer compared to students who thought that the ball twice as heavy will strike the floor at exactly half the distance of the lighter ball, whereas after instruction the percentages of students who held these alternate conceptions are about the same. Table 7.11 shows these alternate conceptions, the questions in which they occur, the percentage of incorrect answers both in the pre-test and in the post-test along with the incorrect answer choices corresponding to the most common alternate conceptions and the percentage of students who hold these alternate conceptions.

2) Two metal balls roll off a horizontal table

The other shift in alternate conceptions occurred in question 2. In the pre-test, more students thought that the heavier ball hits the floor considerably closer than the lighter ball, but not necessarily half the horizontal distance, compared to students who thought that it hits the floor at half the distance of the lighter ball (37% compared to 21% – see Table 7.11). In the post-test however, the percentages of students who selected these choices are about the same (21% and 25%).

S.4. On which questions do calculus-based students perform better than algebra-based students? Are there any questions in which the alternate conceptions of algebra-based students are different from the alternate conceptions of calculus-based students?

Due to the large population sizes, any difference of 5% or more turned out to be statistically significant by means of chi-square tests [59]. However, a difference of 5% in performance from the pre-test to the post-test does not have much practical significance. Instead, questions which were answered correctly by 20% or more of calculus-based students compared to algebra-based students were chosen as a heuristic by the researchers to be indicative of significantly better performance of calculus-based students compared to algebra-based students. The question about whether the alternate conceptions of algebra-based students are different from the alternate conceptions of calculus-based students was answered by investigating whether there were any questions in which the most common incorrect answer choice(s) of algebra-based students was (were) different from the most common incorrect answer choice(s) of calculus-calculus based students. It turned out that there were no such questions. For all questions which included only one common incorrect answer choice, this was most common for both algebra-based and calculus-based students (in addition, the fraction of calculus-based students who selected that particular incorrect answer choice was always smaller than the fraction of algebra-based students – see Tables B3 and B4 included in Appendix B). Similarly, for the questions which included two common incorrect answer choices, they were common for both algebra-based and calculus-based students (see Tables B3 and B4). Moreover, only one question had three common incorrect answer choices and these three answer choices were the most common incorrect answers for both algebra-based and calculus-based students. It therefore appears that in the pre-test, the algebra-based students harbor the same alternate conceptions as calculus-based students. However, algebra-based students hold the same alternate conceptions more strongly than calculus-based students.

1) Pre-test comparison of performance of algebra-based and calculus-based students

Calculus-based students correctly answered every single question on the FCI more frequently than algebra-based students in the pre-test. Differences of 10% or more occurred on 26 questions and differences of 20% or more occurred on 8 questions (see Tables B3 and B4 included in). We will focus on the questions in which the differences were of 20% or more (questions 1, 3, 12, 13, 14, 20, 22 and 28). Table 7.12 shows the percentages of algebra-based and calculus based students who answer these questions incorrectly, the incorrect answer choices which uncover an alternate conception and the percentages of algebra-based and calculus-based students who harbor these alternate conceptions.

Table 7.12. Questions in which calculus-based students outperformed algebra-based students in the pre-test, the most common alternate conceptions/difficulties uncovered by these questions, percentage of incorrect answers for both algebra-based (% overall incorrect algebra) and calculus-based (% overall incorrect calculus) introductory students, incorrect answer choices which correspond to the most common alternate conceptions/difficulties (incorrect answer choices) and percentages of algebra-based (Alg. alt.) and calculus-based (Calc. alt.) students who harbor/have these alternate conceptions/difficulties.

FCI item #	**Pre-test introductory student alternate conceptions**	% overall incorrect algebra	% overall incorrect calculus	Incorrect answer choices	Alg. alt.	Calc. alt.
1	Time it takes an object to fall freely through a certain distance is proportional to mass	47%	18%	C	25%	8%
3	Freely falling objects reach terminal velocity a short time after release	60%	34%	A	31%	17%
12	An object fired horizontally will not immediately descend and continue to move horizontally for some time	41%	16%	C	32%	14%
13	Ball thrown vertically in the air: on the way up - steadily decreasing upward force and gravity, on way down, only gravity	88%	67%	C	64%	5`%
14	Relative velocity and reference frame difficulties	64%	37%	A	35%	21%
20	Confusing velocity with acceleration	68%	46%	C	36%	22%
22	Constant net force implies constant velocity	70%	49%	A	37%	27%
28	Newton's third law: the active object exerts more force on the passive than vice versa	76%	56%	D	61%	45%

Questions
1. Two metal balls, same size, one twice as heavy as the other are dropped from the same height. The time it takes the balls to fall is (A) half as long for heavier ball, (B) half as long for lighter ball (C) same, etc.
3. The two balls from question 1 roll off a horizontal table. (A) distance same for both balls, (B) distance of heavier ball is half the distance of lighter ball, etc.
12. Ball fired horizontally from cannon. Determine the path it follows.
13. Ball thrown vertically in the air, no air resistance. Find the forces acting on the ball while in the air.
14. Bowling ball rolls off a plane while plane is travelling horizontally. Find the path of the ball.
20. Diagrams of positions of two blocks at regular, successive time intervals. Both blocks move at constant velocities, one smaller than the other. Compare the accelerations.
22. Rocket drifting horizontally, constant thrust applied vertically. Speed of rocket during this time (constant, increasing, etc.)
28. Student "a" puts his feet on student "b" and pushes against student "b".

Significantly better performance of calculus-based students compared to algebra-based students is context dependent

An interesting finding suggested by Table 7.12 is that calculus-based students answer questions involving particular force and motion concepts significantly better than algebra-based students (by 20% or more) in some contexts, but not in others. For example, 20% more calculus-based students than algebra-based students correctly interpreted Newton's third law in the context of problem 28 (one student pushing another). However, in the other two questions with the alternate conception that the active object exerts more force on the passive object than vice versa (question 15 – car pushing truck and accelerating and question 16 – car pushing truck at constant speed) calculus-based students did not outperform algebra-based students by more than 20%. In fact, the difference in question 15 is merely 4% (see Tables B3 and B4 in Appendix B).

A similar observation can be made by examining the questions related to the alternate conception that a constant net force implies a constant velocity (questions 17, 21, 22, 24, 25, 26 and 27). We find that 21% more calculus-based students than algebra-based students answered question 22 correctly. However, in the other questions involving the same concept, the smallest difference in the performance of calculus-based and algebra-based students was 11% (i.e., calculus-based students always performed better, but not always by 20% or more).

Understanding of freely falling objects

The better performance of calculus-based students compared to algebra-based students in questions 1 and 3 (by 29% and 26% respectively) in the pre-test indicates that calculus-based students have a better understanding of the physics of freely falling objects before instruction.

An object fired horizontally will not immediately descend, but continue to move horizontally for some time.

The data in Table 7.12 show that algebra-based students performed worse than calculus-based students on question 12, which uncovered this alternate conception, by 25%.

Relative velocity and reference frame difficulties

In question 14, it appears that algebra-based students find it difficult to view the motion of the bowling ball falling from the airplane from the correct frame of reference. Many introductory students thought that the path of the ball falling from the plane arches backwards because they have difficulty viewing the path of the ball as ground observers [1-3].

Impetus view of motion

Question 13 indicates that many algebra-based (64%) and calculus-based (51%) students have an impetus view of motion before instruction (ball thrown vertically in the air, on the way up will be acted upon by a steadily decreasing upward force and the force of gravity, and on the way down, will only be acted upon by the force of gravity), however, 21% more calculus-based than algebra-based students answer this question correctly.

Confusion between velocity and acceleration

Question 20, which was answered correctly by 22% more calculus-based students than algebra-based students, indicates that calculus-based students are more likely to correctly interpret acceleration from strobe diagrams of motion. The most common incorrect answer choice for both groups is choice C (acceleration of "b" is greater than that of "a") which indicates that many introductory students confuse acceleration with velocity of an object.

2) Post-test comparison of performance between algebra-based and calculus-based students

Similar to the pre-test, calculus-based students outperformed algebra based students on all but one question (item 15) after instruction (post-test). Differences of 10% or more occurred on 26 questions and differences of 20% or more occurred on 14 questions (more than in the pre-test for which differences of 20% or more occurred on only 8 questions). These questions are 5, 9, 10, 11, 13, 17 through 23, 25 and 26. Question 10 is not included in Table 7.13 because although calculus-based students performed better than algebra-based students by 20%, there were no incorrect answer choices selected by 19% or more of either calculus-based or algebra-based students in the post-test, and therefore no strong or medium level common alternate conceptions were uncovered by this question. Table 7.13 shows the percentages of algebra-based and calculus based students who answer these questions incorrectly, the incorrect answer choices which uncover an alternate conception and the percentage of algebra-based and calculus-based students who harbor these alternate conceptions in the pre-test.

Identifying all of the distinct forces that act on an object

Questions 5, 11 and 18 all ask students to identify all of the distinct forces acting on an object. Comparison of algebra-based students' alternate conception shifts in these questions indicated that on the pre-test many of them had failed to identify contact forces, while on the post-test,

Table 7.13. Questions on which calculus-based students outperformed algebra based students in the post-test, the most common alternate conceptions/difficulties uncovered by these questions, percentage of incorrect answers for both algebra-based (% overall incorrect algebra) and calculus-based (% overall incorrect calculus) introductory students, incorrect answer choices which correspond to the most common alternate conceptions/difficulties (Incorrect answer choices) and percentages of algebra-based (Alg. alt.) and calculus-based (Calc. alt.) students who have these alternate conceptions/difficulties.

FCI item #	**Post-test introductory student alternate conceptions/difficulties**	% overall incorrect algebra	% overall incorrect calculus	Incorrect answer choices	Alg. alt.	Calc. alt.
5	Moving objects have a distinct force in the direction of motion	76%	53%	C, D	56%	42%
11		65%	39%	B, C	56%	30%
18		72%	45%	C, D, E	71%	46%
17	Constant net force implies constant velocity (also: zero net force implies decreasing velocity)	76%	56%	A, D	72%	52%
21		67%	43%	C	38%	23%
22		55%	33%	A	33%	20%
25		77%	49%	D	53%	38%
26		86%	58%	A, B	73%	48%
9	After performing an action on an object, its speed depends only on the action, not the previous motion	47%	27%	B, C	39%	23%
13	Ball thrown vertically in the air: on the way up - steadily decreasing upward force and gravity, on way down, only gravity	65%	39%	C	50%	31%
19	Confusing position with velocity	49%	25%	D	29%	12%
20	Confusing velocity with acceleration	51%	29%	C	27%	16%
23	If a constant force acts on an object for some time and then it is removed, the object will eventually return to the direction in which it was originally moving	61%	36%	D	23%	14%

5. Identify the forces acting on a ball while moving in a frictionless, circular channel.
9. Puck sliding horizontally with speed "v_0". Kicked vertically (if at rest kick would give the puck speed "v_k". What is the speed of the puck just after the kick?
11. Identify the forces acting on a puck while moving on a frictionless surface.
13. Ball thrown vertically in the air, no air resistance. Find the forces acting on the ball while in the air.
17. Elevator being pulled up by a cable at constant speed.
18. Identify the forces acting on a boy while swinging on a rope.
19. Diagrams of positions of two blocks at regular, successive time intervals. One block is accelerating, the other has constant velocity. Do they ever have the same speed?
20. Diagrams of positions of two blocks at regular, successive time intervals. Both blocks move at constant velocities, one smaller than the other. Compare the accelerations.
21. Rocket drifting horizontally, constant thrust applied vertically, find path followed by the rocket.
22. Speed of the rocket during this time (constant, increasing, etc.)?
23. Rocket moving horizontally. Constant thrust applied vertically for some time, then removed. Find the path of the rocket after thrust is removed.
25. Constant horizontal force exerted on a box which causes it to move at constant speed.

26. Force in question 25 is doubled, what happens to speed of box?

they do identify them. However, they retain the alternate conception that moving objects are acted upon by a distinct force in the direction of motion. Comparison of performance of algebra-based students with calculus-based students for the post-test indicates that more algebra-based than calculus-based students, even after instruction, still claim that if an object is moving in a certain direction, a distinct force must be acting on the object in the same direction. Question 30 is similar, and in this question as well, more algebra-based students than calculus-based students think that there is a force of the "hit" that continues to act on the ball even when the tennis ball loses contact with the racquet. In particular, the calculus-based students outperformed the algebra-based students by 18% on this question.

Constant net force implies constant velocity and zero net force implies decreasing velocity

Calculus-based students outperformed algebra-based students by at least 20% in almost all questions on the FCI in which this alternate conception is uncovered (see Table 7.13). Furthermore, the largest discrepancies between students in the calculus-based and algebra-based courses on all FCI questions occurred in questions 25 and 26 (28%). It appears that calculus-based students are better than algebra-based students at discarding the alternate conception that constant net force implies constant velocity and improving their performance in questions dealing with Newton's 2nd law. In particular, on the pre-test, calculus-based students outperformed algebra-based students on only one question (question 22), which dealt with the alternate conception that constant net force implies constant velocity, but in the post-test, on all these questions, they improved more than algebra-based students, both in the percentage of correct answers and in the percentage of students who hold this alternate conception (see Tables B3 and B4 in Appendix B).

After applying a force on an object, its speed depends only on the applied force, and not on the previous motion.

On question 9, more algebra-based students retained the alternate conceptions that the speed of the puck after receiving the kick would be the same as the speed the kick would impart if the puck was stationary and independent of the original speed of the puck. Calculus-based students performed better than algebra-based students on this question by 20% (see Table 7.13).

Impetus view of motion

On question 13, the performance of calculus-based students was better in the pre-test (by 21%) as well. The discrepancy in performance is slightly higher in the post-test (26%).

Interpreting strobe diagrams of motion

Calculus-based students outperformed algebra-based students in both questions 19 and 20, which assess students' ability to extract information about velocity and acceleration from strobe diagrams of motion. In particular, algebra-based students are more likely than calculus-based students to confuse position with velocity (in question 19) and velocity with acceleration (in question 20).

3) Post-test comparison of alternate conceptions between algebra-based and calculus-based students

Similar to the pre-test, in the post-test, the most common alternate conceptions were the same for algebra-based and calculus-based students. However, on almost all questions, algebra-based students held these alternate conceptions more strongly than calculus-based students (see Tables B3 and B4 in Appendix B).

7.5 DISCUSSION AND SUMMARY

7.5.1 Instructor and graduate student performance in identifying common introductory student alternate conceptions related to force and motion as revealed by the FCI

Awareness of introductory physics students' difficulties and being able to understand the way they reason about physics is an important aspect of pedagogical content knowledge because instruction can take advantage of students' initial knowledge and pedagogical approaches and curricula can explicitly account for these difficulties. Our investigation used the FCI to evaluate this aspect of the pedagogical content knowledge of both instructors and Teaching Assistants (TAs) with varying degrees of teaching experience. For each item on the FCI, the instructors and TAs were asked to identify the most common incorrect answer choice of introductory physics students. We also discussed the responses individually with a few instructors and in a class discussion with the graduate students.

The ability to identify common introductory students' alternate conceptions in the FCI does not appear to be dependent on teaching experience or familiarity with US teaching practices

We find that the instructors, who on the average had significantly more teaching experience as lecturers, did not perform better at identifying common introductory student alternate conceptions than graduate students, who had limited teaching experience as lecturers. We note however, that graduate students had taken introductory physics only four years prior to this study as undergraduate students and a majority of them were TAs in an introductory recitation or laboratory class, graded quizzes, homework and exams, and held office hours in which they helped introductory students individually (or in small groups). These experiences may have, on average, improved their ability to identify introductory physics students' difficulties related to force and motion. Among both instructors and graduate students, some of them performed very well, while others performed poorly. Moreover, the ability to correctly identify students' difficulties was not correlated with the teaching experience of the physics instructors in introductory algebra-based and calculus-based mechanics courses. In particular, the performance among instructors was not better for those who had taught these courses recently (last seven years) compared to those who had not taught recently. One possible reason for why there was no statistically significant difference between the two groups of instructors is that all instructors who taught introductory mechanics employed traditional methods, most had minimal contact with students in the large introductory classes, and did not grade introductory students' homework and quizzes which may have provided some insight into students' common difficulties (the grading was done by the TAs). Moreover, even instructors in the other group who had not taught introductory mechanics had taught other introductory courses in which force concepts were relevant and many of these instructors had taught introductory mechanics more than seven years ago.

We also investigated whether the ability of American graduate students to identify introductory students' alternate conceptions was better than that of foreign graduate students and found that this was not the case. The numbers of graduate students in the different groups (American – 9, Chinese – 9 and other foreign – 7) were too small to perform meaningful statistics, but it appears that their average performance in identifying common student alternate conceptions is very similar. The discussions with graduate students from different countries in the TA training class about this FCI related PCK task suggested that foreign students were similar to American students in this regard, but it is difficult to justify why their performance in identifying student difficulties are comparable despite their different backgrounds.

Instructors appear to identify 'strong' student alternate conceptions better than 'medium' level ones, while graduate students exhibit similar performance in identifying 'strong' and 'medium' alternate conceptions

An alternate conception was considered 'strong' if it is held by more than 1/3 of introductory students. 'Medium' level alternate conceptions were connected to incorrect answer choices selected by a percentage of introductory students between 19% and 33%. We found that instructors were able to identify the strong alternate conceptions somewhat more often than the medium level ones while graduate students exhibited similar performance.

Discussions among graduate students improved their PCK performance in identifying common student alternate conceptions.

The graduate students identified what they thought to be the most common introductory student alternate conceptions first individually and then in groups of two or three. Their group performance was statistically significantly better than their individual performance. In addition, when the individual answers of graduate students working in a group disagreed, discussions more often shifted towards the more common alternate conception (74% of the time) than on the less common one. This implies that discussing student difficulties with other TAs/instructors leads to a better understanding of students' initial knowledge state (and difficulties). Therefore, exercises which encourage such discussions in the context of conceptual assessments could be beneficial and should be incorporated into teacher preparation and/or training courses.

For most alternate conceptions which appear in more than one question, the ability of both instructors and graduate students to identify them is context dependent.

We find that while both physics instructors and TAs, on average, performed better than random guessing at identifying introductory students' alternate conceptions related to force and motion, they did not identify many common difficulties that introductory physics students have even after traditional instruction and their ability to identify them was context dependent. For example, for Newton's third law alternate conceptions, the vast majority of both instructors (97%) and graduate students (84%) identified the most common alternate conception in the typical context (truck colliding with car), but fewer identified it in other contexts (for example car pushing truck and accelerating – 60% of instructors and 40% of graduate students identified the most common student alternate conception that the car exerts the larger force).

Similarly, identifying the common alternate conception that a constant force implies a constant velocity was also context dependent. For example, questions 17 (elevator being pulled up by a cable at constant speed) and 25 (constant horizontal force applied on a box which causes it to move at constant speed) are similar. However, both instructors and graduate students correctly selected the alternate conception in question 25 much less frequently than in question 17 (90% compared to 57% for instructors and 88% compared to 44% for graduate students). Similar observations can be made while examining the other five questions involving this alternate conception.

For alternate conceptions related to identifying all distinct forces that act on an object, there was no context dependence in the ability of both instructors and graduate students to identify the most common student alternate conceptions; however, their PCK performance leaves a lot of room for improvement. In particular, the largest percentage of instructors who identified the most common alternate conception related to identification of distinct forces in any of these questions was 40% and for graduate students it was 60%.

Alternate conceptions for which the PCK performance of instructors and graduate students leaves a lot of room for improvement

As noted earlier, neither instructors nor graduate student TAs performed well at identifying student alternate conceptions related to identifying distinct forces (questions 5, 11, 18, 29 and 30). This is because in almost all these questions (all except for question 30), a sizeable majority of instructors and graduate students selected answer choices which did not include contact forces or any forces, which was inconsistent with introductory student choices. In question 28, for example, (chair at rest on the floor), 44% of instructors and 45% of graduate students thought that the most common student alternate conception is that no forces act on the chair because it is at rest, an answer choice selected by only 1% of introductory students.

For introductory student difficulties related to interpreting strobe diagrams of motion, the majority of instructors did not identify that introductory students confuse position with velocity (only 38% of instructors identify this difficulty), and only half of the graduate students identify that introductory students confuse velocity with acceleration.

Alternate conceptions related to Newton's third law are identified by both instructors and graduate students in a typical context (truck colliding with car), but not in less typical contexts (questions 15, 16 and 28) for which the largest percentage of instructors who identify the most common alternate conception is 60% and for graduate students 52%. A similar observation can be made for the alternate conception that constant force implies constant velocity. However, in these questions instructors and graduate students perform reasonably well in more than half of them (5/7 questions for instructors and 4/7 questions for graduate students).

There are three other alternate conceptions/difficulties which are not identified by the majority of instructors or graduate students (for two of them, the largest percentage of instructors or graduate students who identify them is 24%). These occur on questions 13 (ball thrown vertically in the air on which students have to identify all the forces), question 14 (bowling ball rolls off a plane while the plane is moving horizontally, on which students have to determine the path of the ball as viewed from the ground) and question 23 (rocket moving horizontally, with constant thrust applied vertically for some time and then removed, on which students have to determine the path of the rocket after the thrust is removed).

In summary, there were many alternate conceptions held by more than 19% of introductory students (strong or medium level) that were not identified very often by both instructors and graduate students. Even instructors who teach introductory courses on a regular basis struggled to identify some common alternate conceptions. In addition, some instructors and graduate students explicitly noted that this task was challenging and it was difficult for them to think about physics questions from a student's perspective and expressed concern about their performance (a few noted that they were confident that they have performed poorly on this task).

7.5.2 Introductory student FCI performance – most prevalent difficulties

The performance of introductory physics students is discussed at length in the results section, which discusses introductory student FCI performance (section 7.5.2); here we will summarize the most important results.

Introductory students have a very difficult time abandoning the alternate conception that a constant force implies constant velocity. In all the questions which can be used to test for this alternate conception (questions 17, 21, 22, 24, 25, 26 and 27), either the normalized gain was less than 0.175 or the percentage of introductory students who hold this alternate conception did not decrease by more than 5% from pre-test to post-test.

The introductory students' performance on questions which can be used to uncover the alternate conception that moving objects are acted upon by a distinct force in the direction of motion (questions 5, 11, 18 and 30) improved on some questions, but not on others. Two of the questions had normalized gains less than 0.173 and the other two had larger normalized gains, but not by much (they were 0.19 and 0.24).

Confusing between position and velocity (question 19) was the difficulty most resistant to change, and 3% more students had this difficulty in the post-test. In the other question which required interpretation of strobe diagrams of motion (question 20), the normalized gain was 0.25.

Introductory student shifts in alternate conceptions from pre-test to post-test

There was only one major shift in alternate conceptions which occurred in more than one question (in questions 5, 11, 18 and 29, which asked students to identify all the distinct forces that act on an object). In the pre-test, many students were unaware of contact forces and believed that moving objects are acted upon by a distinct force in the direction of motion, while in the post-test, most students were aware of contact forces. However the alternate conception that there must be a distinct force in the direction of motion was still present.

Comparison of performance and alternate conceptions of algebra-based with calculus-based students

In the pre-test, calculus-based students answered every question correctly more frequently than algebra-based students and in the post-test, they answered every question correctly more frequently except for one (question 15). The differences appeared to get larger between these two populations in the post-test compared to the pre-test. In particular, in the pre-test, there were 8 questions in which differences were 20% or larger while in the post-test there were 14 such questions.

In addition, the better performance of students in the calculus-based courses compared to the algebra-based courses was context dependent. For example, calculus-based students answered a Newton's third law question better (by 20%) than algebra-based students in the context of question 28 (in which one student was pushing another), but they did not perform better in the other three questions involving the same concept. In fact, in question 15 (a car pushing a truck and accelerating) the percentages of calculus-based and algebra-based students who answered correctly are identical. A similar observation can be made while examining the alternate conception that a constant net force implies constant velocity, in that the better performance of calculus-based students compared to algebra-based students in questions which can be used to uncover this alternate conception is context dependent. In the post-test, algebra-based students consistently answered most questions involving two common alternate conceptions correctly less often (by 20% or more) than calculus-based students. These questions are related to the alternate conceptions that moving objects are acted upon by a distinct force in the direction of motion and that constant net force implies constant velocity. There are other common alternate conceptions which occur on only one question (questions 9, 13, 19 and 20) which are held more strongly by algebra-based students than calculus-based students.

We also investigated whether algebra-based students hold different alternate conceptions than calculus-based students. We found that this was not the case both on the pre-test and on the post-test. On all of the questions, the most common alternate conceptions of algebra-based students and calculus-based students were the same; the difference was that on almost all the questions, more algebra-based students than calculus-based students have these common alternate conceptions both on the pre-test and on the post-test.

Previous studies have found that calculus-based students are more adept than algebra-based students at performing identical tasks that are primarily conceptual [8,61,62,64,65]. The present study corroborates this result because the performance of calculus-based students on the FCI, which is a conceptual assessment, is better than the performance of algebra-based students. It is possible that the better mathematical preparation of calculus-based students helps them develop a better conceptual understanding. In particular, while learning physics, one must process information both about the conceptual and mathematical aspects of physics. A student with a better mathematical preparation can use fewer cognitive resources while engaged in problem solving and learning to process the mathematical aspects, and allocate more cognitive resources to the conceptual aspects. Since working memory is finite, the mathematical facility can reduce the cognitive load [66, 67] and provide more opportunities to build a robust knowledge structure of physics. In contrast, a student lacking the requisite mathematical preparation might spend a significant portion of his/her cognitive resources in processing mathematical information, both while engaged in problem solving and while examining problem solutions. This increased cognitive load can hinder reflection and building of good knowledge structure. Therefore, a better mathematical preparation can help improve conceptual understanding of physics; however, more research is needed to understand the connection between mathematical preparation and conceptual understanding.

7.6 CHAPTER REFERENCES

1. D. Hestenes, M. Wells and G. Swackhammer (1992). "Force Concept Inventory." Phys. Teach. 30, 141-158.

2. I. Halloun, R.R. Hake, E.P. Mosca, and D. Hestenes (1995). "Force Concept Inventory." (Revised, 1995); online (password protected) at http://modeling.la.asu.edu/R&E/Research.html and also printed in Mazur (1997).

3. J. Halloun and D. Hestenes (1985) "The initial knowledge state of college physics students." Am. J. Phys. 53, 1043-1055.

4. F. Reif (1974). "Educational challenges for the university." Science 184, 537-542.

5. J. Clement (1982). "Students' preconceptions in introductory mechanics." Am. J. Phys. 50, 66-71.

6. L. C. McDermott (1984). "Research on conceptual understanding in mechanics." Phys. Today 37, 24-32.

7. A. Arons (1983). “Student patterns of thinking and reasoning.” Phys. Teach. 21, 576.

8. R. Thornton and D. Sokoloff (1998). “Assessing student learning of Newton’s laws: The Force and Motion Conceptual Evaluation.” Am. J. Phys. 66(4), 228-351.

9. D. Hestenes and M. Wells (1992). “A Mechanics Baseline Test.” Phys. Teach. 30, 159-166.

10. R. Beichner (1994). “Testing student interpretation of kinematics graphs.” Am. J. Phys. 62(8), 750-762.

11. G. L. Gray, D. Evans, P. J. Cornwell, B. Self and F. Constanzo (2005) “The Dynamics Concept Inventory Assessment Test: A Progress Report.” Proceedings of the 2005 American Society for Engineering Education Annual Conference, Portland, OR.

12. J. Mitchell, J. Martin and T. Newell (2003). “Development of a Concept Inventory for Fluid Mechanics.” Proceedings, Frontiers in Education Conference, Boulder, CO, USA, T3D 23-28.

13. D. Huffman and P. Heller (1995). “What does the Force Concept Inventory actually measure?” Phys. Teach. 33, 138-143.

14. D. Hestenes and I. Halloun (1995). “Interpreting the Force Concept Inventory: A response to march 1995 critique by Huffman and Heller.” Phys. Teach. 33, 502-506.

15. P. Heller and D. Huffman (1995). “Interpreting the Force Concept Inventory: A reply to Hestenes and Halloun.” Phys. Teach. 33, 503-511.

16. I. Halloun and D. Hestenes (2001). “The search for conceptual coherence in FCI data.” Modeling Instruction Workshop website at Arizona State University. http://modeling.asu.edu/R&E/CoherFCI.pdf (14 December 2001).

17. G. A. Morris, L. Branum-Martin, N. Harshman, S. D. Baker, E. Mazur, T. Mzoughi and V. McCauley (2006). “Testing the test: Item response curves and test quality.” Am. J. Phys. 74(5), 449-453.

18. G. A. Morris, N. Harshman, L. Branum-Martin, E. Mazur, T. Mzoughi, and S. D. Baker (2012). “An item response curve analysis of the Force Concept Inventory.” Am. J. Phys. 80(9), 425-431.

19. J. Wang and L. Bao (2010). “Analyzing force concept inventory with item response theory,” Am. J. Phys. 78, 1064–1070.

20. M. Planinic, L. Ivanjek and A. Susac (2010). “Rasch model based analysis of the Force Concept Inventory.” Phys. Rev. ST Phys. Educ. Res. 6, 010103.

21. R. Hake (1998). “Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses.” Am. J. Phys. 66, 64-74.

22. N. Lasry, E. Mazur and J. Watkins (2008). “Peer Instruction: From Harvard to the two-year college.” Am. J. Phys. 76, 1066-1069.

23. A. P. Fagen, C. H. Crouch and E. Mazur (2002). “Peer Instruction: Results from a range of classrooms.” Phys. Teach. 40, 206-209.

24. E. Mazur (1997). Peer Instruction: A User’s Manual (Prentice-Hall, Engelwood Cliffs).

25. D. W. Johnson, R. T. Johnson, and K. A. Smith (1991). Cooperative learning: Increasing college faculty instructional productivity (George Washington U. P., Washington DC).

26. P. Heller, R. Keith, and S. Anderson (1992). “Teaching problem solving through cooperative grouping. 1. Group vs. individual problem solving.” Am. J. Phys. 60, 627–636 (1992).

27. P. Heller and M. Hollabaugh (1992). “Teaching problem solving through cooperative grouping. 2. Designing problems and structuring groups.” Am. J. Phys. 60, 637–644 (1992).

28. I. A. Halloun and D. Hestenes (1987). “Modeling instruction in mechanics.” Am. J. Phys. 55, 455–462.

29. D. Hestenes, “Toward a modeling theory of physics instruction”, Am. J. Phys. 55, 440–454 (1987).

30. M. Wells, D. Hestenes, and G. Swackhamer (1995). “A modeling method for high school physics instruction.” Am. J. Phys. 63, 606–619.

31. R. F. Tinker (1989). “Computer Based Tools: Rhyme and Reason”, Proceedings of the Conference on Computers in Physics Instruction, edited by E. Redish and J. Risley (Addison-Wesley, Reading, MA, 1989) pp. 159–168.

32. R. K. Thornton and D. R. Sokoloff (1990). “Learning motion concepts using real-time microcomputer-based laboratory tools.” Am. J. Phys. 58, 858–867.

33. D. R. Sokoloff, P. W. Laws, and R. K. Thornton (1995). “Real Time Physics, A new interactive introductory lab program.” AAPT Announcer 25(4), 37.

34. A. Van Heuvelen (1991). “Learning to think like a physicist: A review of research-based instructional strategies.” Am. J. Phys. 59, 891.

35. A. Van Heuvelen (1991). Overview, Case Study Physics, Am. J. Phys. 59, 898.

36. R. R. Hake (1987). “Promoting student crossover to the Newtonian world.” Am. J. Phys. 55, 878–884.

37. R. R. Hake (1991). “My conversion to the Arons-advocated method of science education”, Teach. Educ. 3(2), 109–111.

38. J. Docktor and K. Heller (2008). “Gender differences in both Force Concept Inventory and introductory physics performance.” AIP Conference Proceedings 1064, 15-18.

39. S. Bates, R. Donnelly, C. MacPhee1, D. Sands, M. Birch, and R. W. Niels (2013). “Gender differences in conceptual understanding of Newtonian mechanics: a UK cross-institution comparison.” Eur. J. Phys. 34, 421.

40. M. Lorenzo, C. H. Crouch and E. Mazur (2006). “Reducing the gender gap in the physics classroom.” Am. J. Phys. 74, 118-122.

41. S. J. Pollock, N. D. Finkelstein and L. E. Kost (2007). “Reducing the gender gap in the physics classroom: how sufficient is interactive engagement?” Phys. Rev. ST Phys. Educ. Res. 3, 010107.

42. L. McCullough and D. Meltzer (2001). “Differences in male/female response patterns on alternative versions of FCI items.” AIP Conference Proceedings 103-106.

43. L. J. Rennie and L. H. Parker (1998). “Equitable measurement of achievement in physics: high school students’ responses to assessment tasks in different formats and contexts.” J. Women and Minorities in Sci. Eng. 4(2-3), 113-127.

44. L.E. McCullough and T. Foster (2000). “A Gender context for the Force Concept Inventory.” AAPT Announcer 30(4), 105.

45. R. D. Dietz, R. H. Pearson, M. R. Semak, and C. W. Willis (2012). “Gender bias in the force concept inventory?” AIP Conf Proceedings 1413, 171-174.

46. M. H. Dancy, “Investigating animations for assessment with an animated version of the Force Concept Inventory (2000).” Ph.D. dissertation, N.C. State University.

47. V. P. Coletta and J. A. Phillips (2005). “Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability.” Am. J. Phys. 73, 1172.

48. K. L. Malone (2008). “Correlations among knowledge structures, force concept inventory, and problem solving behaviors.” Phys. Rev. ST. Phys. Educ. Res. 4, 020107.

49. P. M. Pamela and J. M. Saul (2006). “Interpreting FCI normalized gain, pre-instruction scores, and scientific reasoning ability.” AAPT Announcer 36, 89.

50. D. Meltzer (2002). “The relationship between mathematics preparation and conceptual learning gains in physics: A possible 'hidden variable'.” Am. J. Phys. 70(12), 1259-1268.

51. V. P. Coletta, J. A. Phillips, and J. A. Steinert (2007) “Interpreting force concept inventory scores: Normalized gain and SAT scores.” Phys. Rev. ST Phys. Educ. Res. 3, 010106.

52. P. Nieminen, A. Savinainen, and J Viiri (2012). “Relations between representational consistency, conceptual understanding of the force concept, and scientific reasoning.” Phys. Rev. ST Physics Educ. Res. 8, 010123.

53. L. S. Shulman (1986). “Those who understand: Knowledge growth in teaching.” Educ. Res. 15(2), 4- 31.

54. L.S. Shulman (1987). “Knowledge and teaching: Foundations of the new reform.” Harv. Educ. Rev. 57(1), 1-22.

55. J. H. van Driel, N. Verloop, and W. de Vos (1998). Developing science teachers’ pedagogical content knowledge.” J. Res. Sci. Teach. 35, 673.

56. P. L. Grossman (1991). “What are we talking about anyhow: Subject matter knowledge for secondary English teachers.” Advances in Research on Teaching, Vol. 2: Subject Matter Knowledge, edited by J. Brophy (JAI Press, Greenwich, CT, 1991), pp. 245–264.

57. J. Gess-Newsome and N. G. Lederman (2001). Examining Pedagogical Content Knowledge, (Kluwer Academic Publishers, Boston, 2001).

58. J. Loughran, P. Mulhall, and A. Berry (2004). “In search of Pedagogical Content Knowledge in science: Developing ways of articulating and documenting professional practice.” J. Res. Sci. Teach. 41, 370.

59. G. V. Glass and K. D. Hopkins (1996). Statistical Methods in Education & Psychology, (Allyn & Bacon, Boston, MA).

60. C. Singh (2002). “Effectiveness of group interaction on conceptual standardized test performance.” Proceedings of the Phys. Ed. Res. Conference, Boise (Eds. S. Franklin, J. Marx, and K. Cummings), 67-70.

61. C. Crouch and E. Mazur (2001). “Peer Instruction: Ten years of experience and results”, Am. J. Phys. 69(9), 970-977.

62. C. Hieggelke, D. Maloney, A. Van Heuvelen, T. O’Kuma (2001) ”Surveying students’ conceptual knowledge of electricity and magnetism.” Am. J. Phys. 69, S12-S23.

63. E. Redish (2005) “Changing Student Ways of Knowing: What should our students learn in physics class?” to be published in Proceedings of the Conference, World View on Physics Education in 2005: Focusing on Change, Delhi (Aug 21-26, 2005).

64. A. Mason and C. Singh (2011). “Assessing expertise in physics using categorization task.” Phys. Rev. ST. Phys. Educ. Res. 7, 020110.

65. N. L. Nguyen, and D. Meltzer (2003). “Initial understanding of vector concepts among students in introductory physics courses.” Am. J. Phys. 71(6), 630-638.

66. J. Sweller (1988). “Cognitive load during problem solving: Effects on learning.” Cog. Sci. 12, 257-285.

67. J. Sweller (1994). “Cognitive load theory, learning difficulty, and instructional design.” Learn. Instruct. 4, 295-312.

8.0 FUTURE OUTLOOK

The studies presented in this thesis can be extended in several different ways. In the studies discussed in chapters two, three and four, students in different recitations were given different instructions regarding diagrams (draw a diagram, given a diagram, no instructions) in each quiz problem. However, the interventions were not matched to particular recitations and the same group of students received different instructions regarding diagrams from week to week. Therefore, we did not expect cumulative effects due to the same group of students being given the same instruction regarding diagrams in each quiz problem. It would be interesting to investigate whether these cumulative effects do arise. For example, the midterm and final exams could have no instruction to draw a diagram (nor provide one) and it is possible that the students who are always asked to draw diagrams in the quizzes end up drawing more diagrams than the other students even when the instruction to draw a diagram is omitted. If that is indeed the case, one could also investigate whether the students asked to draw a diagram in the quizzes exhibit improved performance compared to the other students in midterm and final exams in which the instructions are omitted. If that is not the case and students who are asked to draw diagrams in quizzes do not end up drawing more diagrams than the other students when the instruction is omitted, one could implement one or two more involved interventions than the ones in these studies to investigate the extent of scaffolding needed for positive cumulative effects to arise.

In the study discussed in chapter three, we found that students provided with diagrams that looked very similar to what most experts would initially draw in the conceptual planning stage performed worse than students who were not given those diagrams. This outcome was found to be strong for two problems which involved considerations of initial and final situations. A future study could investigate if the same effect happens in other problems which also involve initial and final situations. If the same outcome is found, the study could investigate in depth why this outcome is strong in this type of problems but not others.

In the study in chapter four, we found that students who drew more detailed diagrams performed better than students who did not. In the future, a more in depth study could investigate which students benefit most from drawing more detailed diagrams, the high, mid or low achieving students. The categories of high, mid and low achieving could be made either by conceptual pre- or post-tests (such as FCI, MCE, CSEM etc.), by performance on exams, or by other considerations.

In the study discussed in chapter five, we found that the additional instructions (higher level of scaffolding compared to the first level of scaffolding), while intended to improve students' representational consistency, had the opposite effect. Interviews suggested that students did not discern the relevance of the additional instructions and to them, completing those instructions was another chore which increased their cognitive load. In a future study, the second level of scaffolding could be improved by adding hints intended to reduce cognitive load, for example, by helping students focus on only a few pieces of information at a time and gradually go through the process of solving the problem. In addition, hints intended to help students perceive the relevance of the additional instructions could be provided. Or, since the first level of scaffolding was found to be beneficial, one could start there and use cognitive task analysis to determine how it could be improved. Then, the extent to which the newly developed level of scaffolding is beneficial in improving students' representational consistency could be investigated. In addition, the study in chapter five found evidence (although not very strong) that the effects of the first level of scaffolding which was used only once in one quiz stuck with some students for quite some time because they exhibited somewhat improved representational consistency in the same problem in the final exam (in the multiple-choice format). It is possible that if this intervention (scaffolding level 1) is used multiple times during the course, the effects of this intervention could be much stronger in the long term. Therefore, a future study could implement this intervention multiple times during the semester and investigate whether long term effects become stronger (in the final exam, or perhaps several months after the course).

In the study in chapter six, TAs were asked to identify the most common student difficulty on each item on the TUG-K. This study could be extended to include instructors who have taught introductory physics courses recently and instructors who have not. Then, it could be investigated if similar results to the study in chapter seven occur (experience teaching independent courses does not correlate with better performance on the task and neither does recent teaching experience, instructors do not perform better than TAs at identifying common student difficulties, etc.)

In the study in chapter seven, instructors and TAs were asked to identify the most common student difficulties. However, this task did not provide information about how knowledgeable instructors are about the difficulty of the questions from the point of view of the students. Instructors and TAs could be asked to also predict the percentage of students who answer each question correctly, in addition to being asked to identify the most common incorrect answer choice. This would make the task more challenging, but it would provide richer data.

However, many questions include more than one common incorrect answer choice. Therefore, an even more difficult problem solving task could be given to instructors and TAs: predict the percentages of introductory students who would select *each* answer choice, *including the correct one*. This would be a more in-depth way to investigate the knowledge instructors and TAs have of student difficulties (the study in chapter six could be extended in this manner as well). However, even in the study in chapter seven in which instructors and TAs were asked to select only the most common incorrect answer choice, it took some instructors a long time (several weeks) to complete the task and many were sent reminders repeatedly. Comments from instructors indicate that many of them found this task challenging, which might partly account for the slow response. We suspect that if the task is modified to ask instructors to predict the percentage of introductory physics students who would select each answer choice, this task will be extremely challenging for many instructors which may drastically reduce their response rate.

APPENDIX A

MATHEMATICAL DESCRIPTION OF THE CALCULATION OF THE TUG-K RELATED PCK SCORES AND DATA TABLES FOR THE TUG-K STUDY (CHAPTER SIX)

Mathematical description of how the TUG-K related PCK scores were calculated

We define indices *i*, *j* and *k* that correspond to the following:

- *i*: index of graduate student (25 graduate students; it takes values from 1 to 25);
- *j*: TUG-K question number (21 questions; it takes values from 1 to 21);
- *k*: incorrect answer choice number (4 incorrect answer choices; it takes values from 1 to 4).

Then, let F_{jk} be the fraction of introductory physics students who select incorrect answer choice *k* on item *j* (e.g. F_{11} = 0.4, F_{12} = 0.04, F_{13} = 0.22, F_{14} = 0.17). Let GS_{ijk} correspond to whether graduate student *i* selected incorrect answer choice *k* on item *j* (for a given graduate student *i* and TUG-K item *j*, GS_{ijk}=1 only for the incorrect answer choice *k*, selected by graduate student *i* on item *j*, otherwise GS_{ijk}=0). Then, the PCK score of the *i*th graduate student on item *j* (referred to GS_{ij}) is: $GS_{ij} = \sum_{k=1}^{4}(GS_{ijk} * F_{jk})$. Then, the PCK score of the *i*th graduate student on the whole survey (GS_i) can be obtained by summing over all the questions:

$$GS_i = \sum_{j=1}^{21} GS_{ij} = \sum_{j=1}^{21}\left[\sum_{k=1}^{4}(GS_{ijk} * F_{jk})\right].$$

Also, the average score of all the graduate students on item *j* (referred to as $\overline{GS_J}$) is:

$$\overline{GS_J} = \frac{1}{25}\sum_{i=1}^{25} GS_{ij} = \frac{1}{25}\sum_{i=1}^{25}\left[\sum_{k=1}^{4}(GS_{ijk} * F_{jk})\right].$$

A similar approach can be adopted for random guessing:

- RG_{ij} = PCK score of *i*th random guesser on item *j*;
- RG_i = PCK score of *i*th random guesser;
- $\overline{RG_J}$ = PCK score of random guessing on item *j*).

The PCK score of each graduate student and random guesser (GS_i, RG_i as described above) were used to obtain averages and standard deviations in order to perform *t*-tests to compare the performance of graduate students with random guessing on the whole survey (and to compare different subgroups of graduate students).

In order to compare the performance of these different groups on individual items, the averages and standard deviations of the PCK scores on that particular item (e.g., for question *j* on the TUG-K: GS_{ij}, RG_{ij}) were used to perform *t*-tests.

Table A1. Questions on the TUG-K for which at least 20% of introductory students selected on incorrect answer choice in a post-test, percentages of introductory physics students (Intro. stud. choices) who selected each answer choice A through E in a post-test (they were asked to select the correct answer for each question) and graduate students (Grad student choices) who selected each answer choice A through E (they were asked to select the most common incorrect answer for each question if introductory students did not know the correct answer). The first column of the table lists the TUG-K question numbers and the second column titled >RG indicates whether the graduate students on average performed better than random guessing.

TUG-K item #	>RG	Intro stud. choices						Grad student choices				
		A	B	C	D	E		A	B	C	D	E
1	Yes	41	16	4	22	17	1	36	0	0	60	4
2	Yes	2	10	24	2	63	2	0	40	52	4	4
3	Yes	8	0	20	62	10	3	24	0	72	0	4
4	Yes	2	14	23	28	32	4	0	16	40	0	44
6	No	46	26	6	6	16	6	20	4	20	20	36
7	No	31	20	10	28	10	7	0	28	28	36	8
8	No	11	11	37	37	5	8	40	40	8	4	8
9	No	7	57	5	7	24	9	40	28	16	12	4
10	Yes	30	2	62	3	3	10	12	4	56	28	0
11	No	28	17	11	36	8	11	8	64	8	8	12
14	No	25	48	15	9	3	14	16	0	28	56	0
15	No	29	24	13	8	26	15	8	8	16	16	52
16	Yes	1	39	31	22	7	16	4	16	68	4	8
17	No	21	46	8	7	19	17	4	16	16	20	44
18	Yes	7	46	32	4	10	18	17	4	58	0	21
19	No	19	9	37	12	23	19	21	13	4	13	50
21	Yes	18	73	2	5	2	21	4	79	8	8	0

x Correct answer
x x > 33 – major difficulty (more than 1/3 of introductory students chose it)
x 20 ≤ x ≤ 33 – moderate difficulty

Table A2. Questions on the TUG-K on which at least 20% of introductory students selected one incorrect answer choice after instruction, percentages of introductory students who answered each question correctly (% intro. correct), minimum possible TUG-K related PCK score (Min. pos. PCK score), maximum possible TUG-K related PCK score (Max. pos. PCK score), graduate students' average PCK score (GS avg. PCK score), graduate students' normalized average PCK score on a scale from 0 to 100 (Norm. GS avg. PCK score).

TUG-K item #	% intro. correct	Min. pos. PCK score	Max. pos. PCK score	GS avg. PCK score	Norm. GS avg. PCK score
1	16	0.04	0.41	0.29	68
2	63	0.02	0.24	0.17	68
3	62	0.00	0.20	0.17	85
4	28	0.02	0.32	0.26	80
6	26	0.06	0.46	0.17	28
7	31	0.10	0.28	0.19	50
8	37	0.05	0.37	0.12	22
9	24	0.05	0.57	0.20	29
10	30	0.02	0.62	0.36	57
11	36	0.08	0.28	0.15	35
14	48	0.03	0.25	0.13	45
15	29	0.08	0.26	0.19	61
16	22	0.01	0.39	0.28	71
17	21	0.07	0.46	0.18	28
18	46	0.04	0.32	0.22	64
19	37	0.09	0.23	0.18	64
21	18	0.02	0.73	0.58	79

#	Question in which there was a moderate difficulty
#	Question in which there was a major difficulty
x	Grad students' TUG-K related PCK score is less than 1/2 of maximum possible
x	Grad students' TUK-K related PCK score is between 1/2 and 2/3 of maximum possible
x	Grad students' TUG-K related PCK score is more than 2/3 of maximum possible

APPENDIX B

MATHEMATICAL DESCRIPTION OF THE CALCULATION OF THE FCI RELATED PCK SCORES, DATA TABLES AND COPY OF THE FCI FOR THE FCI STUDY (CHAPTER SEVEN)

Mathematical description of how the FCI related PCK scores were calculated

We define indices *i*, *j* and *k* that correspond to the following:

- *i*: index of instructor (30 instructors; it takes values from 1 to 30);
- *j*: FCI question number (30 questions; it takes values from 1 to 30);
- *k*: incorrect answer choice number for each question (4 incorrect answer choices; it takes values from 1 to 4).

Then, we let F_{jk} be the fraction of introductory physics students who selected incorrect answer choice *k* on item *j* (e.g. $F_{21} = 0.44$, $F_{22} = 0.06$, $F_{23} = 0.21$, $F_{24} = 0.04$). We let I_{ijk} correspond to whether instructor *i* chose incorrect answer choice *k* on item *j* (for a given *i* and *j*, I_{ijk}=1 only for the incorrect answer choice *k*, selected by instructor *i* on item *j*, otherwise I_{ijk}=0). Then, the PCK score of the *i*th instructor on item *j* (referred to I_{ij}) is: $I_{ij} = \sum_{k=1}^{4}\left(I_{ijk} \cdot F_{jk}\right)$. Then, the total PCK score of the *i*th instructor (I_i) on the whole survey can be obtained by summing over all of the questions:

$$I_i = \sum_{j=1}^{30} I_{ij} = \sum_{j=1}^{30} \left[\sum_{k=1}^{4} (I_{ijk} * F_{jk}) \right].$$

Also, the PCK score of all of the instructors on item j (referred to as $\overline{I_J}$) can be obtained by summing over the instructors:

$$\overline{I_J} = \sum_{i=1}^{30} I_{ij} = \sum_{i=1}^{30} \left[\sum_{k=1}^{4} (I_{ijk} * F_{jk}) \right].$$

A similar approach can also be adopted for the graduate students (GS_{ij} – PCK score of the ith graduate student on item j; GS_i – PCK score of the ith graduate student on the whole survey; $\overline{GS_J}$ – PCK score of all graduate students on item j) and for random guessers (RG_{ij} – PCK score of ith random guesser on item j; RG_i – PCK score of ith random guesser; $\overline{RG_J}$ – PCK score of random guessers on item j). The PCK scores of each (i) instructor/graduate student/random guesser (I_i, GS_i, RG_i as described above) were used to obtain averages and standard deviations in order to perform *t*-tests to compare the FCI related PCK performance of instructors with that of the graduate students and random guessers on the whole survey (and to compare different subgroups of instructors and graduate students). In order to compare the PCK performance of these different groups on individual items, the averages and standard deviations of the PCK scores on that particular question (e.g., for question j on the FCI: I_{ij}, GS_{ij}, RG_{ij}) were used to perform *t*-tests.

(Note, Tables B3 and B4 also list normalized gains for each question which are defined as (%correct post - %correct pre)/(100% - %correct pre).

Table B1. Questions on the FCI in which at least 19% of introductory algebra-based students selected one incorrect answer choice in a post-test, percentages of introductory algebra-based physics students who selected each answer choice A through E in a post-test (they were asked to select the correct answer for each question), instructors and graduate students who selected each answer choice A through E (they were asked to select the most common incorrect answer for each question if introductory physics students did not know the correct answer). The first column of the table lists the FCI question numbers and the second column titled > RG shows an "I" when the instructors on average performed better than random guessing, "GS" when the graduate students on average performed better than random guessing; and "I, GS" when both instructors and graduate students performed better than random guessing.

FCI #	>RG	Intro student choices					Instructor choices					Grad student choices				
		A	B	C	D	E	A	B	C	D	E	A	B	C	D	E
2	I, GS	44	25	6	21	4	13	53	7	27	0	4	68	0	24	4
4	I, GS	39	1	0	0	60	97	0	0	3	0	84	0	4	12	0
5		3	24	12	44	17	20	0	27	30	23	0	0	32	40	28
9	I, GS	4	20	19	5	53	0	40	54	3	3	12	8	76	0	4
11	I, GS	5	8	48	35	4	17	23	40	0	20	20	20	60	0	0
12	I, GS	1	77	19	2	1	10	3	64	3	20	28	0	68	0	4
13	I	4	11	50	35	0	20	30	47	0	3	36	44	16	0	4
14		19	10	9	61	0	17	57	23	0	3	20	44	36	0	0
15	I	44	7	48	1	0	10	7	60	20	0	12	20	40	28	0
16		73	2	19	2	4	3	17	37	23	20	8	12	16	28	36
17	I, GS	62	24	1	10	3	87	3	0	3	7	72	0	0	16	12
18	GS	1	28	4	42	25	16	0	47	30	7	4	0	16	52	28
19	GS	12	3	5	29	51	24	14	21	38	3	8	4	12	76	0
20	I, GS	16	4	27	49	4	7	3	72	3	14	40	4	56	0	0
21	I	7	13	38	9	33	7	40	43	7	3	0	20	44	36	0
22	I	33	45	3	16	2	67	0	7	26	0	28	24	4	40	4
23		15	39	18	23	5	20	7	47	23	3	20	4	28	24	24
24	I	70	2	22	2	5	7	3	70	0	20	0	16	68	12	4
25	I	3	9	23	53	12	10	17	0	57	16	24	24	0	44	8
26	I	41	32	3	9	14	60	27	0	13	0	52	32	0	12	4
27	I, GS	26	13	58	2	0	63	27	0	7	3	68	16	4	4	8
28	GS	1	2	6	32	59	7	45	10	38	0	4	36	8	52	0
29		2	71	3	23	1	24	3	0	28	45	20	8	0	28	44
30	I, GS	3	10	26	2	59	3	0	0	3	94	4	4	0	4	88

x Correct answer

x x > 33 – strong alt conception (more than 1/3 of intro students chose it)

x 19 ≤ x ≤ 33 – medium level alternate conception

Table B2. Questions on the FCI in which at least 19% of introductory algebra-based students selected one incorrect answer choice in a post-test, percentages of introductory algebra-based students who answer each question correctly (% intro. alg. correct), normalized gain (Intro. alg. norm. gain), maximum possible FCI related PCK score (max. pos. PCK score), Average FCI related PCK score of instructors (Ins. avg. PCK score) and graduate students (GS avg. PCK score), percentage of maximum possible score of the instructors' average FCI related PCK score (Ins. % max score) and of the graduate students average FCI related PCK score (GS % max score).

FCI item #	% intro. alg. correct	Intro. alg. norm. gain	Min. pos. PCK score	Max. pos. PCK score	Instructors		Graduate students	
					Ins. avg. PCK score	Norm Ins. avg. PCK score	GS avg. PCK score	Norm GS avg. PCK score
2	44	0.23	0.04	0.25	0.19	71	0.22	86
4	60	0.46	0	0.39	0.38	97	0.33	85
5	24	0.16	0.03	0.44	0.21	~~44~~	0.26	56
9	53	0.17	0.04	0.20	0.18	88	0.17	81
11	35	0.24	0.04	0.48	0.23	~~43~~	0.31	61
12	77	0.43	0.01	0.19	0.12	61	0.13	67
13	35	0.27	0	0.50	0.27	54	0.14	~~28~~
14	61	0.39	0	0.19	0.11	58	0.11	58
15	44	0.25	0	0.48	0.30	63	0.21	~~44~~
16	73	0.40	0.02	0.19	0.09	~~41~~	0.05	~~18~~
17	24	0.17	0.01	0.62	0.54	87	0.47	75
18	28	0.19	0.01	0.42	0.16	~~37~~	0.30	71
19	51	-0.06	0.03	0.29	0.15	~~46~~	0.24	81
20	49	0.25	0.04	0.27	0.21	74	0.22	78
21	33	-0.02	0.07	0.38	0.23	52	0.23	52
22	45	0.22	0.02	0.33	0.26	77	0.16	~~45~~
23	39	0.14	0.05	0.23	0.17	67	0.15	56
24	70	0.20	0.02	0.22	0.16	70	0.16	70
25	23	0.13	0.03	0.53	0.34	62	0.27	~~48~~
26	14	0.11	0.03	0.41	0.34	82	0.33	79
27	58	0.09	0	0.26	0.20	77	0.20	77
28	59	0.6	0.01	0.32	0.14	~~42~~	0.18	55
29	71	0.50	0.02	0.23	0.07	~~24~~	0.07	~~24~~
30	26	0.16	0.02	0.59	0.55	93	0.53	89

#	Question in which there was a medium level alternate conception
#	Question in which there was a strong student alternate conception
~~x~~	Ins./grad students' FCI related PCK score is less than 50% of maximum possible
x	Ins./GS score FCI rel. PCK score is between 50% and 67% of maximum possible
x	Ins./GS FCI related PCK score is more than 67% of maximum possible

Table B3. Percentages of algebra-based introductory physics students who selected each answer choice for each item on the FCI when it was given in a pre-test and in a post-test and normalized gain (Norm. gain) on each item on the FCI. The percentages on the pre-test are based on data from 601 students taught by two different instructors in two different semesters and the percentages on the post-test are based on data from 899 students taught by 4 different instructors over several years. The green shaded boxes indicate correct answers. All the courses were taught in a traditional manner which did not incorporate PER based teaching strategies.

FCI	Pre-test algebra					Post-test algebra					Norm.
item #	A	B	C	D	E	A	B	C	D	E	gain
1	13	6	53	25	4	10	4	78	8	1	0.53
2	27	21	7	37	8	44	25	6	21	4	0.23
3	31	16	40	3	10	15	13	63	5	5	0.38
4	73	0	0	1	26	39	1	0	0	60	0.46
5	4	10	31	25	29	3	24	12	44	17	0.16
6	25	68	5	2	0	16	79	3	1	1	0.33
7	17	57	9	5	11	12	74	6	3	4	0.40
8	20	47	1	13	18	14	66	0	8	11	0.36
9	4	26	20	6	43	4	20	19	5	53	0.17
10	54	1	11	19	15	71	2	7	13	7	0.36
11	7	31	45	14	3	5	8	48	35	4	0.24
12	1	59	32	5	2	1	77	19	2	1	0.43
13	4	21	64	12	0	4	11	50	35	0	0.27
14	35	18	11	36	0	19	10	9	61	0	0.39
15	25	10	61	3	0	44	7	48	1	0	0.25
16	55	3	37	4	1	73	2	19	2	4	0.40
17	60	8	1	22	9	62	24	1	10	3	0.17
18	2	12	14	27	46	1	28	4	42	25	0.19
19	14	3	3	26	54	12	3	5	29	51	-0.06
20	19	6	36	32	8	16	4	27	49	4	0.25
21	7	21	23	14	35	7	13	38	9	33	-0.02
22	37	30	4	26	2	33	45	3	16	2	0.22
23	16	29	21	28	7	15	39	18	23	5	0.14
24	63	1	25	3	6	70	2	22	2	5	0.20
25	3	8	12	58	19	3	9	23	53	12	0.13
26	41	42	4	11	3	41	32	3	9	14	0.11
27	31	13	54	1	0	26	13	58	2	0	0.09
28	0	6	8	61	24	1	2	6	32	59	0.46
29	15	42	1	37	4	2	71	3	23	1	0.50
30	1	7	12	1	79	3	10	26	2	59	0.16
						Avg. normalized gain					0.26

x Correct answer

x x > 33 – strong alternate conception (more than 1/3 of intro students chose it)

x 19 ≤ x ≤ 33 – medium level alternate conception.

Table B4. Percentages of calculus-based introductory physics students who selected each answer choice for each item on the FCI when it was given in a pre-test and in a post-test. The percentages on the pre-test are based on data from 364 students taught by three different instructors over several semesters and the percentages on the post-test are based on data from 296 students taught by two different instructors during two different semesters. The green shaded boxes indicate correct answers. All the courses were taught in a traditional manner which did not incorporate PER based teaching strategies.

FCI	Pre-test calculus					Post-test calculus					Norm.
item #	A	B	C	D	E	A	B	C	D	E	gain
1	8	3	82	6	1	7	1	86	6	1	0.24
2	39	24	4	26	7	55	27	4	15	1	0.25
3	17	8	66	4	5	15	4	76	3	2	0.30
4	59	1	0	0	39	23	0	1	0	76	0.60
5	5	23	20	27	25	3	47	13	29	9	0.31
6	17	79	3	1	1	12	86	2	1	0	0.32
7	9	76	4	3	7	7	85	3	1	5	0.39
8	11	62	0	12	17	8	74	0	8	10	0.33
9	2	20	20	6	53	1	9	14	3	73	0.44
10	70	3	6	12	8	91	1	3	3	3	0.69
11	8	16	42	30	4	6	2	28	61	3	0.44
12	0	84	14	2	0	0	90	10	0	0	0.41
13	4	12	51	33	0	3	5	31	61	0	0.42
14	21	12	5	63	0	14	9	5	72	0	0.25
15	29	6	63	2	0	42	5	52	0	1	0.19
16	72	1	23	2	3	86	2	7	1	4	0.50
17	63	21	1	12	4	46	44	2	6	2	0.29
18	3	27	9	34	29	0	55	2	33	11	0.38
19	11	3	3	20	62	8	2	3	12	75	0.35
20	11	4	22	54	6	9	2	16	71	3	0.36
21	6	11	26	11	46	4	6	23	11	57	0.20
22	27	51	2	19	1	20	67	2	11	1	0.34
23	9	48	16	24	3	5	64	13	14	3	0.32
24	74	1	18	2	4	87	1	9	1	3	0.48
25	2	9	29	51	10	3	6	51	38	2	0.31
26	35	29	1	14	20	20	28	3	7	42	0.28
27	22	7	66	4	1	14	6	75	3	1	0.28
28	2	4	6	45	44	1	3	3	20	72	0.51
29	6	61	2	27	3	1	83	1	13	2	0.56
30	3	7	30	1	59	3	7	44	1	45	0.20
						Avg. normalized gain					0.36

x Correct answer

x x > 33 – strong alternate conception (more than 1/3 of intro students chose it)

x 19 ≤ x ≤ 33 – medium level alternate conception

Force Concept Inventory

Please:

Do not write anything on this questionnaire.
Mark your answers on the scantron sheet.
Make only one mark per item.
Do not skip any question.
Avoid guessing. Your answers should reflect what you personally think.

On the scantron sheet:

Use a No. 2 pencil only, and follow marking instructions.
Fill in your ID number.

Plan to finish this questionnaire in 30 minutes.

Thank you for your cooperation.

1. Two metal balls are the same size but one weighs twice as much as the other. The balls are dropped from the roof of a single story building at the same instant of time. The time it takes the balls to reach the ground below will be:
   (A) about half as long for the heavier ball as for the lighter one.
   (B) about half as long for the lighter ball as for the heavier one.
   (C) about the same for both balls.
   (D) considerably less for the heavier ball, but not necessarily half as long.
   (E) considerably less for the lighter ball, but not necessarily half as long.

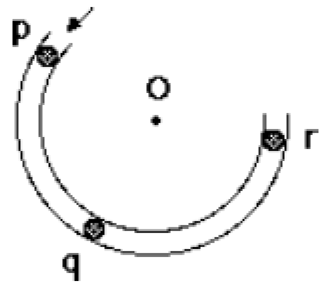
2. The two metal balls of the previous problem roll off a horizontal table with the same speed. In this situation:
   (A) both balls hit the floor at approximately the same horizontal distance from the base of the table.
   (B) the heavier ball hits the floor at about half the horizontal distance from the base of the table than does the lighter ball.
   (C) the lighter ball hits the floor at about half the horizontal distance from the base of the table than does the heavier ball.
   (D) the heavier ball hits the floor considerably closer to the base of the table than the lighter ball, but not necessarily at half the horizontal distance.
   (E) the lighter ball hits the floor considerably closer to the base of the table than the heavier ball, but not necessarily at half the horizontal distance.

3. A stone dropped from the roof of a single story building to the surface of the earth:
   (A) reaches a maximum speed quite soon after release and then falls at a constant speed thereafter.
   (B) speeds up as it falls because the gravitational attraction gets considerably stronger as the stone gets closer to the earth.
   (C) speeds up because of an almost constant force of gravity acting upon it.
   (D) falls because of the natural tendency of all objects to rest on the surface of the earth.
   (E) falls because of the combined effects of the force of gravity pushing it downward and the force of the air pushing it downward.

4. A large truck collides head-on with a small compact car. During the collision:
   (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 gets in the way of the truck.
   (D) the truck exerts a force on the car but the car does not 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.

USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT TWO QUESTIONS (5 and 6).

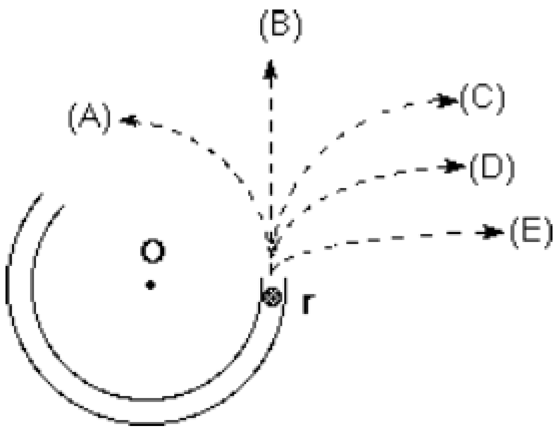
The accompanying figure shows a frictionless channel in the shape of a segment of a circle with center at "O". The channel has been anchored to a frictionless horizontal table top. You are looking down at the table. Forces exerted by the air are negligible. A ball is shot at high speed into the channel at "p" and exits at "r."



5. Consider the following distinct forces:
   1. A downward force of gravity.
   2. A force exerted by the channel pointing from q to O.
   3. A force in the direction of motion.
   4. A force pointing from O to q.

   Which of the above forces is (are) acting on the ball when it is within the frictionless channel at position "q"?

   (A) 1 only.
   (B) 1 and 2.
   (C) 1 and 3.
   (D) 1, 2, and 3.
   (E) 1, 3, and 4.

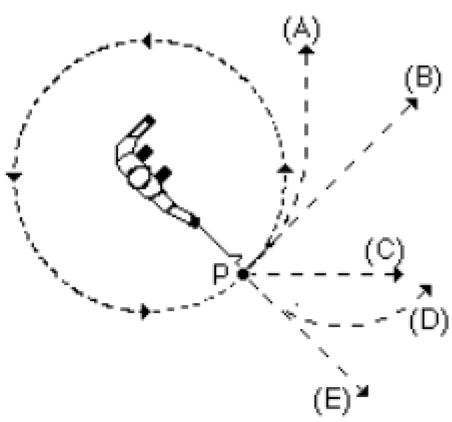


6. Which path in the figure at right would the ball most closely follow after it exits the channel at "r" and moves across the frictionless table top?

7. A steel ball is attached to a string and is swung in a circular path in a horizontal plane as illustrated in the accompanying figure.

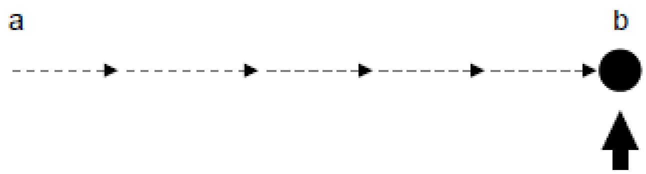
   At the point P indicated in the figure, the string suddenly breaks near the ball.

   If these events are observed from directly above as in the figure, which path would the ball most closely follow after the string breaks?

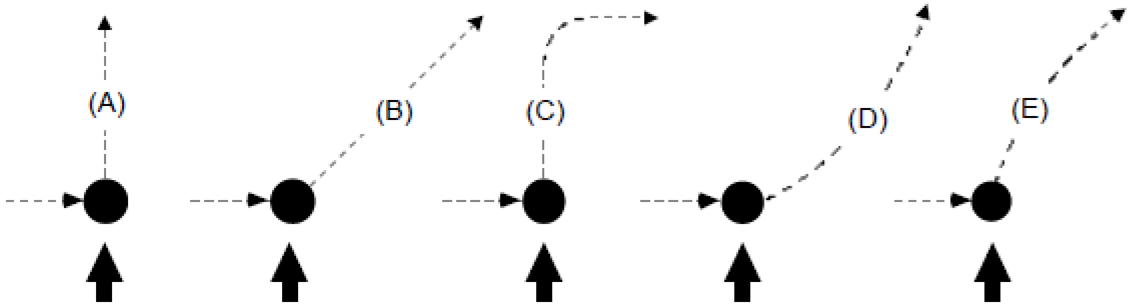


USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT FOUR QUESTIONS (8 through 11).

The figure depicts a hockey puck sliding with constant speed v_o in a straight line from point "a" to point "b" on a frictionless horizontal surface. Forces exerted by the air are negligible. You are looking down on the puck. When the puck reaches point "b," it receives a swift horizontal kick in the direction of the heavy print arrow. Had the puck been at rest at point "b," then the kick would have set the puck in horizontal motion with a speed v_k in the direction of the kick.



8. Which of the paths below would the puck most closely follow after receiving the kick ?

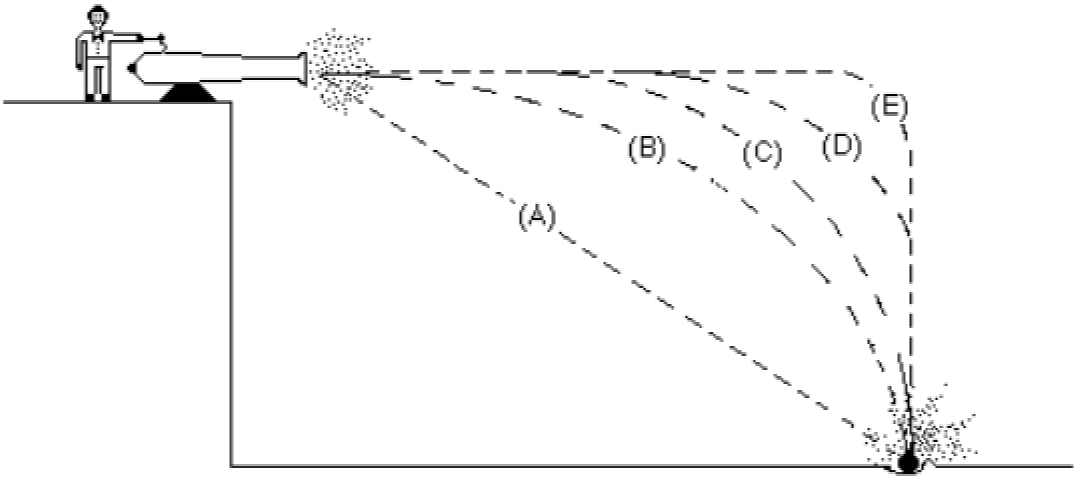


9. The speed of the puck just after it receives the kick is:
   (A) equal to the speed "v_o" it had before it received the kick.
   (B) equal to the speed "v_k" resulting from the kick and independent of the speed "v_o".
   (C) equal to the arithmetic sum of the speeds "v_o" and "v_k".
   (D) smaller than either of the speeds "v_o" or "v_k".
   (E) greater than either of the speeds "v_o" or "v_k", but less than the arithmetic sum of these two speeds.

10. Along the frictionless path you have chosen in question 8, the speed of the puck after receiving the kick:
   (A) is constant.
   (B) continuously increases.
   (C) continuously decreases.
   (D) increases for a while and decreases thereafter.
   (E) is constant for a while and decreases thereafter.

11. Along the frictionless path you have chosen in question 8, the main force(s) acting on the puck after receiving the kick is (are):
   (A) a downward force of gravity.
   (B) a downward force of gravity, and a horizontal force in the direction of motion.
   (C) a downward force of gravity, an upward force exerted by the surface, and a horizontal force in the direction of motion.
   (D) a downward force of gravity and an upward force exerted by the surface.
   (E) none. (No forces act on the puck.)

12. A ball is fired by a cannon from the top of a cliff as shown in the figure below. Which of the paths would the cannon ball most closely follow?



13. A boy throws a steel ball straight up. Consider the motion of the ball only after it has left the boy's hand but before it touches the ground, and assume that forces exerted by the air are negligible. For these conditions, the force(s) acting on the ball is (are):

(A) a downward force of gravity along with a steadily decreasing upward force.

(B) a steadily decreasing upward force from the moment it leaves the boy's hand until it reaches its highest point; on the way down there is a steadily increasing downward force of gravity as the object gets closer to the earth.

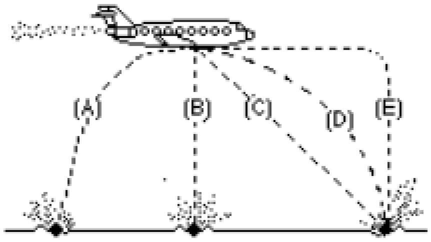
(C) an almost constant downward force of gravity along with an upward force that steadily decreases until the ball reaches its highest point; on the way down there is only a constant downward force of gravity.

(D) an almost constant downward force of gravity only.

(E) none of the above. The ball falls back to ground because of its natural tendency to rest on the surface of the earth.

14. A bowling ball accidentally falls out of the cargo bay of an airliner as it flies along in a horizontal direction.

As observed by a person standing on the ground and viewing the plane as in the figure at right, which path would the bowling ball most closely follow after leaving the airplane?



USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT TWO QUESTIONS (15 and 16).

A large truck breaks down out on the road and receives a push back into town by a small compact car as shown in the figure below.

15. While the car, still pushing the truck, is speeding up to get up to cruising speed:
    (A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
    (B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
    (C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
    (D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
    (E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

16. After the car reaches the constant cruising speed at which its driver wishes to push the truck:
    (A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
    (B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
    (C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
    (D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
    (E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

17. An elevator is being lifted up an elevator shaft at a constant speed by a steel cable as shown in the figure below. All frictional effects are negligible. In this situation, forces on the elevator are such that:

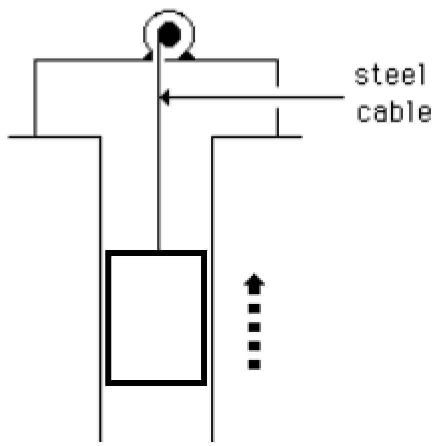
(A) the upward force by the cable is greater than the downward force of gravity.

(B) the upward force by the cable is equal to the downward force of gravity.

(C) the upward force by the cable is smaller than the downward force of gravity.

(D) the upward force by the cable is greater than the sum of the downward force of gravity and a downward force due to the air.

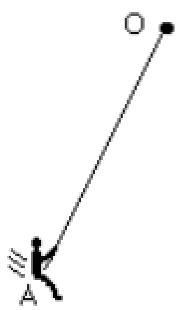
(E) none of the above. (The elevator goes up because the cable is being shortened, not because an upward force is exerted on the elevator by the cable).



Elevator going up
at constant speed

18. The figure below shows a boy swinging on a rope, starting at a point higher than A. Consider the following distinct forces:

1. A downward force of gravity.
2. A force exerted by the rope pointing from A to O.
3. A force in the direction of the boy's motion.
4. A force pointing from O to A.



Which of the above forces is (are) acting on the boy when he is at position A?

(A) 1 only.

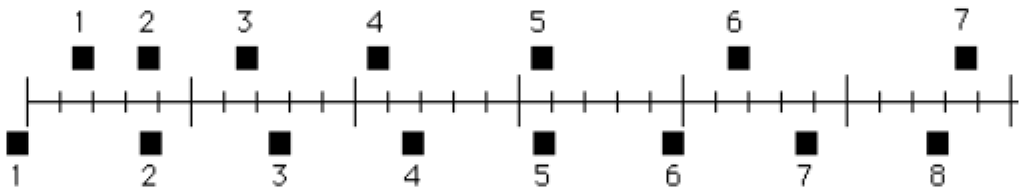
(B) 1 and 2.

(C) 1 and 3.

(D) 1, 2, and 3.

(E) 1, 3, and 4.

19. The positions of two blocks at successive 0.20-second time intervals are represented by the numbered squares in the figure below. The blocks are moving toward the right.



Do the blocks ever have the same speed?

(A) No.

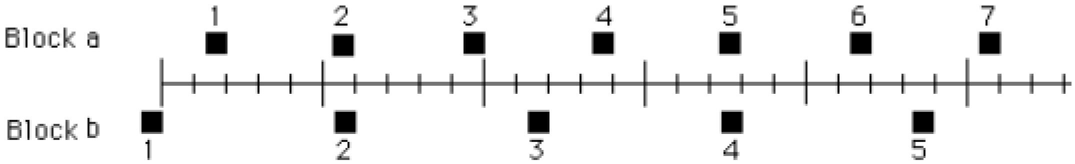
(B) Yes, at instant 2.

(C) Yes, at instant 5.

(D) Yes, at instants 2 and 5.

(E) Yes, at some time during the interval 3 to 4.

20. The positions of two blocks at successive 0.20-second time intervals are represented by the numbered squares in the figure below. The blocks are moving toward the right.



The accelerations of the blocks are related as follows:

(A) The acceleration of "a" is greater than the acceleration of " b".

(B) The acceleration of "a" equals the acceleration of "b". Both accelerations are greater than zero.

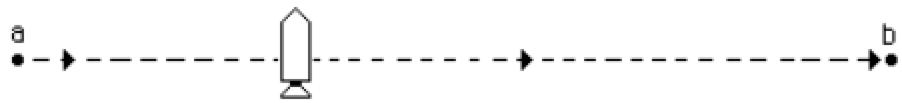
(C) The acceleration of "b" is greater than the acceleration of "a".

(D) The acceleration of "a" equals the acceleration of "b". Both accelerations are zero.

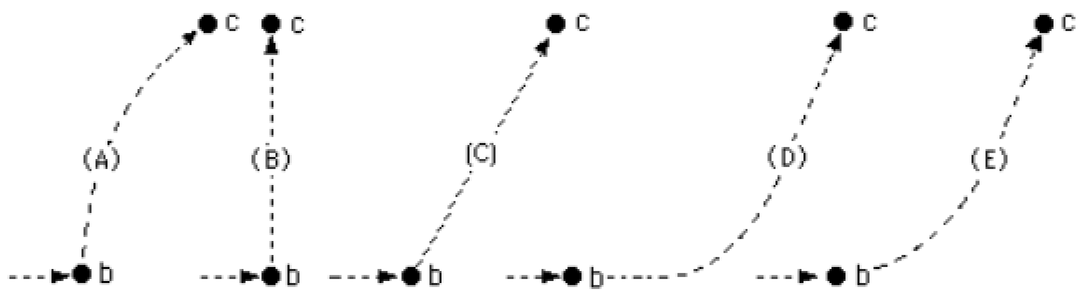
(E) Not enough information is given to answer the question.

USE THE STATEMENT AND FIGURE BELOW TO ANSWER THE NEXT FOUR QUESTIONS (21 through 24).

A rocket drifts sideways in outer space from point "a" to point "b" as shown below. The rocket is subject to no outside forces. Starting at position "b", the rocket's engine is turned on and produces a constant thrust (force on the rocket) at right angles to the line "ab". The constant thrust is maintained until the rocket reaches a point "c" in space.

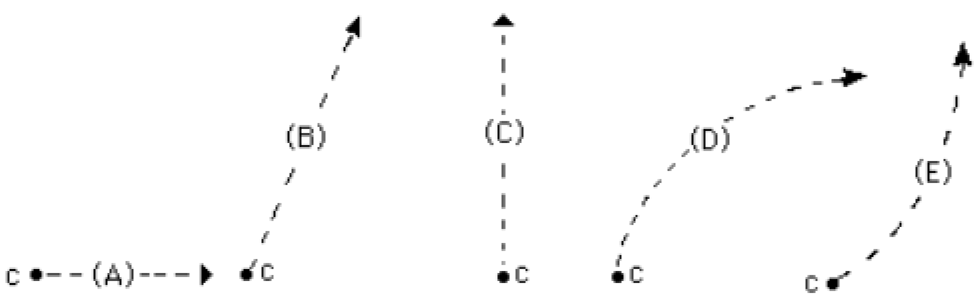


21. Which of the paths below best represents the path of the rocket between points "b" and "c"?



22. As the rocket moves from position"b" to position "c" its speed is:
    (A) constant.
    (B) continuously increasing.
    (C) continuously decreasing.
    (D) increasing for a while and constant thereafter.
    (E) constant for a while and decreasing thereafter.

23. At point "c" the rocket's engine is turned off and the thrust immediately drops to zero. Which of the paths below will the rocket follow beyond point "c"?



24. Beyond position "c" the speed of the rocket is:
    (A) constant.
    (B) continuously increasing.
    (C) continuously decreasing.
    (D) increasing for a while and constant thereafter.
    (E) constant for a while and decreasing thereafter.

25. A woman exerts a constant horizontal force on a large box. As a result, the box moves across a horizontal floor at a constant speed "v_o".

The constant horizontal force applied by the woman:

(A) has the same magnitude as the weight of the box.
(B) is greater than the weight of the box.
(C) has the same magnitude as the total force which resists the motion of the box.
(D) is greater than the total force which resists the motion of the box.
(E) is greater than either the weight of the box or the total force which resists its motion.

26. If the woman in the previous question doubles the constant horizontal force that she exerts on the box to push it on the same horizontal floor, the box then moves:

(A) with a constant speed that is double the speed "v_o" in the previous question.
(B) with a constant speed that is greater than the speed "v_o" in the previous question, but not necessarily twice as great.
(C) for a while with a speed that is constant and greater than the speed "v_o" in the previous question, then with a speed that increases thereafter.
(D) for a while with an increasing speed, then with a constant speed thereafter.
(E) with a continuously increasing speed.

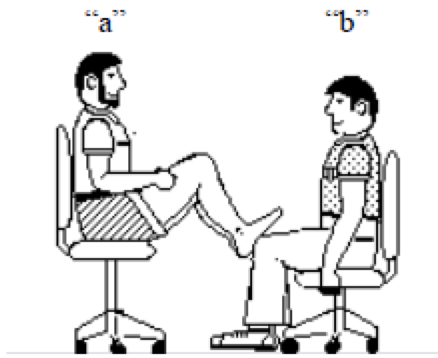
27. If the woman in question 25 suddenly stops applying a horizontal force to the box, then the box will:

(A) immediately come to a stop.
(B) continue moving at a constant speed for a while and then slow to a stop.
(C) immediately start slowing to a stop.
(D) continue at a constant speed.
(E) increase its speed for a while and then start slowing to a stop.

28. In the figure at right, student "a" has a mass of 95 kg and student "b" has a mass of 77 kg. They sit in identical office chairs facing each other.

Student "a" places his bare feet on the knees of student "b", as shown. Student "a" then suddenly pushes outward with his feet, causing both chairs to move.

During the push and while the students are still touching one another:



(A) neither student exerts a force on the other.

(B) student "a" exerts a force on student "b", but "b" does not exert any force on "a".

(C) each student exerts a force on the other, but "b" exerts the larger force.

(D) each student exerts a force on the other, but "a" exerts the larger force.

(E) each student exerts the same amount of force on the other.

29. An empty office chair is at rest on a floor. Consider the following forces:

1. A downward force of gravity.
2. An upward force exerted by the floor.
3. A net downward force exerted by the air.

Which of the forces is (are) acting on the office chair?

(A) 1 only.

(B) 1 and 2.

(C) 2 and 3.

(D) 1, 2, and 3.

(E) none of the forces. (Since the chair is at rest there are no forces acting upon it.)

30. Despite a very strong wind, a tennis player manages to hit a tennis ball with her racquet so that the ball passes over the net and lands in her opponent's court.

Consider the following forces:

1. A downward force of gravity.
2. A force by the "hit".
3. A force exerted by the air.

Which of the above forces is (are) acting on the tennis ball after it has left contact with the racquet and before it touches the ground?

(A) 1 only.

(B) 1 and 2.

(C) 1 and 3.

(D) 2 and 3.

(E) 1, 2, and 3.