UNDERGRADUATE AND GRADUATE STUDENTS' ATTITUDES AND APPROACHES TO PROBLEM SOLVING

by

Melanie L. Good

B.S. in Music Education, Lebanon Valley College, 1999

B.S. in Physics and Astronomy, University of Pittsburgh, 2009

M.S. in Physics, 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

PhD in Physics

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

THE KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Melanie L. Good

It was defended on

July 11, 2018

and approved by

Chandralekha Singh, PhD, Professor

Robert Devaty, PhD, Associate Professor

Roger Mong, PhD, Assistant Professor

Russell Clark, PhD, Senior Lecturer

Larry Shuman, PhD, Professor

Dissertation Director: Chandralekha Singh, PhD, Professor

UNDERGRADUATE AND GRADUATE STUDENTS' ATTITUDES AND APPROACHES TO PROBLEM SOLVING

Melanie L. Good, PhD

University of Pittsburgh, 2018

Student attitudes and approaches to problem solving in physics and astronomy may influence their development of expertise, as well as their engagement and perception of physics and astronomy as an academic endeavor. Introductory physics and astronomy undergraduate classes, which are gateways to a major in physics and astronomy, are foundational experiences in physical science education and development of problem solving skills. Understanding undergraduate students' attitudes and approaches may shed light on these formative experiences and what may be done to improve such experiences. On the other end of the spectrum, physics graduate students are expected to have developed significant problem solving expertise, and are potential future faculty. In their role of teaching assistants (TAs) and/or in a future capacity as instructors, graduate students may be responsible for making decisions on the types of problems used to shape their introductory students' experiences. These decisions by TAs may be crucial in the development of introductory student problem-solving expertise. Therefore, graduate students' attitudes about the instructional merits of different physics problems are worthy of examining in order to inform professional development programs for graduate teaching assistants. Investigating both undergraduate and graduate student perceptions about problem solving, we analyzed data related to gender, course and method of instruction, and type of problems preferred. Our data suggest that female introductory students and introductory students instructed in an evidence-based active engagement man-

ner have more favorable attitudes and approaches to problem solving compared with male students and traditionally-instructed students. Similarly, introductory astronomy students were found to have more favorable attitudes than introductory physics students. Moreover, it was found that graduate students' preferences regarding the types of problems they prefer to use with their introductory students does not always reflect the potential instructional benefits afforded by those problems. These findings illuminate pathways toward improving both teaching and learning of problem solving in college physics and astronomy courses.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgements

Though it is impossible to thank everyone who has helped me along the way in my research, I would like to acknowledge support from NSF and the following key people: Dr. Robert Devaty and Dr. Arthur Kosowsky for and supporting me with numerous letters of recommendation. Dr. Robert Devaty, Dr. Roger Mong, Dr. Russell Clark, and Dr. Larry Shuman for being a part of my thesis committee. Dr. Emily Marshman for her mentoring and collaborating on research. Dr. Andrew Mason and Dr. Alexandru Maries for their collaboration on the AAPS project. Dr. Gurudev Dutt for his letter of support and nomination of me for teaching awards. Zeynep Yasemin Kalender for her help with statistical analysis of the AAPS data.

My students for the opportunity to see the world of physics through their eyes.

My fellow graduate students who agreed to participate in one-on-one interviews.

My friends who never doubted me on my academic journey, even when I doubted myself.

My family: Dr. Joe Meyers, my husband, for his academic insight and techinical support, and most of all for his love and inspiration. My children, Sarah and Ashley Good-Lang and Rigel and Segin Meyers, for motivating me to seek ways of better educating the students of the future and for being patient with me through countless hours spent working behind my computer. Debra Metzger, my mother, for sharing her prior research into educational issues related to gender and for providing childcare to allow me distraction-free time to write my thesis. Kenneth Good, my father, for inspiring me to aim for the stars.

Most of all, a very special thanks to Dr. Chandralekha Singh who has helped shape and guide my academic and personal growth in innumerable and priceless ways.

1.0 INTRODUCTION

1.1 OVERVIEW

Problem-solving is central to the teaching and learning of physics. The way in which problem-solving is approached by the problem-solver is tied to many factors. Investigating those factors which relate to the way in which physics problems are perceived can shed light on ways to improve upon those perceptions. Learning does not take place in a vacuum. There is reason to believe that our perceptions shape our cognition [1]. Thus in order to optimize the experience of learning physics, it is useful for educators to understand their students' perceptions, as well as to be cognizant of the perceptions of those who may be in a current or future instructional role. The research presented here illuminates ways in which perceptions, attitudes, and problem-solving approaches intertwine. A general overview of learning theories and pedagogies, and a summary of background research into problem-solving give context to the implications of the research results that will be presented in subsequent chapters.

1.2 COGNITIVE THEORIES AND PEDAGOGICAL APPROACHES

At the intersection of the fields of cognitive psychology and education, an abundance of research has been conducted to investigate the mechanisms behind learning and a multitude

of theories about learning and pedaogies have been developed. Studies have sought to uncover ways of understanding and improving the process of learning and transfer of knowledge to new situations [2, 3, 4, 5]. It has been found that prior knowledge, context, expertise and organization of knowledge, as well as motivation may all impact learning success [6, 7, 8, 9, 10]. As a springboard for investigating problem-solving approaches, it is instructive to review some highlights from the learning theories and pedagogies that have emerged. In particular, those theories and pedagogies which are built from a constructivist perspective (i.e., a view that regards learning as an active process in which learners construct their knowledge) offer valuable insight into the issues surrounding problem-solving approaches.

1.2.1 Zone of Proximal Development

The importance of prior knowledge in learning was central to Vygotsky's concept of the "Zone of Proximal Development," which is the idea that learning will be most effective when the level of challenge of the task lies between what a learner can do alone and what a learner cannot do even with guidance [11]. A learner's prior knowledge helps define the zone of proximal development, as does the assumption of guided learning. The zone of proximal development expands on the Piagetian concept of "optimal mismatch," which represents an optimized discrepancy between new and existing knowledge. Building on the idea of optimal mismatch, the idea of the zone of proximal development posits that when learning tasks are reasonable enough that they can be accomplished with support and guidance, but challenging enough that they would be difficult to accomplish without guidance, learning is taking place within the zone of proximal development, which fosters optimal development of knowledge and skills [11]. As a learner develops understanding, the zone of proximal development will change, as what was once a task that could only be accomplished with guidance becomes a task a learner can accomplish independently. One important instructional implication of the fact that the zone of proximal development evolves with time is that it is important to frequently and formatively assess the level of understanding and mastery students have developed at a given time in the learning process.

1.2.2 Preparation for future learning

Learning tasks may require the learner to efficiently apply routine skills or to demand new innovations of the learner's current knowledge state and skill set. "Preparation for future learning" combines the ideas of innovation and efficiency in a two-dimensional model of expertise development [7]. With efficiency defining the horizontal axis and innovation defining the vertical axis, optimal expertise development takes place in this model along a corridor referred to as the "optimal adaptability corridor" which balances the vertical (innovation) and horizontal (efficiency) trajectories of learning [7]. The idea of "adaptability" is derived from the concept of "adaptive expertise," which is a kind of expertise in which one is not only adept at a skill but also able to adapt skill proficiency to new situations [12]. The upshot of the preparation for future learning model is that too much of either innovation or efficiency inhibits effective learning [7]. When learning involves mainly efficiency tasks, then learning becomes routine but not adaptable to new situations [7]. On the other hand, when innovation demands more than what can be accomplished given some level of mastered task proficiency, it can become cognitively overloading and unproductive [7]. Thus it is optimal to design learning tasks so that there is a balance between innovation and efficiency. This balance is not a straight line, but more of a swath in the two-dimensional innovation-expertise space. Within this swath, there is some evidence that, early on in the learning process, mastery is fostered better by more of an emphasis on innovation over efficiency within the optimal adaptability corridor [7]. For example, problem-solving which requires students to formulate the question before answering it demands more innovation than problem-solving in which the question is explicit. Striking a balance between innovation and efficiency shares a similar motivation to striving for optimal mismatch and aiming for the zone of proximal development, in that existing skills and knowledge inform the degree to which a learner may be challenged to form new knowledge effectively.

1.2.3 Cognitive apprenticeship

As Vygotsky's concept of the zone of proximal development emphasizes, learners may need to support in the learning process, and the cognitive apprenticeship model expands on how to effectively provide support and scaffolding. In the cognitive apprenticeship model, learning takes place through a guided process in which students gradually develop self-reliance. To facilitate this process, the framework includes three aspects: modeling to demonstrate the criteria for good performance, coaching and scaffolding to provide immediate feedback, and weaning to build autonomy [13]. In many traditional instructional approaches, modeling is provided by lectures and in-class examples, but students may not receive much coaching and scaffolding or gradual weaning before they are asked to independently demonstrate understanding and mastery of skills. When, instead, instructional approaches engage students in ways that offer opportunities for coaching and feedback and provide appropriate scaffolding, learning is being supported in the spirit of the cognitive apprenticeship model.

1.2.4 Motivation in Learning

Motivation can impact the learning process, and motivation may be multi-faceted, including aspects related to achievement goals, interest, assumptions about intelligence, and self-efficacy [10, 9, 14]. Moreover some of these aspects may be intertwined. For example, achievement goals may impact self-efficacy [15]. Motivational goals can be thought of as oriented towards one of two broad categories: performance orientation or mastery orientation. Mastery oriented goals are intrinsically motivated by a desire to develop genuine understanding [10, 9, 8, 16, 17, 18, 19]. By contrast, performance oriented goals are extrinsic and focused on achieving recognition (often in the form of high grades in the context of classroom learning) and may involve comparing one's performance to that of others [10, 9, 8, 16, 17, 18, 19]. Performance goals are sometimes further broken down into a performance approach orientation or a performace avoidance orientation, depending on whether a learner is more concerned with demonstrating performance or avoiding failure [20]. Moreover, motivation

and student engagement may be interconnected [19]. That is, high interest-level, engagement in learning, and motivation to learn may all go hand-in-hand. Thus it is not surprising that active engagement instructional methods are becoming more pervasive as awareness of these methods grows [21].

1.3 BACKGROUND RESEARCH IN PROBLEM SOLVING

Learning theories and pedagogies which emphasize contructivism, summarized in the previous section, suggest that engaging students in the learning process is an approach which fosters knowledge acquisition and development of problem-solving skills. Actively engaging students in a physics classroom requires opportunities for guided practice, feedback, and construction of understanding. The way in which problems are posed and the way in which instruction takes place can influence the degree to which students are engaged in the learning process, and this can shape their attitudes and approaches to problem solving, and expertise development in solving problems.

1.3.1 The Role of Expertise in Problem Solving

Constructivist models such as preparation for future learning are built upon the notion that a desired goal of instruction is expertise development. The way in which knowledge is structured in the mind depends upon one's level of expertise [22]. Although level of expertise exists along a continuum, this knowledge structure is revealed when experts' thought processes at one end of the spectrum are contrasted with those of novices at the other end of the spectrum. When asking problem-solvers to think aloud while solving problems, Chi found that the knowledge structure of experts, such as physics faculty members, is connected coherently and organized hierarchically [22, 23]. The knowledge structure of novices, such as introductory students, has been shown to be fundamentally different from that of ex-

perts, in that their knowledge can be much less organized and connected [22, 23]. Novices may think of problem-solving as an exercise in "plugging and chugging" equations that are disconnected from each other and from an underlying conceptual framework [22, 23]. As expertise develops, the process of problem solving becomes more efficient, because experts can draw upon an organized knowledge structure built on a coherent conceptual framework [22, 23].

1.3.1.1 Metacognition and Systematic Problem Solving Approaches A hallmark of an expertlike approach to problem-solving is the use of metacognition and a systematic strategy [23, 24]. In working with students of mathematics, Schoenfeld encouraged students' metacognition by asking guiding questions such as "Why are you doing what you are doing?" [24]. Schoenfeld's students were also instructed on how to take a systematic approach to problem solving by beginning with a careful qualitative analysis and planning before implementing and assessing the solution and reflecting on the answer [24]. Schoenfeld found that this approach resulted in improved student problem-solving strategies and metacognition, decreasing the number of students who dove directly into implementation before analysis and planning from 60% down to 20% [24].

In a similar way, Reif found positive results when students were given a deliberate problem solving strategy that incorporated reflection upon their solution [23]. In particular, Reif insisted that students conduct a conceptual analysis of the problem before constructing a solution and checking their answers [23]. When this method of problem-solving was employed and students were matched in pairs with those who were not given instruction on an explicit problem-solving strategy, it was found that final exam scores were improved and students had fewer alternate conceptions that interfered with their understanding [23].

1.3.2 Sensemaking and Motivation During Problem Solving

From a sensemaking framework, problem-solving broadly involves constructing a representation of the context of the problem (which includes framing the problem, activation of

knowledge components, and converging on a representation), and generating a problem solution, the end-point of which is arrived when satisficing (i.e., generating a solution that meets the problem-solver's goals) is achieved [8]. This framework is particularly useful for novice problem-solvers [8], and as such, it is an appropriate framework for understanding problem-solving of introductory college students. The sensemaking framework, as the name implies, emphasizes sensemaking and satisficing, but instead of including only a classical cognitivist approach, this framework also incorporates other important factors in the process of sensemaking and satisficing, such as motivational goals and beliefs [8]. For example, when motivation is focused mainly on performing well in a class without much focus on achieving understanding, this has been identified as a "performance" motivation, which can be contrasted with a "mastery" motivation that focuses on achieving understanding and mastery of the material [9, 10]. It has been proposed that when motivational goals focus on mastery rather than performance, knowledge transfer should more readily be promoted [8, 9, 10]. In addition to driving the framing of the problem, motivational goals and beliefs also drive the criteria for satisficing [8]. As such, motivational goals and beliefs can be thought of as attitudinal factors which may shape problem-solving because of their influence on sensemaking and satisficing.

1.4 UNDERGRADUATE AND GRADUATE STUDENTS' ATTITUDES AND APPROACHES TO PROBLEM SOLVING

A variety of attitudinal and epistemological surveys have been designed to measure the views of students regarding physics, science, learning, and problem solving [14, 25, 26, 27, 28, 29, 30, 31]. Many of the studies which have utilized these surveys have been conducted on introductory undergraduate students, and it has been found that attitudes towards physics and physics learning often decline during the course of instruction (i.e., when scores at the

end of a semester are compared with those at the beginning of the semester) [25, 27, 32]. Factors that may influence the amount of decline may include self-efficacy, gender, and sense of belonging [14]. It has also been found that graduate students' attitudes and beliefs about physics learning and practice may be more expertlike than undergraduate students' but less expert-like than physics faculty [31, 33, 34, 35, 36, 37]. However, even physics faculty may not always utilize instructional strategies or problem features to promote expertise development of their students in a way that is evidence-based [38, 39]. Since graduate students are potential future faculty members and limited professional development occurs in the time between graduate studies and commencing of a faculty position, identifying needed areas of growth can inform professional development efforts for graduate teaching assistants.

The following chapters explore the connections between the attitudes and approaches to physics problem solving and factors that may influence these attitudes and approaches. Chapters 2-4 report on the findings regarding the views of graduate student teach assistants towards the ways in which an introductory physics problem may be posed. A similar study was conducted previously with faculty [39], which allows for useful comparison. The results suggest ways in which graduate student teaching assistants' views about instructionally beneficial problem-solving could be improved. In particular, growth of graduate student teaching assistants' awareness of the importance of providing feedback by way of formative assessment and offering opportunities for collaborative construction of knowledge are warranted by the findings presented in these chapters. Likewise, a tendency towards overreliance on guiding students through problem-solving may be another area to address the professional development of graduate student teaching assistants.

Chapters 5 and 6 focus on the views of undergraduate introductory students. In chapter 5, issues regarding gender, method of instruction, and change of attitude throughout a semester are explored. The evidence suggests benefit of instructional methods which are evidence-based and actively engage students. In addition, important questions are raised about gender. In particular, female students were found to exhibit more enduring and favorable attitudes and approaches to problem solving, which is intriguing given the performance

gap that persists between male and female students. Finally, the chapter 6 investigates the differences between introductory physics and astronomy students' attitudes and approaches to problem solving. In comparing these two groups of students, more favorable attitudes were found among the introductory astronomy students, which may have implications for instructional methods for physics classes that may generate more interest, motivation, and engagement in the classroom.

2.0 PHYSICS GRADUATE TEACHING ASSISTANTS' PERCEPTIONS OF THE INSTRUCTIONAL BENEFITS OF A CONTEXT-RICH INTRODUCTORY LEVEL PROBLEM

2.1 INTRODUCTION

2.1.1 The role of problem-solving in the achievement of learning goals

The desired learning goals for students in many introductory physics courses often include learning physics concepts and developing expertise in problem-solving and reasoning skills [38, 40, 41, 42, 43]. Physics experts, e.g., physics faculty members, organize their physics knowledge hierarchically so that underlying concepts are connected in a meaningful and structured way and they exhibit positive attitudes towards scientific problem solving [40, 22, 23, 44, 45, 38]. Experts' knowledge, including how the knowledge is structured in well-organized schema, and their positive attitudes and approaches to problem solving can facilitate an effective approach to problem solving [40, 46, 47, 48]. By contrast, many introductory students, view physics as a collection of disconnected facts and equations and have less expertlike attitudes and approaches to problem-solving [22, 23, 40]. One strategy to achieve the goals related to the development of expertise of introductory physics students and improving their attitudes and approaches to problem-solving is to actively-engage them in the learning process using research-based approaches.

Different problem "types" (i.e., different ways of posing the same underlying physics

problem) can be used in different ways to actively engage students in order to meet the instructional goals. Moreover, there is evidence of a growing awareness among faculty of the instructional benefits of research-based methods that emphasize active engagement [49]. Depending on the instructional goals, active engagement methods can include a wide variety of options to meet those goals [50], and different types of problems can be utilized to support these goals. An example is the use of context-rich problems (problems posed in a realistic, narrative manner which may include extraneous information and may not explicitly ask a question–i.e., an explicit problem may need to be formulated by the problem-solver). A problem posed in a context-rich manner can engage students in learning effective problem-solving strategies when used as part of collaborative group problem-solving while the instructor or teaching assistant facilitates the process by providing feedback and support as needed. Moreover, group problem-solving with context-rich problems can promote both positive inter-dependence amongst students and individual accountability [51]. While many other problem types exist, we focus here on the context-rich problem type, its role in helping students learn physics, and the way this type of problem is perceived by graduate teaching assistants.

2.1.2 The role of the teaching assistant in student problem-solving

Physics graduate teaching assistants (TAs) are often employed, especially at large research universities, to carry out duties such as instructing the recitation/discussion sections related to introductory physics courses. It has been noted that TAs are often responsible for a significant portion of undergraduate instruction, and that their training for this role is often limited [52, 53]. TAs may be responsible for choosing the types of problems to use with students, e.g., in designing quizzes for their students to take during recitation/discussion or creating example problems to discuss. Since their training may be limited, these choices may be based upon TAs' perceptions about different types of problems. Moreover, as potential future faculty, TAs may have an ongoing decision-making responsibility about the types of problems to use with their future students. Even though a small minority of physics

departments in the U.S. provide semester-long TA professional development, the majority of physics departments provide only very short training (i.e., a few hours) to prepare them for these various teaching responsibilities. In addition, little in the way of guidance or supervision is typically provided to TAs to support them in their teaching activities [52]. As such, TAs' teaching practices are often affected by the expectations of their supervising instructors and also by their beliefs and workload as graduate students [54, 55, 56, 57, 58, 59, 60, 61, 30, 31, 62, 63]. Thus, with limited opportunities for professional development and training in the intervening time between the TA role and the faculty role, TAs perceptions may also shed light on their future perceptions as potential faculty members. Moreover, the perceived instructional value may affect choices about the use of various types of problems and thus impact the degree to which different types of problems are exploited for their effectiveness in actively engaging students and their ability to facilitate instructional goals. As such, TAs' perceptions of different types of problems are worthy of examination to inform professional development courses and programs.

2.1.3 Focus of our research

In the study presented here, TAs in a TA professional development course were asked to reflect upon five problem types that are appropriate for an introductory mechanics problem scenario. These same problem types had been used in an earlier study with instructors [39]. The example problem types were meant to generate a broader discussion and reflection upon the types of problems TAs might choose to use in their teaching Here we focus on TAs' initial views, after some amount of experience as a TA, about a context-rich introductory physics problem and investigate the following research questions: (1) How challenging and instructionally beneficial do TAs perceive context-rich types of problems? (2) If TAs had complete control of an introductory course, how likely would they be to use context-rich problems compared with other types of problems, and for what purpose might they use them? (3) Why do TAs perceive context-rich problem types the way they do?

2.2 BACKGROUND

2.2.1 Physics faculty views about different problem types

A prior study regarding physics instructors' views about different problem types in which they were presented with the same variations of a physics problem (including the context-rich problem) given to the TAs in the current study [39]. It was found that the instructors generally valued different problem types intended to develop expert-like problem-solving but they were not as likely to use certain problem types. Regarding context-rich problems, it has been found that physics instructors instructors generally valued context-rich problems and felt that such problems supported the goal of developing students' ability to plan and explore solution paths. However, they were not very likely to use context-rich problems to avoid stressful situations for students since they are complex and ill-structured and that such problems lacked clarity [39].

2.2.2 Prior work on TAs' professional development and their views about teaching and learning

Several studies have investigated TAs' views about teaching and learning [33, 49, 34, 35]. Prior research suggests that there are discrepancies between physics graduate TAs' perceptions of what teaching strategies are beneficial for students' learning and many of the findings of physics education research [33, 49, 34, 35, 36, 37]. For example, TAs have been found to struggle with the idea that effective grading practices can be a formative assessment tool, e.g., grading practices that encourage students to show their work can improve their learning from problem-solving and encourage them to learn from their mistakes [33, 49, 34, 36, 37]. Another study involving TAs' beliefs about example solutions provided to students shows that many TAs were unlikely to identify features in the problem solutions that the research literature describes as supporting learning goals for students [35].

It has been found that TAs' beliefs affect their teaching practices [54, 55]. Because of

the role of the TA in decision-making on use of various problems, both in the TAs' current capacity and in possible future roles as faculty, their beliefs about the pros and cons of posing an introductory physics problem in different ways and in different instructional contexts can affect the ways in which they use various types of problems. Thus, identifying the views of the TAs' about the way in which a problem is posed can be useful in developing activities to improve their professional development and help them recognize the pedagogical value of posing the same problem in various ways.

2.2.3 Context-rich problems as tools for promoting problem-solving skills

Context-rich problems have been shown to be effective in helping introductory physics students become good problem solvers [64, 65]. Physics problems which are context-rich are often complex and lacking in structure, frequently provide redundant information or are missing information and have real-life contexts [64, 65]. One prior investigation suggests that students who worked in groups were more likely to use effective problem-solving strategies and show positive inter-dependence when working on context-rich problems than when working on analogous traditional textbook problems [64, 65]. Research also suggests that students who engage with context-rich problems are more likely to think about the concepts first, use diagrams in their problem solving process and have a more positive attitude about problem solving [66]. As students become more experienced in solving context-rich problems, they show progress towards expert-like problem solving [67]. Because context-rich problems require students to formulate the question and make inferences, a systematic, expert-like problem-solving approach is more effective [23]. Thus, context-rich problems can facilitate progression towards expertlike problem-solving [23] such as executing a conceptual analysis and planning of the problem solution before implementing the solution. Reflection and metacognition, as well as utilizing a well-organized knowledge structure, also play key roles in solving the problem and learning from the problem-solving process. [23, 22, 68, 69, 24].

2.3 METHODOLOGY AND DATA COLLECTION

Participants and description of TA professional development program: A total of 97 TAs from a typical large research university participated in this study during 4 different years. Participants were physics graduate students who had teaching responsibilities (recitation or lab instruction, and a majority were also assigned to help students in a physics tutoring center) and were concurrently enrolled in a mandatory TA professional development course that met once per week for 2 hours for an entire semester. The TAs were expected to do approximately one hour of homework each week pertaining to the professional development course, in which various activities took place throughout the semester. During the course, initial activities related to grading practices occurred near the beginning of the semester, followed by discussions of pedagogy, including the use of tutorials and clicker questions as learning tools. After this, discussions turned to how different problem types (e.g., multiple choice problems, context rich problems, problems that are broken into sub-problems, and traditional textbook style problems) and different example solutions to problems can help students learn physics. TAs were involved in evaluating the effectiveness of multiple choice questions on conceptual surveys, and predicting which choices students might pick. This gave TAs the chance to reflect on the design of conceptual multiple-choice questions, anticipate challenges their students might encounter, and make judgements about how effectively a well-designed question might uncover where students are struggling. TAs also were given a physics problem and asked to present the solution to the TA professional development class as they would in their recitations. These presentations were video-recorded so that they could reflect on their teaching and also receive feedback from other TAs and the instructor. Thus the problem-type activity was one of a number of activities all aimed at the professional development of the TAs while investigating their assumptions about physics teaching and learning.

Data collection tools and artifacts: The data collection tools consisted of five introductory physics problem types that had been developed previously [39] and served as

guiding examples for the activities. The problem types were designed for the same introductory physics problem scenario in mechanics. They included two different versions of a problem which was broken into sub-parts (one was framed in a more conceptual manner than the other), a multiple-choice problem, a context-rich problem and a traditional textbook version of the problem. Here we focus on the context-rich problem, namely Problem C (see Figure 1) and the discussion related to problems posed in a context-rich manner. The example context-rich problem requires that students first construct a concrete question and then solve the problem using the relevant information provided (extraneous information is also included and the problem requires explicit calculation after formulating it, as is typical of this type of problem). While the example presented to TAs was a quantitative problem, TAs were engaged in discussion about the merits of context-rich problems in general.

Based upon our research questions, the TAs were asked to answer questions about these problem types on a worksheet, a partial sample of which can be found in Figure 2. Among the entries shown in Figure 2, TAs were directed to list pros and cons for each problem type. Specifically, in the instructions, TAs were asked to list at least one pro and one con for each problem type based upon the features each of the five problem examples contained. Data were collected over four different years. In the most recent year's worksheet, TAs were also asked what they would change about the example problems. In addition they were asked to rank the features of problem types on their instructional benefit (i.e., how instructionally beneficial the TAs judged each problem type to be), and to rank the problem types in terms of the level of challenge (i.e., how difficult the TAs judged each problem type to be for students), how much they liked the problem types, and the likelihood that they would use the problem if they had complete control of the choice of the problem types to use. For example, a TA who ranked a problem 1 for "challenging" judged this problem to be the least challenging for students; a 5 for "challenging" indicates that the TA perceived it to be the most challenging among the five problem types. The rankings allowed us to investigate research questions 1 and 2. Furthermore, throughout all four years, TAs were asked to list pros and cons of the problem types. These pros and cons were useful for investigating why

Problem C

You are working at a construction site and need to get a 3 lb. bag of nails to your co-worker standing on the top of the building (60 ft. from the ground). You don't want to climb all the way up and then back down again, so you try to throw the bag of nails up. Unfortunately, you're not strong enough to throw the bag of nails all the way up so you try another method. You tie the bag of nails to the end of a 2 ft. string and whirl the string around in a vertical circle. You try this, and after a little while of moving your hand back and forth to get the bag going in a circle you notice that you no longer have to move your hand to keep the bag moving in a circle. You think that if you release the bag of nails when the string is horizontal to the ground that the bag will go up to your co-worker. As you whirl the bag of nails around, however, you begin to worry that the string might break, so you stop and attempt to decide before continuing. According to the string manufacturer, the string is designed to hold up to 100 lbs. You know from experience that the string is most likely to break when the bag of nails is at its lowest point.

Figure 1: The context-rich problem posed to the TAs

Problem Type	What are the **FEATURES** of the problem?	What are the **PROS/CONS** of the problem?	Would you use this problem as . . .			
			a **HW** problem? Why or why not?	a **QUIZ** problem? Why or why not?	an **EXAM** problem? Why or why not?	a small **GROUP** problem solving session? Why/why not?
Problem C	Practical application problem.; No diagram.	Pros: Connect knowledge to practical application. Cons: Too long description.	No, too many unnecessary descriptions.	No, takes too much time to read.	No, takes too much time.	No, takes too much time

Figure 2: Part of a sample worksheet given to the TAs

TAs ranked the problem types the way they did (research question 3).

Data collection in the TA professional development course and in individual interviews: TAs were given the problem types and worksheets in the professional development course in the middle of the semester when they had some teaching experience in order to elicit their initial ideas about different problem types. They were asked to answer worksheet questions under the assumption that they had complete control over the intro-

ductory physics class, including control over problem types chosen for various purposes. The worksheet (see a part of the worksheet in Figure 2) was completed as part of a homework assignment. Later on, 12 participants who had taken the TA professional development course earlier volunteered to be interviewed in a one-on-one setting using a think-aloud protocol. These interviews took place at least one semester after the initial activity described here in the TA professional development course and were audio-recorded. TAs who participated in the interviews were asked questions both about the example context-rich problem and about context-rich problems, in general (similar to the broader in-class discussion about the instructional benefits and pros and cons of well-designed context-rich problems). Thus, the interviews served to more deeply probe the TAs' reasoning behind their written responses and to explore such questions as the use of the context-rich format as a general problem feature.

Coding TA responses: Two of the researchers met weekly to identify appropriate coding categories for pros/cons for all four years of data; agreements on these were reached through discussion. The researchers used open coding of the data from the individual homework assigned in the middle of the semester regarding the TAs' views of the problem types. The categories were created over several weeks based on emergent themes. Some categories were merged if they were found to be sufficiently similar. The inter-rater reliability the coding of the pros/cons for a subset of the coded data (i.e., one of the four years) was examined, and the average Cohen's kappa [70] was calculated to be $\kappa = 0.982$. Table 1 shows the three most common pros and cons for the context-rich problem, along with their definitions and examples from TA worksheets. We note that for the context-rich problem, some TAs' responses could not always be appropriately coded as a particular pro or con because they were negative but non-specific in nature. For example, one TA response was, "never assign this." We were unable to code this response, and others of this nature, into a specific pro/con category.

Code	Definition	Examples	Percentage of TAs
(Pro) real	real life scenario; can help students relate physics to their lives	"gives insight to how physics laws and principles are ingrained in our lives"	47 %
(Con) time	time-consuming	"takes up a lot of time" "can take too long"	33 %
(Con) unclear	problem is vague and/or confusing	"can be ambiguous; not sure what to do" "hard to decipher"	42 %

Table 1: The most commonly listed pros/cons of a context-rich problem and the percentages of TAs who listed them. Some TAs listed more than one of the following or other pros/cons not listed here. However, other pros were negligible and the overlap of the two cons listed below represents only 5% of all TAs.

2.4 RESULTS

2.4.1 Rankings

Context-rich problem ranked lowest for like, use, and instructional benefit, but highest for level of challenge for students: Figure 3 summarizes the rankings for the context-rich problem in all four categories. n the category of "instructional benefit" TAs were asked to rank the problem *features* and the example problem types themselves. In response to research questions (1) and (2), on average, the context-rich problem received the lowest ranking of all problem types in the three categories "like," "use," and "instructional benefit". On the other hand, the context-rich problem was ranked, on average, the highest for "challenge" out of all five problem types the TAs were given. Given the extreme rankings of the context-rich problem for the most recent year's data that included ranking information, we examined the pros and cons the TAs listed for all four years of data, as well as the feedback obtained during individual interviews to look for reasons for why they perceived the context-rich problem the way they did, in response to research question (3).

It is important to note that while written responses were guided by the five example problem types presented to TAs, interviewed TAs were asked to think not only about the specific

example of the context-rich problem that they were given, but also to think about context-rich problems generally-speaking, and to imagine any possible way of posing a context-rich type of problem. The interviewed TAs responded in a similar manner to those who gave written responses for all of the preceding categories (i.e., "instructional benefit," "challenge," "like," and "use") even when asked more generally about context-rich problems. In other words, rather than the above rankings holding true only for the specific context-rich problem given, interviewed TAs saw little instructional benefit and were unlikely to use context-rich of any kind. Thus the rankings appear to generalize to the posing of problems in a context-rich manner.

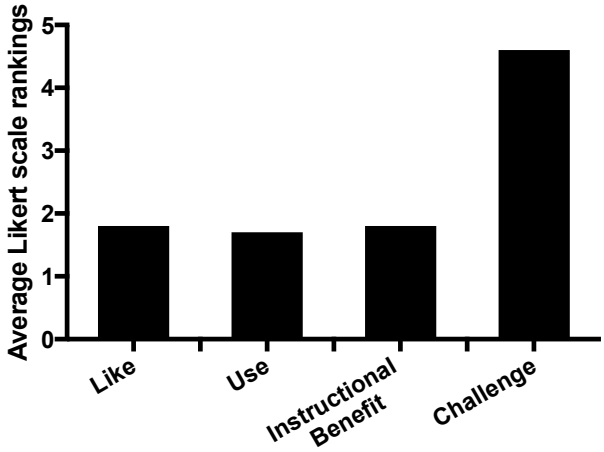


Figure 3: Average rankings of the context-rich problem. Rankings were included only on the most recent year's worksheet. Interview data and prior years' worksheets were further analyzed for their pros/cons and other information.

2.4.2 Reasons for rankings

The pros and cons in written responses as well as the interview data elucidate possible reasons for why the TAs ranked the context-rich problem low for "like," "use," and "instructional benefit," and high for "challenge." Table 1 shows the most common pros and cons mentioned by TAs in written responses.

TAs viewed the context-rich problem as unclear: Table 1 shows that the most commonly stated con was coded by the researchers to be in the category "unclear." "Unclear" refers to TA responses which described the problem feature as confusing (e.g.,"Overly confusing and frustrating") or lacking in clarity or explicit question (e.g.,"The point of the problem is not clear"). In fact, the con "unclear" appears to be one major reason why the TAs thought that the context-rich problem type was highly challenging and may have contributed to the TAs' reluctance to use this problem type or recognize its instructional benefit. The majority of TAs who ranked the context-rich problem as the most challenging and low in instructional benefit also mentioned a con categorized as "unclear." Likewise the majority of TAs who indicated they would be unlikely to use a context-rich problem by ranking it low for the category "use," also mentioned a con coded as "unclear."

The TAs who listed the con "unclear" had a variety of reasons for why they perceived this type of problem to be lacking in clarity, and often the same TA stated more than one reason for why he/she felt the problem is unclear. Written responses and interviews suggest that one reason the context-rich problem was often viewed as unclear by the TAs was that there was a lack of an explicit question in it (which is common for problems posed in a context-rich manner). For example, one interviewed TA explained that the lack of an explicit question made the problem confusing: *"This one is very vague. It's not asking any question, so the student might get confused about what they're supposed to do...."* This type of sentiment about the lack of explicit question and the problem being "vague" or a source of confusion for students was commonly mentioned by the TAs both in written responses and interviews. For example, another interviewed TA who described the problem as "unclear" explained, *"at least if you give a student a problem where they know what they need to do, they can ask you questions about that, but I can imagine students sitting there saying 'I don't know where to start."'* Moreover, discussions with this TA suggest that he felt that his students will not be able to proceed. He noted that without being able to discern a clear goal, the students will neither be able to make sense of the problem nor be able to reach out for support from the TA or instructor, as they will not know what questions to ask. Similar to this TA, many TAs

connected a perceived lack of clarity to the fact that the context-rich problem does not ask an explicit question. Another TA stated, *"It is challenging... because the wording of the problem is vague. I don't even really see an explicit question that it's asking."* This TA (and many others) explicitly made a connection between the level of challenge and the lack of a concrete question posed. It was implied that it did not make sense to give this type of problem to students under any circumstance. This TA appeared not to recognize the benefit of requiring students to determine the question that is being asked in context-rich problems, which is the rationale for not including a question, a common feature of context-rich problems.

TAs viewed the context-rich problem as unclear, having too much of a narrative: Another reason many TAs found this problem to be "unclear" was that the TAs felt it was verbose with too much extraneous descriptive information and that their students will have difficulty interpreting it (note that wordiness and possible extraneous information are common features of context-rich problems). This reason for perceived lack of clarity also seems to be contributing to the high ranking for challenge and low ranking for instructional benefit and likelihood to use. For example, one interviewed TA first described the problem as having "too much detail," and then went on as follows, *"it's very good that it has a story to it but the story is too much and the science is not enough. Because this is a physics class and... it looks more like a story than a physics problem."* It is interesting that the TA appeared to view the narrative aspect of the problem (i.e., the "story") as separate from the "science" of the problem. This TA appeared to be of the opinion that a context-rich problem may not hold benefit for physics students, since he felt that this sort of problem does not belong as part of a physics class. Further discussion with the TA suggests that he preferred textbook problems and viewed the detailed descriptive narrative in a physics problem as being outside of the scope of what should be part of a "usual" physics curriculum. This could explain why this TA reported that he would not be likely to use a problem such as this. Another TA remarked, *"I think... intuition in being given a word problem and knowing how to translate that into math is the hardest part."* Further discussions with this TA suggest that he agreed that formulating the problem (i.e., translating the problem from

words to quantitative expressions) may potentially be useful for students. However, he then stated that he would not use this type of problem in his own class due to the fact that it is unclear and challenging for students and his students would not know how to solve this type of problem. Similarly, another TA who ranked the context-rich problem the lowest for likelihood to use it said: *"It's not about the physics; it's just about the wording... I find this question itself is kind of testing how you understand a paragraph of description, but the physics itself might [not] be..."* This TA felt that the wording of the problem as a descriptive paragraph was something separate from the physics involved, almost as if implying that the paragraph form obfuscates the underlying physics. Indeed this TA went on to say that *"It's kind of confusing...I took a long time to understand it...It's not quite clear."* This TA mentioned both the lack of an explicit question and the lengthy narrative as problematic for the problem's clarity, and did not appear to see instructional benefit in using a context-rich problem.

TAs viewed the context-rich problem as overly time-consuming for the student: In addition to the con "unclear," another common con category was "time". "Time" was the category used for TA responses that indicated that solving this problem would be too time consuming for students ("It may take a while for the student to interpret the problem into a mechanics problem") and/or did not make the best use of time ("Takes time reading things that are not directly helpful for solving the problem"). The con "time" may also have contributed to TAs' hesitation in using this problem in their own classes. Of the TAs who ranked the context-rich problem the lowest in terms of "use", the majority listed "time" as a con. A TA's reluctance to use a context-rich problem due in part to time constraints is evident in the following comment during the interview: *"The student will take much more time to solve this... I think the student has to read this question so many times more to understand it. ...I won't use this at all."* We note that this TA identified both the time it will take the student to solve the problem and also the time required to read and understand the problem statement as being problematic with regard to the use of this problem. In a similar manner, another TA who reported that he would never use a context-rich problem

stated: *"The students have to spend too much time trying to figure out what they are supposed to answer...It does not help the student at all."* This TA identified the time needed to interpret the problem and construct the question as excessive and went on to say that this time requirement makes the problem unhelpful to the student, which appears to imply that this negatively impacts the instructional benefit of the problem in the mind of this TA.

No commonly-identified ways of using a context-rich problem: It is interesting that many TAs did not envision any situation in which a context-rich problem could be used effectively, e.g., in a collaborative group problem solving session or as part of a homework assignment. That the ranking for "use" was the lowest of all problem types indicates a reluctance to use a context-rich problem. Furthermore, even though the instructions regarding the worksheet completion included a directive to think of at least one way in which each problem type could be used, over 20% of the TAs stated that they would never use a context-rich problem in any way and listed no pro even when asked for at least one pro. Moreover, of the few TAs who mentioned group work as a potential use, most did not mention that the group work could help actively engage students in co-constructing knowledge and learning problem-solving skills.

Similarly during interviews, when asked how they might use the context-rich problem, many TAs noted that they would not use a context-rich problem at all and only one of the TAs suggested the idea of group work as a possibility. Exhibiting this reluctance to use a context-rich problem, one TA struggled to think of a possible way to use a context-rich problem, saying, *"I think this might not be the kind of question you ask in a recitation...maybe this is the outline for a lab or something, but it doesn't seem right [even then]."* This TA felt that the context-rich problem was not well-suited to use in a recitation. Furthermore, even in a lab setting where time may not be as big a factor and where students might collaborate, the TA was hesitant to endorse using it.

Pros do not outweigh cons: Many TAs did not list any pros for the context-rich problem even though they were asked to list at least one pro for each problem type. The inability to come up with a pro for the context-rich problem supports the fact that TAs did

not perceive this type of problem to be instructionally beneficial. Moreover, the inability to come up with pros for such a problem may explain the low average rankings of TAs on "like" and "use" for the context-rich problem. Of the pros TAs did mention, the most common pro was "real" (i.e., relatable to a real-life scenario). One TA put it succintly: "Connects to daily life". We note that the TAs seldom mentioned any other pro. Interviews suggest that the pro "real" may not be perceived as compelling enough to outweigh the negative light in which the TAs viewed the context-rich problem overall. For example, one interviewed TA stated, *"There's some redeeming elements to it, like I like that this frames it from the perspective of the student, so they can think about what they see and feel while they're whirling the string around. So it's not all bad, but...I would tend not to use it at any level. I just don't think it's an effective problem."* This TA appeared to recognize some "redeeming elements" in the real-life aspects of the problem, but qualified this pro by stating that this is not enough of a reason for him to use a problem like this in his own classes in any way. Additionally, the remark about it not being effective appears to speak to his low opinion of the problem in terms of its instructional benefit. Other TAs had similar views.

Thus, written responses and interviews suggest that overall, the TAs appear to have a negative opinion of the context-rich problem type. The total percentage of TAs who listed one or more significant cons to the context-rich problem is 80% and other negative responses could not be coded since they were not specific. Indeed, it is interesting to note that the context-rich problem type elicited a strong negative response from many TAs. Some TAs were extremely negative with statements such as: "Absolutely does not help the students at all", "meaningless", and "It sucks". In summary, for the context-rich problem, there were more cons listed than pros and the cons were often strongly worded, which appears to be consistent with the very low rankings in categories related to use, like, and instructional benefit. In addition, the aspects that the TAs perceived as negatives are typically intentional design aspects of context-rich problems (e.g., context-rich problems, by design, are challenging, wordy, and require interpretation in order to ensure that students focus on formulating the problem and doing a conceptual analysis and planning of the problem solution). Based upon

TAs' negative views about the features of the context-rich problems in written responses and individual interviews, it appears that they did not discern instructional benefit in these types of problems in general and were unlikely to use them in their classes.

2.5 SUMMARY AND CONCLUSIONS

Our findings suggest that TAs did not, in general, perceive a context-rich problem type in a positive light. Written responses and interviews suggest that there were several reasons for this negativity, none of which appear to be unique to this particular context-rich problem, and would apply to the general features of context-rich problems.

Regarding our research question (1), we find that many TAs felt that the context-rich problem type was too challenging to be of instructional benefit for their students, with the majority of TAs ranking the context-rich problem the lowest for instructional benefit and highest in challenge, and with cons outweighing pros. Regarding research question (2), TAs were not likely to use a context-rich problem in their classes if given complete control of making decisions. This was evidenced by both the low "use" ranking and the written and interview responses indicating they would be unlikely to use a context-rich problem. Finally, regarding research question (3), the reasons for TAs' views often pointed towards the context-rich problem type being unclear and difficult for their students to interpret as well as being time-consuming. Many TAs explicitly mentioned the lack of a clear-cut question, wordiness, and extraneous information in the context-rich problem as being problematic. The TAs also identified the time required to parse the information and formulate the problem as being major drawbacks to using a context-rich problem.

However, many of the perceived drawbacks of context-rich problems mentioned by TAs are identified as positive aspects of context-rich problems in the research literature [64, 65, 66]. Indeed, context-rich problems are usually purposely designed with these features

because of the benefit that those features afford for helping students actively engage in learning effective problem solving strategies, including the importance of doing a qualitative analysis and planning of the problem before jumping into the implementation phase. For example, the lengthy narrative without a clear question can develop students' ability to differentiate what information is important, identify the relevant concepts, and formulate the problem. Likewise, the length of time needed to solve a context rich problem is partly due to the fact that it is realistic. Not every problem one encounters in real life can be solved quickly, and the time required to solve a context-rich problem can also help students learn the importance of perseverance in problem-solving involving realistic situations. This is a lesson that could be beneifical for students, since research has shown that students often are likely to give up if they cannot solve a problem in 10 minutes [71, 30, 31]. Yet these beneficial aspects were seen by a majority of the TAs in a negative light.

It is important to note that much of the written data was guided by specific examples problem types; however discussion in the TA professional development class and interview data were meant to elicit more general perceptions. Moreover, the qualitative interview data provided reasons for responses given in the written data. Thus, while an important limitation to our findings might be the ability to generalize our results since a single example of each type of problem was given, the agreement with the interview data suggests that the written results may hold true for TA perceptions, in general. In addition, some aspects of the written data suggest some ability to generalize the findings. For example, the most recent year's worksheet included "instructional benefit" rankings that instructed TAs to rank the *features* of the problems first before ranking the example problems, and the feature rankings agreed with the rankings found when TAs were considering the specific examples given. In other words it was the *feature* of these problems that appeared to rank low for instructional benefit in the opinion of the TAs. While the written data that contain these additional questions is only present for the most recent year, it is reasonable to assume that the subset of TAs who responded to these additional questions is representative of the larger sample.

These findings appear to agree with findings from previous studies regarding TAs' beliefs

about assessment and example problem solutions, wherein TAs struggled to identify aspects that are supported by the research literature as beneficial for students [33]. With regard to studies about the perception of problem types, our results for TAs' views of the context-rich problem type differ somewhat from instructors', who were found to value the context-rich problem for its capacity to develop students' abilities to plan and explore solution paths [39]. Nevertheless, while faculty appear to discern the benefits of a context-rich problem type more readily than the TAs in our study, the TAs do share some similarites of perception with those of faculty. Specifically, like the faculty, the TAs in our investigation were not likely to use a context-rich problem in their classes, both mentioning issues of clarity as problematic [39]. (Faculty also mentioned avoiding student stress as a reason to avoid using context-rich problems [39], which was also mentioned by some TAs, but in smaller numbers than the issues of time and clarity). This suggests that while faculty may value a context-rich problem more than TAs (or at least expressed in one-on-one interviews that such problems could be valuable for introductory students), a majority of faculty and TAs appear unlikely to utilize a context-rich problem for instructional purposes, which represents a missed opportunity for introductory students (in terms of how much they could benefit from such problems, particularly because other problem types do not provide the same benefits). TAs and faculty, who are both in a position to select physics problems for introductory students, share a consistent reluctance to use a context-rich problem suggesting a common struggle in identifying the instructional value of such a problem.

Leaders of TA professional development programs can build on these findings to help TAs reflect on the benefits and effective uses of context-rich problems. For example, professional development programs can help TAs reflect on the use of context-rich problems in collaborative group problem-solving settings, since this setting was not identified as one which can be used to help students become good problem-solvers making use of context-rich problems. The use of a context-rich problem in collaborative problem-solving engages students actively in the problem-solving process, and has been shown to benefit student learning [64, 65].

In addition to reflection on how to effectively use a context-rich problem, TAs' strong

negative perceptions about the benefits of this type of problem are dissimilar to the findings of the analogous problem-type study with faculty [39]. Helping TAs become more aware of research that supports context-rich problems as instructionally beneficial for students could help TAs come to discern the perceived cons as pros. Ultimately, the goal would be to mitigate negative perceptions TAs might have about problems posed in a context-rich manner so that aspects such as requiring the student to construct the question and make inferences are viewed as instructionally beneficial.

3.0 PHYSICS TEACHING ASSISTANTS' PERCEPTIONS OF MULTIPLE CHOICE PROBLEMS: OVERLOOKING THE BENEFITS OF FORMATIVE ASSESSMENT

3.1 INTRODUCTION

3.1.1 Problem-solving expertise and active engagement learning

The desired learning goals for students in many introductory physics courses often include learning physics concepts and developing expertise in problem-solving and reasoning skills [38, 40, 41, 42, 43]. Physics experts, e.g., physics faculty members, organize their physics knowledge hierarchically so that underlying concepts are connected in a meaningful and structured way and they exhibit positive attitudes towards scientific problem solving [40, 22, 23, 44, 45]. Experts' knowledge, including how the knowledge is structured in well-organized schema, and their positive attitudes and approaches to problem solving facilitate an effective approach to problem solving [40, 46, 47, 48]. By contrast, many introductory students, view physics as a collection of disconnected facts and equations and they have less expertlike attitudes and approaches to problem solving [22, 23]. One strategy to achieve the goals of developing the expertise of introductory physics students and improving their attitudes and approaches to problem-solving is to actively-engage them in the learning process.

Different problem "types" (i.e., different ways of posing the same underlying physics problem) can be used in different ways to actively engage students using evidence-based

approaches in order to meet the instructional goals. Moreover, there is evidence of a growing awareness among faculty of the instructional benefits of research-based methods that emphasize active engagement [49]. Depending on the instructional goals, active engagement methods can include a wide variety of options to meet those goals [50], and different types of problems can be utilized to support these goals. For example, even in large-enrollment classes, multiple-choice questions can be used, e.g., they can be administered via clickers to help to provide formative assessment opportunities and engage students in discussion with peers. Such formative assessment opportunities afforded by multiple-choice questions can help students take ownership of their learning [72, 51]. Moreover, group problem-solving with "Think-Pair-Share" activities using multiple-choice clicker questions can promote both positive inter-dependence amongst students and individual accountability [51]. While many other problem types exist, we focus here on the multiple-choice problem type, its role in helping students learn physics, and the way this type of problem is perceived by graduate teaching assistants.

3.1.2 The role of the teaching assistant in student problem-solving

Physics graduate teaching assistants (TAs) are often employed, especially at large research universities, to carry out duties such as instructing the recitation/discussion sections related to introductory physics courses. TA professional development programs may be the only opportunity for groth as an instructor that TAs may have as potential future faculty. It has been noted that even though TAs are often responsible for a significant portion of undergraduate instruction, their training for this role is often limited [52, 53]. TAs may be responsible for choosing the types of problems to use with students, e.g., in designing quizzes for their students to take during recitation/discussion or creating example problems to discuss. Moreover, in particular, as potential future faculty, TAs may have an ongoing decision-making responsibility about the types of problems to use with their future students. Even though a small minority of physics departments in the U.S. provide semester-long TA professional development, the majority of physics departments provide only very short

training (i.e., a few hours) to prepare them for these various teaching responsibilities. Thus with limited opportunities for professional development and training in the intervening time between the TA role and the faculty role, TAs perceptions may also shed light on their future perceptions as potential faculty members. Moreover, the perceived instructional value may affect choices about the use of various types of problems and thus impact the degree to which different types of problems are exploited for their effectiveness in actively engaging students and their ability to facilitate instructional goals. As such, TAs' perceptions of different types of problems are worthy of examination to inform professional development courses and programs.

3.1.3 Focus of our research

In the study presented here, TAs in a TA professional development course were asked to reflect upon five problem types and their features that are appropriate for an introductory mechanics problem scenario. These same problem types had been used in an earlier study with physics instructors [39]. The example problem types were meant to generate a broader discussion and reflection upon the types of problems TAs might choose to use in their teaching. Here we focus on TAs' initial views, after some experience in their first semester as a TA, about a multiple choice introductory physics problem and investigate the following research questions: (1) How challenging and instructionally beneficial do TAs perceive a multiple choice type of problem? (2) If TAs had complete control of an introductory course, would they be likely to use multiple choice problems, and for what purpose might they use them? (3) Why do TAs perceive multiple choice problems the way they do?

3.2 BACKGROUND

3.2.1 Physics faculty views about different problem types

A prior study regarding physics instructors' views about different problem types in which they were presented with the same variations of a physics problem given to the TAs in the current study [39]. It was found that the instructors generally valued different problem types intended to develop expert-like problem-solving but they were not as likely to use certain problem types. In particular, instructors' views about multiple-choice problems for introductory physics were not typically positive–the majority of faculty reported that they would never or rarely use a multiple-choice problem, and that its only reported use was for high stakes exams [39]. Many faculty reported their reluctance to use multiple-choice problems was because it hindered their ability to monitor their students' thinking because they could not see their students' work. This finding regarding multiple-choice questions agrees with other research indicating that many faculty members never used multiple-choice questions, even for formative assessment purposes [21].

3.2.2 Prior work on TAs' professional development and their views about teaching and learning

Discussions with faculty members about TAs at the Graduate Education in Physics Conference jointly sponsored by the American Physical Society and the American Association of Physics Teachers suggests that introductory physics courses at large research universities typically employ graduate TAs whose responsibilities include e.g., grading of homework, quizzes, and parts or all of exams, as well as implementing and/or designing quizzes, examples, and other supplementary material in course recitations and/or labs [52]. Moreover, even though a small minority of physics departments in the U.S. provide semester-long TA professional development, the majority of physics departments provide only very short training (i.e., a few hours) to prepare them for these various teaching responsibilities. The majority

of conference participants further noted that little in the way of guidance or supervision is provided to TAs to support them in their teaching activities [52]. As such, TAs' teaching practices are often affected by the expectations of their supervising instructors and also by their beliefs and workload as graduate students [54, 55, 57, 58, 73, 59, 60, 61]. Thus, it is important to investigate TAs' views to inform their professional development activities.

Several studies have investigated TAs' views about teaching and learning [33, 49, 34, 35]. Prior research suggests that there are discrepancies between physics graduate TAs' perceptions of what teaching strategies are beneficial for students' learning and many of the findings of physics education research [33, 49, 34, 35, 36, 37]. For example, TAs have been found to struggle with the idea that effective grading practices can be a formative assessment tool, e.g., grading practices that encourage students to show their work can improve their learning from problem-solving and encourage them to learn from their mistakes [33, 49, 34, 36, 37]. Another study involving TAs' beliefs about example solutions provided to students shows that many TAs were unlikely to identify features in the problem solutions that the research literature describes as supporting learning goals for students [35].

It has been found that TAs' beliefs affect their teaching practices [54, 55]. Because of the role of the TA in decision-making on use of various problems, both in the TAs' current capacity and in possible future roles as faculty, their beliefs about the pros and cons of posing an introductory physics problem in different ways and in different instructional contexts can affect the ways in which they use various types of problems. Thus, identifying the views of the TAs' about the way in which a problem is posed can be useful in developing activities to improve their professional development and help them recognize the pedagogical value of posing the same problem in various ways to meet different instructional goals.

3.2.3 Multiple choice problems as a formative assessment tool

Multiple-choice questions may be used in both summative and formative assessments. Assessments that measure the extent to which students have learned and the goals of a course have been achieved at the end of a course, and nothing more, are called "summative." On

the other hand, "formative" assessment is an assessment in which both instructors and students receive feedback on students' understanding and their skills at a given point in time and there is opportunity to address student difficulties and help them learn those concepts better and improve their problem solving, reasoning and meta-cognitive skills. Formative assessments are often "low-stakes;" they are used frequently to actively engage students in the learning process, but have little impact on a student's final course grade. When multiple-choice problems are used as a formative assessment tool, they are often implemented in class as a low-stakes assessment and include "distractor" options among the given answer options. "Distractors" are choices that are meant to be selected by someone who is not knowledgeable about the correct answer and for good multiple-choice questions focus on the common student difficulties found via research [74, 75, 76, 77, 78, 79]. The presence of distractor choices reduces the chance that students can narrow down the correct answer based upon test-taking strategies rather than based on sound understanding of the content and problem-solving process [75]. If these distractor choices are based upon common student difficulties, the multiple-choice assessment can serve as a diagnostic tool to measure student understanding at a given point of time so that the instructor can address those difficulties using suitable pedagogical approaches [80].

There are many ways in which multiple-choice questions can be used in class to actively engage students and provide formative assessment feedback. For example, the use of multiple choice questions in the form of clicker questions has been shown to enhance students' conceptual and quantitative problem-solving and reasoning skills [81] even in very large classes. When clicker questions are combined with active discussion, a majority of students have been found to have a positive attitude about the usefulness and enjoyment of the clicker questions and recognize the role of those questions in supporting their learning, in addition to being able to co-construct knowledge with peers [82, 83, 84, 85]. If carefully sequenced, clicker questions have been shown to yield significantly higher conceptual understanding and help students feel actively involved in the learning process [86]. In using clicker questions during in-class formative assessment, instructors can award partial or full credit to students simply

for their participation in the clicker question activity (and not on whether the answers are correct). Such a way of handling clicker questions has been shown to enhance discussion among students and more accurately assess student understanding [87, 88].

Peer Instruction is an example of a method which utilizes conceptual multiple-choice questions as formative assessment tools and may positively impact students' conceptual understanding and problem-solving skills [81]. However, clicker questions may be conceptual or quantitative. Another application of multiple-choice questions in formative assessment is in the context of the "inverted" or "flipped" classroom, in which there are opportunities to incorporate in-class group problem-solving into lectures because the content which is normally covered in lectures to assigned videos and/or tutorials outside of the class [89, 90, 91]. Such in-class problem-solving could include quantitative clicker questions which are designed to reinforce the content learned outside of class and provide efficient low-stakes formative assessment of students' mastery of relevant material [89].

Low-stakes use of clicker questions has been found to reduce the achievement gap between underrepresented students (or students with lower levels of prior knowledge) and the majority students [92, 93, 69]. This is because while all students benefit, those from an underrepresented group and/or with lower level of prior preparation benefit disproportionately more compared with other students [92, 93, 69]. Thus, even in large enrollment classes, carefully-designed multiple-choice questions can be convenient for gathering efficient feedback regarding students' current knowledge and difficulties and can facilitate effective formative assessment.

Furthermore, interactive learning experiences can involve multiple-choice questions to keep all students actively engaged in the learning process. For example, in using interactive lecture demonstrations (ILDs), students are asked to predict what will happen before a demonstration takes place [94]. Quick feedback on students' predictions for such ILDs can be gathered, even in very large classes, via clickers, flashcards, or even a show of hands if they are asked for such predictions in the form of a multiple-choice question [95]. The ILDs have been associated with conceptual learning gains [94, 95, 96]. Multiple-choice questions may

also be used as part of self-paced learning tools, e.g., self-paced learning tutorials, to give instant feedback to students, and the results could also be followed up via in-class or online instruction that takes into account the students' difficulties found through their responses to the questions [97, 98, 99, 100, 101]. When students engage with such an approach to self-paced tutorials, student understanding is enhanced [102, 103]. Similarly, multiple-choice questions can help facilitate a "Just-in-Time-Teaching" (JiTT) approach wherein pre-lecture feedback is gathered just before instruction takes place and serves to guide the instructor in addressing student difficulties. Such ways of using multiple-choice questions as a low-stakes formative assessment tool have been found to enhance student learning [104].

3.3 METHODOLOGY

Participants and description of TA professional development program: A total of 97 TAs from a typical large research university participated in this study during 4 different years. Participants were physics graduate students who had teaching responsibilities (recitation or lab instruction, and a majority were also assigned to help students in a physics tutoring center) and were concurrently enrolled in a mandatory TA professional development course that met once per week for 2 hours for an entire semester. The TAs were expected to do approximately one hour of homework each week pertaining to the professional development course, in which various activities took place throughout the semester. During the course, initial activities related to grading practices occurred near the beginning of the semester, followed by discussions of pedagogy, including the use of tutorials and clicker questions as learning tools. After this, discussions turned to how different problem types (e.g., multiple choice problems, context rich problems, problems that are broken into sub-problems, and traditional textbook style problems) and different example solutions to problems can help students learn physics. TAs were involved in evaluating the effectiveness of multiple choice

questions on conceptual surveys, and predicting which choices students might pick. This gave TAs the chance to reflect on the design of conceptual multiple-choice questions, anticipate challenges their students might encounter, and make judgements about how effectively a well-designed question might uncover where students are struggling. TAs also were given a physics problem and asked to present the solution to the TA professional development class as they would in their recitations. These presentations were video-recorded so that they could reflect on their teaching and also receive feedback from other TAs and the instructor. Thus the problem-type activity was one of a number of activities all aimed at the professional development of the TAs while investigating their assumptions about physics teaching and learning.

Data collection tools and artifacts: The quantitative data collection tools consisted of five introductory physics problem types that had been developed previously [39] and served as a guiding example for the activities. The example problem types were designed for the same introductory physics problem scenario in mechanics. They included two different versions of a problem which was broken into sub-parts (one was framed in a more conceptual manner than the other), a multiple-choice problem, a context-rich problem and a traditional textbook version of the problem. Here we focus on the multiple-choice problem type (Problem B was the example problem given for reference) as can be seen in Figure 4. The dicussion regarding these types of problems was aimed at probing the TAs' views about the problem types in a general sense, using the example problems as an illustration of just one example of each type of problem.

The example multiple-choice problem is posed in a standard "textbook" style with choices that include common student difficulties as strong distractors. There is a note at the end of the problem that the TAs could see and which was pointed out in discussion, explicitly stating that the choices are based on common student difficulties. Discussion included explanation of how distractor choices could help reveal what students may be misunderstanding. While the multiple-choice example presented to TAs was a quantitative problem, the discussion in the TA professional development class focused on the merits of well-written multiple-choice

problems in general, including the use of qualitative multiple-choice questions geared towards probing conceptual understanding.

Based upon our research questions, the TAs were asked to answer questions about these problem types on a worksheet, a partial sample of which can be found in Figure 5. Among the entries shown in Figure 5, TAs were directed to list pros and cons for each problem type. Specifically, in the instructions, TAs were asked to list at least one pro and one con for each problem type based upon the features each of the five problem examples contained. Data were collected over four different years. In the most recent year's worksheet, TAs were also asked what they would change about the example problems. In addition they were asked to rank the features of problem types on their instructional benefit (i.e., how instructionally beneficial the TAs judged each problem type to be), and to rank the problem types in terms of the level of challenge (i.e., how difficult the TAs judged each problem type to be for students), how much they liked the problem types, and the likelihood that they would use the problem if they had complete control of the choice of the problem types to use. For example, a TA who ranked a problem 1 for "challenging" judged this problem to be the least challenging for students; a 5 for "challenging" indicates that the TA perceived it to be the most challenging among the five problem types. The rankings allowed us to investigate research questions 1 and 2. Furthermore, throughout all four years, TAs were asked to list pros and cons of the problem types. These pros and cons were useful for investigating why TAs ranked the problem types the way they did (research question 3).

Data collection in the TA professional development course and later in individual interviews: TAs were given the problem types and worksheets in the professional development course in the middle of the semester, when they had some teaching experience, in order to elicit their ideas about different problem types. They were asked to answer worksheet questions under the assumption that they had complete control over the introductory physics class, including control over problem types chosen for various purposes. The worksheet (see a part of the worksheet in Figure 5) was completed as part of a homework assignment. Later, 12 participants who had taken the TA professional development course

Problem B

You are whirling a stone tied to the end of a string around in a vertical circle having a radius of 65 cm. You wish to whirl the stone fast enough so that when it is released at the point where the stone is moving directly upward it will rise to a maximum height of 23 meters above the lowest point in the circle. In order to do this, what force will you have to exert on the string when the stone passes through its lowest point one-quarter turn before release? Assume that by the time that you have gotten the stone going and it makes its final turn around the circle, you are holding the end of the string at a fixed position. Assume also that air resistance can be neglected. The stone weighs 18 N.

A) 1292 N
B) 1258 N
C) 1248 N
D) 1210 N
E) None of the Above

Note: The choices are based on common student problems.

Figure 4: The multiple choice problem posed to the TAs

Problem Type	What are the **FEATURES** of the problem?	What are the **PROS/CONS** of the problem?	Would you use this problem as . . . a **HW** problem? Why or why not?	a **QUIZ** problem? Why or why not?	an **EXAM** problem? Why or why not?	a small **GROUP** problem solving session? Why/why not?	Are there **OTHER** situations in which you would use this problem? Explain your reasoning.
Problem B	Multiple Choice.	PROS: quick and efficient CONS: can't see the whole solution	No can't see the whole ideas	Yes. quick	No can't assess the students ability by this problem they may randomly chose one.	No. can't see the whole solution.	

Figure 5: Part of a sample worksheet given to the TAs

Code	Definition	Examples	Percentage of TAs
(Pro) check	students can check their answer	"can check your answer"	22
(Pro) time	saves time	"quick way to get answer"	14
(Pro) grade	easy/efficient to grade	"makes problem faster to grade"	9
(Con) no partial credit/no understanding shown	can guess; does not demonstrate understanding	"can't tell if they can do the process, can't prevent guessing"	54

Table 2: The most commonly listed pros/cons of a multiple choice problem and the percentages of TAs who listed them. Some TAs listed more than one of the following or other pros/cons not listed here.

earlier volunteered to be interviewed in a one-on-one setting using a think-aloud protocol. These interviews took place at least one semester after the initial activity described here in the TA professional development course and were audio-recorded. TAs who participated in the interviews were asked questions both about the example problems and about the problem types, in general (similar to the broader in-class discussion about the instructional benefits and pros and cons of well-designed multiple-choice problems). Thus, these interviews served to more deeply probe the TAs' reasoning behind their written responses and to explore such questions as the use of the multiple-choice format as general problem features.

Coding TA responses: Two of the researchers met weekly to identify appropriate coding categories for pros/cons; agreements on these were reached through discussion. The researchers used open coding of the data from the individual homework assigned in the middle of the semester regarding the TAs' views of the problem types. The categories coded were arrived at over several weeks based on emergent themes. Some categories were merged if they were found to be sufficiently similar. The inter-rater reliability was examined for the coding of the pros/cons in the year 3 data set, and the average Cohen's kappa [70] was calculated to be $\kappa = 0.982$. The most common pros and cons for the multiple choice problem, along with their definitions and examples from TA worksheets, are included in Table 2.

3.4 RESULTS

3.4.1 Multiple choice problem ranked low for instructional benefit and use

Figure 6 summarizes the ranking for the multiple choice problem type in all four categories. In the category of "instructional benefit" TAs were asked to rank the problem *features* and the example problem types themselfs. In response to research question (1), the average ranking is 2.0 for "instructional benefit," which indicates that the multiple-choice problem type is seen by TAs as low for instructional benefit. In particular, this ranking for "instructional benefit" is the second-lowest of all the problems ranked in this category. The average ranking of 3 for "challenge" represents a moderate challenge level. This ranking is best understood in the context of the other problem types that the TAs were asked to rate. Two types of broken-into-parts problems were consistently ranked as the easiest two problems, and the context-rich problem was ranked as the most challenging. The multiple-choice problem and a standard "textbook" style problem were ranked in-between these two extremes. In response to research question (2), the average ranking of 2.5 for "use" is the second-lowest of all problems, and TA responses regarding the manner in which they would use a multiple-choice problem indicate limited ways in which TAs would use such problems. In response to research question (3), we examined the pros and cons that the TAs listed, as well as their explanations of how they might use a multiple choice problem shed light on the rankings they gave to this problem.

It is important to note that while written responses were guided by the five example problem types presented to TAs, interviewed TAs were asked to think not only about the specific example of a multiple choice problem that they were given, but also to think about multiple choice problems generally-speaking, and to imagine any possible format for a multiple choice problem. The interviewed TAs responded in a similar manner to those who gave written responses for all of the preceding categories (i.e., "instructional benefit," "challenge," "like," and "use") even when asked more generally about multiple choice problems.

In other words, rather than the above rankings holding true only for the specific multiple choice problem given, interviewed TAs saw little instructional benefit and were unlikely to use multiple choice problems of any kind. Thus the rankings appear to generalize to multiple choice in formats other than the problem seen in Figure 4 (e.g., a conceptual multiple choice question may be regarded by TAs in a similar manner).

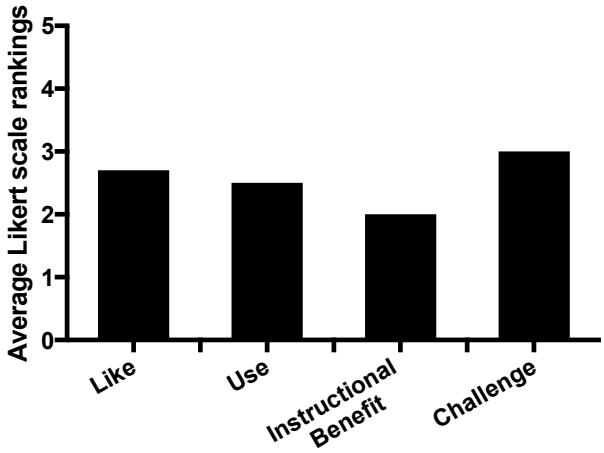


Figure 6: Average rankings of the multiple choice problem

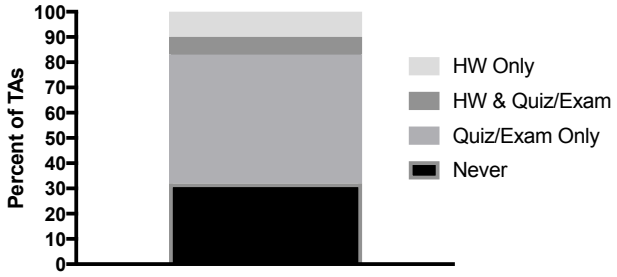


Figure 7: TAs' reported usage of a multiple choice problem type. Yellow indicates the TA would only use this type of problem in homework. Green indicates they would use it in homework or during a quiz or exam. Blue indicates they would only use it in a quiz or exam. Red indicates they would *never* use it for any purpose.

3.4.2 TAs viewed multiple choice problems primarily as a summative rather than formative assessment tool

As seen in Figure 6, the TAs ranked the multiple-choice problem type low for "use." Reasons behind TAs' perceptions about the use of multiple-choice problems can be found by examining their written responses on the worksheet regarding how they would use such a problem. These usages are summarized in Figure 7.

Figure 7 reveals that many TAs would never use a multiple-choice problem at all. Moreover, of those who would use a multiple-choice problem, the majority reported that they would use it only for quizzes or exams. Interviews and written responses suggest that the TAs were hesitant to use multiple choice problems and, even if they would use them, they had a summative assessment view of such problems in mind. One TA explained his reluctance to use multiple choice problems in an interview as follows: *"I think multiple choice takes the focus away from the problem at hand...This introduces more anxiety and confusion...I remember [one multiple-choice test] and I didn't do well, because I was so overthinking all of my answers and changing them multiple times."* This TA identified students' anxiety over getting the correct answer and evaluating the validity of the choices given as a reason for avoiding using multiple-choice problems. It was apparent from the discussion during the interview that this TA was thinking about a high-stakes, summative use of multiple-choice problems, and the impression that those situations left on him as a student for being anxiety-producing. Similarly, when other TAs also spoke about using a multiple-choice problem, their responses focused on a summative assessment type of usage. One TA, when explicitly asked how he might use a multiple choice problem, said, *"Maybe if I want to make the final totally multiple choice,"* but did not offer any other type of use for such a problem. Similarly, another TA, when asked the same question, offered, *"Quiz or exam type things, not homework... I don't think it's necessarily helpful in a homework to have the multiple choices."* This TA also did not offer any other way of using a multiple-choice problem, even when asked explicitly. The main appeal to multiple-choice problems for this TA appeared to be to help him with the task of grading the quiz or exam efficiently.

Indeed, few TAs mentioned using a multiple choice problem as a formative assessment tool such as "clicker questions," pre-lecture electronic assessments, in-class group problem-solving in flipped classes, or self-paced tutorials which can be used even in very large classes (all TAs were aware of clicker questions and other technology since it was discussed in an earlier session in the TA professional development class). It is also important to note that TAs were specifically asked to assume they had full control over teaching the course independently when responding on their worksheets regarding problem types and their instructional benefits, etc. (and this point was also emphasized several times). Thus, the absence of possible formative uses in their responses as well as interview data suggest that the TAs are overlooking these possibilities for overall course instruction, not due to their limited roles as TAs. A lack of valuing of the multiple-choice clicker questions was observed in the comments made by the TAs during the interviews. For example, when the interviewer specifically asked about use of a multiple-choice question with clickers, one TA stated: *"I don't even think about clicker questions...."* Discussions with this TA suggest that he was well aware of the fact that clickers can be utilized in physics classes, but he did not appear to have reflected on the instructional value of using multiple-choice clicker questions as a formative assessment tool.

3.4.3 TAs thought of distractor choices as a "trap"

Other reasons for not using multiple-choice questions included some TAs feeling that such a question constituted a "trap." This theme of "trapping" students is evident in the following interview quote by a TA discussing a multiple choice problem in general: *"If I were to do a problem and get one of the answers... then you look at the others and you doubt yourself and you get started thinking about patterns in the question and what you're off by, or is it 'none of the above?' Is it all a trick?"* This TA expressed the idea that the alternate choices provided may throw a student off by the "patterns" that appear to them to exist in these alternative choices. Likewise, another TA explained that *"Having additional likely answers is majorly serving as a trap and is evil and malicious."* This TA used the strong words "malicious" and "evil" in describing the distractor choices acting in a way he saw as trapping students.

Other TAs expressed similar sentiments. The concern for perceived fairness is definitely laudable, but interviews suggest that it arises from TAs assuming a high-stakes assessment in which an incorrect answer may have a major impact on a student's grade. Interviews suggest that the assumption of high-stakes assessment often influenced the idea that the distractor choices were seen by some TAs as a "trick" rather than having instructional value in assessing whether the student was truly knowledgeable about the correct concepts and problem-solving approach.

3.4.4 TAs did not view multiple choice problems as reflecting student understanding

The pros and cons listed by the TAs in written responses regarding the uses of the multiple-choice problem type as well as the interview data were examined for possible reasons for why the TAs ranked the multiple choice problem type the way they did and for further analysis of why many TAs were not likely to use such a problem. Table 2 shows the most common pros and cons mentioned by TAs in written responses. As can be seen in Table 2, a majority of TAs listed the con "no parital credit/no understanding shown." This con referred to TA responses that indicated that they felt that this problem would be a poor reflection of students' understanding, that the students may not get partial credit, and/or that the students could potentially guess the correct answer. As such, this category encompasses several reasons why TAs might be concerned that the multiple-choice question is not necessarily an accurate measure of students' understanding and/or would not necessarily give students appropriate credit for their level of understanding. The themes of these reasons appeared to center around empathy for students and concerns about fairness in grading.

TAs were concerned that students may guess the correct answer to multiple choice questions: The issue of potentially guessing the correct answer to a multiple choice question was mentioned by many TAs. For example, in an interview, one TA explained the concern about guessing as a reason for why this type of problem may not be valuable: *"There's always a possibility that they could make a guess, and I don't really see any particular value to*

using multiple choice....” The possibility of guessing appeared to influence this TA’s regard for the problem’s perceived value. Guessing is certainly a valid concern and it is encouraging that TAs were concerned with the implications for fair grading when guessing is a possibility. However, it appears that TAs did not think more deeply about how distractor choices or other elements of a well-designed multiple choice problem might discourage guessing (e.g., consistent units and similar numbers might not allow a student to easily “rule out” an answer choice). It is interesting to note that sometimes TAs were reporting concerns that could be considered to be contradictory. In particular, TAs often disliked the idea of guessing, but also disliked the idea of using distractors which can dissuade guessing because they felt the distractors constituted a “trap.” This suggests a possible lack of reflection and/or a lack of awareness of the instructional value of design elements such as distractor choices. It also appeared that TAs did not think about how the formative assessment value of such questions could outweigh the risk of guessing for a low-stakes assessment use, such as clicker questions or as assessment in self-paced learning tools.

TAs disliked the idea of not being able to give students partial credit: In addition to guessing, some TAs mentioned the issue of partial credit. In an interview, one TA expressed this concern as follows: *“If they did good work, and they chose b instead of a, and they got zero points, I wouldn’t like that as a TA. I expect myself to be someone who grades on problem solving merit, and the thought process as opposed to the final answer”*. This TA was empathic in his concern about a student potentially getting a score of zero for a multiple-choice problem, which is commendable. Interviews suggest that such concerns were often based on an assumption that multiple-choice questions would be given in a high-stakes summative assessment. The concern for fairness that this TA and others showed is laudable, but it appears that he did not realize that partial or full credit may be built into low-stakes formative assessments (as is often done, e.g., with clicker question responses). Other TAs expressed similar sentiments about the negative aspect of students not receiving partial credit. Many TAs assumed that multiple-choice problems necessarily mean that credit is not possible if the answer is not correct. However, credit possibilities do exist, such

as participation or completion credit for clicker questions or in online assessments as part of a self-paced learning environment.

TAs felt it might be possible for a student to use incorrect methods to arrive at a correct answer: Other TAs stated that they were concerned that, even if the student was not simply guessing, a correct answer may not indicate that the student understood the problem. One TA illustrated this concern in an interview, when he described what he feared a student might do: *"Maybe just use this number and divide it by any of these numbers and see if any of them gives you an integer, and just choose that one. Even though you know what the student's answer is, you don't know how they got it."* This TA indicated concern about the potential for students to simply plug in numbers based upon the given values and choices and look for a choice that fits rather than truly understanding the physics behind the problem. In a similar vein, another TA stated that he would not use multiple-choice problems because: *"Students can use other methods to arrive at the answer rather than the desired physics."* Once again, this concern is very legitimate for multiple-choice questions in general and indicates the importance of fairness in grading that TAs value. However, interviews suggest that, even when asked to think of any possible multiple choice problem, TAs appear not to have reflected upon the fact that well-designed multiple-choice problems can play a key role in low-stakes formative assessment and that good distractor choices may inhibit students from arriving at the correct answer unless they were truly using the correct methods.

3.4.5 TAs mostly cited practical matters as pros for using multiple choice problems

Although none of the individual pros were mentioned by a majority of TAs (i.e., the most common pro was only mentioned by 22% of TAs), the common pros included several sentiments that appeared to center around practical matters. None of the pros appear to relate to instructional benefits such as the use of multiple-choice questions as a formative assessment tool. In fact, the most common pros had a theme of practical considerations. The perceived

time-efficiency, the ease and quickness of grading, and the idea of students "checking their answers" suggest more utilitarian concerns than perceived instructional benefit. This finding suggests that the positive aspects TAs saw in the multiple-choice problem did not have to do with such problems benefiting student learning when used as a formative assessment tool. Furthermore, it should be noted that none of the pros listed in Table 2 were stated by even one-quarter of the TAs. This is because almost 40% of TAs did not list pros at all, even when explicitly asked to list at least one pro. Other pros that were cited were mentioned by even smaller numbers of TAs than the percentages in Table 2. This suggests that TAs struggled to see value in the multiple-choice problem, and even when they did, this value was more pragmatic than instructional in nature.

TAs viewed multiple-choice problems as easy to grade: Some TAs expressed that ease and/or efficiency of grading a multiple-choice problem was a pro (and most often the only pro). The researchers coded these responses in the category "grade." As an example of this sentiment, one TA explained the pro in an interview as follows: *"I think this problem is easier to grade."* This TA mentioned the ease with which a multiple-choice problem can be graded as a positive aspect of the problem. As mentioned above, the code "grade" included any time-savings on the part of the TA in grading the multiple-choice problem. A TA response indicating this time-efficiency when it comes to grading was *"Time saving to check."* Both the ease and efficiency of grading a multiple-choice problem were commonly considered as pros by the TAs. One TA, who mentioned in an interview that the multiple-choice problem had the perk of being easy to grade, acknowledged this as a practical issue: *"It's not a nice answer but it's pragmatic."* Practicality had obvious appeal to the TAs, but did not imply any deeper instructional benefits of multiple choice problems as a formative assessment tool.

TAs viewed multiple-choice problems as saving time for students despite requiring the same problem-solving process as other types of problems: Another common pro listed by the TAs was coded in the category "time." In this case, "time" means that the TAs felt that a multiple-choice question would be time-efficient or would save time

for the student. We note that the pro "time" specifically applies to only TA responses that indicated that the time-savings would be for the student (as noted earlier; responses that indicated that the multiple-choice problem would be faster or easier to grade, were coded as "grade"). For the pro "time," TAs explained that not needing to show work in order to select a choice would make the problem-solving faster. Expressing the sentiment of time-savings, one TA succinctly noted that the multiple-choice problem would be *"[A] quick way to get answer."* Similarly, another TA stated that it *"Saves time to get correct answer."*. This perceived time-efficiency was often cited as helpful for a quiz situation where time might be limited. One TA expressed this sentiment by stating: *"Saves time for a quiz."* However, the fact that the multiple-choice problem posed was quantitative, involving the same problem to be solved as other problem types (that were posed in other ways) implies that the problem-solving process may still require a comparable amount of time as analogous problems posed in other ways. This was a point not often noted by TAs.

TAs claimed that multiple-choice problems allow students to check their answers even though distractor choices could inhibit them from doing so: The pro category titled "check" refers to TA responses that indicated that they viewed it as feasible to check the correctness of one's answers in the multiple-choice question. For example, one TA explained that having the choices present means that *"[Students] can see that their answer might be correct."* A similar idea was expressed by another TA who stated that: *"If the student makes a calculation error, they will know right away."* And yet another TA simply stated that in a multiple-choice question *"You can check your answer."* TAs such as these regarded this feature as a positive aspect of the problem–they felt that the students could benefit from checking if they had the correct answer. However, TAs did not express or recognize that the presence of common incorrect answers included in the choices could preclude students from checking that an answer was correct, since they may simply be verifying an incorrect answer choice (one which is based on a known common student misunderstanding). It appears that TAs did not look past the obvious feature of one of the choices being correct to realize that carefully-designed distractor choices would make it difficult to check one's

answer.

3.5 DISCUSSION

We find that, in response to our first two research questions, in general, TAs viewed a multiple-choice problem as moderately challenging, but did not perceive it as instructionally beneficial. TAs reported that they would not be likely to use a multiple-choice problem, except in test situations. Both in the written responses and in the interviews, TAs did not mention multiple-choice questions as playing an important role in formative assessment, viewing multiple-choice questions as problematic for various reasons. In response to research question (3), we found several reasons for why TAs viewed multiple-choice problems the way they did. Concerns TAs mentioned were very legitimate, e.g., one does need to think about the possibility of a student guessing or shortcutting a systematic problem-solving approach with all steps shown, when students need only choose an answer. TAs also expressed appropriate concern for fairness when explaining their desire that students receive appropriate credit for good understanding. The majority of TAs were concerned that the results of multiple-choice questions could not be trusted to reliably gauge student understanding or give students fair credit. It is commendable that TAs were thinking about being sure their students understand the concepts and can properly execute the problem-solving process as opposed to simply thinking about the final answer.

However, TAs did not readily identify the use of multiple-choice problems as a low-stakes formative assessment tool, where this type of use lends itself to efficiently identifying the common difficulties students have and provides an opportunity for students to reflect on their own difficulties. It is telling, for example, that a significant number of TAs reported that they would never use a multiple-choice question in any way, and that, among those who would use it, they primarily had summative assessment and practicality in mind. The idea of

utilizing the information one could gather from a multiple-choice question, e.g., using clickers in a large class, about the percentage of students who chose a common incorrect answer to address common difficulties appeared not to occur to the vast majority of TAs. Likewise, TAs did not mention any of the plethora of ways that multiple-choice questions can be used in low-stakes formative assessment, even in very large classes. This use as a formative assessment is not only restricted to clicker questions, which are very versatile for both individual or group questions and both conceptual or quantitative problems, but such questions can also be integrated with interactive lecture demonstrations, self-paced tutorials, pre-lecture videos and corresponding assessment tools and other "Just-in-Time-Teaching" strategies. Even when explicitly asked for at least one pro for the multiple-choice question, the power of such questions as a formative assessment tool was overlooked by a majority of the TAs. Since the TAs were specifically told to assume they had complete control of a hypothetical introductory physics class, and since interview data in which TAs were specifically asked to think more generically about any multiple-choice type of questions corroborates these findings, this lack of recognition of multiple-choice questions as a formative assessment tool appears to suggest a possible oversight. These results agree with prior research into the grading practices used by TAs, in which it was found that TAs struggled with the idea of recognizing that grading practices can play a role in formative assessment [33, 49, 34].

Interviews suggest that the assumption of high-stakes summative assessment rather than low-stakes formative assessments appears to be the main driving force behind the TAs' concerns about utilizing multiple-choice questions with their students under any circumstance. TAs' concerns could be addressed if they were asked to reflect upon the wide range of useful ways in which multiple-choice questions could be implemented as effective formative assessment tools and were made more aware of the ways in which design and implementation elements of multiple-choice questions can make them excellent low stakes assessment tools (such as the benefit of using good distractor choices based upon research on student difficulties, and awarding students participation credit regardless of the correctness of their answers, etc.) [81, 97, 98, 99, 93, 69, 105]. For example, the use of good distractor choices or

sequences of multiple-choice questions which build on each other can help reinforce concepts, encourage a desirable problem-solving approach, and/or dissuade guessing. Yet, to the TAs who discussed the presence of the distractor choices in detail during interviews, the instructional value of these distractors did not appear attractive. In fact, some TAs expressed serious misgivings that these served to simply maliciously "trap" a student. Concern for fairness is certainly a very positive aspect of instruction, but it appears that the TAs had an underlying assumption of a high-stakes assessment use of multiple-choice problems and a lack of awareness of low-stakes formative assessment and the instructional value of distractor choices (e.g., you can award full credit for clicker questions for participation only regardless of the correctness of the answer).

With regard to the instructional benefits of multiple-choice problems, TAs identified either pragmatic benefits, such as making grading easier for them, and/or benefits that were sometimes not real. For example, unlike what many TAs thought, students cannot really check their answers easily when distractors are present, as is the case in the multiple choice question in the activity discussed here, since they might simply be verifying an incorrect answer. Also, although many TAs perceived it to be the case, the time required to solve a multiple-choice problem is not necessarily significantly less than that required for solving other types of problems, especially for a quantitative problem, since one must solve the problem entirely whether or not there are answer choices present (to check which choice is consistent with the answer obtained).

It is important to note that much of the written data was guided by a specific example of a multiple-choice problem; however, interview data was meant to elicit more general perceptions of multiple-choice problems. Moreover, the interview data agreed with the written data. Thus, while an important limitation to our findings might be the ability to generalize our results since a single multiple-choice problem was given as an example, the agreement with the interview data suggests that the written results may hold true for TA perceptions of multiple-choice problems in general. In addition, some aspects of the written data suggest some ability to generalize the findings. For example, the addition of a question asking TAs

what they would change about the problems in the most recent year's worksheet revealed that no TAs mentioned, for example, changing the problem to a conceptual question or otherwise revising it for use as a formative assessment tool such as a clicker question. Likewise, the most recent year's worksheet included "instructional benefit" rankings that instructed TAs to rank the *features* of the problems first before ranking the example problems, and the feature rankings agreed with the rankings found when TAs were considering the specific examples given. In other words it was the *feature* of multiple choice that appeared to rank low for instructional benefit in the opinion of the TAs. While the written data that contain these additional questions is only present for the most recent year, it is reasonable to assume that the subset of TAs who responded to these additional questions is representative of the larger sample. Assuming this to be true, it would appear that the written data suggest some evidence of a possible oversight among the TAs in our study regarding the possibile benefit of multiple-choice questions, generally speaking, as a formative assessment tool.

Our results are consistent with previous findings on faculty views and stated uses of multiple choice problems, in which faculty reported that they did not value multiple-choice problems and that either they would not use multiple-choice problems at all or they would only use them for practical reasons in test situations for convenience in high-stakes summative assessment [39, 21]. Like the faculty in the previous study which used the problem types in our study, we find that TAs assumed that multiple-choice questions offer only practical benefits for time-saving in high-stakes summative assessment. In addition, although the issue of formative assessment was not explicitly addressed by the study on faculty members' views of problem types, the TAs' written and interview responses in our study strongly suggest a lack of awareness of the range of ways in which well-designed multiple choice assessments can be effectively used as a formative assessment tool. As individuals who may be responsible for choosing problems for use in the classroom and as potential future faculty members, TAs' complete oversight regarding using multiple-choice problems as low-stakes formative assessment is an important finding with implications for informing TA professional development programs.

3.6 SUMMARY AND IMPLICATIONS

Overall, TAs had reasonable concerns about multiple-choice problems. Constraints such as guessing or limited student work shown can certainly be a concern with multiple-choice problems, as the TAs pointed out. However, the TAs overlooked the substantial possible instructional benefits of using multiple-choice problems for a wide range of purposes other than high-stakes, summative assessment. The instructional benefits offered by multiple-choice questions transcend the pragmatic pros listed by TAs in this study. Indeed, such problems can be a powerful low-stakes assessment tool for use in classrooms of all sizes, e.g., through the use of clicker questions and in self-paced learning environments. Both written responses and interviews suggest that these possibilities were not commonly recognized by TAs.

A limitation to our findings is that even though it was explained that the TAs should reflect on the instructional benefits and use of well-designed multiple-choice problems in general, they were given one example of each problem type. In addition, the TAs who participated in this study were from a typical large research university in the United States, so another limitation to our findings is that they may not apply to other institutions which are different. However, leaders of TA professional development programs at similar universities can use the findings of this study to help TAs reflect on the benefits of multiple-choice problems as a formative assessment tool to aid students in the development of problem-solving expertise. For example, professional development programs can help TAs reflect on the plethora of potential formative assessment uses of multiple-choice questions, such as "clicker questions," in-class group problem-solving, self-paced learning in an online environment, "Just-in-Time-Teaching," etc. In addition, TAs could be challenged to take their good ideas to the next step by thinking of ways to address their concerns and think "outside the box" to identify ways to use a multiple-choice problem as a low-stakes formative assessment tool rather than assuming that the purpose of such problems is in high-stakes summative assessment.

4.0 GRADUATE TEACHING ASSISTANTS' VIEWS OF BROKEN-INTO-PARTS PHYSICS PROBLEMS: PREFERENCE FOR GUIDANCE OVERSHADOWS DEVELOPMENT OF SELF-RELIANCE IN PROBLEM-SOLVING

4.1 INTRODUCTION

The desired learning goals for students in many introductory physics courses often include learning physics concepts and developing expert-like problem-solving skills [38, 40, 41, 42, 43, 45, 78]. The cognitive apprenticeship model can serve as a useful model to support these goals. In this field-tested framework, learning takes place through a guided process in which students gradually develop self-reliance in solving problems on their own. To facilitate this process, the cognitive apprenticeship model includes three aspects: modeling to demonstrate the criteria for good performance in problem-solving, coaching and scaffolding to provide immediate feedback as students solve problems, and weaning to build autonomous expert-like problem-solving ability [13].

Introductory physics students, who are often novice problem-solvers, frequently may struggle with solving problems in a systematic way. Instead of carefully analyzing and planning the solution, students often attempt to skip straight to implementing a solution by browsing a formula sheet [22]. In other words, many introductory students often employ a "plug and chug" method of searching through equations and formulas to find one that appears to have the same variables as given in the problem instead of starting with a

careful conceptual analysis of the problem. A more expert-like, efficient problem-solving approach involves systematically analyzing a problem and planning a solution path (including decomposing the problem into sub-problems), implementing the solution plan, and checking the results [22, 24, 64, 65, 23]. To develop expertise in problem-solving, students can benefit from being explicitly instructed on how to use effective problem-solving strategies [24, 64, 65, 23, 6, 106]. In particular, studies have shown that when students are deliberately taught to follow a systematic problem-solving approach, they outperform students who are not taught to solve problems in a systematic manner on challenging follow-up problems [64, 65, 23, 6].

Moreover, different problem types, i.e., different ways in which a physics problem is posed, can facilitate various aspects of the cognitive apprenticeship model, e.g., helping students develop expert-like problem-solving skills [100] and learn physics [107]. For example, a problem that is broken into parts may be useful in modeling and coaching in expert-like problem-solving approaches. Alternatively, a problem type that provides less support can help with the weaning aspect if used after modeling and coaching, because it provides opportunities for students to develop self-reliance in expert-like problem-solving.

Because different problem types can support different aspects of the cognitive apprenticeship model, the use of such problem types in physics courses can impact the effectiveness of instruction and, ultimately, student learning. Making choices about the use of different problem types in various instructional situations often is one of the responsibilities of both faculty and graduate teaching assistants (TAs). These choices may depend upon the perceived instructional value and constraints that posing a problem in a certain way may offer. Therefore, it is important to understand the views of those responsible for making decisions about which physics problem types to use in their introductory physics courses. The perceived pros and cons of posing an introductory physics problem in different ways and in different instructional contexts can inform activities designed to improve professional development efforts and to help ensure reflection on and recognition of the pedagogical value of posing the same problem in various ways.

In the study presented here, we focus specifically on the views of physics graduate student TAs about posing problems in broken-into-parts format (problem posed had sub-problems). In particular, TAs in a professional development course were asked to reflect upon five problem types for the same introductory mechanics problem scenario in which two of the five problem types were broken-into-parts problems. Here we summarize the TAs' views about two broken-into-parts problems and investigate the following research questions: (1) How challenging and instructionally beneficial do TAs perceive the broken-into-parts physics problems to be? (2) If TAs had complete control of an introductory physics course, would they be likely to use a broken-into-parts problem? (3) Why do TAs perceive the broken-into-parts problems the way they do?

4.2 BACKGROUND

4.2.1 Broken-into-parts problems and the development of expertise in problem-solving via the cognitive apprenticeship model

Many introductory physics students use novice-like approaches while solving physics problems [23, 22]. It has been found that, without explicit guidance, novices employ a problem-solving process which is not systematic and display an underlying knowledge structure that is fragmented and not well-organized [22]. By contrast, experts employ a systematic problem-solving process and have an underlying knowledge structure that is connected and organized in a hierarchical manner [22]. Organizing their knowledge allows experts to reduce their cognitive load during the problem-solving process and solve problems more effectively and efficiently [22, 108, 23].

In order to help students develop more expert-like problem-solving approaches, it is beneficial to give them explicit instruction in organizing their problem-solving process in a systematic way [23, 64, 65, 24]. This systematic problem-solving approach begins with

carefully analyzing the problem and planning the solution before embarking on the implementation of the plan. Such a systematic approach can be particularly challenging for many introductory students who often resort to a "plug and chug" method of searching for an equation or formula specific to the problem at hand without deeper contemplation and putting given values into the formula in the hopes of obtaining a correct final answer. The deliberate act of performing a careful conceptual analysis and decomposing a problem into more manageable sub-problems that can facilitate the problem solving process does not often come naturally for novice problem-solvers. Therefore, they will benefit from guidance and scaffolding support in learning to use these effective problem solving strategies explicitly [22, 23]. Once a problem solution plan has been constructed, students can then implement the plan and ultimately reflect upon the problem solving process and check the validity of their solutions [23]. When students are explicitly taught such a systematic approach, they perform better than students with similar prior knowledge who are not explicitly taught to follow a systematic approach on increasingly challenging problems [23, 64, 65].

Students' expertise in physics problem solving can be developed by explicit emphasis on using a systematic approach, making use of the cognitive apprenticeship model. All aspects of the cognitive apprenticeship model (i.e., modeling, coaching and scaffolding, and weaning) are crucial to the development of expertise in problem-solving [109, 110]. The choices that instructors make on the types of problems to use with their students may allow for different aspects of the guided process to unfold and can also facilitate different areas of learning a systematic problem-solving approach. For example, students need opportunities to see a systematic approach modeled for them so that they can develop an understanding of what is required in solving problems in an effective way (criteria of good performance). They also need to receive coaching and scaffolding support through the process of systematic problem-solving so that they can practice problem-solving while receiving immediate feedback on how to improve. Problems which provide built-in support and/or modeling, e.g., broken-into-parts problems, may be beneficial for the modeling and coaching aspects of student learning. However, after modeling, coaching and scaffolding, students also need opportunities

to experience removal of the support so that they can be weaned into more independent execution of a systematic problem-solving approach. For the weaning aspect, when self-reliance is being developed, problems which provide less in the way of built-in support can be useful.

Broken-into-parts problems can provide opportunities for modeling and coaching students through a systematic problem-solving approach. Specifically, broken-into-parts problems can be used as a model for students to learn how to decompose a problem into smaller sub-problems and provide them support in managing a complex problem. In addition, broken-into-parts problems afford the opportunity for instructors to provide coaching, scaffoding and immediate support since such problems allow for student difficulties at each step to be readily identified. Then, instructors can provide targeted feedback to students. Thus, broken-into-parts problems are one way to help students adopt a systematic problem-solving approach by modeling the process and coaching them in how to proceed in breaking a problem into sub-problems.

Nevertheless, to ensure development of expert-like problem-solving skills and help students become good problem-solvers, the weaning stage of the cognitive apprenticeship model is also important [13]. In particular, while broken-into-parts problems can serve as a model for decomposing problems into sub-problems and can provide opportunities to coach students, they do not readily offer the opportunity for students to engage in the conceptual analysis and decomposition process of a problem into sub-problems independently. Thus, this type of problem is not effective for the weaning aspect of the cognitive apprenticeship model, an aspect that is crucial for helping students develop self-reliance in solving problems. If students are mostly given problems which are broken into parts, they will not have many opportunities to practice decomposing problems into sub-problems on their own to gain problem-solving independence. Therefore, for the weaning aspect of developing problem solving skills, other problem types for the same physics scenario can be more beneficial [64, 65]. To conclude, while broken-into-parts problems can play an important role in the development of students' problem-solving skills, other problem types that provide less support

are beneficial in helping students develop self-reliance in problem-solving.

4.2.2 Physics faculty views about broken-into-parts problems

A prior study was conducted about physics instructors' views regarding different problems in which they were presented with the same problem types (including the broken-into-parts problems) given to the TAs in the current study [39]. It was found that the instructors generally valued different problem types intended to develop different aspects of expert-like problem-solving but their reported use of different problem types in their classes did not always reflect their beliefs regarding the instructional benefits of various problem types. Instructors had differing opinions about the merits of the broken-into-parts problem type. More than half of the instructors felt that it was important to lead students through a problem by breaking it up into sub-problems for the student, while slightly less than half of the instructors felt that students benefit from not providing such a guide. Nevertheless, the majority of instructors reported widely using broken-into-parts problems in homework, quizzes, and exams, even if they had reservations about such problems, stating that using such problems would help avoid stressful situations for students [39].

4.2.3 Prior work on TAs' professional development and views about teaching and learning

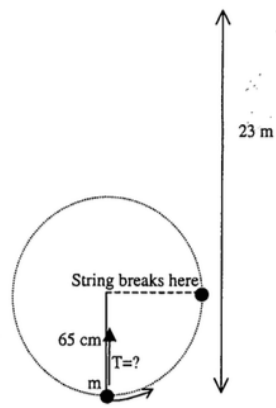
Physics graduate students are often employed as TAs, especially by large universities. Their tasks usually involve helping with grading, instructing recitations, and implementing and/or designing quizzes, examples, and other materials for introductory physics students. Despite the widespread use of TAs and their role in teaching students, limited training is typically provided for TAs [52]. Moreover, TAs often carry out their responsibilities without much guidance or support [52]. Therefore, teaching practices used by TAs are often affected by their workload and their prior beliefs about learning and teaching [55, 56, 57, 59, 30, 31, 62]. Additionally, TAs are potential future faculty members whose teaching roles could expand

in time. For these reasons, it is important to investigate TAs' views in order to inform their professional development activities related to teaching and learning.

Several studies have investigated TAs' views about teaching and learning [33, 49, 34, 36, 37, 35]. Prior research suggests that there are discrepancies between physics graduate TAs' perceptions of the kinds of teaching strategies that are beneficial for student learning and many of the findings of physics education research. For example, TAs have been found to struggle with the idea that effective grading practices that encourage students to show their work can improve their learning from problem-solving and help their students to learn from their mistakes [33, 49, 34, 36, 37]. Moreover, the issue of showing work in problem-solving has particular relevance for introductory student learning [36]. For example, it has been found that, while grading, many TAs do not require that introductory students show the steps of their solution or explain why they are using certain concepts. On the other hand, they *do* expect advanced students to show their work and explain their steps [36, 37]. Another study involving TAs' beliefs about the type of example solutions provided to students shows that many TAs were unlikely to identify features in the problem solutions that the research literature describes as supporting learning goals for students [35]. Because TAs are often responsible for deciding which types of problems to use, both in their current teaching appointments and in possible future roles as faculty, their beliefs about the pros and cons of posing an introductory physics problem in different ways and in different instructional contexts can affect how those types of problems are ultimately used. Thus, identifying the views of the TAs about the way in which a problem is posed can inform TA professional development programs. Here, we describe the findings of an investigation focused on TAs views about the pros and cons of two introductory problems that are broken-into-parts and involve the same physics scenario.

Problem A

A 1.8 kg mass is attached to a frictionless pivot point and is moving in a circle at the end of a 65 cm string. The string breaks when the mass is moving directly upward and the mass rises to a maximum height of 23.0 m. What is the tension in the string one-quarter turn before the string breaks? Assume that air resistance can be neglected.



A) What velocity, v_1, must the stone have when released in order to rise to 23 meters above the lowest point in the circle?

B) What velocity, v_o, must the stone have when it is at its lowest point in order to have a velocity v_1 when released?

C) What force will you have to exert on the string at its lowest point in order for the stone to have a velocity v_o?

Problem D

You are whirling a stone tied to the end of a string around in a vertical circle of radius R. You wish to whirl the stone fast enough so that when it is released at the point where the stone is moving directly upward it will rise to a maximum height, H, above the lowest point in the circle. In order to do this, what force will you have to exert on the string when the stone passes through its lowest point one-quarter turn before release? Assume that by the time that you have gotten the stone going and it makes its final turn around the circle, you are holding the end of the string at a fixed position. Assume also that air resistance can be neglected.

A) For each point labeled in the diagram, circle the symbol(s) that describe how the speed of the stone is changing.

Point	Change in Speed
A	↑ ↓ = max min
B	↑ ↓ = max min
C	↑ ↓ = max min
D	↑ ↓ = max min
E	↑ ↓ = max min

Change of Speed Symbols

↑ Speed is increasing
↓ Speed is decreasing
= Speed is constant
max Speed is at a maximum
min Speed is at a minimum

B) At each point on the diagram, draw and label a vector representing the acceleration of the stone.

C) At each point, draw and label vectors to represent all of the forces acting on the stone.

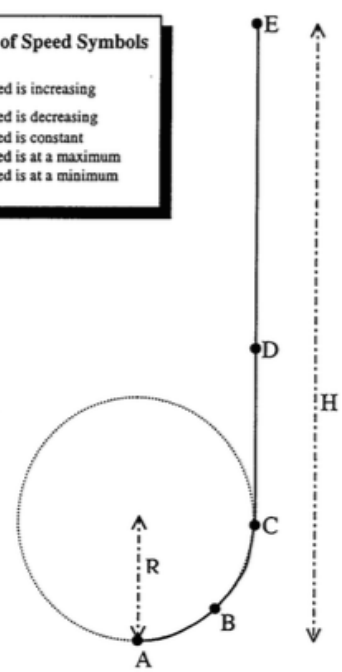


Figure 8: The two broken-into-parts problems posed to the TAs.

Problem Type	What are the FEATURES of the problem type (e.g., includes a diagram, breaks problem into sub-problems, etc.)?	What are the PROS/CONS of the problem type?	Would you use this problem in . . .				Are there OTHER situations in which you would use this problem? Explain your reasoning.
			a HW? Why/why not?	a QUIZ? Why/why not?	an EXAM? Why/why not?	a small GROUP problem solving session? Why/why not?	
Problem A	Breaks problems into sub-problems. Includes a graph	PROS: The question is very clear. CONS: The problem becomes too easy	Yes. Part A, B and C can help students learn different concepts	Yes. Easy to solve. Easy to give partial credits	Yes Easy to solve. Easy to give partial credits	Yes. the three parts should guide students very well	As an example shown during the class it put several concepts together

Figure 9: Part of a sample worksheet given to the TAs in the TA professional development course

4.3 METHODOLOGY

Participants: A total of 97 TAs participated in this study during 4 different years. Participants were physics graduate students who were enrolled in a mandatory TA professional development course that met once per week for 2 hours for an entire semester. The TAs were expected to do approximately one hour of homework each week pertaining to the professional development course. They concurrently had teaching responsibilities involving recitation or lab instruction and a majority were assigned to help students in a physics tutoring center.

Data collection tools and artifacts: The data collection tools consisted of five introductory physics problem types that had been developed previously [39]. The problem types were designed to focus on the same introductory physics problem scenario in mechanics. These problem types included two different versions of broken-into-parts problems, a multiple-choice problem, a context-rich problem, and a traditional textbook version of the problem. Here, we focus on two broken-into parts problems, namely Problem A and Problem D (see Figure 8). Problem A is broken-into-parts, includes a figure, and requires an explicit calculation. Problem D is similar to problem A in that it is broken-into-parts and includes

a figure, but it does not require explicit calculation. Based upon our research questions, the TAs were asked to answer questions about the five problem types on a worksheet, a partial sample of which can be found in Figure 9. Data were collected over four different years. In the most recent year's worksheet, TAs were also asked to rank the problem types on their instructional benefit (i.e., how instructionally beneficial the TAs judged each problem type to be), the level of challenge (i.e., how difficult the TAs judged each problem type to be for students), how much they liked the problem type, and the likelihood of them using the problem type if they had complete control of the choice of the problem types to use. For example, a TA who ranked a problem 1 for "challenging" judged this problem to be the least challenging for students; a 5 for "challenging" would mean that the TA perceived it to be the most challenging out of the five problem types. The rankings allowed us to investigate research questions 1 and 2. Furthermore, throughout all four years, TAs were asked to list pros and cons of the problem types. These pros and cons were useful for investigating why TAs ranked the problem types the way they did (research question 3).

Data collection in the TA professional development course and later in individual interviews: TAs were given the problem types and worksheets in the professional development course in the middle of the semester when they had some teaching experience in order to elicit their initial ideas about different problem types. They were asked to answer worksheet questions under the assumption that they had complete control over the introductory physics class, including control over problem types chosen for various purposes. The worksheet (see a part of the worksheet in Figure 8) was completed as part of a homework assignment. Later on, 12 participants who had been enrolled in the TA professional development course volunteered to be interviewed in a one-on-one setting using a think-aloud protocol. The interviews took place at least one semester after the initial activity described here in the TA professional development course. The purpose of the interviews was to more deeply probe the TAs' reasoning behind their responses, and the researchers had the opportunity to discuss with TAs their beliefs about broken-into-parts problems in general.

Coding TA responses: Two of the researchers met weekly to identify appropriate coding categories for pros/cons; agreements on these were reached through discussion. The researchers focused on coding the data from the individual homework assigned in the middle of the semester regarding the TAs' views of the problem types. The inter-rater reliability was examined for the coding of the pros/cons in a subset of the data (encompassing one year of data), and the average Cohen's kappa [70] was calculated to be $\kappa = 0.982$. The most common pros and cons for the broken-into-parts problems, along with their definitions and examples from TA worksheets are included in Table 3.

Code	Definition	Examples	Percentage of TAs
(Pro) guide	walks students through step-by-step; helps students solve harder problems	"parts make the problem more guided"	80
(Con) help	provides too much support or makes the problem too easy	"student does not have to do too much thinking"	37

Table 3: The most commonly listed pros/cons of the broken-into-parts problems and the percentages of TAs who listed them. Some TAs listed more than one of the following or other pros/cons not listed here.

4.4 RESULTS

Broken-into-parts problems ranked high for like, use, and instructional benefit, but low for challenge: As shown in Figure 10, the average rankings for the broken-into-parts problems are consistently high for "like," "use," and "instructional benefit," but low for "challenging." In fact, the broken-into-parts problems received the highest rankings for "like," "use," and "instructional benefit," and the lowest rankings for "challenging" out of the five problem types students were given. These rankings indicate that TAs appear to value this type of problem, although they find it not to be challenging for students.

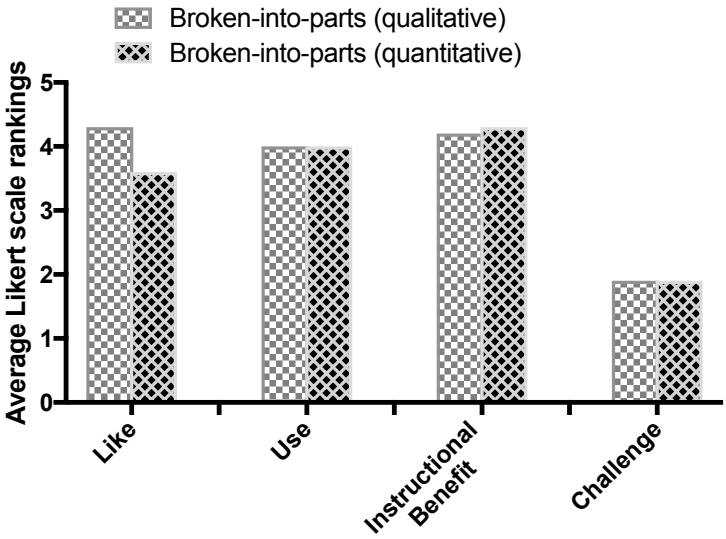


Figure 10: Average rankings of the broken-into-parts problems

TAs reported a wide usage of a broken-into-parts problems and preferred to use this problem type over other more challenging problem types: As seen in Figure 10, the TAs ranked the broken-into-parts problems highly for "use." This ranking is the highest ranking of all problem types the TAs considered. Both of the broken-into-parts problems received an average ranking of 4 out of 5, indicating that TAs were far more likely to use such a problem compared to other problem types, the next highest ranking for which was only a 2.9 out of 5 in the category of "use."

Figure 11 summarizes the TAs' stated use of the broken-into-parts problems, and shows that TAs reported that they would readily use such a problem for homeworks, quizzes, and exams. It appears that the broken-into-parts problem type was one type of problem that TAs would readily use for many purposes. However, written responses and interview data hint at possibly excessive valuing and use of this type of problem, indicating a potential overreliance on broken-into-parts problems. For example, one TA stated in an interview: "I always prefer sub-questions" whether it is for homework, quizzes, or exams. Further discussion with the

TA suggests that he would almost exclusively use broken-into-parts problems and was not likely to use other problems which may provide less support and more of a challenge for the introductory physics students. Similarly, regarding the broken-into-parts problems, another TA said: "I will use it everywhere." The idea of almost-exclusive preference for using this type of problem was conveyed by many TAs during the interviews, as well as in written responses. Below, we discuss reasons behind the rankings and stated uses, based upon interview data and the written responses in the columns of the worksheet which asked for explanations and/or reasons for their responses (see Figure 8).

Figure 11: TAs reported usage of the broken-into-parts problem type. These are the reported usages as averaged over both examples of the broken-into-parts problems since there was no significant difference between the two examples. Yellow indicates the TA would only use this type of problem in homework. Green indicates they would use it in homework or during a quiz or exam. Blue indicates they would only use it in a quiz or exam. Red indicates they would *never* use it for any purpose.

TAs viewed the pro of guiding students as outweighing the con of providing too much help: The pros and cons and written responses regarding the broken-into-parts problems, as well as the interview data, were examined for possible reasons for why the TAs ranked the broken-into-parts problems the way they did. Table 3 shows the most common pros and cons mentioned by TAs in written responses.

The most common pro stated for broken-into-parts problems was "guide," which was mentioned by 80% of TAs. "Guide" was the category used for TAs' responses which included what they judged to be an opportunity to guide the student in the problem solving process. Some examples include: "Guides students to understand how to solve the problem," and "Leads the student to solve the problem step-by-step."

The prevalence of the pro "guide" could explain why TAs ranked the broken-into-parts problems highly in terms of instructional benefit. One TA stated: "Very instructional ... break-down helps students solve problem." The TAs with this type of response appreciated the support provided in these problems because it was perceived as a way to guide students in solving the problem. Moreover, of all the TAs who mentioned the pro "guide" for Problems A and D, all but one TA ranked Problems A and D as the two highest out of all the problem types they were given, in terms of their instructional benefit.

One TA described his reason for judging the broken-into-parts problems as being highly instructionally beneficial by stating that he prefers to give "questions that the student can just jump in and start immediately. And later, they look at their results from one part and can get some inspiration for the next part." This TA further explained how a broken-into-parts problem can make the problem-solving process more manageable for introductory physics students by providing them with support to get from one step to the next through the solution.

By contrast, TAs did not list very many cons. Even though they were specifically asked to list at least one pro and one con, many TAs failed to list any cons for the broken-into-parts problems at all. Table 3 shows that the only commonly stated con for the broken-into-parts Problems A and D was "help," and that this con was mentioned by only 37% of TAs. The category "help" contained TA responses in which the TA expressed reservations that it was making the problem too easy (for example: "Not difficult enough") or helping the student too much.

Although the con "help" suggests that some TAs had reservations about the broken-into-parts problems potentially providing too much help to students, this con is mentioned

by only about one-third of TAs. Even in interviews, any con for broken-into-parts problems was rarely mentioned even when TAs were explicitly asked for at least one con. Moreover, interview data suggests that the possibility that broken-into-parts problems may not be difficult enough was a minor concern to TAs and would not deter them from giving priority to this type of problem over other types on homework, quizzes and exams. For example, one TA reported in an interview that "solution D might be too easy." However he went on to say that "... actually, Solution D is good in the sense that it's like a good way of testing the students' conceptual understanding ... ." He went on to discuss the merits of breaking a problem into parts for a student, stating: "It kind of leads you through all the steps ... so [for an introductory class] I would probably give them this," indicating that he appreciated the guidance the problem provides. He further suggested that his concern about these problems being "too easy" was outweighed by the value he saw in them for the introductory physics students. Another TA expressed his reservation by stating that the broken-into-parts problems "may be setting up a problem a little too much ... ." But this same TA also ranked the broken-into-parts problems the highest in the category "like" and cited the guidance offered by the problem as a pro. He further explained that such problems were appealing to him because of the "... multiple parts to guide students" and that he would give priority to using such problems on homework, quizzes, and exams in introductory physics. Further discussion suggests that the TA's overall impression of the broken-into-parts problems was positive, and his concern about helping students too much appears to be overshadowed by the positive aspects he perceived.

Other TAs expressed similar sentiments about the pros being more compelling than the cons of broken-into-parts problems (if they mentioned a con at all, which was only roughly one-third of the TAs). Both written and interview data suggest that the con "help" may not be viewed as a major drawback to TAs. Indeed, TAs were often reluctant to report downsides to broken-into-parts problems, sometimes using superlative language to describe such problems. For example, several TAs went as far as to use the word "perfect" in describing broken-into-parts problems for homework, quizzes, and exams in introductory

physics, apparently not detecting any drawbacks to such problems.

Although TAs mentioned the asset of a broken-into-parts problem guiding the introductory physics students in solving the problem at hand, the nature of the responses did not usually indicate the idea that such problems could be used to train students to solve future problems that are not broken into parts. Furthermore, TAs rarely mentioned in interviews or written responses that the scaffolding support provided by these type of problems should gradually be removed to help students develop self-reliance in problem-solving.

TAs had different expectations of introductory and advanced students: Despite the low ranking for "challenging," the majority of TAs reported a preference for these problems, ranking them highly in the category "like." Our follow-up interviews indicated that many TAs felt that introductory physics students should not be given problems that were overly challenging (e.g., a problem scenario which is not broken-into-parts) and they did not expect introductory students to be able to break up the problem into sub-problems on their own. On the other hand, when TAs were asked during interviews whether they would use broken-into-parts problems for advanced students if they were teaching an advanced physics course, many reported that such problems would be "too easy" for advanced students, and therefore such broken-into-parts problems should not be used in advanced courses. It is important to note that, in the interviews, TAs were specifically asked whether appropriate advanced topic broken-into-parts problems should be used for advanced students. Thus, it was not the introductory physics topic that TAs felt was "too easy" for advanced students. Rather, it was the support provided by the problem being decomposed into sub-problems that TAs felt was not appropriate for advanced physics students, even if the problem dealt with an advanced-level topic. Thus, there was a discrepancy between the TAs' expectations of introductory students with regard to introductory-level physics problem-solving compared to advanced students with regard to advanced-level physics problem-solving.

Exhibiting a relatively low expectation level for introductory physics students, one TA said, "My gut feeling is that you don't want to knock the students out with a tough problem." This TA further suggested that challenging problems should be avoided, and that he felt that

introductory students may not be able to handle difficult problems. Another TA who ranked the broken-into-parts problems the highest in the category "like" explained that: "This would be, for me as an intro student, the ideal problem, because I have to use my skills to be able to translate what I'm being asked into the mathematical formulas, but at the same time I have enough guidance..." This TA felt that broken-into-parts problems provide the appropriate level of difficulty for introductory physics students. Likewise, in explaining his preference for a broken-into-parts problem, another TA stated: "Trying to unpack a problem into different parts is half the battle sometimes with solving these problems. When it's done for you that's helpful, and it's really important on a test too..." This TA felt that making the problem less difficult for introductory students was important even in a test situation. He believed that the challenge of "unpacking" the problem should be done for the introductory student, in homework, quizzes, and exams, rather than asking that the students do this unpacking themselves. He did not express confidence that introductory students could accomplish this task on their own. In a similar way, another TA who ranked the problem low for its level of challenge reported: "I like that the problem is broken down into small questions that help you solve the original problem. For introductory classes, [even] this may be a bit difficult." This TA expressed that while these problems are easier than problems which are not broken-into-parts, even this level of challenge could potentially be difficult for introductory students, indicating a relatively low expectation level regarding the types of problems introductory physics students should be expected to handle in homework, quizzes, and exams.

Most TAs did not feel that it was necessary or identify when introductory physics students should be expected to learn how to break physics problems into sub-problems themselves. They did not report that problems with less support would be important to use in order to help introductory students gain more independence in problem solving by developing skill in decomposing a problem into sub-problems on their own. One TA who reported that he would widely use a broken-into-parts problem explained: "Students can be instructed while doing the problem and, while it doesn't have them connect the steps themselves, that probably isn't the point." While this TA momentarily considered the fact that this type of problem did

not require students to connect the steps of the problem-solving process themselves, further discussion suggested that he did not regard "connecting the steps" as a crucial component of the types of problems introductory students should be expected to solve. He also did not indicate during the discussions that the independent practice of "connecting the steps" and learning to break a problem into sub-problems is a critical step towards developing expert-like problem-solving skills.

Given that many interviewed TAs did not find a broken-into-parts problem to be appropriate for advanced students, it appears as though the TAs believed that advanced physics students should be able to decompose a problem into sub-problems on their own. However, most interviewed TAs did not mention that introductory students should independently practice the skill of decomposing introductory problems into sub-problems in order to develop expertise in problem solving and independently solve advanced problems. For example, one TA explained the discrepancy between expectations for introductory vs. advanced students as follows: "[I would use it] very much, because the steps lead the students to solve the problem, but this knowledge is basic and too easy for an advanced level student." This TA expressed an expectation that advanced students should be able to break a problem down on their own, but that introductory students should not be expected to do so. TAs with this type of response did not identify using problems with less support for introductory students as a way to cultivate expert-like problem-solving skills and had relatively low expectations of what types of problems introductory students could be expected to solve.

TAs' preference for broken-into-parts problems may be influenced by introductory students' preferences: The broken-into-parts problems were strongly preferred by TAs compared with other problem types, as evidenced by the average rankings in the category "like." Both written and interview data suggest that one of the reasons for this preference could be the TAs' interest in what they believed introductory students will like. In particular, introductory students might prefer problems that are easier for them to solve, and this preference appeared to be on the minds of the TAs. For example, one interviewed TA who had ranked the broken-into-parts problems as the highest for "like" expressed that:

"Breaking it into parts is reducing the workload for the students and I think they're going to appreciate that." This TA indicated that lightening the workload for introductory students entered into his thinking, suggesting that pleasing students in this way may be at least part of why he valued these types of problems for homework, quizzes, and exams. Another TA who ranked the broken-into-parts problems highly in the category "like" and noted that he likes these problems for homework, quizzes and exams stated in the interview: "I think students would like this one [Problem A] most." Again the idea of what students might prefer appeared to contribute to the TA's preference for this type of problem. Other interviewed TAs had similar views about introductory students' preferences. Likewise, another interviewed TA tried to explain why he liked broken-into-parts problems by stating that, for such problems, "If the basics are clear, they [introductory physics students] will sail through." While this TA acknowledged that this problem type is easier than problems which are not broken into parts, he thought they were "nice" problems for use in homework, quizzes, and exams so that introductory students do not have to struggle too much and he ranked it highly for "like" and "use." It appears that the ease with which students could solve such problems was a factor in his preference for a broken-into-parts problem. Other interviewed TAs appeared to convey similar sentiments regarding making problems easier for the introductory physics students by way of breaking a problem into parts.

4.5 DISCUSSION AND CONCLUSIONS

Most TAs highly valued broken-into-parts problems and stated that they would use such problems often on homework, quizzes, and exams because such problems facilitate the problem-solving process for introductory students. Discussion during interviews suggests that TAs may overuse broken-into-parts problems partly due to their preference to guide introductory physics students through the problem-solving process, their relatively low expectation level for introductory students, and their consideration of introductory students'

preferences and the desire to reduce their stress while solving physics problems. In the cognitive apprenticeship model, appropriate coaching and scaffolding support can help develop expertise and train a student to eventually gain independence in solving complex physics problems [13]. This type of long-term goal was not typically mentioned or implied by TAs' responses in written or interview data. Instead, the use of broken-into-parts problems was regarded by TAs as beneficial for helping guide students in solving the problem at hand, and that asset alone appears to be a major reason for why the TAs would be likely to frequently use broken-into-parts problems in homework, quiz and exam situations. However, TAs did not indicate that introductory students should also practice more independent problem solving via problems in which the scaffolding support is removed after the modeling and coaching part of the cognitive apprenticeship process.

These findings partly agree with a similar study involving physics instructor's views of various problem types, in that, like instructors, TAs reported a wide use of broken-into-parts problems despite any reservations they might have about them [39]. However, the TAs appear to have an even stronger preference for broken-into-parts problems than did the faculty in that fewer TAs expressed a concern that such problems may provide too much help to students (even when explicitly asked to state a con of a broken-into-parts problem) compared with the number of faculty who expressed similar concerns. While nearly half of faculty identified independent problem solving without guidance as an important goal in teaching problem-solving [39], few TAs mentioned that using problems which do not provide introductory students with guiding support was important because they can help introductory physics studetns develop self-reliance in problem solving. Additionally, interviews and written data suggest that, even among those TAs who had a concern about a broken-into-parts problem potentially providing too much help, this concern was not strong and did not outweigh the benefit of guiding a student through a problem by breaking it into parts in homework, quizzes, and exams. Moreover, while both TAs and faculty reported copious use of broken-into-parts problems with their introductory students whether or not they had concerns about such problems, most TAs overlooked the need to challenge introductory stu-

dents by offering them opportunities to solve problems which do not have the steps already broken-down for them so that they can develop self-reliance in problem-solving.

Interviews also suggest that TAs expected more advanced students to be capable of decomposing an advanced problem independently. Yet, if introductory students do not practice breaking down an introductory physics problem into sub-problems on their own, they may not be able to develop this skill on their own and become effective problem-solvers. TAs' perspectives appeared to be missing the crucial bridge between highly supported problem-solving and independent problem-solving. Without this step, introductory students can be severely hindered in their development of expertise in problem solving, reasoning, and meta-cognitive skills. This missing puzzle piece is similar to the discrepancy in TAs' grading practices in a prior investigation in that TAs did not demand that introductory students explicate the steps in their solutions (and would not penalize them for neglecting to show work), but expected advanced students to do so and would penalize them if they did not do it [36, 37]. In particular, in a previous study related to TAs' views about grading practices, many TAs felt that advanced students should be required to show steps and reasoning in their advanced physics problem solutions but claimed that introductory students need not show steps or do conceptual reasoning in their introductory physics solutions and should not be penalized for omitting such steps [36, 37].

This study suggests that TAs had a relatively low expectation of introductory students' problem solving skills and had not reflected on ways in which introductory students can be provided guidance to cultivate independent problem solving skills. Such skills (e.g., breaking a problem into sub-problems on one's own) are unlikely to develop spontaneously and must be explicitly cultivated by incorporating them into instructional design and having high expectations of introductory students while helping them develop self-reliance in physics problem solving. TAs' preference for continually providing problems for introductory students which are broken into parts represents an important oversight in the steps required for introductory students to learn independent expert-like problem-solving. In particular, introductory students must be given opportunities to practice bridging the gap between solving

problems that are broken-into-parts and solving problems with less built-in support, a point that most TAs appear to have missed.

Leaders of TA professional development programs can use the findings of this study to help TAs elucidate their teaching and learning goals for both introductory and advanced physics students and reflect on instructional approaches that support their goals. For example, our findings indicate that TAs had relatively low expectations for introductory students. This finding suggests that TAs may not have thought about the goal of helping introductory students become independent problem solvers. To help TAs readjust their expectations of introductory students, TAs can be asked to reflect on and clarify their learning goals for introductory students. They can also examine how different problem types can support (or hinder) achievement of different learning goals. In particular, TAs may be given opportunities to discuss how broken-into-parts problems support the goal of helping introductory students learn, e.g., how to decompose problems into sub-problems in the modeling and coaching phases of the cognitive apprenticeship model to help students develop expertise. In addition, TAs can reflect upon and discuss as a group how other problem types which do not decompose the problem into sub-problems may support the goal of helping introductory students develop self-reliance in problem solving. They can also reflect on their differing expectations of advanced physics students with regard to advanced physics problem solving and introductory students with regard to introductory physics problem solving and why having a lower expectation of what introductory students can learn and be able to do while solving an introductory problem can be detrimental to their overall learning. In the professional development programs, TAs can be asked to contemplate and discuss how the use of different introductory problem types can help introductory students progress toward expert-like problem solving approaches. In this way, TAs may begin to appreciate that, while broken-into-parts problems are an important stepping stone in the development of expert-like problem solving, other problem types that do not provide scaffolding support are also critical in the development of students' self-reliance in problem solving.

5.0 IMPACT OF INSTRUCTION ON INTRODUCTORY FEMALE AND MALE STUDENTS' ATTITUDES AND APPROACHES TO PHYSICS PROBLEM SOLVING

5.1 INTRODUCTION

5.1.1 Expert vs. novice attitudes and approaches to problem-solving

Instructional goals of an introductory physics course may include developing problem-solving skills and conceptual understanding [45, 39]. These goals may be facilitated by approaches and attitudes which reflect the way an expert might think about physics [111, 29, 25, 26]. Physics experts, e.g., physics faculty members, organize their physics knowledge hierarchically so that underlying concepts are connected in a meaningful and structured way and they exhibit positive attitudes towards scientific problem solving [22, 40, 23, 44, 45, 38]. Experts' knowledge, including how the knowledge is structured in well-organized schema, and their positive attitudes and approaches to problem solving facilitate an effective approach to problem solving [23, 46]. By contrast, novices, e.g., many introductory students, view physics as a collection of disconnected facts and equations [22, 40, 23]. They often have less-expert like attitudes towards problem solving and approach physics problem solving in a haphazardous manner [22, 40, 23].

5.1.2 Evaluating grown among introductory physics students along various dimensions using surveys before (pre) and after (post) instruction

One can measure changes along various dimensions from the beginning to the end of a course as a result of instruction, e.g., growth in content knowledge, or approaches to problem solving. Students' epistemological belief is one dimension and it can impact problem solving and learning in a particular discipline. [112, 111, 25, 26, 22, 23]. Several surveys have been developed to evaluate both students' attitudes about physics and physics learning as well as their conceptual understanding of physics[26, 27, 32, 28, 25]. Attitudinal surveys which focus on students' epistemological beliefs about physics include the Maryland Physics Expectation Survey (MPEX) and the Colorado Attitudes about Science Survey (CLASS) [26, 27]. When these attitudinal surveys have been given both at the beginning and end of a semester of instruction, they typically show declines in students' attitudes towards physics and physics learning over the course of instruction compared to what would be considered an expert-like response [32]. This has been found for both male and female students, even when students could correctly identify how their instructors would have answered the survey questions [27].

The Attitudes toward Problem Solving Survey (APSS) was developed with inspiration from the MPEX survey with a focus only on attitudes about problem solving [28, 25]. The APSS reveals a similar decline in student attitude from the beginning to the end of the semester after instruction to that found via the MPEX and CLASS surveys, especially in large enrollment traditionally taught classes [25].

A modified version of the APSS, the Attitudes and Approaches to Problem Solving (AAPS) survey, was developed and validated to include questions regarding the approaches students take when solving physics problems [30]. The AAPS survey is unique in its focus compared with other broader attitudinal survey because, like the APSS, it hones in specifically on problem solving, but unlike the APSS, the AAPS survey also probes approaches to problem solving in addition to attitudes towards problem solving. The added dimension of approaches to problem solving is important for assessing growth in problem solving expertise. To investigate evolution of expert-like response, the AAPS survey was validated

based upon introductory student, graduate student and faculty responses [30]. Since physics experts, e.g., physics faculty members, organize their physics knowledge hierarchically with underlying concepts connected in a meaningful and structured way [22, 40, 23, 44, 45, 38], their attitudes and approaches towards problem solving, as reflected by their response to the AAPS survey establishes the criteria for an expert-like response. By contrast, many introductory students, have less-expert like attitudes and approaches to physics problem solving [30]. After initial validation, the AAPS survey has been used to investigate students' attitudes and approaches to problem solving in high school and university physics classes in different countries [113, 114]. However, until now, the AAPS survey has not been used to investigate changes in introductory physics students' attitudes towards problem solving before and after instruction in an introductory course.

In addition to attitudinal surveys, several conceptual surveys have been developed and validated to assess students' conceptual understanding. Conceptual surveys include, among others, the Force Concept Inventory (FCI) [115, 78, 116, 92], which is usually given to first semester introductory physics students learning mechanics, and the Conceptual Survey of Electricity and Magnetism (CSEM) [117], which is typically administered to second semester introductory physics students learning electricity and magnetism.

The Force Concept Inventory (FCI) was developed to assess students' conceptual understanding of introductory Newtonian mechanics [115, 78, 116]. The FCI has been administered to introductory physics students in various universities in the United States and elsewhere for decades. It has been well-established that there is a performance gap based upon gender of the students who take the FCI, with male students scoring higher, on average, than female students [115, 78, 116, 92].

The Conceptual Survey of Electricity and Magnetism (CSEM) was developed to assess students' conceptual understanding of introductory electricity and magnetism topics broadly [117]. The CSEM is a particularly difficult survey for introductory students, with typical scores being less than 50% [117] after instruction in relevant concepts. Like the FCI, male and female students may perform differently on the CSEM, but the "gender gap" is typically

not as large or consistent for the CSEM as it is for the FCI [92].

5.1.3 Evidence-based active engagement (EBAE) methods

Many introductory physics classes in the U.S. are taught primarily in a lecture-based manner. Instruction in which regular class time is primarily focused on the instructor lecturing and the students taking notes is often referred to as "traditional." However, some faculty members implement physics education research-based instructional strategies (RBIS) in their teaching [21]. There is evidence of a growing awareness of RBIS among faculty, especially among those full-time faculty who attend teaching-related workshops and/or read teaching-related journals [118, 119]. We prefer to refer to RBIS as "evidence-based active engagement" (EBAE) methods for clarity and specificity. EBAE methods may vary significantly, but they share a common goal: facilitate an environment in which students take a more active role in their learning using physics education research-based approaches than is afforded by traditional approaches [50].

Peer Instruction is an example of an EBAE method in which in-class clicker questions along with student discussion are interspersed throughout instruction. This method has been shown to promote both conceptual understanding and better problem-solving skills, and may also positively impact students' attitudes about physics and physics learning [81, 62, 63]. Another EBAE method which has been associated with conceptual learning gains [94] involves the use of interactive lecture demonstrations (ILDs) in which students are asked to predict what will happen before a demonstration is shown during class, and after the demonstration, the instructor leads a class discussion guiding students to build a coherent knowledge structure of the concepts involved. Collaborative group work involving the use of context-rich problems is yet another example of active engagement in the physics classroom [64, 65]. Use of context-rich problems in collaborative group work has been associated with the development of more expert-like problem solving approaches [64, 65]. Moreover, "inverting" or "flipping" instruction is becoming increasingly popular and provides an opportunity to move part of the content which is normally covered in lectures to assigned videos and/or tutorials

outside of the class [89, 90, 91]. This allows more class time to be used for EBAE activities (with students often working in small groups) which are designed to reinforce the content presented outside of class using videos, textbooks, and other resources [89].

Findings are mixed as to the degree to which the use of these types of EBAE methods impacts the "gender gap" on conceptual surveys such as the FCI and CSEM. A gender gap appears to remain even when active engagement is used [120, 121, 90, 91]. However, there is some evidence that female students benefit more from EBAE methods of instruction when they are carefully designed [120, 90, 91].

5.1.4 Gender differences in introductory physics performance

Prior research suggests that male students in introductory physics often outperform female students on standardized conceptual assessments such as the Force Concept Inventory (FCI) [115] or the Conceptual Survey of Electricity and Magnetism (CSEM) [117]. The discrepancy between male and female students performance is commonly referred to as the "gender gap" [90, 91, 122, 123] and has been found even after controlling for factors such as different prior preparation or coursework of male and female students [122, 123]. Findings are mixed as to the degree to which the use of the EBAE methods impact the "gender gap" on conceptual surveys such as the FCI and CSEM. Some prior research has also found that using carefully designed evidence-based pedagogies can reduce the gender gap [92], but other studies suggest that a gender gap remains even when EBAE methods are used [120, 121, 90, 91].

The origins of gender gap on the FCI both at the beginning and end of a physics course has been a subject of debate, raising questions about whether the test itself may be gender-biased [124]. Some of the origins of the gender gap can be attributed to societal gender stereotypes [125, 126] that begin from an early age. For example, research suggests that even six year old boys and girls have gendered views about intelligence in which they view boys as smarter [126]. Such stereotypes can impact female students' self-efficacy [127], their beliefs about their ability to perform well in disciplines such as physics in which they are underrepresented and which have been associated with "brilliance".

5.1.5 Focus and framework for our investigation

Our investigation focused on examining the changes in the student response to the AAPS survey from the beginning to the end of a semester in introductory calculus-based physics courses at two large state-related research universities in the Unites States. Furthermore, we investigated these changes for different instructional approaches (EBAE vs. traditional lecture-based) and for female and male students separately. Moreover, we investigated student performance on standardized conceptual surveys and the correlation between student performance on conceptual surveys and the AAPS survey.

It is also important to investigate how the attitudes and approaches to problem solving using the AAPS survey are impacted by whether primarily traditional lecture or EBAE methods are employed. The overarching framework that inspired the comparison of the instructional methods in this study is that the EBAE methods focused on a cognitive approach and building on students' prior knowledge to facilitate learning and on developing of a robust knowledge structure. In these EBAE methods, there was an effort to align learning goals and objectives, instructional design, and assessment of learning with each other and there was a focus on evaluating whether the pedagogical approaches employed have been successful in meeting the goals and enhancing student learning. The instructors using the EBAE approaches were employing the cognitive apprenticeship model [13] which focuses on "modeling", "coaching and scaffolding" and "weaning". In particular, providing opportunities to coach students and scaffold their learning was a central aspect of the EBAE methods. It is important to investigate how these EBAE methods impact student attitudes and approaches to problem solving compared to teaching focused primarily on traditional lecture.

Moreover, the way in which physics is instructed has been found to be connected to students' beliefs about physics [128]. It is also important to explore if attitudes and approaches to problem-solving in physics are likewise connected to method of instruction. While other attitudinal surveys have investigated broader questions regarding epistemology and beliefs about physics [26, 27, 32], and how these beliefs may be related to method of instruction [122], no study has focused on the aspect of student perspectives related specifically to at-

titudes and approaches to problem-solving, and whether or not the method of instruction is related to these attitudes and approaches. Our hypothesis is that EBAE instructional strategies which focus on helping students develop effective problem-solving skills may encourage better attitudes and approaches to problem solving. We therefore investigated the impact of instruction on the AAPS scores. The AAPS survey was designed and validated [30] to investigate questions such as these.

Moreover, the broader CLASS survey shows that female students overall score worse than male students but it would be beneficial to investigate whether similar or different type of gender differences are found on the AAPS survey which focuses specifically on the attitudes and approaches to problem solving. We note that prior research suggests that female students in introductory physics courses for science and engineering majors, on average, have lower self-efficacy and intelligence mindset than male students even after controlling for performance [129]. Moreover, activation of a stereotype, i.e., stereotype threat, about a particular group in a test-taking situation can alter the performance of that group in a way consistent with the stereotype [129] and some researchers have argued [125] that female students, when working on a physics test, undergo an implicit stereotype threat due to the prevalent societal stereotypes. However, since differences in attitudes and approaches to problem-solving of female and male students in introductory physics courses has never been investigated, it is worthy of investigation and may prove to be useful in improving the learning environments for all students, regardless of their gender, in introductory physics.

In order to evaluate the robustness of the findings for similar institutions (large research universities in the US), the investigation in calculus-based introductory physics I was carried out at two institutions so that results can be compared. Furthermore, investigation of correlation between students' AAPS survey scores with their actual performance on the concetual assessment or final exam can be a valuable measure of students' developing physics expertise.

5.1.6 Research questions of our investigation

Our research questions are as follows:

RQ1. How do the pre-survey (i.e., before instruction in relevant concepts) and post-survey (i.e., after instruction in relevant concepts) scores on the AAPS survey compare at a typical large state-related research university in the United States?

RQ2. How do the AAPS survey scores compare for EBAE vs. traditional instruction without separating students by gender and how do they compare when male and female students are considered separately?

RQ3. Is there any difference between AAPS survey pre and post scores in introductory physics I (mechanics) and II (mainly electricity and magnetism)?

RQ4. Are students' AAPS survey scores in a given class correlated with their conceptual survey (FCI or CSEM) or final exam performance?

RQ5. How do the pre-/post-test scores on the AAPS survey and FCI compare at two different large research universities in the United States?

5.2 METHODOLOGY

5.2.1 Courses and participants

After matching was done to ensure that each student's responses included both pre- and post-test data (we note that when initial analysis was conducted without matching, no significant differences were found), a total of 784 calculus-based introductory physics students from 9 classes from two separate large state-related research universities were included in this study. This group included 528 first semester students (content was mainly mechanics) and 256 second semester students (content was mainly electricity and magnetism). Four sections of first semester physics were taught in a traditional manner and three sections of first semester

physics were taught using EBAE methods. The two sections of second semester physics were taught in a traditional manner. With the exception of one section that could have contained possible overlap, the concurrent timing of when the data were collected enforced that second semester students were different students from the first semester students, but represented comparable cohorts of students. A majority of students in these courses are engineering, physical science or mathematics majors and are typically in their first year of college and had taken at least one physics course in high school, although the content and quality of these high school courses can vary greatly.

Traditional classes were operationally defined as classes in which evidence-based active engagement methods were used either very infrequently or not at all. The primary method of instruction for classes defined as traditional was the use of lectures by the course instructor for the majority of instruction. Some classes defined as traditional included limited use of group work, but this use was constrained to recitations or labs run by teaching assistants and only accounted for a small fraction of instructional time. The EBAE classes were operationally defined as classes in which active engagement methods were used frequently. In most cases, this involved course instruction taking place in a fully flipped manner [89, 90, 91]. In addition, one class designated as EBAE was not fully flipped but involved a significant amount of Peer Instruction, in which students were frequently engaged with clicker questions and discussion throughout the duration of class time [81], and included frequent group problem solving (at least once/week). No instructor taught both types of classes–traditional methods and EBAE.

The average class size was 87 students (an average class size for first semester physics was 75 students, and an average class size for second semester physics was 128 students). At the university which offered some traditional and some EBAE sections of classes, students were free to register for whichever section (of the many sections available) they preferred, and teaching reputations and styles of the different instructors were generally known in advance by the students, or were publicly available from online websites. A summary of the courses can be found in Table 4, along with the number of participants included in each class.

Because no significant difference was found in initial analysis of unmatched participants, only those who took both the pre and post surveys (i.e., "matched") were included in the full analysis. The average loss of students participants due to the missing pre or post survey data was 19%, which includes losses from students who dropped, withdrew, were absent for the pre or post survey, or otherwise chose not to participate in the post-test. Occasionally, a participant was excluded from the matched sets because they added the class late and did not take the pre-survey. Because of the many possible reasons for a student missing pre and/or post, it is not possible to infer whether the loss from unmatched to matched implies interest level in participation in the surveys.

5.2.2 Data collection tools and artifacts

For the first semester physics classes, both the FCI and the AAPS were administered twice–once near the beginning of the semester (pre) and once near the end of the semester (post). Similarly, in the second semester classes, both the CSEM and AAPS were administered twice (again, once near the beginning and once near the end of the semester). The AAPS survey can be found in the appendix of Ref [30]. It consists of 33 questions on a five point likert scale spanning "strongly agree," "agree," "neutral," "disagree," and "strongly disagree." For some questions, a favorable or expert-like response to a survey question (based upon faculty responses) is "agree" or "strongly agree," while for other questions a favorable or expert-like response may be "disagree" or "strongly disagree." In addition to the survey data, information was gathered about the teaching methods used at both universities, and gender of the students for all classes at University 1.

5.2.3 Data collection and analysis methods

Students took the surveys either in lecture or in recitation and were often offered a small amount of bonus points as an incentive for completion. A "dummy" question was included on the survey for all but one class, which was designed to allow the researchers to identify

Class	Description	$N_{Matched}$	N_{male}	N_{female}	University	Level
EBAE 1	flipped classroom instruction	46	31	15	1	first semester
EBAE 2	peer instruction/group problem-solving	81	55	26	1	first semester
EBAE 3	flipped classroom instruction	53	32	20	1	first semester
Trad. 1	primarily lecture-based instruction	87	69	18	1	first semester
Trad. 2	primarily lecture-based instruction	78	65	13	1	first semester
Trad. 3	primarily lecture-based instruction	77	63	11	1	first semester
Trad. 4	primarily lecture-based instruction	106	N/A	N/A	2	first semester
Trad. A	primarily lecture-based instruction	116	90	25	1	second semester
Trad. B	primarily lecture-based instruction	140	105	33	1	second semester

Table 4: A summary of the descriptions of the manner in which each class was taught. An initial analysis was conducted without matching students from pre to post and then further analysis was conducted with matching and found not to be statistically significantly different from the unmatched analysis. Thus the values listed for $N_{Matched}$ are the students included in further analysis. Average loss of students from unmatched to matched in each class was 19%, and could include losses due to students who dropped or withdrew from the class, were absent for the pre or post-test surveys, added the class and missed the pre-test survey, or chose not to complete the post-test surveys. The number of matched male and female students is also included. Note that gender information was not available for University 2, so the data from this university were not included in analysis of differences by gender. Also note that sometimes gender was not indicated by a student. In that case, provided the student had both pre- and post-test data, they would have been included in the overall analysis, but not included in the analysis of differences by gender. Thus $N_{male} + N_{female}$ may not equal $N_{Matched}$.

surveys in which the students were not reading the questions carefully, but simply marking down answers randomly. This question asked students who were reading the questions to indicate a specific choice on their bubble sheets for that question. If that choice was not indicated, the students' survey results were not included in the data analysis.

Once all data were collected, normalized scores (which we will simply call scores) for each question were computed by assigning a +1 to a favorable response (regardless of whether it

was strongly favorable or favorable), a -1 to an unfavorable response (regardless of whether it was strongly unfavorable or unfavorable), and a 0 to a neutral response. This same convention was used in earlier analysis when the AAPS survey was originally validated [30] and was adopted here for consistency with the validation study. We computed the average normalized score for each question based upon the student responses of +1, 0, or -1, and then averaged these values for every question to arrive at an average overall normalized score on the AAPS for each class.

In addition, we made use of the factors identified from a principal component analysis in the original validation study, Ref. [30]. These factors and the questions associated with each one can be seen in 8. In particular, there were 9 factors identified from a principal component analysis. Factor 1 involves questions related to metacognition and enjoyment of physics problem solving. Factor 2 involves questions about the use of drawing and scratch-work while problem solving. Factor 3 involves questions about perception of problem solving approach. Factor 4 involves questions that distinguish between general expert-novice differences in problem solving. Factor 5 involves questions about solving problems symbolically. Factor 6 involves questions about problem solving confidence. Factor 7 involves questions about solving different problems using the same principle. Factor 8 involves questions related to sense-making. And finally, factor 9 involves questions related to problem solving sophistication. The scores for various subsets of students were compared by these factors, in addition to comparison of scores on individual questions in order to better identify any differences by theme.

In our analysis, the AAPS survey data were compiled and examined for changes in scores from pre to post first without matching students (which includes all students who took a pre or post-survey regardless of whether they took both surveys) and then again for matched students (which includes only those students who took both the pre and post-survey). No significant differences were observed with or without matching students. Therefore, the remainder of the analysis was carried out with matched sets, so that students with a missing pre-survey or post-survey were excluded from the analysis presented here.

Moreover, follow-up interviews were conducted with 12 introductory physics students in order to gather qualitative data to complement the quantitative survey data. Interviews were conducted using a think-aloud protocol. Each interview lasted for approximately one hour. During the interviews, students first answered the AAPS survey questions along with providing reasoning for their answers on their own. Follow up questions were then asked for clarification of points not made clear. The interviewer asked further questions regarding the way in which students perceived their classes were taught. The interviews served to more deeply probe the students' reasoning behind their responses to the AAPS survey, and to understand their experiences with different methods of instruction.

5.3 RESULTS

5.3.1 RQ1. How do the pre-survey (i.e., before instruction in relevant concepts) and post-survey (i.e., after instruction in relevant concepts) scores on the AAPS compare at a typical large state-related research university in the United States?

We find that, in all cases, the overall average AAPS scores exhibited decline from pre to post, consistent with other types of attitudinal surveys (e.g., those focused on general epistemological beliefs of students), for both first semester and second semester physics instruction as seen in Figures 12 and 14 . When each section is considered separately, each section's decline was statistically significant except for one flipped and one traditional section, as shown in in Table 5. Table 5 also displays effect sizes (Cohen's d) [70] for the decline, which are small to medium. Moreover, pre survey scores were moderately correlated with post survey scores on the AAPS ($R \approx 0.5$). Table 9 in the appendix contains the raw scores showing average percentages of favorable, unfavorable, and neutral responses for EBAE and traditional classes.

Class	Semester	p-value pre vs. post	Effect size
EBAE 1	1	0.099	-0.16
EBAE 2	1	0.043	-0.21
EBAE 3	1	0.017	-0.28
Traditional 1	1	0.005	-0.39
Traditional 2	1	0.146	-0.15
Traditional 3	1	< 0.001	-0.34
Traditional 4	1	0.023	-0.22
Traditional A	2	< 0.001	-0.49
Traditional B	2	< 0.001	-0.44

Table 5: p-values obtained via the t-tests comparing average pre and post AAPS scores, and effect sizes for the decline in scores from pre to post AAPS survey.

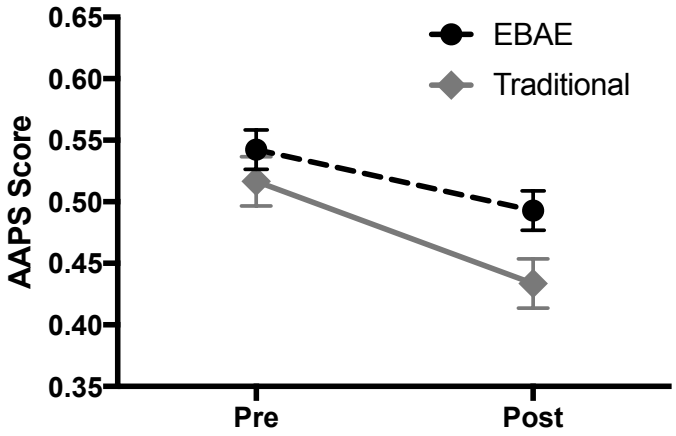


Figure 12: Raw data show the overall average scores for both the pre and post AAPS survey in the calculus-based first semester classes at University 1. The error bars represent standard error.

5.3.2 RQ2. How do AAPS scores compare for EBAE vs. traditional instruction without separating students by gender and how do they compare when male and female students are considered separately?

EBAE vs. traditional instruction: We find that, on average, students in the EBAE classes scored systematically higher than those in the traditional classes on the AAPS survey before instruction, as can be seen in Figure 12. The difference in pre-survey scores between

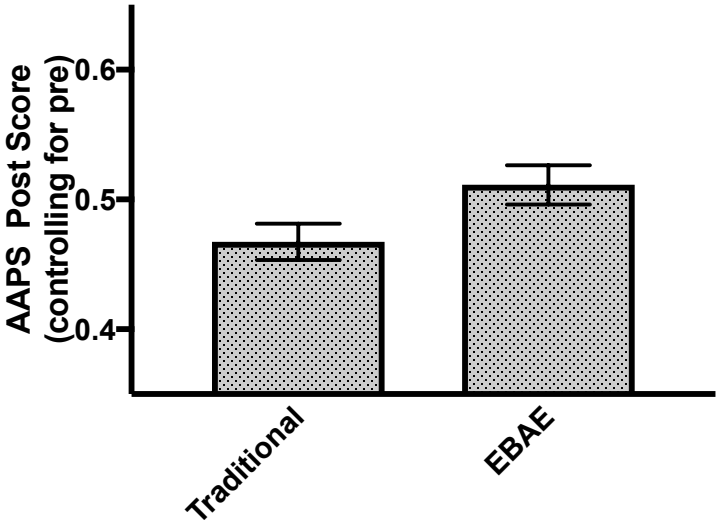


Figure 13: Controlling for the AAPS pre-test score in first semester classes for method of instruction, the AAPS post-test score for traditional classes is significantly lower than that of EBAE classes ($p = 0.027$). The estimated post-test score for traditional classes, controlling for pre-test score, is 0.467 and the estimated AAPS post-test score for EBAE classes is 0.511. Error bars represent standard error.

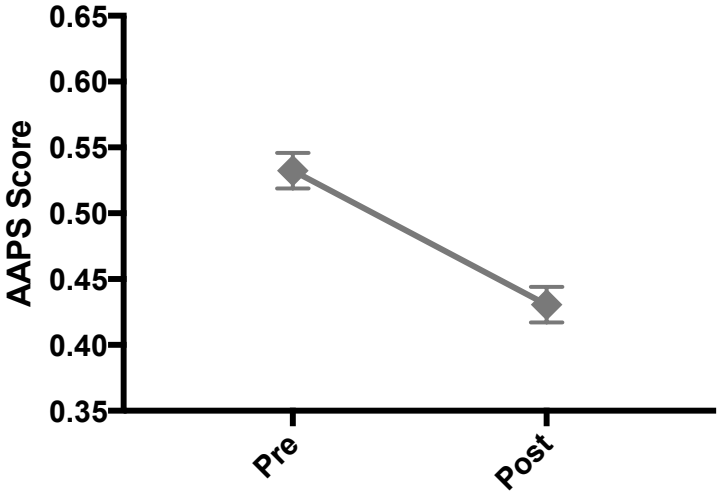


Figure 14: Raw data show the overall average scores for both pre and post AAPS survey in the calculus-based second semester classes (scores for two sections were averaged) taught in a traditional manner at University 1. The error bars represent standard error.

the traditional and EBAE classes was statistically significant ($p = 0.004$) suggesting that students who had a more positive attitude may have chosen instructors who use EBAE instruction over those who use traditional instruction when they enrolled (because they had knowledge of which section of the course would be taught by whom, and this information can be used to find out each instructor's teaching approach). It can also be seen in Figure 12 that the decline in average scores on the AAPS survey from pre to post was smaller for the active learning classes compared with the traditional classes. To examine whether there are differences in the post-test AAPS survey score, controlling for pre-test score, we performed a between-subjects ANOVA in which the dependent variable was the average AAPS survey post-test score, and the independent variable was method of instruction (i.e., EBAE vs. traditional). Controlling for the pretest score, we found a statistically significant difference in post-test score by method of instruction ($F(3, 413) = 54.56, p = 0.027$) [130], with the average AAPS survey score being significantly higher for the EBAE classes. The estimated post-test scores, controlling for the pre-test score, appears in Figure 13. The estimated post-test score for traditional classes, controlling for the pre-test score, is 0.467 with a 95% confidence interval of [0.439,0.495], and the estimated AAPS survey post-test score for EBAE classes is 0.511 with a 95% confidence interval of [0.481,0.540] [130].

The detailed breakdown of average percentages of favorable, unfavorable, and neutral responses on the AAPS survey by question and instructional method (traditional vs. EBAE) can be found in Table 9 of the Appendix. A comparison of this breakdown with the factors identified in the factor-analysis performed in Ref. [30] shows some patterns in the scores on individual questions. In particular, most of the questions included in Factor 4 show a 5% or more difference in average favorable scores for EBAE instruction compared with traditional instruction. This factor describes general expert-novice differences in problem solving. For example, awareness of more than one way to solve a problem, applying the same principle to different contexts, and perseverance in problem solving are all addressed by Factor 4. On this factor, students in EBAE courses, on average, scored higher than traditionally taught students by 5% or more in 4 out of 5 of the questions. In interviews, some

students from classes instructed in an EBAE manner verbalized some reasoning regarding the questions categorized into Factor 4 that demonstrated attitudes and approaches that are in line with general expert-like differences in problem solving. For example, regarding the statement "After I have solved several physics problems in which the same principle is applied in different contexts, I should be able to apply the same principle in other situations," one student from an EBAE class who responded favorable went on to give an expertlike explanation that, "We did force and Newton's laws, and then we kind of apply those same principles now, so it's the same way of going about the problem." This EBAE student saw connections to early problems to problems he encountered later, in terms of how the same principles apply, suggesting a more expert-like attitude about underlying principles when problem-solving.

Another factor for which students in EBAE courses outperformed traditionally taught students on the majority of the questions is Factor 1, which deals with metacognition and enjoyment in physics problem solving. Thinking about the underlying concepts and principles, considering if your answer is reasonable, and enjoying challenging problems are themes included in this factor. The students who had EBAE instruction outperformed traditionally taught students by 5% or more on 7 out of 12 of these questions on the post-test. Moreover, these differences in performance appear in every post-survey result (in addition to several pre-survey results), suggesting that these are areas in which attitudes and approaches were changed during the course of instruction. These findings suggest that students in EBAE classes not only started with better attitudes regarding the importance of metacognition in problem solving compared with traditionally taught students, but they also exhibited more expert-like attitudes about metacognition in problem solving after a semester of instruction compared with their traditionally taught peers. Moreover, there was an awareness of the importance of metacognition that students from EBAE classes demonstrated during the interviews. For example, one EBAE student who answered favorably to the statement, "When I solve physics problems, I always explicitly think about the concepts that underlie the problem," went on to emphasize how important thinking about the concepts is to

problem solving. He stated, "I think you need to know those concepts to understand what's going on... I think the more conceptual understanding that's there, the better." This EBAE student exhibited an expert-like perspective on the importance of carefully thinking about the concepts during problem-solving.

Higher performance by students in EBAE classes was not just in these two factors. Indeed, Table 9 shows that on many questions, EBAE sections had more favorable responses than traditionally taught students. The questions for which the average difference in score between EBAE and traditional sections was 5% or more appear in boldface in Table 9.

Gender differences: When we separated our data by gender and analyzed the responses, we found that unlike gender differences in content-based surveys there is often a gender gap in which female students' performance is lower than that of male students [116, 92], the average AAPS survey scores for female students were higher than those for male students, and they they remained higher at the end of instruction. Figures 15 - 18 show that female students exhibited less of a decline on the AAPS survey pre to post compared to male students regardless of whether they were instructed traditionally or with the EBAE methods. Since the decline among female students was negligible in the first semester, this suggests that the overall decline in the course was mainly driven by the male students because male students outnumbered female students (as can be see in Table 4) and male students exhibited a significant decline in scores. The normalized gain for first semester students was significantly lower for male students in traditional classes compared to female students ($p = 0.022$), but not in the EBAE classes ($p = 0.100$). Controlling for the pre-test score in first semester classes, there is a statistically significant difference in post-test score by gender with female students scoring significantly higher than male students ($F(2, 414) = 78.63, p = 0.033$). The estimated post-test scores for first semester classes, controlling for the pre-test score, appears in Figure 16. Controlling for pre-test score, the estimated AAPS survey post-test score for males in traditional classes is significantly lower than 0.444 with a 95% confidence interval of [0.415,0.469], the estimated AAPS survey post-test score for female students in traditional classes is 0.496 with a 95% confidence interval of [0.438,0.555], the estimated AAPS survey

post-test score for male students in the EBAE classes is 0.489 with a 95% confidence interval of [0.454,0.524], and the estimated AAPS survey post-test score for female students in the EBAE classes is 0.531 with a 95% confidence interval of [0.482,0.579] [130].

Similar to first semester classes, in the second semester classes (all taught traditionally), male students exhibited a significantly larger decline than female students on the AAPS survey ($p = 0.026$). However, female students' attitudes towards problem solving in second semester classes declined more than in first semester classes. Nevertheless when for pre-test scores are controlled for, there are statistically significant differences by gender in the second semester classes, with female students outperforming male students [$F(2, 250) = 59.867, p = 0.023$]. The estimated post-test scores for first semester classes, based on controlling for the pre-test score appears in Figure 18. Controlling for pre-test score, the estimated AAPS survey post-test score for male students in the second semester is 0.414 with a 95% confidence interval of [0.386, 0.442], while the estimated AAPS survey post-test score for female students in the second semester is 0.483 with a 95% confidence interval of [0.431,0.534].

Interviews revealed some further gender differences in problem-solving attitudes and approaches, particularly when it came to working with other students. Female students who were interviewed often reported feeling less comfortable and/or confident when working with male peers; whereas all interviewed male students reported feeling comfortable with either gender. Some reasons female students felt less comfortable working with male peers is revealed in many interview quotes. For example, one female student reported that she feels discouraged when she works with male peers: "It's usually guys that say you're totally wrong, but they don't always know themselves, so I think it's unfortunate that their first response is to say that." Another female student explained that she felt intimidated by working with male students: "I think I go to females first. I think I get more intimidated if I ask a guy because I think they're a lot more condescending about it and I don't like it." Both of these female students appear to be suggesting that their male peers may sometimes be dominant and/or overconfident in their interactions with female students.

General findings: A detailed breakdown of average favorable, unfavorable, and neutral

percentages on the AAPS survey by gender and method of instruction can be found for the first semester courses in the Appendix Table 10, and for the second semester courses in the Appendix Table 11. Questions which exhibit differences in score by gender of 10% or more appear in boldface, and questions in which the difference is statistically significant ($p \leq 0.01$ as found by a chi-square analysis) appear in Figures 19, 20, and 21. (Note that several additional questions exhibit statistically significant differences at the $p = 0.05$ significance level; however, we adopted a more stringent criteria to increase the probability that these differences are meaningful).

A comparison of this breakdown with the factors identified in Ref. [30] allows us to identify which individual questions might address similar themes. Every question in Factor 2 (usefulness of drawings and scratchwork in problem solving) shows a gender difference in favorable response for first semester students, and all but one question shows a gender difference in favorable response for second semester students, with female students having a more favorable response, on average, than male students by 10% or more in both EBAE and traditional classes. Most female students indicated in interviews that they would even do scratchwork or drawings for multiple choice questions. For example, one female student in an interview explained, "I like to do scratch work no matter what...I think I'll always draw a picture, because even for multiple choice it's just as important because it helps yield a correct answer." This student indicated that scratch-work and drawing are crucial to achieving the goal of obtaining a correct result when solving problems.

5.3.3 RQ3. Is there any difference between AAPS scores in introductory physics I (mechanics) and II (mainly electricity and magnetism)?

As can be seen in Figures 12 and 14, at University 1, the AAPS average scores were similar in the second semester compared with the first semester. The second semester classes in this investigation were all instructed in a traditional manner. We find that the average pre AAPS survey scores are approximately 0.50 for both first and second semester students, and about 0.44 for both first and second semester students. This suggests that students' attitudes

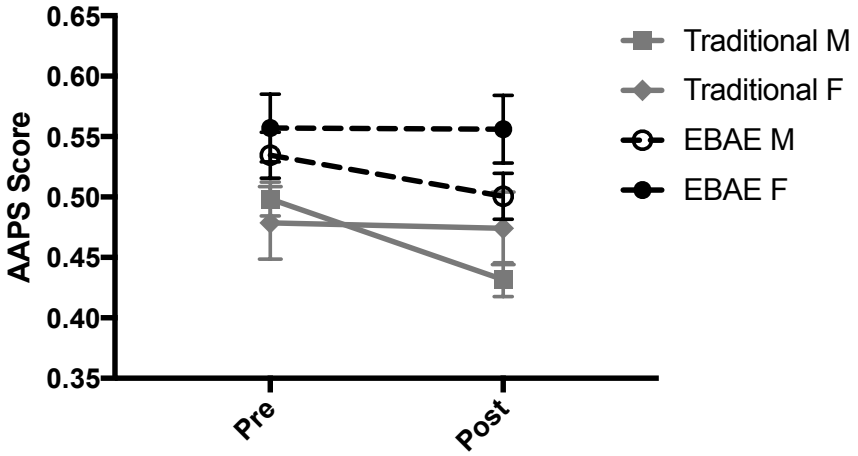


Figure 15: Raw data show the first semester average AAPS scores by gender and method of instruction at University 1. Average male scores (M) are indicated by a square for traditional male students (total number n = 197) and an open circle for EBAE male students (n = 118). Average female scores (F) are indicated by a diamond for traditional female students (n = 42) and a closed circle for EBAE female students (n = 61). Error bars represent standard error. In both traditional and EBAE classes, female students in their first semester of physics show very little decline in AAPS scores, whereas male students show a noticeable decline. The difference in normalized gain by gender is statistically significant for traditional instruction ($p = 0.022$), but not for the EBAE method of instruction ($p = 0.100$). Controlling for pre-test score, there are statistically significant differences in post-test score ($p = 0.03$).

may rebound when they begin their second semester and/or those who start the second semester have better attitudes at the beginning of the course than those in classes taught traditionally in the first semester (both at the beginning and end of the semester). Another possibility is that some students may not continue into the second semester of physics and these students may be ones with less favorable attitudes. Yet another possibility is that the second semester physics courses were all taught in a traditional manner, so students from first semester EBAE courses who have more favorable average attitudes and approaches (who continue into the second semester course) could inflate the pre-survey scores in the second semester traditionally taught course. However, regardless of the reason for the "rebound" at the beginning of the second semester, by the end of the second semester, students' attitudes suffer a similar loss to that which was observed in the first semester.

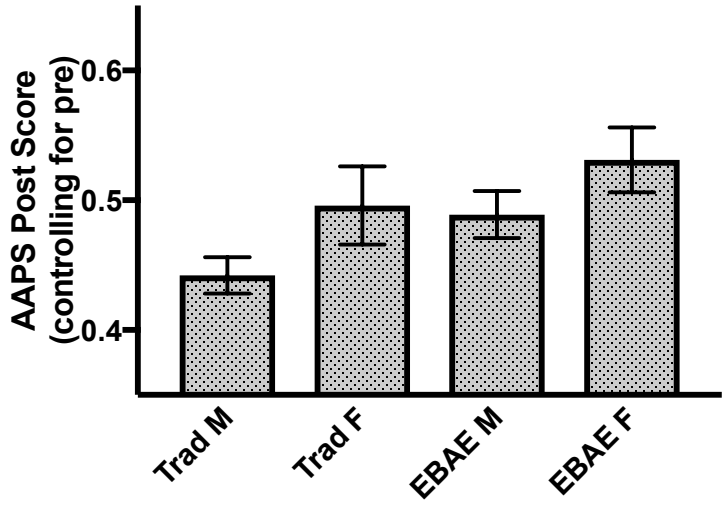


Figure 16: When controlling for the AAPS pre-test score in first semester classes for gender and method of instruction, the AAPS post-test scores show statistically significant differences ($F(3, 414) = 54.56, p = 0.033$) with female students scoring higher than male students in both traditional and EBAE classes. Error bars represent standard error.

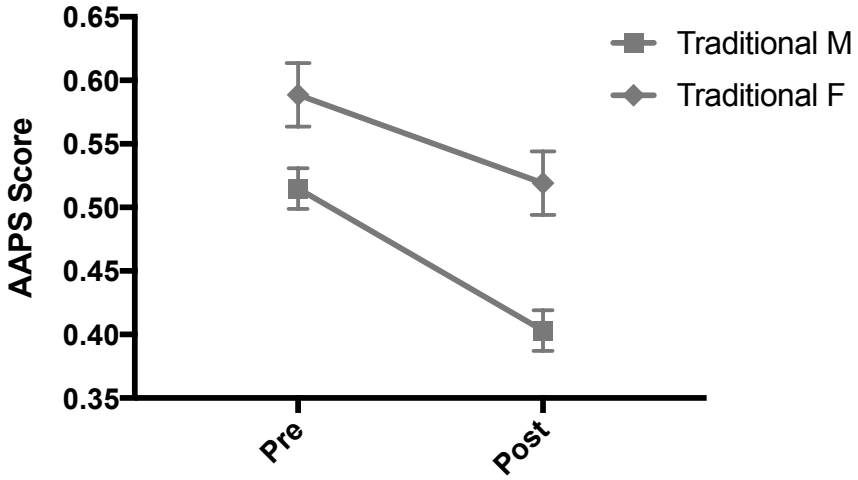


Figure 17: Raw data show the second semester AAPS survey scores by gender at University 1. Average male scores (M)) are indicated by a square (n = 195) and average female scores (F) are indicated by a diamond (n = 58). Error bars represent standard error. Female students in their second semester of physics show less of a decline on AAPS survey scores compared with male students. The difference in normalized loss by gender is statistically significant ($p = 0.026$). Controlling for pre-test scores, there are statistically significan differences by gender in the second semester classes ($p = 0.023$).

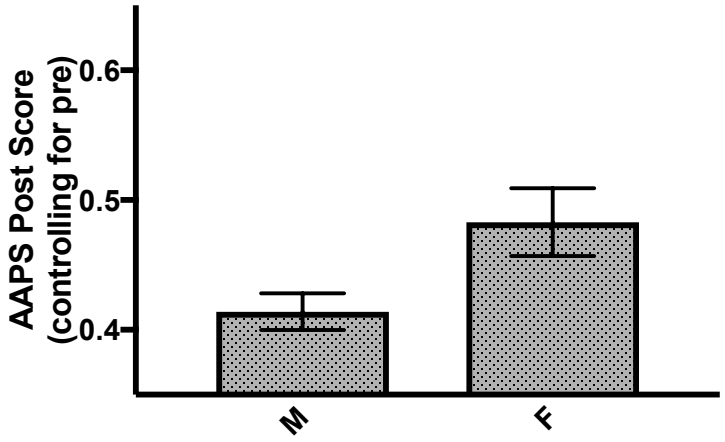


Figure 18: When controlling for the AAPS survey pre-test score in second semester classes for gender, there are statistically significant differences in AAPS survey post test score with female students outperforming male students ($F(2, 250) = 59.867, p = 0.023$). Controlling for the pre-test score, the estimated AAPS survey post-test score for male students in second semester classes is 0.414, while the estimated AAPS survey post-test score for female students in second semester classes is 0.483. Error bars represent standard error.

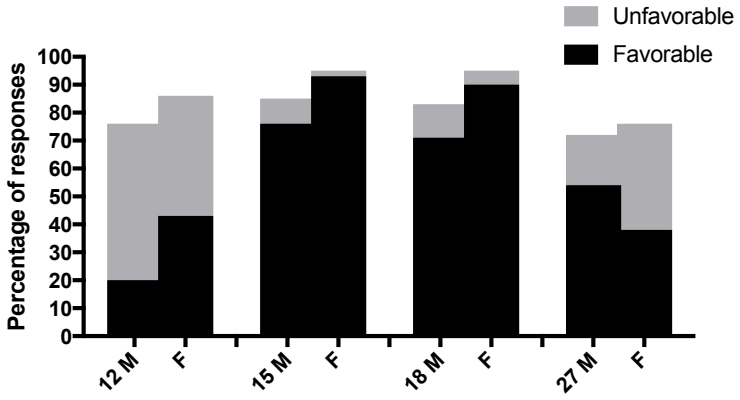


Figure 19: Individual questions in post AAPS survey that show statistically significant differences in male and female students' average favorable and unfavorable responses in **traditional first** semester classes at University 1. Favorable responses are the solid portion of each bar, unfavorable responses are the hashed portion, and the remainder from the top of the bar to 100% represents the portion of neutral responses. The p-values obtained using the chi-squared test were all $p \leq 0.01$.

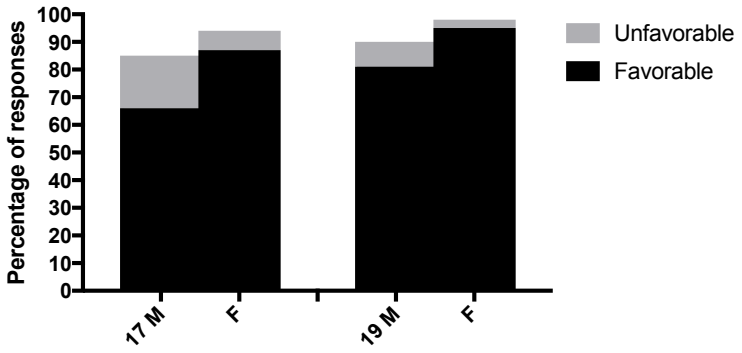


Figure 20: Individual questions in the post-AAPS survey that show statistically significant differences in male and female students' average favorable and unfavorable responses in **EBAE first** semester classes at University 1. Favorable responses are the solid portion of each bar, unfavorable responses are the hashed portion, and the remainder from the top of the bar to 100% represents the portion of neutral responses. The p-values obtained using the chi-squared test were all $p \leq 0.01$.

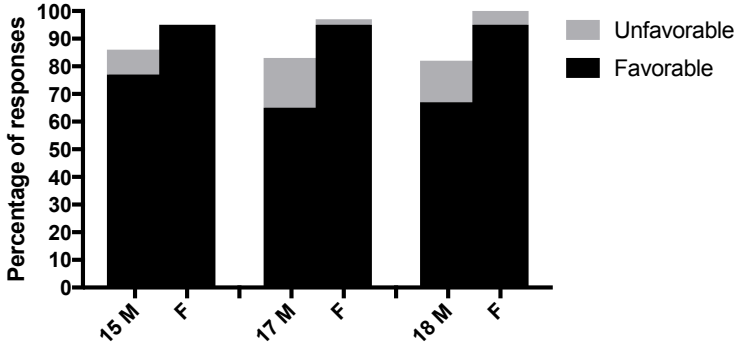


Figure 21: Individual questions in the post-AAPS survey that show differences in male and female students' average favorable responses in **traditional second** semester classes at University 1. Favorable responses are the solid portion of each bar, unfavorable responses are the hashed portion, and the remainder from the top of the bar to 100% represents the portion of neutral responses. The p-values obtained using the chi-squared test were all $p \leq 0.01$.

5.3.4 RQ4. Are students' AAPS scores in a given class correlated with their conceptual survey or final exam performance?

Before examining possible correlation between AAPS survey score and conceptual survey or final exam performance, it is helpful to first put the conceptual survey scores into perspective.

As seen in Figures 22 and 23, the FCI and CSEM performances were similar to typical values in other studies at large research universities [26, 27, 32, 90, 91, 116]. However, the normalized gain for the FCI was statistically significantly higher ($p = 0.033$) for EBAE classes compared with traditional classes. The FCI scores exhibited a gender gap which was present in both EBAE and traditional classes. For traditionally instructed first semester students, the gender gap was statistically significant at $p < 0.001$ for both the average pre and average post survey scores. Similarly, for first semester students who received EBAE methods of instruction, the gap was statistically significant at $p < 0.001$ for average pre and post survey scores. The gap did not decrease or increase statistically for either EBAE classes or traditional classes, based upon comparison of male and female normalized gain. However, the difference in normalized gain for females in EBAE classes compared to females in traditionally taught classes was statistically significant at $p < 0.001$. This suggests that female students may have gained more conceptual understanding in EBAE classes compared with traditional classes in their first semester of instruction.

For second semester students, there was a gender gap in CSEM performance that was not statistically significant for pre survey scores ($p = 0.060$), but was statistically significant in the post-survey scores ($p = 0.017$). There was no statistically significant change in the gender gap based on comparison of normalized gain for males vs. females in the second semester course.

When we examine these scores for possible correlation with the AAPS survey scores, we find that the overall AAPS survey scores and FCI/CSEM scores were not well correlated. All sections have correlation coefficients of $R < 0.3$ when considering all students. However, some small to medium correlations were found when separating the data by gender in first semester classes (for second semester classes the correlations by gender were weaker). In particular, performance on FCI post-test and on final exams is correlated, with the AAPS survey post-test performance for female students in EBAE classes, and is of medium effect ($R = 0.40 - 0.46$). Also final exam performance for female students in first semester traditional classes is correlated with the AAPS survey post-test score (of medium effect, $R = 0.48$). This suggests

that performance on content and/or problem-solving may be more correlated to performance on AAPS for female students.

In addition, since we saw a correlation in the first semester classes by gender, we further examined the correlation between one individual question on the AAPS survey and FCI scores in first semester classes. The motivation to explore this particular question for possible correlation with the FCI scores was twofold: 1. There appeared to be possible differences in this question by method of instruction. 2. It appeared possible that this question may imply a direct connection to performance on conceptual surveys because of the way it is worded. In particular, this question (question 16) states: "When answering conceptual physics questions, I mostly use my 'gut' feeling rather than using the physics principles I usually think about when solving quantitative problems." A favorable response would be to disagree with this statement. When correlation was examined for this question with FCI score it was found that there was a small positive correlation for female students between their score on this question in the post-test and their overall post-test FCI score. In particular, for female students in traditional classes, the correlation coefficient was found to be $R = 0.27$, and for female students in EBAE classes, the correlation coefficient was found to be $R = 0.47$. These correlation coefficients are summarized in Table 7. The small to medium correlations found suggest a possible connection between how one's attitude and approach is towards conceptual questions and performance on conceptual surveys, particularly for female students.

5.3.5 RQ5. How do the pre-/post-test scores on the AAPS and FCI compare at two different large research universities in the United States?

As shown in Figure 24, AAPS scores for first semester calculus-based introductory physics were similar for traditionally taught classes at two different universities. This finding suggests a level of consistency in average pre/post AAPS scores for similar student populations at comparable universities (both large state-related research universities) in traditionally taught large-enrollment introductory physics classes. However, we cannot conclude that AAPS performance would necessarily be similar at smaller universities or when class size is

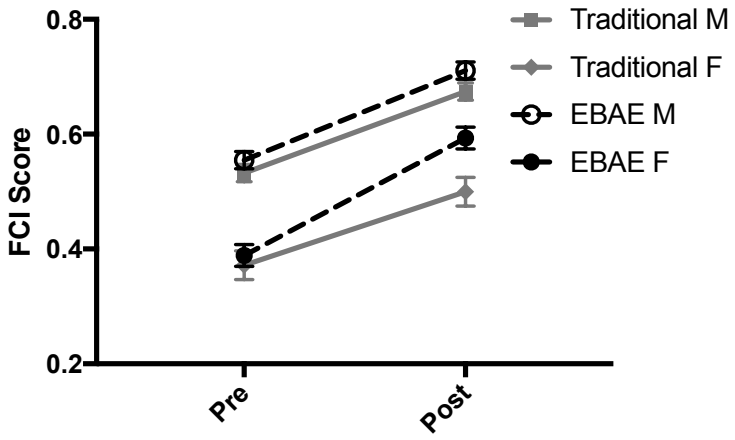


Figure 22: FCI performance by gender and type of instruction at University 1. Average male scores (M) are indicated by a square for traditional male students (total number n = 197) and an open circle for EBAE male students (n = 118). Average female scores (F) are indicated by a diamond for traditional female students (n = 42) and a closed circle for EBAE female students (n = 61). A gender gap which is significant is seen for both traditional instruction and EBAE instruction. Normalized gain for female students in EBAE classes compared with female students in traditional classes is statistically significant ($p < 0.001$).

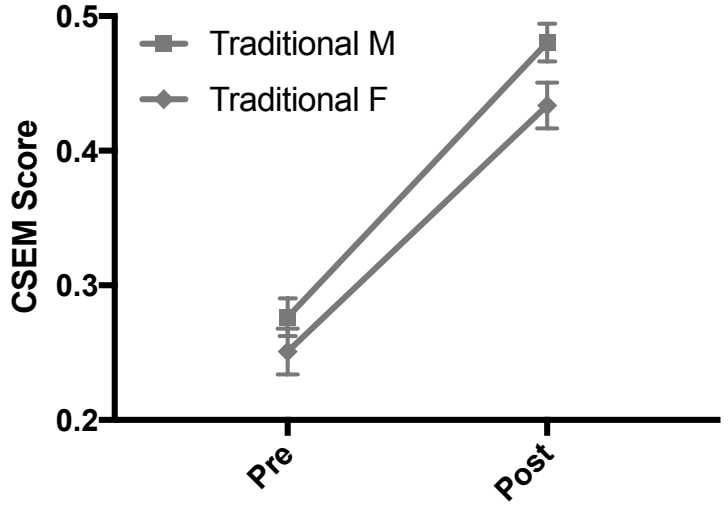


Figure 23: Traditionally-instructed University 1 second semester CSEM performance shows a gender gap which is not initially statistically significant ($p = 0.060$) in pre-survey scores but becomes statistically significant in post-survey scores ($p = 0.017$). The total number n = 195 for male students (M) and n = 58 for female students (F).

Class	Gender	R for AAPS and FCI post	R R for AAPS post and Final
Traditional 1st semester–University 1	M	0.24	0.22
Traditional 1st semester–University 1	F	0.21	**0.48**
EBAE 1st semester–University 1	M	0.13	0.11
EBAE 1st semester–University 1	F	**0.40**	**0.46**

Table 6: Correlation coefficients (R) for the AAPS survey scores and final exam or FCI scores for different types of courses at University 1 in first semester classes, broken down by gender. $R > 0.3$ appears in boldface

significantly smaller than the classes in this study even for the same type of course.

5.4 DISCUSSION

5.4.1 Findings regarding EBAE compared to traditional instruction

Our findings indicate that the post AAPS survey scores were better and more consistent for female students than for male students. Female students exhibited almost no decline in the AAPS survey score from pre to post survey in first semester physics and less of a decline than male students in second semester physics. Moreover, correcting for pre-test scores, significant differences in the AAPS survey post-test score were found by gender, as seen in Figure 16. Indeed, it appears that male student are driving the overall decline in AAPS survey scores from pre to post survey. This is somewhat unlike other attitudinal surveys (e.g., those focused on epistemological beliefs about physics) in which females show as much or more of a decline in scores from pre to post survey [32]. Because the AAPS survey

Class	Gender	R for AAPS question 16 and FCI post
Traditional 1st semester–University 1	M	0.06
Traditional 1st semester–University 1	F	0.27
EBAE 1st semester–University 1	M	0.02
EBAE 1st semester–University 1	F	**0.47**

Table 7: Correlation coefficients (R) for post test scores for the AAPS survey question 16 (which asks about using principles or "gut" when answering conceptual questions) and FCI for different types of courses at University 1 in first semester classes. $R > 0.3$ appears in boldface

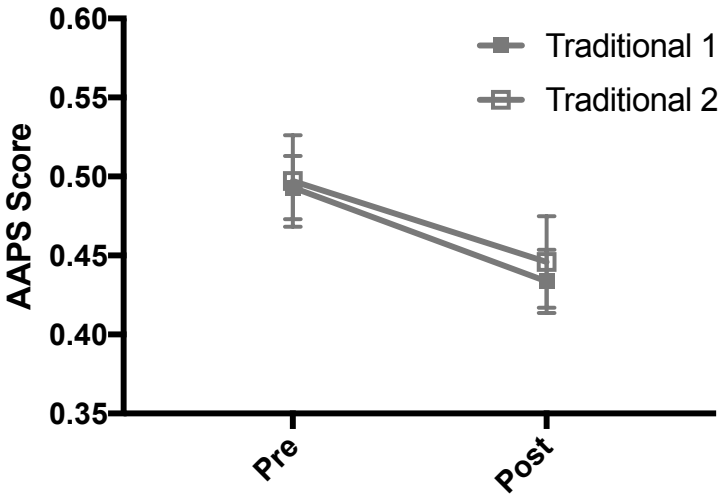


Figure 24: Overall average scores for both pre and post AAPS survey in the calculus-based first semester traditionally taught introductory physics classes at two large research universities.

incorporates questions related to attitudes and *approaches* to problem solving, not just epistemological beliefs of students about physics and physics learning like other surveys, it is possible that female students may be answering these questions (particularly those related to approaches to problem solving) differently than questions in previous surveys. The factor related to scratchwork is an example of a cluster of questions more related to approaches than attitudes, and female students score signficantly higher in questions from this factor. Individual interviews conducted with students appear to support this explanation.

5.4.2 Connections between gender, attitude, and instructional method

Our findings suggest that female introductory students have attitudes and approaches to problem-solving that are more expert-like than their male peers. Moreover, female students' attitudes and approaches do not suffer the same decline that male students' do. This finding suggests that instructional design should incorporate strategies to take advantage of female students' attitudes and approaches when it comes to solving physics problems and their persistence in maintaining more expert-like attitudes and approaches from the beginning to the end of the semester.

Yet it is equally well-known that women are significantly underrepresented in calculus-based introductory physics courses and in physics overall. Introductory physics classes give students their first experience with physics, and could open doors to further pursuit of physics. Although gender gaps, in which female students' performance lags behind that of male students are known to exist in tests of conceptual knowledge and/or problem solving, the correlations we have found among female students between attitude and approach to problem-solving and scores on conceptual surveys and exams suggest a possible path forward. If female students' attitudes and approaches to problem-solving show promise, as we have found, and can be supported during their experience in introductory physics classes, their conceptual performance and performance on high stakes exams may be impacted, since these two things appear to have some meaningful correlation for female students.

In addition to a correlation between attitude and approach to problem-solving and con-

ceptual performance, we also found that with regards to approaches to problem solving, female students engage in the use of drawings and scratchwork more than male students, and these two findings could be connected. In particular, it is possible that drawing could be an important means of promoting conceptual understanding for female students. This possibility is further suggested by interview data. For example, as one female student explained "I tend to draw a picture or diagram to help me visualize and understand conceptually what is going on." This student explicitly made a claim regarding drawing being a tool to help her understand problems conceptually. Responses such as these illuminate a possible strategy that could help promote conceptual understanding for female students, which could serve to reduce the gender performance gap that is often seen in conceptual surveys.

With regard to the way in which a class is taught, we find some evidence that female students make greater strides in conceptual understanding during the course of a semester under EBAE instruction, as evidenced by their higher normalized gain in FCI score compared with females in traditional classes. This finding is consistent with other studies that have shown that female students may develop better conceptual understanding in evidence-based active engagement classes compared to traditional classes [120, 90, 91]. Further studies into possible connections between different instructional approaches and differences in performance by gender on attitudinal and conceptual surveys would shed light on these issues.

In addition, at least part of the answer may lie in the extent to which instructors address women's sense of belonging, self-efficacy and mindset in a physics class in which they are often underrepresented [14, 90, 91]. Although we did not investigate these issues, our study did uncover some suggestion in interviews that female students may feel less comfortable working with male students, which could affect their sense of belonging in a class in which they are outnumbered by male students.

5.5 SUMMARY

We investigated the impact of instruction on attitudes and approaches to problem solving as measured by the AAPS survey administered both at the beginning and end of semester long introductory physics courses at two different large research universities in the United States with 784 students combined. We also compared the AAPS survey scores based upon method of instruction and gender of students, and examined data for correlations between the AAPS survey scores and performance on conceptual surveys (FCI and CSEM) or final exams. We found that groups of students at two different large state-related research universities exhibited similar scores on the AAPS survey in similar types of classes, and that the use of evidence-based active engagement methods can have a positive impact on student attitudes and approaches to problem solving. Moreover, there was a decline in students' scores on the AAPS survey from the beginning to the end of the semester, but female students generally showed less decline than male students. The correlation between the AAPS survey scores and performance on conceptual surveys or final exams appeared stronger for female students compared with male students, and similarly small correlations were found for an individual AAPS survey question and FCI performance, that were stronger for female students than for male students. These findings can be useful for instructors concerned with the type of instructional method that may promote favorable attitudes and approaches to physics problem solving among their students. In addition, those concerned with the under-representation of females in physics may find these findings illuminating, given the promising responses that female students provide on the AAPS survey.

Factor (% of variance explained)	Item	Loading	Description
Factor1 (13)	13	.77	Metacognition and enjoyment in physics problem solving
	14	.73	
	7	.70	
	10	.69	
	22	.63	
	29	.60	
	20	.51	
	4	.57	
	25	.38	
	27	.27	
	6	.28	
	21	.29	
Factor2 (8)	18	.90	Utility of pictures, diagrams or scratch work in physics problem solving
	17	.88	
	15	.75	
	19	.65	
Factor3 (6)	5	.69	Perception of problem solving approach
	11	.68	
	12	.55	
	8	.33	
	26	.28	
	9	-.28	
Factor4 (5)	8	.76	General expert-novice differences in physics problem solving
	28	.56	
	21	.43	
	24	.43	
	29	.30	
Factor5 (5)	31	.84	Difficulty in solving problems symbolically
	30	.83	
Factor6 (5)	1	.75	Problem solving confidence
	24	-.71	
	23	.50	
	6	.30	
Factor7 (5)	33	.88	Solving different problems using the same principle
	32	.80	
Factor8 (4)	16	.76	Sense-making
	2	-.56	
	5	.30	
Factor9 (3)	3	.77	Problem solving sophistication
	20	-.35	
	25	-.32	
	9	.29	

Table 8: Principal component analysis results featuring 9 primary factors and description, reproduced from Mason [30].

Problem number	1		2		3		4		5		6		7	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad	56%	55%	59%	66%	**27%**	**31%**	69%	74%	**40%**	**31%**	65%	**67%**	**81%**	77%
Favorable--EBAE	59%	59%	59%	70%	**42%**	**44%**	73%	71%	**50%**	**46%**	68%	**76%**	**87%**	81%
Neutral--Trad	17%	14%	27%	22%	14%	21%	20%	18%	15%	16%	24%	21%	12%	16%
Neutral--EBAE	10%	12%	30%	15%	17%	18%	15%	19%	18%	15%	17%	16%	9%	14%
Unfavorable--Trad	27%	32%	14%	13%	59%	49%	11%	8%	46%	53%	11%	12%	7%	7%
Unfavorable--EBAE	31%	29%	11%	15%	41%	37%	11%	11%	32%	39%	15%	8%	4%	5%
Normalized--Trad	*0.29*	*0.23*	*0.45*	*0.53*	*-0.33*	*-0.18*	*0.59*	*0.67*	*-0.06*	*-0.22*	*0.54*	*0.56*	*0.74*	*0.70*
Normalized--EBAE	*0.28*	*0.30*	*0.48*	*0.55*	*0.02*	*0.07*	*0.62*	*0.60*	*0.18*	*0.06*	*0.53*	*0.68*	*0.83*	*0.76*
Problem number	**8**		**9**		**10**		**11**		**12**		**13**		**14**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad	**75%**	**78%**	31%	54%	81%	**71%**	**38%**	**32%**	28%	**24%**	84%	75%	69%	**61%**
Favorable--EBAE	**81%**	**87%**	34%	57%	85%	**78%**	**47%**	**41%**	30%	**35%**	87%	79%	69%	**67%**
Neutral--Trad	15%	11%	61%	31%	12%	20%	39%	29%	19%	22%	11%	16%	20%	25%
Neutral--EBAE	14%	5%	52%	26%	12%	16%	33%	31%	16%	21%	8%	13%	18%	20%
Unfavorable--Trad	10%	12%	8%	15%	6%	9%	24%	39%	53%	54%	5%	9%	13%	14%
Unfavorable--EBAE	5%	8%	14%	18%	3%	6%	19%	29%	54%	44%	4%	8%	12%	13%
Normalized--Trad	*0.65*	*0.66*	*0.22*	*0.38*	*0.75*	*0.62*	*0.14*	*-0.07*	*-0.26*	*-0.30*	*0.79*	*0.66*	*0.56*	*0.48*
Normalized--EBAE	*0.76*	*0.79*	*0.20*	*0.39*	*0.82*	*0.72*	*0.28*	*0.12*	*-0.24*	*-0.09*	*0.83*	*0.71*	*0.57*	*0.53*
Problem number	**15**		**16**		**17**		**18**		**19**		**20**		**21**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad	87%	79%	**50%**	**40%**	73%	72%	84%	75%	92%	85%	34%	30%	85%	**77%**
Favorable--EBAE	88%	83%	**56%**	**58%**	74%	73%	83%	77%	92%	86%	36%	30%	89%	**86%**
Neutral--Trad	9%	15%	28%	31%	9%	12%	9%	15%	2%	10%	34%	35%	13%	18%
Neutral--EBAE	8%	12%	23%	22%	10%	12%	11%	16%	5%	7%	37%	31%	7%	9%
Unfavorable--Trad	4%	6%	22%	29%	18%	17%	7%	10%	6%	5%	31%	35%	2%	5%
Unfavorable--EBAE	4%	6%	21%	20%	15%	15%	6%	7%	3%	7%	27%	39%	4%	4%
Normalized--Trad	*0.84*	*0.72*	*0.28*	*0.11*	*0.55*	*0.55*	*0.77*	*0.65*	*0.86*	*0.80*	*0.03*	*-0.05*	*0.83*	*0.72*
Normalized--EBAE	*0.84*	*0.77*	*0.35*	*0.38*	*0.59*	*0.58*	*0.78*	*0.70*	*0.88*	*0.78*	*0.09*	*-0.09*	*0.86*	*0.82*
Problem number	**22**		**23**		**24**		**25**		**26**		**27**		**28**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad	84%	**69%**	63%	**49%**	84%	**82%**	83%	72%	73%	74%	73%	**51%**	90%	87%
Favorable--EBAE	86%	**77%**	64%	**56%**	88%	**76%**	87%	73%	77%	76%	73%	**58%**	92%	87%
Neutral--Trad	12%	22%	23%	29%	10%	12%	14%	19%	21%	17%	20%	27%	9%	9%
Neutral--EBAE	13%	15%	21%	26%	8%	14%	10%	18%	15%	14%	21%	27%	6%	8%
Unfavorable--Trad	5%	10%	13%	23%	6%	7%	4%	8%	6%	9%	7%	22%	1%	4%
Unfavorable--EBAE	2%	7%	15%	18%	3%	9%	3%	8%	8%	9%	7%	15%	2%	4%
Normalized--Trad	*0.79*	*0.60*	*0.50*	*0.26*	*0.78*	*0.75*	*0.79*	*0.64*	*0.67*	*0.65*	*0.66*	*0.29*	*0.89*	*0.83*
Normalized--EBAE	*0.84*	*0.70*	*0.49*	*0.38*	*0.85*	*0.67*	*0.84*	*0.66*	*0.69*	*0.67*	*0.66*	*0.43*	*0.90*	*0.82*
Problem number	**29**		**30**		**31**		**32**		**33**		**Overall**			
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post		Pre	Post	
Favorable--Trad	95%	**84%**	29%	35%	**33%**	45%	79%	**84%**	55%	**60%**		65%	**62%**	
Favorable--EBAE	92%	**91%**	33%	39%	**41%**	41%	76%	**89%**	53%	**76%**		68%	**67%**	
Neutral--Trad	4%	11%	34%	24%	26%	22%	16%	13%	33%	28%		19%	20%	
Neutral--EBAE	6%	7%	33%	26%	24%	20%	17%	7%	33%	14%		17%	17%	
Unfavorable--Trad	1%	5%	36%	41%	41%	33%	4%	3%	11%	12%		16%	18%	
Unfavorable--EBAE	2%	1%	33%	34%	36%	38%	6%	3%	14%	9%		14%	16%	
Normalized--Trad	*0.94*	*0.78*	*-0.06*	*-0.06*	*-0.08*	*0.12*	*0.75*	*0.81*	*0.44*	*0.48*		***0.49***	***0.44***	
Normalized--EBAE	*0.90*	*0.90*	*0.00*	*0.06*	*0.05*	*0.03*	*0.70*	*0.86*	*0.39*	*0.67*		***0.54***	***0.51***	

Table 9: Breakdown of average favorable, unfavorable, and neutral responses to individual questions of the AAPS by method of instruction (trad means traditionally taught) in first semsester classes at University 1. Favorable responses that differ by 5% or more appear in boldface.

Problem number	1		2		3		4		5		6		7	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad M	56%	**58%**	60%	**68%**	24%	29%	73%	**78%**	42%	**28%**	65%	68%	83%	77%
Favorable--Trad F	55%	**43%**	48%	**57%**	36%	36%	52%	**60%**	34%	**45%**	64%	69%	74%	76%
Favorable--EBAE M	63%	59%	61%	**73%**	46%	43%	72%	69%	45%	44%	71%	76%	85%	**76%**
Favorable--EBAE F	52%	60%	57%	**63%**	37%	48%	78%	75%	59%	48%	61%	75%	89%	**90%**
Neutral--Trad M	19%	13%	26%	21%	14%	21%	17%	16%	15%	16%	26%	21%	12%	16%
Neutral--Trad F	9%	19%	33%	21%	14%	19%	33%	26%	12%	12%	17%	21%	17%	17%
Neutral--EBAE M	9%	13%	27%	14%	14%	16%	18%	22%	21%	17%	15%	17%	9%	16%
Neutral--EBAE F	11%	12%	33%	17%	23%	22%	8%	12%	11%	12%	21%	13%	10%	10%
Unfavorable--Trad M	25%	29%	13%	11%	61%	50%	10%	6%	44%	55%	9%	12%	6%	7%
Unfavorable--Trad F	35%	38%	19%	21%	50%	45%	14%	14%	52%	43%	19%	10%	10%	7%
Unfavorable--EBAE M	27%	28%	12%	13%	40%	40%	10%	9%	34%	39%	14%	7%	5%	8%
Unfavorable--EBAE F	36%	28%	10%	20%	39%	30%	13%	13%	30%	40%	18%	12%	2%	0%
Problem number	**8**		**9**		**10**		**11**		**12**		**13**		**14**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad M	74%	77%	33%	53%	83%	70%	39%	32%	27%	**20%**	83%	75%	70%	61%
Favorable--Trad F	79%	83%	19%	60%	74%	74%	33%	31%	29%	**43%**	88%	74%	50%	64%
Favorable--EBAE M	80%	86%	37%	58%	85%	80%	47%	**35%**	28%	36%	87%	79%	69%	64%
Favorable--EBAE F	84%	90%	30%	57%	85%	77%	48%	**53%**	34%	32%	87%	80%	67%	73%
Neutral--Trad M	16%	12%	59%	33%	11%	21%	38%	27%	17%	24%	12%	16%	19%	26%
Neutral--Trad F	9%	5%	71%	24%	17%	19%	48%	36%	29%	14%	7%	19%	26%	21%
Neutral--EBAE M	14%	8%	51%	25%	12%	14%	36%	36%	17%	20%	8%	13%	17%	25%
Neutral--EBAE F	13%	10%	52%	27%	11%	17%	30%	20%	15%	23%	8%	13%	21%	10%
Unfavorable--Trad M	10%	12%	8%	14%	6%	9%	23%	40%	55%	56%	6%	9%	11%	12%
Unfavorable--Trad F	12%	12%	10%	16%	10%	7%	19%	33%	43%	43%	5%	7%	24%	14%
Unfavorable--EBAE M	6%	6%	12%	18%	3%	6%	16%	30%	55%	43%	4%	8%	13%	12%
Unfavorable--EBAE F	3%	0%	18%	17%	3%	7%	23%	27%	51%	45%	5%	7%	11%	17%
Problem number	**15**		**16**		**17**		**18**		**19**		**20**		**21**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad M	86%	**76%**	51%	41%	70%	**68%**	82%	**71%**	91%	**83%**	34%	31%	86%	77%
Favorable--Trad F	93%	**93%**	48%	40%	90%	**86%**	93%	**90%**	98%	**95%**	38%	29%	83%	81%
Favorable--EBAE M	84%	**80%**	54%	56%	67%	**66%**	78%	**71%**	89%	**81%**	34%	28%	89%	86%
Favorable--EBAE F	97%	**90%**	57%	62%	89%	**87%**	93%	**90%**	97%	**95%**	39%	35%	90%	90%
Neutral--Trad M	10%	17%	27%	30%	10%	13%	9%	17%	2%	10%	37%	35%	13%	19%
Neutral--Trad F	2%	5%	31%	36%	2%	5%	7%	5%	2%	5%	26%	38%	12%	14%
Neutral--EBAE M	10%	14%	26%	24%	15%	14%	15%	19%	7%	9%	39%	29%	7%	11%
Neutral--EBAE F	3%	5%	18%	18%	2%	7%	3%	8%	2%	2%	33%	35%	7%	5%
Unfavorable--Trad M	4%	7%	22%	29%	20%	19%	9%	12%	7%	7%	30%	35%	2%	5%
Unfavorable--Trad F	5%	2%	21%	24%	7%	10%	0%	5%	0%	0%	36%	33%	5%	5%
Unfavorable--EBAE M	6%	6%	20%	20%	18%	19%	7%	10%	4%	9%	27%	43%	4%	3%
Unfavorable--EBAE F	0%	5%	25%	20%	10%	7%	3%	2%	2%	3%	28%	30%	3%	5%
Problem number	**22**		**23**		**24**		**25**		**26**		**27**		**28**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad M	85%	69%	62%	48%	83%	81%	83%	**71%**	73%	73%	74%	**54%**	92%	88%
Favorable--Trad F	76%	69%	68%	57%	86%	83%	83%	**83%**	71%	81%	71%	**38%**	81%	81%
Favorable--EBAE M	85%	76%	64%	58%	87%	**73%**	85%	**69%**	76%	75%	75%	61%	91%	86%
Favorable--EBAE F	87%	81%	66%	53%	90%	**86%**	90%	**81%**	80%	81%	69%	54%	93%	88%
Neutral--Trad M	10%	22%	24%	28%	11%	12%	14%	22%	21%	17%	18%	27%	8%	9%
Neutral--Trad F	19%	19%	20%	26%	5%	10%	12%	7%	21%	17%	24%	24%	14%	12%
Neutral--EBAE M	15%	16%	18%	24%	9%	17%	11%	22%	17%	16%	22%	25%	7%	9%
Neutral--EBAE F	10%	12%	23%	31%	8%	10%	8%	12%	10%	10%	16%	29%	5%	7%
Unfavorable--Trad M	5%	9%	13%	24%	6%	7%	4%	7%	6%	10%	8%	18%	0%	4%
Unfavorable--Trad F	5%	12%	13%	17%	10%	7%	5%	10%	7%	2%	5%	38%	5%	7%
Unfavorable--EBAE M	1%	8%	17%	18%	4%	10%	3%	8%	7%	9%	3%	14%	3%	4%
Unfavorable--EBAE F	3%	7%	11%	17%	2%	3%	3%	7%	10%	8%	15%	17%	2%	5%
Problem number	**29**		**30**		**31**		**32**		**33**			**Overall**		
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post		Pre	Post	
Favorable--Trad M	95%	84%	30%	36%	32%	47%	79%	83%	53%	**57%**		64%	62%	
Favorable--Trad F	95%	86%	24%	31%	36%	38%	81%	90%	68%	**76%**		64%	65%	
Favorable--EBAE M	88%	92%	32%	38%	40%	42%	76%	92%	54%	76%		67%	66%	
Favorable--EBAE F	98%	92%	34%	42%	41%	39%	77%	86%	50%	78%		70%	70%	
Neutral--Trad M	4%	11%	35%	23%	27%	22%	16%	14%	36%	29%		19%	20%	
Neutral--Trad F	5%	10%	36%	26%	24%	21%	14%	10%	22%	17%		19%	18%	
Neutral--EBAE M	9%	6%	33%	28%	24%	20%	17%	5%	32%	15%		18%	18%	
Neutral--EBAE F	0%	8%	33%	24%	25%	20%	16%	8%	35%	12%		16%	14%	
Unfavorable--Trad M	1%	5%	35%	42%	41%	31%	4%	3%	11%	13%		16%	18%	
Unfavorable--Trad F	0%	5%	40%	43%	40%	40%	5%	0%	10%	7%		16%	18%	
Unfavorable--EBAE M	3%	2%	34%	34%	36%	38%	6%	3%	13%	8%		14%	16%	
Unfavorable--EBAE F	2%	0%	33%	34%	34%	41%	7%	5%	15%	10%		14%	15%	

Table 10: Breakdown of average favorable, unfavorable, and neutral responses to individual questions of the AAPS by gender and method of instruction in first semester classes at University 1. Post-survey favorable responses that differ by gender by 10% or more appear in boldface and are underlined.

Problem number	1		2		3		4		5		6		7	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad 2 M	64%	**58%**	60%	62%	38%	46%	77%	69%	42%	38%	68%	**64%**	87%	79%
Favorable--Trad 2 F	47%	**47%**	59%	60%	41%	43%	74%	74%	50%	40%	67%	**53%**	88%	72%
Neutral--Trad 2 M	9%	12%	25%	24%	20%	15%	11%	15%	22%	19%	18%	18%	9%	12%
Neutral--Trad 2 F	7%	14%	24%	16%	22%	17%	16%	7%	16%	9%	22%	24%	7%	17%
Unfavorable--Trad 2 M	27%	29%	15%	14%	42%	38%	12%	16%	36%	42%	14%	17%	5%	9%
Unfavorable--Trad 2 F	41%	34%	16%	19%	31%	33%	9%	14%	29%	47%	7%	17%	3%	7%
Problem number	**8**		**9**		**10**		**11**		**12**		**13**		**14**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad 2 M	89%	**69%**	49%	**45%**	77%	65%	39%	32%	40%	**31%**	85%	70%	64%	55%
Favorable--Trad 2 F	97%	**84%**	55%	**55%**	83%	74%	47%	38%	36%	**47%**	90%	74%	62%	59%
Neutral--Trad 2 M	7%	15%	36%	38%	13%	24%	33%	24%	11%	17%	7%	17%	23%	25%
Neutral--Trad 2 F	2%	2%	22%	28%	10%	19%	29%	33%	26%	12%	3%	9%	29%	22%
Unfavorable--Trad 2 M	4%	15%	14%	16%	10%	11%	28%	43%	36%	51%	8%	13%	13%	20%
Unfavorable--Trad 2 F	0%	9%	17%	12%	5%	3%	22%	24%	34%	36%	5%	14%	7%	14%
Problem number	**15**		**16**		**17**		**18**		**19**		**20**		**21**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad 2 M	88%	**77%**	55%	**46%**	69%	**65%**	80%	**67%**	89%	85%	33%	32%	86%	78%
Favorable--Trad 2 F	98%	**91%**	71%	**57%**	86%	**91%**	91%	**91%**	90%	88%	40%	31%	91%	81%
Neutral--Trad 2 M	8%	13%	25%	26%	9%	16%	12%	8%	4%	8%	35%	24%	10%	16%
Neutral--Trad 2 F	0%	5%	17%	21%	5%	3%	2%	0%	2%	3%	40%	33%	3%	10%
Unfavorable--Trad 2 M	4%	9%	21%	28%	22%	18%	8%	15%	7%	7%	32%	44%	4%	6%
Unfavorable--Trad 2 F	0%	0%	9%	17%	7%	2%	5%	5%	7%	3%	17%	31%	2%	5%
Problem number	**22**		**23**		**24**		**25**		**26**		**27**		**28**	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Favorable--Trad 2 M	83%	**62%**	70%	59%	82%	75%	79%	71%	68%	**61%**	66%	42%	96%	82%
Favorable--Trad 2 F	78%	**72%**	71%	62%	86%	81%	84%	71%	74%	**72%**	2%	45%	95%	84%
Neutral--Trad 2 M	10%	20%	18%	26%	10%	12%	15%	21%	21%	24%	21%	27%	4%	13%
Neutral--Trad 2 F	16%	14%	12%	16%	7%	5%	12%	19%	16%	14%	26%	19%	3%	9%
Unfavorable--Trad 2 M	7%	17%	12%	15%	8%	13%	6%	8%	11%	16%	13%	30%	1%	5%
Unfavorable--Trad 2 F	5%	9%	14%	17%	5%	10%	2%	5%	7%	9%	9%	31%	0%	2%
Problem number	**29**		**30**		**31**		**32**		**33**		**Overall**			
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post		Pre	Post	
Favorable--Trad 2 M	90%	88%	33%	31%	37%	45%	83%	84%	63%	70%		68%	61%	
Favorable--Trad 2 F	97%	91%	28%	36%	41%	45%	79%	86%	62%	75%		70%	66%	
Neutral--Trad 2 M	8%	10%	28%	25%	21%	18%	13%	11%	29%	21%		16%	19%	
Neutral--Trad 2 F	2%	3%	26%	21%	21%	16%	16%	9%	26%	16%		15%	14%	
Unfavorable--Trad 2 M	2%	3%	38%	44%	42%	37%	4%	5%	7%	8%		16%	20%	
Unfavorable--Trad 2 F	0%	0%	45%	38%	34%	36%	3%	2%	9%	3%		12%	15%	

Table 11: Breakdown of average favorable, unfavorable, and neutral responses to individual questions of the AAPS by gender in second semester traditionally taught classes at University 1. Favorable responses that differ by gender by 10% or more appear in boldface and are underlined.

6.0 COMPARING INTRODUCTORY PHYSICS AND ASTRONOMY STUDENTS' ATTITUDES AND APPROACHES TO PROBLEM SOLVING

6.1 INTRODUCTION

6.1.1 Differences between introductory physics and introductory astronomy classes and students

A typical calculus-based or algebra-based introductory physics course may offer a different experience for students compared to a typical introductory astronomy course. Although the difficulty level of these courses is comparable for undergraduate students and they cover similar physics principles underlying the relevant content, the content of the two courses is generally organized around very different themes. For example, the introductory physics course often focuses on a linear progression in terms of introducing students to various concepts and fundamental laws one by one, often in a bottom-up approach. In particular, in a typical algebra-based or calculus-based physics course for biological science, physical science, and engineering majors, students are typically introduced to vectors and various kinematic and dynamic variables. This is generally followed by coverage of Newton's laws of motion, implulse and momentum, and work and energy. Then students learn about rotational kinematics and dynamics followed by simple harmonic motion, gravitation and waves. Meanwhile in a typical introductory astronomy course for physics and astronomy majors, students learn about observational techniques, stars and stellar evolution, and the

interstellar medium. This is followed by study of galaxies and cosmology. In a typical introductory astronomy course similar to the one we focus on here, the material is taught in a quantitative fashion with astrophysical problems requiring the same level of mathematical skill and rigor as introductory physics classes.

From the manner in which an introductory astronomy course for physics and astronomy majors is organized, it is clear that the focus is on understanding the large-scale structure of the universe and the rules that govern the past and future evolution of astrophysical objects. In an introductory physics course, the focus is often on helping students learn to apply the laws of physics in simplified contexts, e.g., applying Newton's laws to blocks on inclined planes or masses connected via ropes and pulleys.

While requiring the same level of rigor, the classes are often taught differently. Partly due to the focus of the course, astronomy classes often involve the use of many photographic images of objects in space; whereas physics classes often rely upon sketches, cartoons, or other abstract representations of the objects being studied. Introductory physics is often mandatory for science and engineering majors (who constitute a majority of students in the class), but in introductory astronomy, while many students take it as a science elective, a majority of students are physical science or engineering majors (since the course discussed here is listed as a course for science and engineering majors and others are advised to enroll in another course with similar content but lower mathematical rigor). Also, since all physical science and engineering majors at the University of Pittsburgh are required to take a two-course sequence in introductory calculus-based physics in their freshman year, some students in the astronomy class may have already taken introductory physics.

6.1.2 The role of expertise in problem solving

Experts, such as physics faculty members, organize their knowledge hierarchically, such that underlying concepts which are related are connected in a meaningful and structured way [22, 23, 131, 44, 132, 133]. This knowledge structure allows experts to efficiently approach the problem-solving process [23]. By contrast, novices, who are trying to develop expertise, such

as introductory students, may view physics as a collection of disconnected facts and equations [22, 23]. The lack of organization to their knowledge structure can result in introductory students approaching each problem as a unique challenge in which they search for the correct formula or resort to "plug and chug" [23]. The approaches and attitudes students use in their learning and problem-solving can impact the extent to which they take the time to organize their knowledge structure [111, 112, 25, 26, 22, 23]. This can, in turn, influence the acquistion of conceptual understanding and problem-solving skills [111, 112, 25, 26].

6.1.3 Assessing introductory physics students' views using attitudinal surveys

Understanding students' attitudes and approaches to problem solving is important because it can have instructional implications. Attitudinal surveys have been developed to assess students' beliefs about physics. The Maryland Physics Expectation Survey (MPEX) and the Colorado Attitudes about Science Survey (CLASS) are similar in that they focus on assessing attitudes students have about physics [26, 27]. The Epistemological Beliefs Assessment for Physics Science (EBAPS) survey was designed and validated to probe purely epistimological stances of students of physical sciences along multiple dimensions [29]. The Attitudes toward Problem Solving Survey (APSS) was developed through inspiration from the MPEX survey, with a focus on attitudes about problem solving [28, 25].

A modified version of the APSS, the Attitudes and Approaches to Problem Solving (AAPS) survey, was developed to include questions regarding the approaches students take when solving problems [30]. This survey was validated based upon faculty, graduate student, and introductory student responses [30]. During validation, expected "favorable" responses agreed with faculty responses to a high degree, and it was found to be the case that, on average, introductory students responded differently from graduate students and faculty, in that introductory student responses were found to be less "favorable" compared to graduate students and faculty [30].

6.1.4 Focus of our research

The focus of our research presented here was on analyzing, comparing, and interpreting the Attitudes and Approaches to Problem Solving (AAPS) survey responses for introductory algebra-based and calculus-based physics and introductory astronomy students. All the courses involved in this investigation (whether physics or astronomy) were for science and engineering majors. Our research questions are as follows:

RQ1. Are there differences in average overall performance on the AAPS survey for introductory astronomy students compared with introductory physics students?

RQ2. Are there differences in performance on specific clusters of questions on the AAPS survey for introductory astronomy students compared with introductory physics students?

RQ3. When presented with an isomorphic problem pair (problems with the same underlying physics principle but different contexts–one written in an astronomy context and the other written in a physics context), are introductory physics and introductory astronomy students equally likely to solve both problems correctly and do they find either more or less interesting?

6.2 METHODOLOGY

6.2.1 Courses and participants

The courses in our study consisted of both algebra-based and calculus-based introductory physics, and introductory astronomy at the University of Pittsburgh. Note that no statistically significant difference was found in the AAPS scores for algebra-based introductory physics compared with calculus-based introductory physics, so these two types of classes were combined. The introductory astronomy course discussed here is intended primarily for science and engineering majors, and is a required part of the curriculum for undergradu-

ate students who are majoring in "physics and astronomy" (although many students enroll in this course as an elective course). The algebra-based introductory physics classes are required for some science majors, such as biology. The calculus-based physics classes are required for engineering and physics majors (note that physics majors do have the option to take an honors version of these classes instead, which were not included in this study, so not all physics majors take the classes included in this study). A total of 606 students participated in this study. This includes 541 introductory physics students (including algebra-based and calculus-based courses) and 65 introductory astronomy students. The AAPS survey responses from graduate students and faculty are included in order to serve as a benchmark of more expertlike responses [30]. A total of 42 graduate students and 12 faculty responses to the AAPS survey are included. A subset of the total sample of students participated in individual follow-up interviews, each lasting approximately 1 hour. This included 12 introductory physics students and 8 introductory astronomy students.

6.2.2 Data collection

The AAPS survey was given to students in this study near the end of the semester. The items in the AAPS survey are statements which can elicit agreement or disagreement, and responses are given on a 5 point Likert scale (strongly agree, agree, neutral, disagree, or strongly disagree). When developed and validated, the items were designed so that a favorable response is not always "agree" or "disagree." Once completed, spreadsheets of the results were produced and scores for each question were computed by assigning a +1 to a favorable response, a -1 to an unfavorable response, and a 0 to a neutral response, in which "favorable" means that the response reflects that which an expert in the field (such as a faculty member) would give. We will refer to this convention (of scoring between -1 and +1) as "normalized" data, which we adopted for consistency with the convention used when the AAPS survey was originally validated [30]. We then averaged these values across each group of interest (e.g., introductory physics or introductory astronomy). This average score will be referred to as the "normalized" score. In addition, we also computed the percentage of the total

responses that were favorable, neutral, and unfavorable responses for each question, and we also averaged these over all questions to obtain the average scores for each question on the AAPS survey. These percentages (of favorable, unfavorable, and neutral responses) will be referred to as "unnormalized" data.

In order to gather qualitative data, the written survey data collection was followed by individual hour-long interviews with introductory physics and introductory astronomy students. The interviews utilized a "think aloud" protocol in which students answered the AAPS survey questions along with providing reasoning for why they answered the way they did. They were not disturbed as they answered the survey questions while thinking aloud, but later we asked them for clarifications of the points that had not been made clear. Follow up questions were asked when necessary in order to probe deeper into students' reasoning. In addition, at the interviews, both introductory physics and introductory astronomy students were presented with an isomorphic pair of problems. As can be seen in Figures 25 and 26, one was written in a non-astronomy context and the other written in the context of astronomy. They were asked to solve both problems and were asked about whether either problem was more difficult or enjoyable. Both problems required the students to solve for the speed of an object based upon the assumption of uniform circular motion with centripetal acceleration. Thus, the problems required the same concepts and formulas, but contained different contexts–the astronomy problem involved the motion of the Earth around the Sun and the physics problem involved the circular motion of a yo-yo whirled in a horizontal circle. The centripetal acceleration in each case was due to different mechanisms (gravity for the astronomy problem and tension for the yo-yo problem), but the manner of solving the problems is essentially identical. The purpose was to determine if both physics and astronomy students were equally proficient in solving these problems, whether one of the problems was more challenging than the other and whether there was something about the context of astronomy or physics that produced different attitudes or approaches to the problem-solving process. We also looked at the pre-/post- test data for the Force Concept Inventory (FCI), a standardized conceptual survey [78], to investigate how the performance of physics and

astronomy students compares on this topic, which is taught in both classes.

I Horizontal spinning yo-yo

A yo-yo is spinning in a horizontal circle. The force of tension on the string of the yo-yo is 2 Newtons and there is 0.5 m of yo-yo string between the yo-yo body and center of the circle. If the yo-yo's mass is 0.15 kg, how fast must the yo-yo's speed be?

Figure 25: Non-astronomy context in isomorphic problem pair. Note that students were told to assume no other forces were present other than the force of tension.

II Earth-Sun orbit

The Sun exerts an impressive 3.4×10^{22} Newtons of force on the Earth, as the Earth executes a nearly perfectly-circular orbit around the Sun. The Earth's mass is approximately 6.0×10^{24} kg, and its average distance from the Sun is approximately 1.5×10^{11} m. How fast is the Earth's orbital speed?

Figure 26: Astronomy context in isomorphic problem pair.

6.3 RESULTS

6.3.1 RQ1. Are there differences in average overall performance on the AAPS survey for introductory astronomy students compared with introductory physics students?

Figure 27 shows that introductory physics and astronomy students both have less expertlike attitudes and approaches to problem solving, compared with graduate students and faculty. However, astronomy students have more expertlike attitudes and approaches than physics

students. The comparison to graduate students and faculty is based upon the results from the validation study [30]. All differences in overall scores were found to be statistically significant on the t-tests, with $p < 0.001$ in all cases. The effect sizes, as defined by Cohen's d, is given by: $d = \frac{\mu_1 - \mu_2}{\sigma_{pooled}}$ where $\mu_1 - \mu_2$ is the difference in means between two groups and σ_{pooled} is the pooled standard deviation, given by $\sigma_{pooled} = \sqrt{\frac{(n_1-1)\sigma_1^2+(n_2-1)\sigma_2^2}{n_1+n_2-2}}$ where n_1, n_2, σ_1, and σ_2 are the sample sizes and standard deviations of the two groups being compared [70] in Table 12.

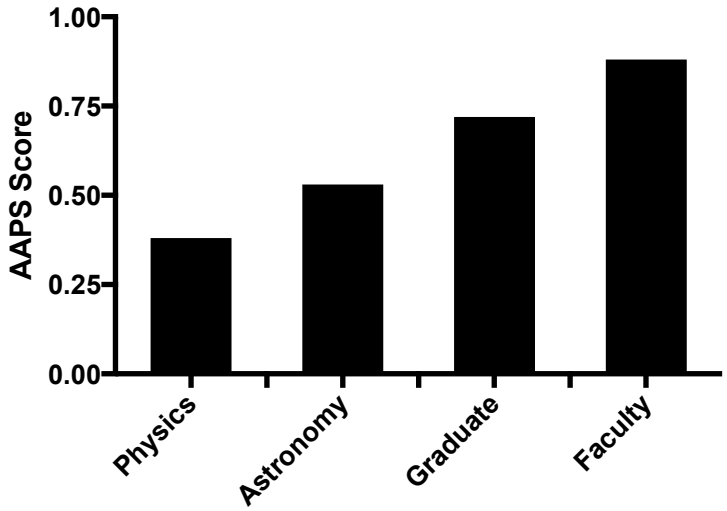


Figure 27: Average AAPS scores of introductory physics students, introductory astronomy students, graduate students, and faculty.

As noted, while the content is presented differently in introductory physics and astronomy classes, the underlying physics principles and level of rigor involved in the instruction are equivalent. Also, while the courses are presented very differently, foundational topics in physics, such as forces and energy are part of the content of both courses. Indeed, when we compare the performance on the Force Concept Inventory (FCI), we find that students in introductory astronomy had very similar pre- and post-test scores (taken at the beginning and end of the semester, respectively) compared to introductory physics students at the same institution, for a typical cohort of introductory physics and astronomy students included in this investigation. In particular, introductory physics students scored an average of 47% on the FCI pre-test, and an average of 58% on the FCI post-test. Introductory astronomy

students scored an average of 46% on the FCI pre-test, and an average of 62% on the FCI post-test. Thus, the scores on force concepts is comparable. Therefore, the difference in attitudes and approaches to problem-solving in physics and astronomy is not likely to be due to the differences in rigor of treatment but some other reason. For example, there may be something about the manner and context in which introductory astronomy classes are taught which might at least partly account for this difference. In other words, while introductory astronomy students learned just as many physics concepts pertaining to forces and Newton's laws as introductory physics students as measured by the FCI, introductory astronomy students appear to have more favorable attitudes than introductory physics students, perhaps owing to a more engaging context in which the content is learned and the manner in which the entire course curriculum is framed.

Some evidence for the lack of an engaging context was seen in individual student interviews with physics and astronomy students. For example, one student expressed that he would enjoy physics classes more if he could relate them to something interesting or realistic. He went on to point out a common type of introductory physics problem as a counter-example to an interesting or realistic type of problem, stating "How many times are you realistically doing that with a pulley?" Another student, in an interview, discussed at length how she had the opportunity to study modern physics topics in high school, and lamented that introductory physics was not presented in an interesting way, stating: "I feel like there's definitely more about physics, and that this is just the rut of it... before you can get to more interesting things." Although this student appeared to assume that the topics would likely get more interesting once you progress beyond introductory physics (based upon the modern physics she learned in high school), it was clear from the discussion that, as it stands, introductory physics was not engaging or interesting in her opinion.

Similarly, a student who had taken both introductory physics and introductory astronomy expressed excitement when comparing the two experiences, stating: "Astronomy is more exciting. I just read this story today about a galaxy that doesn't have any dark matter. That's freaking awesome... Physics doesn't do that. Anyone can learn about Newton's

laws of motion." Further discussions suggested that not only did the astronomy course generate his interest in astronomy, but he took the initiative to read news stories related to astronomy on his own. He also hinted at the contrasting way in which introductory physics is taught, in which new discoveries and exciting events were not part of typical instruction.

6.3.2 RQ2. Are there differences in performance on specific clusters of questions on the AAPS survey for introductory astronomy students compared with introductory physics students?

There are a few AAPS survey questions in which physics students outperformed astronomy students. However, for most of the AAPS survey questions, astronomy students outperformed physics students. This can be seen in Table ??, which contains normalized scores on each question, and in Table 14 as raw percentages of favorable, neutral, and unfavorable responses. A factor analysis was performed earlier in the AAPS validation study [30]. Table ?? compares the factors identified in the validation study [30] with responses from introductory physics and astronomy students. Astronomy students generally outperformed physics students in all factors except for Factor 2, which is related to drawings and scratchwork while problem solving.

Among the AAPS survey questions on which introductory astronomy students outperformed introductory physics students, many of the questions center around Factor 1, which is related to metacognition and enjoyment in problem solving. In fact, Factor 1 accounted for more of the variance than any other factor in Ref. [30]. For example, question 10 on the AAPS states "If I am not sure about the correct approach to solving a problem, I will reflect upon the principles that may apply and see if they yield a reasonable solution." Students are supposed to say whether or not they agree with the statements, and a favorable response to this question would be to agree, as 89% of introductory astronomy students did. By comparison, only 65% of introductory physics students gave a favorable response to this question, which can be seen in Figure 28. As one introductory physics student said in an interview, "I find it hard, if I was stuck on a problem, to think back on the principles, because if I

am struggling then my conceptual understanding isn't that strong." Further discussion with this student suggests that this student appeared to believe that it would not be possible to think about the principles if one's understanding of them was not strong to begin with. On the other hand, all but one of the interviewed astronomy students gave a favorable response, stating that they reflect upon principles when they are unsure about the correct approach to solving an astronomy problem, and that the interesting context of these problems is a motivational factor that keeps them engaged even if they are stuck.

Another question from Factor 1, question 27 states, "I enjoy physics/astronomy even though it can be challenging at times." (Note that physics students only saw the word "physics" in this question and astronomy students only saw the word "astronomy"). As can be seen in Figure 29, while 72% of astronomy students agreed with this statement, only 40% of physics students agreed. One physics student said in an interview, "To just sit there and do problems over and over is not really fun," conveying that she did not find physics to be enjoyable because it involves many problems to solve (that are not intrinsically interesting). Another student identified the uninteresting nature of physics problems as being why she answered that she did not enjoy physics, stating "I think it's just like 'here's a problem, just solve it' and not really making it an interesting, this could be in the real world thing." On the other hand, some interviewed astronomy students explicitly reported that they enjoyed astronomy problems more than physics problems, even if they are just as challenging. As one student, who had taken both introductory physics and introductory astronomy, put it, "I think physics and astronomy are strongly aligned for a lot of both of those introductory courses but where they start to split, I just find that subject matter [astronomy] more interesting. So even though the physics involved can be basically the same, the setup is what drives me to like astronomy more." This student indicated that the physics behind the courses is similar, but that the subject matter of astronomy was more enjoyable and kept him engaged and persistent despite the challenging nature of many of the problems. Other interviewed students expressed similar sentiments.

Related to the metacognitive aspect of Factor 1, some questions asked about conceptual

thinking over "plug-and-chug" approaches. For example, question 4 states "In solving problems in physics/astronomy, I always identify the physics/astronomy principles involved in the problem first before looking for corresponding equations." A favorable response to this question would be to agree with the statement, as 80% of the introductory astronomy student did. Fewer introductory physics students (62%) responded favorably by comparison, as seen in Figure 30. Likewise, question 14 states, "When I solve physics/astronomy problems, I always explicitly think about the concepts that underlie the problem." Figure 31 shows that while 78% of introductory astronomy students indicated a favorable response to this question (i.e., they agreed with the statement), only 55% of introductory physics students indicated a favorable response.

Factor 6 was another one in which introductory astronomy students gave more favorable responses than introductory physics students. This factor has to do with problem solving confidence, and relates to students' willingness to persist in working through problems in the face of difficulty. For example, question 1 says "If I'm not sure about the right way to start a problem, I'm stuck unless I go see the teacher/TA or someone else for help." As can be seen in Figure 32, only 49% of physics students had favorable responses to this question (i.e., a favorable response would be to disagree with the statement); wheras, 68% of astronomy students gave favorable responses. Similarly, question 23 states "If I cannot solve a problem in 10 min, I give up on that problem." As shown in Figure 33, a favorable response of disagree was given by 82% of astronomy students but only 61% of physics students.

Another factor in which introductory students performed more favorably than introductory physics students was Factor 8, which deals with sensemaking. For example, question 2 states, "When solving physics/astronomy problems, I often make approximations about the physical world." The majority (65%) of introductory astronomy students agreed with this statement, constituting a favorable response. On the other hand, only 43% of introductory physics students gave a favorable response, as can be seen in Figure 34. Another question in this factor was question 16, which states, "When aswering conceptual physics/astronomy questions, I mostly use my 'gut' feeling rather than the physics/astronomy principles

I usually think about when solving quantitative problems." Figure 35 shows that while 50% of introductory physics students gave a favorable response (i.e., they disagreed with the statement), the majority of astronomy students (71%) gave a favorable response.

The only factor in which introductory astronomy students gave *less* favorable responses compared with introductory physics students was Factor 2, which is related to the use of drawings and scratchwork while solving problems. For example, question 15 states "When solving physics/astronomy problems, I often find it useful to first draw a picture or diagram of the situations described in the problems." As can be seen in Figure 36, while 82% of introductory physics students answered favorably (i.e., agreed with the statement) to this question, only 68% of introductory astronomy students answered favorably. However, in follow-up interviews, one possible reason for less favorable responses was uncovered. For example, one student described certain topics that come up in astronomy problems to be more difficult to depict in a drawing, stating: "For example, involving things like a spectrum, or EM [electromagnetic waves] in general. Im not going to draw a light wave." Another interviewed student touched on a similar theme in making the following distinction about when drawing is helpful in solving astronomy problems, "For questions about celestial movement, I think that's where the bulk of the drawing comes from... But like right now were learning about magnitudes, I don't know how I would draw different magnitudes...because I can't draw light...unless it's reflected onto something like the moon. Phases of the moon I really need to draw." This student indicated that there are only some problems in astronomy that appear to lend themselves well to drawing and, similar to other introductory astronomy students who discussed this issue, there are many astronomy problems that were either difficult to depict in a sketched diagrams and/or were problems for which a drawing may not have offered much benefit. Since introductory physics problems often focus more on simpler physical objects that lend themselves well to drawing, this dichotomy could account for the less favorable responses of astronomy students regarding the use of drawing and scratchwork in problem solving.

6.3.3 RQ3. When presented with an isomorphic problem pair (problems with the same underlying physics principle but different contexts–one written in an astronomy context and the other written in a physics context), are introductory physics and introductory astronomy students equally likely to solve both problems correctly and do they find either more or less interesting?

At the end of the interviews, introductory astronomy and physics students were presented with a pair of isomorphic problems (one written in and one without the context of astronomy). Both problems required the use of the concept of centripetal force and the assumption of a uniform circular motion in order to solve for the speed. Altogether 81% of all students (both introductory physics and introductory astronomy) who were presented with these problems were able to solve both problems entirely correctly. This includes 75% of introductory physics students and 86% of introductory astronomy students. Other students in both groups were also able to solve a significant fraction of both isomorphic problems correctly but made some mistakes along the way, e.g., writing centripetal force as $\frac{v^2}{r}$ as opposed to $\frac{mv^2}{r}$. Like the FCI results, this similarity in performance on the isomorphic problems suggests that student performance on similar content is comparable in both classes. In particular, while the FCI reveals similar conceptual performance for the physics and astronomy students, the performance on isomorphic problems reveals similar facility in quantitative problem solving.

While performance of students appears to be similar, differences were found when subjects were asked which of the two isomorphic problems (non-astronomy context or astronomy context) they enjoyed more. In particular, the majority of both introductory physics and introductory astronomy students reported that they found the astronomy context more interesting. For example, one physics student, when asked which she found to be more enjoyable reported, "The Sun and the Earth one, because I think space is cooler than a yo-yo." This student indicated that there is something more captivating about space than about everyday objects. Another physics student stated, "The Earth one, yes, was more interesting," conveying a similar sentiment that the context of the astronomy problem was more interesting

than the context of the physics problem.

It is interesting to note that several students expressed more interest in the astronomy problem, even if they found it to be more difficult to solve or challenging to think about. For example, one physics student stated, "[The non-astronomy] one was less annoying to solve because I could think about what the answer could be, but [the astronomy] one, everything's scientific notation so I can't really think about... But the [astronomy] one is definitely more interesting than the first one because it's an application of all these things." Similarly, an interviewed astronomy student reported that the physics problem was easier to solve, but that she enjoyed the astronomy problem more. When asked why the astronomy problem was more enjoyable, she stated, "I'm just into the subject matter of astronomy. It's more interesting than a yo-yo's tension," again conveying a more motivating scenario and an interest in the astronomy context.

One astronomy student in an interview discussed at length how the astronomy problem sparked more curiosity and interest and described how it got him thinking, "That's cool to me because then you could think about well what if all of a sudden we stop orbiting and we go off on a tangent, and at that speed, well that sucks, that's going to be rough." The context of the astronomy problem appeared to cause him to think further about centripetal force, and how the removal of this force would result in tangential movement at a constant speed. When further asked if the problem with non-astronomy context would inspire him to think about what might happen if the yo-yo string broke and it, similarly, went off on a tangent, he indicated that this thought was not as captivating or interesting, stating that, "That just means it's hitting my TV and I'm getting yelled at," Like this student, most of the interviewed students expressed less interest in the physics problem. This interview finding suggests that, given an identical physics content and equivalent performance, a more favorable attitude and approach to problem-solving might arise if the problem is written in a context that students find more interesting, which in this case was the astronomy context.

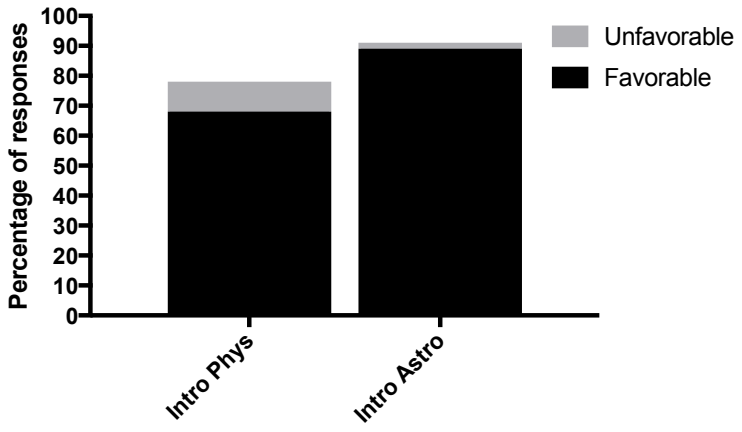


Figure 28: Question 10: Astronomy students exhibited more favorable attitudes than physics students when asked if they reflect on principles when they are stuck.

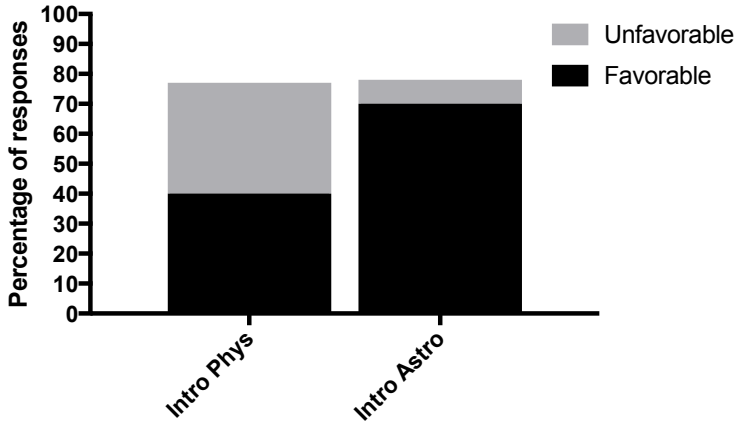


Figure 29: Question 27: Astronomy students exhibited more favorable attitudes than physics students when asked if they enjoy solving problems.

6.4 DISCUSSION

Similar students, different attitudes: Our results indicate that introductory astronomy students, overall, had more expertlike attitudes and approaches than introductory physics students. An explanation for this difference would have to account for the fact that astron-

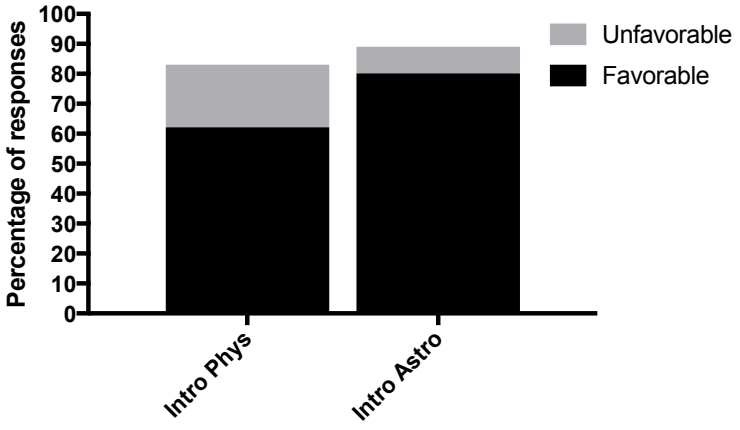


Figure 30: Question 4: Astronomy students exhibited more favorable attitudes than physics students when asked if they identify principles before looking for equations.

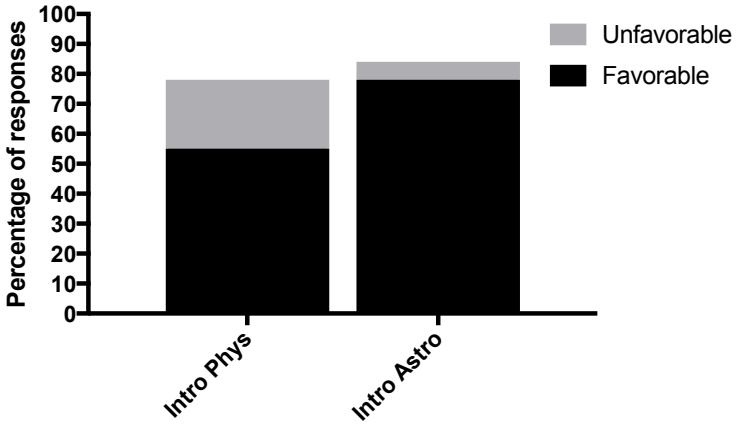


Figure 31: Question 14: Astronomy students exhibited more favorable attitudes than physics students when asked if they think about the concepts underlying the problem.

omy students had comparable physics content knowledge as measured by the FCI performance and also comparable performance on two isomorphic problems (posed in an astronomy and non-astronomy context). They consistently gave much more favorable responses to AAPS survey questions involving metacognition and enjoyment in problem solving. An explanation for the difference between the AAPS scores of introductory physics and intro-

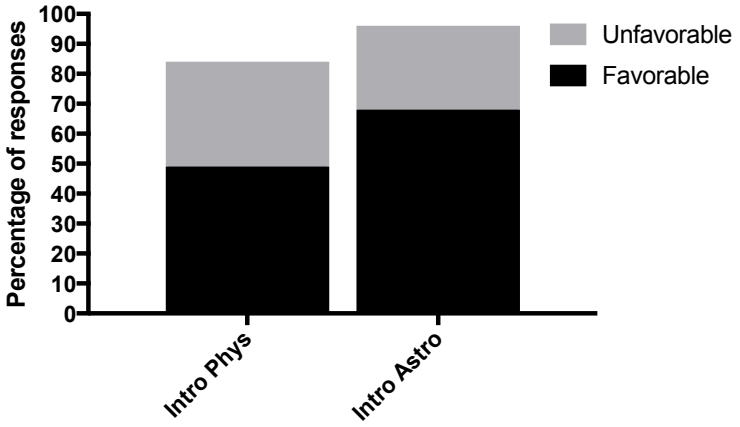


Figure 32: Question 1: Astronomy students exhibited more favorable attitudes than physics students when asked if they are stuck unless they go see the teacher or TA.

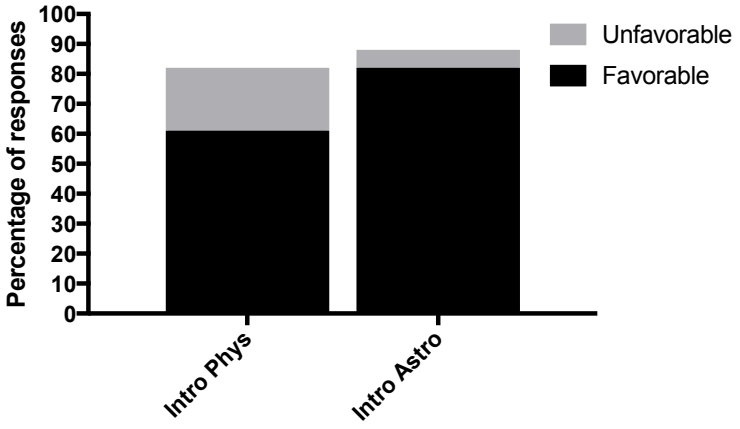


Figure 33: Question 23: Astronomy students exhibited more favorable attitudes than physics students when asked if they give up on a problem after 10 minutes.

ductory astronomy students would also have to account for the fact that both algebra-based and calculus-based introductory physics students score nearly identically on the AAPS (so much so that their results were combined) [30, 31, 113]. This is noteworthy because students who enroll in algebra-based physics tend to have different reasons for enrolling in introductory physics than calculus-based physics students, and these two courses contain students

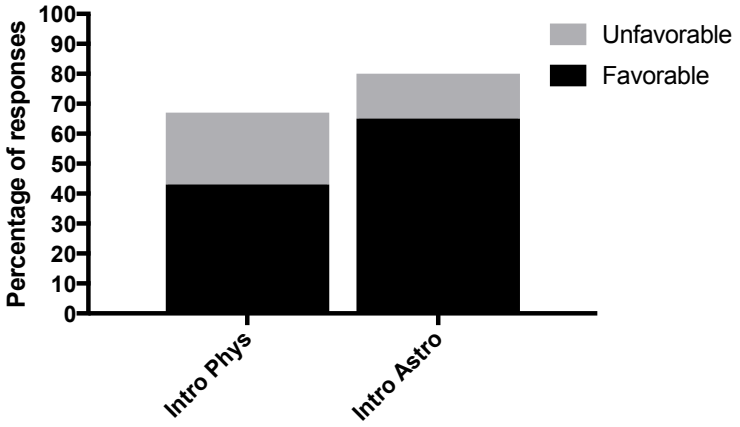


Figure 34: Question 2: Astronomy students exhibited more favorable attitudes than physics students when asked if they often make approximations.

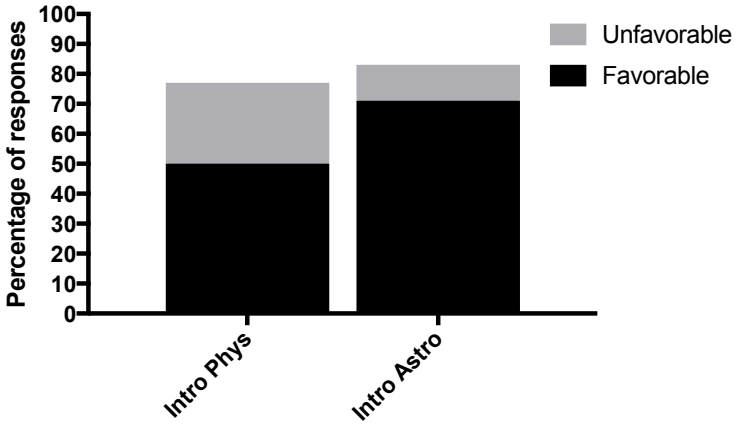


Figure 35: Question 16: Astronomy students exhibited more favorable attitudes than physics students when asked if they mostly use their gut feeling when answering conceptual questions.

which could be argued to be more dissimilar to each other in their prior preparation in mathematics, physics, and abstract thinking, in general, than introductory physics students are to introductory astronomy students in the astronomy course discussed here. If differences in AAPS score should be expected to be tied to differences in motivation for taking the class, one would expect differences between algebra-based and calculus-based physics classes' aver-

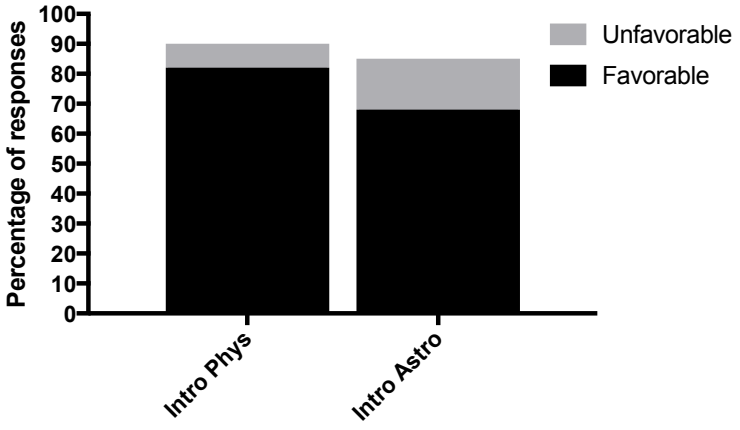


Figure 36: Question 15 Astronomy students exhibited *less* favorable attitudes than physics students when asked if they draw pictures or diagrams when solving problems.

age scores to show similar differences, but such a difference was not found. Different scores at the introductory level appear only when comparing physics students to astronomy students.

Captivating content: The differences between AAPS scores of introductory physics and introductory astronomy appear to suggest that there may be something about the subject matter in astronomy that might promote metacognition and may be more interesting and engaging for students compared to what they learn in a typical physics class. For example, it is possible that learning about galaxies and black holes and viewing images of outer space is more captivating than learning about pulleys and inclined planes. One interviewed student described the appeal of real astronomical images compared with images one might encounter in a physics class as follows, "For astronomy there are all these real images... you can't take a picture of a projectile the same way, but with astronomy you can take basically any image from the Hubble telescope and be like 'there' and that's all you need... What our teacher did was pull from the astronomy picture of the day and it was really cool. Just having that image really set the tone." Even the context of the problems themselves appears to yield more student interest when the problems are written in the context of astronomy, with many students who were interviewed indicating that they enjoyed the astronomy problem

more than the isomorphic physics problem. If it is the case that students find astronomy and astronomy problems more interesting, it is possible that more captivating instruction of physics classes, including more engaging contexts such as those used in astronomy classes, may improve physics students' attitudes and approaches to problem solving. If introductory physics classes can emphasize aspects of physics that resonate with students and capture their interests, perhaps their attitudes would more closely resemble those of introductory astronomy students. Using suitable astronomy contexts in physics courses would be one effective approach that physics instructors could consider.

Conceptually-rich instruction: Some students identified the conceptually-rich nature of astronomy as a reason they found the instruction interesting. On the other hand, typical traditional physics course is often taught in a dry manner and problems often emphasize algorithmic problem-solving rather than promoting fuctional conceptual understanding with engaging content. An informal survey of introductory physics faculty at the University of Pittsburgh suggests that a majority of the physics homework, quiz, and exam questions instructors assign focus on quantitative problem solving. Discussions with some instructors suggest that conceptual learning is regarded by physics instructors as "easier" or less rigorous than quantitative problem-solving, and that making problems interesting is secondary to developing effective problem-solving skills. For example, in one qualitative study in which the researcher followed students through the experience of an introductory physics course, it was concluded that when the subject is taught in a way that is "boring, dull, or simply not fun," students will often become disengaged, and that this sort of "boring" teaching described the way introductory physics is often taught [134]. If the content of problems were more engaging through the use of a conceptually rich, interesting context, then perhaps students' attitudes regarding problem-solving would improve. One student who was interviewed hit on this theme when he said "Algorithmic learning is kind of terrible," as he described the way in which physics is taught in his opinion. He went on to contrast this with the way astronomy is often taught, saying "...conceptual learning really builds foundation for the subject." This student then identified ways in which physics problem-solving would be improved, "if there

were pointed questions that went along with the mathematics, like even a mathematics question and then at the very end "why?" He felt that physics courses only focused on "plug and chug" and there was a de-emphasis on integrating the conceptual aspects of the problem in the student learning process.

Another interviewed student expressed that physics courses could engage students more in problem-solving through better use of demonstrations to build conceptual thinking. As this student explained, "I think in physics there's more possibility for demonstrations... The one thing though was that the demos...some of them were really cool like when we were learning about angular momentum and you're spinning in a chair, but some of them tended to be so technical that you lost interest. But I think if physics took a lot more advantage of the demos, the ability to do those kind of demos, if students could really see what is going on, in a really spectacular way, that would be like 'oh yeah!"' This sort of active engagement has been shown to be a benefit of interactive demonstrations [94]

After discussing the idea of combining conceptual instruction with problem-solving, one student explained that physics instruction could be made more interesting by including problems that would engage students in something fun, after conceptual instruction, stating: "Here's an idea. You take famous scenes from movies and you try to figure out what's actually happening in them. So if you have a James Bond car chase scene and you see the car spinning in a circle and you're teaching friction forces, you could say if the car spins out 10 meters that way and assuming there's no outside forces, what is the friction force on the tires?" This student saw the power behind bringing concepts to life in the use of exciting problem-solving. Indeed, combining conceptual reasoning with problem-solving may help promote problem-solving expertise, [81, 135] which should translate to more favorable attitudes and approaches to problem-solving. Therefore, incorporating more conceptually-rich instruction into the physics classroom could be an important piece of the puzzle.

Impact on motivation in learning: When consdering the role of motivation in learning, it is possible that, our results may have implications for motivating students. Motivation is often considered to be either focused mainly on performing well in a class without much

focus on achieving understanding–i.e., a "performance" motivation, or focused on achieving understanding and mastery of the material–i.e., a "mastery" motivation [9, 10]. Motivational goals and beliefs may also shape problem-solving processes through sensemaking [8, 9, 10]. Because many AAPS survey questions tap into issues of motivational goals, the more favorable responses from astronomy may imply that the way astronomy is taught encourages a more mastery-oriented motivational attitude. If this is the case, then it could mean that motivational goals are malleable, and that if instructors can captivate and interest their students in the way they teach, it may have a positive impact on their motivational attitudes and goals, which could benefit introductory physics students, particularly those from underrepresented groups [14].

6.5 SUMMARY AND IMPLICATIONS

We have found that written survey responses and interviews suggest that introductory astronomy students had more favorable attitudes and approaches to problem solving overall, and in the majority of clusters of individual questions. In light of the results presented here pertaining to the AAPS survey scores in introductory physics and introductory astronomy courses and individual interviews with students, an important instructional implication of the more favorable attitudes and approaches to problem-solving found among introductory astronomy students compared to introductory physics students is for instructors of physics to consider incorporating conceptually-rich and engaging content in their courses. For example, the problems that physics students are asked to solve could include more exciting and realistic contexts, or involve objects that can relate to or find interesting. Preceding problem-solving with interactive demonstrations to build conceptually-rich instructional design in which quantitative and conceptual aspects of learning are integrated may also motivate students to relate to subject matter involved in the problems they are asked to

Cohen's d	**Physics**	**Graduate**	**Faculty**
Astronomy	0.81	1.06	1.78
Physics		1.60	2.18
Graduate			1.19
Faculty			

Table 12: Effect sizes between different groups

solve. Additionally, introductory physics instructors could utilize the fascination students may have with the cosmos in their teaching of introductory physics by incorporating astronomical objects and examples into problems when appropriate. Moreover, problems which contain real images rather than simply cartoons or sketches may bring the physics content to life for students. From in-class examples, to homework problems, to exam questions, if students can more effectively engage with the content and questions asked of them, as it appears astronomy students do, then perhaps the attitudes and approaches to problem solving for introductory physics students would improve.

Problem number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Introductory physics	0.14	0.19	0.15	0.41	0.16	0.24	0.61
Introductory astronomy	0.40	0.49	-0.15	0.71	0.23	0.51	0.83
Graduate students	0.71	0.42	-0.04	0.83	0.17	0.75	0.83
Faculty	0.83	1.00	0.50	0.92	0.92	1.00	0.83
Problem number	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Introductory physics	0.67	0.24	0.58	-0.03	-0.06	0.56	0.32
Introductory astronomy	0.66	0.31	0.88	0.26	0.13	0.75	0.72
Graduate students	0.83	0.46	0.88	0.67	0.54	0.88	0.88
Faculty	0.92	0.58	0.92	0.67	0.83	1.00	0.50
Problem number	**15**	**16**	**17**	**18**	**19**	**20**	**21**
Introductory physics	0.74	0.23	0.55	0.69	0.77	-0.19	0.71
Introductory astronomy	0.51	0.58	0.28	0.48	0.86	0.26	0.86
Graduate students	0.96	0.50	0.79	0.96	0.88	0.38	0.92
Faculty	1.00	0.67	1.00	0.92	1.00	0.75	1.00
Problem number	**22**	**23**	**24**	**25**	**26**	**27**	**28**
Introductory physics	0.52	0.40	0.43	0.56	0.37	0.03	0.75
Introductory astronomy	0.69	0.75	0.15	0.71	0.60	0.65	0.89
Graduate students	1.00	1.00	0.21	0.54	0.71	0.67	0.96
Faculty	1.00	0.92	0.42	0.92	1.00	0.92	1.00
Problem number	**29**	**30**	**31**	**32**	**33**	**Overall**	
Introductory physics	0.74	-0.04	0.08	0.70	0.46	0.38	
Introductory astronomy	0.82	0.14	0.12	0.80	0.63	0.53	
Graduate students	1.00	0.92	0.92	1.00	0.83	0.72	
Faculty	1.00	1.00	1.00	1.00	1.00	0.88	

Table 13: Average normalized scores by question

Problem number	1	2	3	4	5	6	7
Favorable--Intro Phys	49%	43%	49%	62%	48%	52%	73%
Favorable--Intro Ast	68%	65%	29%	80%	55%	66%	85%
Favorable--Graduate	79%	67%	38%	88%	50%	79%	92%
Favorable--Faculty	83%	100%	58%	92%	92%	100%	92%
Neutral--Intro Phys	16%	32%	18%	18%	18%	20%	15%
Neutral--Intro Ast	5%	20%	26%	11%	12%	18%	14%
Neutral--Graduate	13%	8%	21%	8%	17%	17%	0%
Netural--Faculty	17%	0%	33%	8%	8%	0%	0%
Unfavorable--Intro Phys	35%	24%	33%	21%	33%	28%	12%
Unfavorable--Intro Ast	28%	15%	45%	9%	32%	15%	2%
Unfavorable--Graduate	8%	25%	42%	4%	33%	4%	8%
Unfavorable--Faculty	0%	0%	8%	0%	0%	0%	8%
Problem number	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Favorable--Intro Phys	78%	43%	68%	31%	36%	71%	55%
Favorable--Intro Ast	75%	45%	89%	48%	43%	83%	78%
Favorable--Graduate	88%	67%	92%	75%	75%	92%	88%
Favorable--Faculty	92%	67%	92%	75%	92%	100%	67%
Neutral--Intro Phys	10%	39%	22%	34%	21%	14%	22%
Neutral--Intro Ast	15%	42%	9%	31%	25%	9%	15%
Neutral--Graduate	8%	13%	4%	17%	4%	4%	13%
Netural--Faculty	8%	25%	8%	17%	0%	0%	17%
Unfavorable--Intro Phys	12%	19%	10%	34%	42%	15%	23%
Unfavorable--Intro Ast	9%	14%	2%	22%	31%	8%	6%
Unfavorable--Graduate	4%	21%	4%	8%	21%	4%	0%
Unfavorable--Faculty	0%	8%	0%	8%	8%	0%	17%
Problem number	**15**	**16**	**17**	**18**	**19**	**20**	**21**
Favorable--Intro Phys	82%	50%	71%	79%	85%	25%	78%
Favorable--Intro Ast	68%	71%	57%	65%	92%	45%	88%
Favorable--Graduate	96%	63%	88%	96%	88%	54%	92%
Favorable--Faculty	100%	75%	100%	92%	100%	75%	100%
Neutral--Intro Phys	10%	23%	12%	10%	8%	30%	15%
Neutral--Intro Ast	15%	17%	14%	18%	2%	37%	11%
Neutral--Graduate	4%	25%	4%	4%	13%	29%	8%
Netural--Faculty	0%	17%	0%	8%	0%	25%	0%
Unfavorable--Intro Phys	8%	27%	16%	11%	8%	45%	7%
Unfavorable--Intro Ast	17%	12%	29%	17%	6%	18%	2%
Unfavorable--Graduate	0%	13%	8%	0%	0%	17%	0%
Unfavorable--Faculty	0%	8%	0%	0%	0%	0%	0%
Problem number	**22**	**23**	**24**	**25**	**26**	**27**	**28**
Favorable--Intro Phys	66%	61%	63%	68%	57%	40%	81%
Favorable--Intro Ast	77%	82%	43%	78%	71%	72%	91%
Favorable--Graduate	100%	100%	54%	71%	79%	79%	96%
Favorable--Faculty	100%	92%	67%	92%	100%	92%	100%
Neutral--Intro Phys	20%	18%	17%	21%	23%	22%	12%
Neutral--Intro Ast	15%	12%	29%	14%	18%	20%	8%
Neutral--Graduate	0%	0%	13%	13%	13%	8%	4%
Netural--Faculty	0%	8%	8%	8%	0%	8%	0%
Unfavorable--Intro Phys	14%	21%	20%	12%	20%	37%	6%
Unfavorable--Intro Ast	8%	6%	28%	8%	11%	8%	2%
Unfavorable--Graduate	0%	0%	33%	17%	8%	13%	0%
Unfavorable--Faculty	0%	0%	25%	0%	0%	0%	0%
Problem number	**29**	**30**	**31**	**32**	**33**	**Overall**	
Favorable--Intro Phys	81%	36%	45%	83%	68%	60%	
Favorable--Intro Ast	86%	43%	49%	86%	71%	68%	
Favorable--Graduate	100%	96%	96%	100%	88%	82%	
Favorable--Faculty	100%	100%	100%	100%	100%	90%	
Neutral--Intro Phys	11%	24%	17%	2%	5%	18%	
Neutral--Intro Ast	9%	28%	14%	8%	7%	17%	
Neutral--Graduate	0%	0%	0%	0%	8%	9%	
Netural--Faculty	0%	0%	0%	0%	0%	7%	
Unfavorable--Intro Phys	8%	40%	38%	16%	29%	22%	
Unfavorable--Intro Ast	5%	29%	37%	6%	21%	15%	
Unfavorable--Graduate	0%	4%	4%	0%	4%	9%	
Unfavorable--Faculty	0%	0%	0%	0%	0%	3%	

Table 14: Breakdown of average favorable, unfavorable, and neutral responses to individual questions of the AAPS.

	Problem Number	Physics Score	Astronomy Score
Factor 1 (Metacognition and enjoyment of problem-solving)	13	0.56	**0.75**
	14	0.32	**0.72**
	7	0.61	**0.83**
	10	0.58	**0.88**
	22	0.52	**0.69**
	29	0.74	**0.82**
	20	-0.19	**0.26**
	4	0.41	**0.71**
	25	0.56	**0.71**
	27	0.03	**0.65**
	6	0.24	**0.51**
	21	0.71	**0.86**
Factor 2 (use of drawings and scratch-work)	18	**0.69**	0.48
	17	**0.55**	0.28
	15	**0.74**	0.51
	19	**0.77**	0.86
Factor 3 (Perception of problem-solving approach)	5	0.16	**0.23**
	11	-0.03	**0.26**
	12	-0.06	**0.13**
	8	0.67	0.66
	26	0.37	**0.60**
	9	0.24	**0.31**
Factor 4 (General expert-novice differences in problem-solving)	8	0.67	0.66
	28	0.75	**0.89**
	21	0.71	**0.86**
	24	**0.43**	0.15
	29	0.74	**0.82**
Factor 5 (Use of symbolic notation)	31	0.08	**0.12**
	30	0.70	**0.80**
Factor 6 (Problem-solving confidence)	1	0.14	**0.40**
	24	**0.43**	0.15
	23	0.40	**0.75**
	6	0.24	**0.51**
Factor 7 (Solving different problems using the same principles)	33	0.70	**0.80**
	32	0.46	**0.63**
Factor 8 (Sensemaking)	16	0.23	**0.58**
	2	0.19	**0.49**
	5	0.16	**0.23**
Factor 9 (Problem-solving sophistication)	3	**0.15**	-0.15
	20	-0.19	**0.26**
	25	0.56	**0.71**
	9	0.24	**0.31**

Table 15: Average scores by question and factor. Order of question numbers reflects that in Ref. [30].

7.0 CONCLUSIONS AND FUTURE DIRECTIONS

The investigations presented here underscore the importance of considering students' views about problem-solving in physics. Introductory students start with their initial attitudes and approaches to problem-solving that will evolve during their courses and may shape their experiences and success in their physics courses. Views held by graduate students about problem-solving may or may not align with physics education research findings about the teaching and learning of problem-solving. Few opportunities currently exist in physics classrooms or in TA professional development programs to shape the views of either introductory students or graduate students. Introductory students' attitudes and approaches to problem-solving may be shaped by their experience in the classroom, and whether or not they find instruction interesting and engaging. Graduate students' views about problem-solving and awareness of research regarding instructional strategies may be shaped by professional development opportunites they receive. As such, the findings presented here help inform ways in which faculty and leaders of TA professional development can improve upon the commonly-held student attitudes and approaches to problem-solving.

With regard to graduate students' views, teaching assistants were asked to reflect upon a variety of ways in which introductory physics problems could be posed. It was found that TAs valued problem types that offered guidance and support to students but were reluctant to identify the instructional benefits to other problem types that offered less support. In particular, multiple-choice and context-rich problem types were ranked low for instructional benefit by TAs.

While well-designed multiple-choice problems can readily serve as a formative assessment

tool, even for large-enrollment classe, this potential use of multiple-choice problems did not occur to the majority of TAs. It appeared, instead, that TAs had a tendency to assume that multiple-choice problems were only useful for utilitiarian reasons, such as time-saving or convenience in grading high stakes assessments. Moreover, TAs felt that the drawbacks to multiple-choice problems (such as "trapping" students) outweighed these pragmatic assets. The idea of utilizing multiple-choice questions for low-stakes formative assessment purposes was a use for which TAs lacked an apparent awareness.

Similarly, context-rich problems have been shown to be highly beneficial to the development of students' conceptual understanding and problem-solving skills, especially when such problems are incorporated into cooperative group work [64, 65]. However, TAs did not appear to be aware of this benefit. Indeed the very features of a typcial context-rich problem type, such as a lack of explicit question, wordiness, and extranneous information, that are intentionally designed to generate deep understanding and strategic problem-solving processes, were identified by the majority of TAs as drawbacks to using such problems. TAs often held such a negative opinion about context-rich problems that this negativity impacted their ability to view such problems as potentially beneficial for students.

Even for the problem type that TAs had highly positive opinions about (the broken-into-parts problem type), their views appeared to be limited in terms of instructional implications of such problems for students. While the TAs easily identified that broken-into-parts problems could offer guidance to students, the majority of TAs did not identify the potential pitfall of providing too much help to students if guidance is always built into the problems they solve. In the cognitive-apprenticeship framework, the fading of support and weaning to develop autonomy requires that students gradually learn to break problems into sub-problems independently [13]. However, this long-term goal did not appear to impact the thinking of the majority of the TAs, and they reported that they would highly prefer a broken-into-parts problem nearly exclusively above any other problem types for homework, quizzes, and exams. Thus, while TAs appeared to understand that introductory students may benefit from guided support in their problem-solving, their awareness of the importance

of weaning off that support appeared to be lacking.

Given the findings regarding problem types, there are relevant implications and future directions for the professional development of graduate student teaching assistants. Helping TAs reflect upon the importance of formatively assessing students and gradually weaning their support would be important goals of professional development efforts. In addition, professional development could focus on building an awareness of the benefits to context-rich problem solving in collaborative group work, and the effectiveness of multiple-choice questions as formative assessment tools. One possible way in which reflection on these ideas may be facilitated in professional development instruction would be to offer TAs an opportunity to work collaboratively on creating problems of their own design and thinking critically about the instructional design of such problems. Moreover, future investigations into the views of TAs regarding problem types could further explore more example problem types both before and after professional development instruction that focused on the goals identified here.

Undergraduate introductory students' views were examined for the favorability of their attitudes and approaches to problem-solving using the AAPS survey. Differences in favorable attitudes were found depending on instructional methods and gender. In particular, female students and students of both genders who were instructed using evidence-based active engagement (EBAE) methods were found to have more favorable attitudes and approaches to problem solving than corresponding groups that were instructed primarily using traditional methods. In addition, introductory astronomy students were found to have more favorable attitudes than introductory physics students, and many interviewed students identified the context of astronomy as more interesting.

Female introductory physics students not only exhibited more favorable attitudes and approaches to problem solving, regardless of the way in which they were instructed, but their attitudes remained more favorable over the course of a semester of instruction. Considering that there is a gender gap in performance in introductory physics class, this result is intriguing. A future question to explore might be why more favorable attitudes and approaches

to problem solving do not appear to reduce this gender gap. It may be possible that issues of low self-efficacy that are known to be problematic for female students in physics classes [14] could "counter-intuitively" lead to more favorable scores on the AAPS survey due to more careful, cautious approaches to problem solving taken by female students with lower self-efficacy than male students. Moreover, it has been shown that female students may exhibit superior study habits compared with male students which could connect to improved attitudes and approaches to problem solving [136, 15, 137, 138]. Further investigating this connection would help inform efforts to reduce the gender gap in performance in physics courses and representation in the field of physics.

When introductory students' AAPS scores were examined for differences based upon method of instruction, it was found that average scores for classes taught using evidence based active engagement (EBAE) methods were higher than those of traditionally taught classes. This was true for both pre-test (at the beginning of the course) and post-test (at the end of the course) AAPS survey scores. The initially-higher scores suggest that there may be some self-selection effect, since students were free to choose which course to enroll in. It appears that students with more favorable attitudes enrolled in classes that were taught using EBAE methods. However, this initial offset is not enough to explain the post-test score discrepancy between the EBAE classes and the traditionally-taught classes, since the decline for the former was less than the decline for the latter. Moreover, post-test scores in EBAE classes were significantly higher on clusters of questions related to metacognition and problem-solving expertise. A future direction to explore would be to investigate whether other institutions show a similar initial offset in favorable attitudes and approaches to problem solving when EBAE classes are compared to traditionally-taught courses. In addition, an open question to investigate is the reasons for the post-test difference in scores. For example, are initial favorable attitudes and approaches to problem solving more resistent to decline than less favorable attitudes and approaches, or can the difference in scores be ascribed to the success of EBAE methods in encouraging favorable student attitudes and approaches to problem solving?

Interest and motivation may also play key roles in introductory students' attitudes and approaches to problem-solving, as was suggested when introductory physics and astronomy students' AAPS scores were compared and students were interviewed to elicit further discussion about their views. Although the courses are of comparable rigor, introductory astronomy students scored higher on the AAPS survey compared to introductory physics students, and students in both classes spoke about their interest and motivation in these classes. Introductory astronomy was identified as an especially interesting class and the majority of students reported a preference for solving introductory astronomy problems, even if they found them to be more difficult than isomorphic introductory physics problems. Moreover, in almost every cluster of AAPS questions, introductory astronomy students held more favorable attitudes and approaches. One implication may be that captivating students with instruction that generates interest may motivate them and yield positive results regarding their attitudes and approaches to problem solving, which implies a strategy that could be taken by physics instructors who wish to improve their students' attitudes and approaches to problem solving. If introductory physics can be taught in a manner that fascinates students, by taking advantage of their thirst for real-life conceptual inquiry, enthusiasm for active learning demonstrations, and natural interest in the cosmos, then physics students' attitudes and approaches to problem solving may benefit.

The findings of all of the preceding investigations have highlighted many facets to the views of students regarding problem solving in physics. Graduate student TAs exhibited sincerity in their desire to help their students learn, and so professional development to expand their awareness of how to effectively help their students learn offers the potential for them to put that desire into action. Undergraduate introductory students appear to yearn for interesting and engaging instruction, and female students (who are underrepresented in physics) exhibit promising attitudes and approaches to problem-solving. Moreover, engaging instructional methods can be implemented in ways that research has suggested may enhance the learning process. Thus efforts to improve learning experiences for introductory students and encourage more female participation in the field of physics can focus on evidence-based meth-

ods that both enhance student learning and potentially encourage positive attitudes among introductory students. Each of these implications suggests the possibility that improving the learning experience for students may also improve their perspectives and generate more positive attitudes and approaches to problem solving among both undergraduate students in introductory physics courses and graduate student teaching assistants.

BIBLIOGRAPHY

[1] Like J. A. and Kelaher Young A. J. Cognition in context: Students' perception of classroom goal structures and reported cognitive strategies use in the college classroom. *Research in Higher Ed.*, 47:477–490, 2006.

[2] M. Bassok and K. Holyoak. Interdomain transfer betweeen isomorphic topics in algebra and physics. *J. of Experimental Psychology: Learning and Cognition*, 15:153–166, 1989.

[3] K. Duncker. On problem solving. *Psychol. Monogr.*, 58:1, 1945.

[4] J. Mestre. Probing adults' conceptual understanding and transfer of learning via problem solving. *J. Appl. Dev. Psychol*, 23:9, 2002.

[5] P. Adey and Shayer M. An exploration of long-term far-transfer efects following an extended intervention in the high school science curricula. *Cogn. Instr.*, 11:1, 1993.

[6] Dufresne R., Mestre J., Thaden-Koch T., Gerace W., and Leonard W. In Mestre J., editor, *Transfer of learning from a modern multidisciplinary perspective*, page 155. Information Age Publishing, Greenwich, CT, 2005.

[7] Schwartz D., Bransford J., and Sears D. Efficiency and innovation in transfer. In *Transfer of learning from a modern multidisciplinary perspective*, page 1. Information Age Publishing, Greenwich, CT, 2005.

[8] T.J. Nokes-Malach and Mestre J. Toward a model of transfer as sense-making. *Educ. Psych.*, 48:184, 2013.

[9] D. M. Belenky and Nokes-Malach T. J. Mastery-approach goals and knowledge transfer: An investigation into the effects of task structure and framing instructions. *Learning and Individual Differences*, 25:21, 2013.

[10] K. J. Pugh and Bergin D. A. Motivational influences on transfer. *Educ. Psych.*, 41:147, 2006.

[11] Vygotsky L. S. In Cole M., John-Steiner V., Scribner S., and Souberman E., editors, *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA, 1978.

[12] Hatano G. and Inagaki K. Two courses of expertise. In Stevenson H. W., Azuma H., and Hakuta K., editors, *A serious of books in psychology, child development, and education in Japan*, page 262. W. H. Freeman/Times Books/Henry Holt and Co, New York, NY, 1986.

[13] Collins A., Brown J. S., and Newman S. E. *Cognitive Apprenticeship: Teaching the craft of reading, writing, and mathematics*. Lawrence Erlbaum, Hillsdale, NJ, 1989.

[14] Marshman E., Kalender Z. Y., Schunn C., Nokes-Malach T., and Singh C. A longitudinal analysis of underrepresented students' motivational characteristics in introductory physics courses. *Can. J. Phys.*, 96:391, 2017.

[15] Elliot A. J. and Harackiewicz J. M. Approach and avoidance goals and intrinsic motivation: a mediational analysis. *J. of Personality and Social Psych.*, 70:461, 1996.

[16] Pintrich P. R. Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *J. of Educ. Psych.*, 92:544, 2000.

[17] Nicholls J. G. Achievement motivation: Concepts of ability, subjective experience, task choice, and performance. *Psych. Rev.*, 91:328, 1984.

[18] Dweck C. S. Motivational processes affecting learning. *American Psych.*, 41:1040, 1986.

[19] Ames C. Classrooms: Goals, structures, and student motivation. *J. of Educ.*, 84:261, 1992.

[20] Harackiewicz J. M., Barron K. E., Pintrich P. R., Elliot A. J., and Thrash T. M. Revision of achievement goal theory: necessary and illuminating.

[21] Dancy M. and Henderson C. Pedagogical practices and instructional change of physics faculty. *Am. J. Phys*, 78:10, 2010.

[22] Chi M., Feltovich P., and Glaser R. Characterization and representation of physics problems by experts and novices. *Cognitive Science*, 5:121, 1981.

[23] Reif F. Millikan lecture 1994: Understanding and teaching important scientific thought processes. *Am. J. Phys.*, 63:17, 1995.

[24] Schoenfeld A. H. *Mathematical Problem Solving*. Academic Press, 1985.

[25] Cummings K., Lockwood S., and Marx J. D. Attitudes towards problem solving as predictors of student success. *AIP Conf. Proc.*, 720:133, 2004.

[26] Saul J. M. Redish E. F. and Steinberg R. N. Student expectations in introductory physics. *Am. J. Phys.*, 66:212, 1998.

[27] Adams W. K., Perkins K. K., Podolefsky N. S., Dubson M., Finkelstein N. D., and Weiman C. E. New instrument for measuring student beliefs about physics and learning physics: The colorado learning attitudes and science survey. *Phys. Rev. STPER*, 2:010101, 2006.

[28] Marx J. and Cummings K. What factors really influence shifts in students' attitudes and expectations in an introductory physics course? *AIP Conf. Proc.*, 883:101, 2007.

[29] Elby A., Frederiksen J., Schwarz C., and White B. Epistemological beliefs assessment for physical science. 2004.

[30] Mason A. and Singh C. Surveying college introductory students' attitudes and approaches to problem solving. *Euro. J. Phys.*, 37:055704, 2010.

[31] Mason A. and Singh C. Surveying graduate students' attitudes and approaches to problem solving. *Phys. Rev. STPER*, 6:020124, 2010.

[32] Gray K. E., Adams W. K., Wieman C. E., and Perkins K. K. Students know what physicists believe, but they don't agree: A study using class survey. *Phys. Rev. STPER*, 4:020106, 2008.

[33] Yerushlami E., Marshman E., Maries A., Henderson C., and Singh C. Grading practices and considerations of graduate students at the beginning of their teaching assignment. In *Proc. Phys. Ed. Res. Conf.*, Minneapolis, MN, 2015.

[34] Yerushlami E., Sayer R., Marshman E., Henderson C., and Singh C. Physics graduate teaching assistants' beliefs about a grading rubric: Lessons learned. In *Proc. Phys. Ed. Res. Conf.*, Sacremento, CA, 2016.

[35] Lin S-Y., Henderson C., Mamudi W., Singh C., and Yerushalmi E. Teaching assistants' beliefs regarding example solutions in introductory physics. *Phys. Rev.*, 9:010120, 2013.

[36] Marshman E., Sayer R., Henderson C., Yerushalmi E., and Singh C. The challenges of changing teaching assistants grading practices: Requiring students to show evidence of understanding. *Can. J. Phys.*, 96:420, 2018.

[37] Marshman E., Sayer R., Henderson C., and Singh C. Contrasting grading approaches in introductory physics and quantum mechanics: The case of graduate teaching assistants. *Phys. Rev. PER*, 13:01020, 2017.

[38] Yerushalmi E., Henderson C., Heller K., Heller P., and Kuo V. Physics faculty beliefs and values about the teaching and learning of problem solving i. mapping the common core. *Phys. Rev. STPER*, 3:020109, 2007.

[39] Yerushalmi E., Heller K., Heller P., and Henderson C. Instructors' reasons for choosing problem features in a calculus-based introductory physics course. *Phys. Rev. STPER*, 6:020108, 2010.

[40] Elyon B. and Reif F. Effects of knowledge organization on task performance. *Cognition Instruction*, 1:5, 1984.

[41] E. F. Redish. Implications of cognitive studies for teaching physics. *American Journal of Physics*, 62:796–803, 1994.

[42] Scott M., Stelzer T., and Gladding G. Evaluating multiple-choice exams in large introductory physics courses. *Phys. Rev. ST Phys. Educ. Res.*, 2:020102, 2006.

[43] J. Mestre. Is transfer ubiquitous or rare? new paradigms for studying transfer. *AIP Conference Proceedings*, 790:3–6, 2005.

[44] Mashood K. K. and Singh V. A. Large-scale studies on the transferability of general problem-solving skills and the pedagogic potential of physics. *Physics Education*, 48:629, 2013.

[45] Bolton J. and Ross S. Developing students' physics problem-solving skills. *Physics Education*, 32:176, 1997.

[46] Singh C. When physical intuition fails. *American Journal of Physics*, 70:1103–1109, 2002.

[47] Singh C. Assessing student expertise in introductory physics with isomorphic problems. I. Performance on nonintuitive problem pair from introductory physics. *Physical Review Physics Education Research*, 4:010104, 2008.

[48] Singh C. Assessing student expertise in introductory physics with isomorphic problems. ii. effect of some potential factors on problem solving and transfer. *Phys. Rev. ST Phys. Educ. Res.*, 4:010105, 2008.

[49] Henderson C., Marshman E., Maries A., E. Yerushalmi, and Singh C. Instructional goals and grading practices of graduate students after one semester of teaching experience. *Proc. Physics Education Research Conf.*, 2016.

[50] Hake R. R. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66:64–74, 1998.

[51] Faust J. L. and Paulson D. R. Active learning in the college classroom. *Journal on Excellence in College Teaching*, 9:3, 1998.

[52] http://www.aps.org/programs/education/graduate/conf2008, 2008.

[53] Nyquist Jody D., Abbott Robert D., and Wulff Donald H. The challenge of ta training in the 1990s. *New Directions for Teaching and Learning*, 1989:7–14.

[54] Goertzen R. M., Scherr R. E., and A. Elby. Respecting tutorial instructors' beliefs and experiences: A case study of a physics teaching assistant. *Phys. Rev. ST Phys. Educ. Res.*, 6:020125, 2010.

[55] Chini J. and Al-Rawi A. Alignment of tas' beliefs with practice and student perception. In *Physics Education Research Conference 2012*, volume 1513 of *PER Conference*, pages 98–101, Philadelphia, PA, August 1-2 2012.

[56] Maries A. and Singh C. Exploring one aspect of pedagogical content knowledge of teaching assistants using the test of understanding graphs in kinematics. *Phys. Rev. ST Phys. Educ. Res.*, 9:020120, 2013.

[57] Maries A. and Singh C. Performance of graduate students at identifying introductory physics students' difficulties telated to kinematics graphs. In *Proc. Phys. Ed. Res. Conf.*, 2015.

[58] Maries A. and Singh C. Teaching assistants' performance at identifying common introductory student difficulties in mechanics revealed by the force concept inventory. *Phys. Rev. Phys. Educ. Res.*, 12:010131, 2016.

[59] Lawrenz F., Heller P., Keith R., and Heller K. Training the teaching assistant: Matching ta strengths and capabilities to meet specific program goals. *J. Coll. Sci. Teaching*, 22:106–109, 1992.

[60] Singh C. Categorization of problems to assess and improve proficiency as teacher and learner. *Am. J. Phys.*, 77:73, 2009.

[61] Singh C. Rethinking tools for training teaching assistants. In Singh C. Sabella M., Henderson C., editor, *Proceedings of the 2009 Phys. Ed. Res. Conf.*, Ann Arbor, MI, 2009.

[62] Mason A. and Singh C. Helping students learn effective problem solving strategies by reflecting with peers. *Am. J. Phys.*, 78:748–754, 2010.

[63] Mason A. and Singh C. Impact of guided reflection with peers on the development of effective problem solving strategies and physics learning. *The Physics Teacher*, 54:295–299, 2016.

[64] Heller P., Keith R., and Anderson S. Teaching problem solving through cooperative grouping. part 1: Group versus individual problem solving. *American Journal of Physics*, 60:627–636, 1992.

[65] Heller P. and Hollabaugh M. Teaching problem solving through cooperative grouping. part 2: Designing problems and structuring groups. *American Journal of Physics*, 60:637–644, 1992.

[66] Ogilvie C. A. Changes in students' problem-solving strategies in a course that includes context-rich, multifaceted problems. *Phys. Rev. ST Phys. Educ. Res.*, 5:020102, 2009.

[67] Antonenko Pavlo D., Ogilvie C. A., Niederhauser D. S., Jackman J., Kumsaikaew P., Marathe R., and Ryan S. M. Understanding student pathways in context-rich problems. *Education and Information Technologies*, 16:323–342, 2011.

[68] Ding L. Verification of causal influences of reasoning skills and epistemology on physics conceptual learning. *Phys. Rev. ST Phys. Educ. Res.*, 10:023101, 2014.

[69] Beichner R. The student-centered activities for large enrollment undergraduate programs (scale-up) project. In *Research-Based Reform of University Physics*, volume 1. April 2007.

[70] Cohen J. A coefficient of agreement for nominal scales. *Educ. and Psych. Meas.*, 20:37–46, 1960.

[71] Schoenfeld A. Toward a theory of teaching in context. *Issues in Educ.*, 4:1, 1998.

[72] Nicol J. D. and MacfarlaneDick D. Formative assessment and selfregulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31:199–218, 2006.

[73] Maries A. and Singh C. Improving one aspect of pedagogical content knowledge of teaching assistants using the tug-k. *Phys. Rev. STPER*, 9:020120, 2013.

[74] Mashood K. K. and Singh V. An inventory on rotational kinematics of a particle: Unravelling misconceptions and pitfalls in reasoning. *Eur. J. Phys.*, 33:1301–1312, 2012.

[75] Engelhardt P. An introduction to classical test theory as applied to conceptual multiple-choice tests. In *Getting Started in PER*, volume 2. April 2009.

[76] Ding L., Chabay R., Sherwood B., and Beichner R. Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Phys. Rev. ST Phys. Educ. Res.*, 2:010105, 2006.

[77] Li J. and Singh C. Developing and validating a conceptual survey to assess introductory physics students understanding of magnetism. *Euro. J. Phys.*, 38:025702, 2017.

[78] Savinainen A. and Scott P. Using the force concept inventory to monitor student learning and to plan teaching. *Phys. Educ.*, 37:53, 2002.

[79] Fuhrman M. Developing good multiple-choice tests and test questions. *Journal of Geoscience Education*, 44:379–384, 1996.

[80] Treagust D. F. Development and use of diagnostic tests to evaluate students misconceptions in science. *Int. J. Sci. Educ.*, 10:159–169, 1988.

[81] Mazur E. *Peer Instruction: A user's manual*. Prentice Hall, Upper Saddle River, NJ, 1997.

[82] Keller C., Finkelstein N., Perkins K., Pollock S., Turpen C., and Dubson M. Research-based practices for effective clicker use. *AIP Conference Proceedings*, 951:128–131, 2007.

[83] Singh C. Impact of peer interaction on conceptual test performance. *Am. J. Phys.*, 73:446–451, 2005.

[84] Singh C. and Zhu G. Improving students' understanding of quantum mechanics by using peer instruction tools. In *Physics Education Research Conference 2011*, volume 1413 of *PER Conference Invited Paper*, pages 77–80, Omaha, Nebraska, August 3-4 2011.

[85] Sayer R., Marshman E., and Singh C. The impact of peer interaction on the responses to clicker questions in an upper-level quantum mechanics course. In *Physics Education Research Conference 2016*, PER Conference, pages 304–307, Sacramento, CA, July 20-21 2016.

[86] Lee A., Ding L., Reay N. W., and Bao L. Single-concept clicker question sequences. *The Phys. Teach.*, 49:385–389, 2011.

[87] James M. C. The effect of grading incentive on student discourse in peer instruction. *American Journal of Physics*, 74:689–691, 2006.

[88] Willoughby S. D. and Gustafson E. Technology talks: Clickers and grading incentive in the large lecture hall. *Am. J. Phys.*, 77:180–183, 2009.

[89] Zappe S., Leicht R., Messner J., Litzinger T., and Lee H. W. "flipping" the classroom to explore active learning in a large undergraduate course. In *American Society for Engineering Education Annual Conference & Exposition*, 2009.

[90] Karim N. I., Maries A., and Singh C. Do evidence-based active-engagement courses reduce the gender gap in introductory physics? *Eur. J. Phys.*, 39:025701, 2018.

[91] Karim N. I., Maries A., and Singh C. Impact of evidence-based flipped or active-engagement non-flipped courses on student performance in introductory physics. *Canadian Journal of Physics*, 96:411–419, 2018.

[92] Lorenzo M., Crouch C. H., and Mazur E. Reducing the gender gap in the physics classroom. *American Journal of Physics*, 74:118–122, 2006.

[93] Bonham S. W., Beichner R. J., Titus A., and Martin L. Education research using web-based assessment systems. *Journal of Research on Computing in Education*, 33:28–45, 2000.

[94] Sokoloff D. R. and Thornton R. K. Using interactive lecture demonstrations to create an active learning environment. *The Phys. Teach.*, 35:340–347, 1997.

[95] Manivannan K. and Meltzer D. Use of in-class physics demonstrations in highly interactive format. In *Phys. Educ. Research Conf.*, 2001.

[96] Sharma M. D., Johnston I. D., Johnston H., Varvell K., Robertson G., Hopkins A., Stewart C., Cooper I., and Thornton R. Use of interactive demonstrations: A ten year study. *Phys. Rev. STPER*, 6:020119, 2010.

[97] Putt G. D. Testing the mastery concept of selfpaced learning in physics. *American Journal of Physics*, 45:472–475, 1977.

[98] Keebaugh C., Marshman E., and Singh C. Developing and evaluating an interactive tutorial on degenerate perturbation theory. *Physics Education Research Conference proceedings*, pages 184–187, 2016.

[99] Saul J. *Beyond problem solving: Evaluating introductory physics courses through the hidden curriculum.* Dissertation, University of Maryland, January 1998.

[100] Singh C. Interactive video tutorials for enhancing problemsolving, reasoning, and metacognitive skills of introductory physics students. 720:177–180, 09 2004.

[101] Singh C. Computer-based tutorials to develop expertise in introductory physics students. 2008.

[102] DeVore S., Marshman E., and Singh C. Challenge of engaging all students via self-paced interactive electronic learning tutorials for introductory physics. volume 13, page 010127, 2017.

[103] DeVore S. *Using the tutorial approach to improve physics learning from introductory to graduate level*. Dissertation, University of Pittsburgh, 2010.

[104] Sayer R., Marshman E., and Singh C. Case study evaluating just-in-time teaching and peer instruction using clickers in a quantum mechanics course. *Phys. Rev. Phys. Educ. Res.*, 12:020133, 2016.

[105] Lin S-Y. and Singh C. Can multiple-choice questions simulate free-response questions? In *Physics Education Research Conference 2011*, volume 1413 of *PER Conference Invited Paper*, pages 47–50, Omaha, Nebraska, August 3-4 2011.

[106] Simon H. A. *Problem solving and education*. Erlbaum, Hillsdale, NJ, 1980.

[107] Lin S.-Y. and Singh C. Publisher's note: Effect of scaffolding on helping introductory physics students solve quantitative problems involving strong alternative conceptions. *Phys. Rev. ST Phys. Educ. Res*, 11:029901, 2015.

[108] Sweller J. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12:257 – 285, 1988.

[109] Black P. and Wiliam D. Assessment and classroom learning. *Assessment in Education: Principles, Policy, and Practice*, 5:7–74, 1998.

[110] Singh C. What can we learn from per: Physics education research? *Phys. Teach.*, 52:568, 2014.

[111] Schommer M. Effects of beliefs about the nature of knowledge on comprehension. *J. Educ. Psych.*, 82:498, 1990.

[112] Elby A. Helping physics students learn how to learn. *American Journal of Physics*, 69:S54–S64, 2001.

[113] Balta N., , Mason A., and Singh C. Surveying turkish high school and university students' attitudes and approaches to physics problem solving. *Phys. Rev. Phys. Educ. Res.*, 12:010129, 2016.

[114] Reddy L. An insight into the students' attude and approach to problem solving in physics at a south african university. *Proceedings of the 2016 ISTE International Conference on Mathematics, Science and Technology Education*, 2016.

[115] Hestenes D., Wells M., and Swackhamer G. Force concept inventory. *The Phys. Teach.*, 30:141–158, 1992.

[116] Coletta V. P. and Phillips J. A. Interpreting fci scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *Am. J. Phys.*, 73:1172–1182, 2005.

[117] Maloney D. P., O'Kuma T. L., Hieggelke C. J., and Van Heuvelen A. Surveying students conceptual knowledge of electricity and magnetism. *Am. J. Phys.*, 69:S12–S23, 2001.

[118] Henderson C. and Dancy M. Impact of physics education research on the teaching of introductory quantitative physics in the united states. *Phys. Rev. ST Phys. Educ. Res.*, 5:020107, 2009.

[119] Henderson C., Dancy M., and Niewiadomska-Bugaj M. Use of research-based instructional strategies in introductory physics: Where do faculty leave the innovation-decision process? *Phys. Rev. ST Phys. Educ. Res.*, 8:020104, 2012.

[120] Kreutzer K. and Boudreaux A. Preliminary investigation of instructor effects on gender gap in introductory physics. *Phys. Rev. ST Phys. Educ. Res.*, 8:010120, 2012.

[121] Adegoke B. A. Impact of interactive engagement on reducing the gender gap in quantum physics learning outcomes among senior secondary school students. *Phys. Educ.*, 47:462, 2012.

[122] Madsen A., McKagan S., and E. Sayre. Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap? *Phys Rev ST PER.*

[123] Traxler A., Cid X., . Blue J, and Barthelemy R. Enriching gender in physics education research: A binary past and a complex future. *Phys. Rev. Phys. Educ. Res.*

[124] Traxler A., Henderson R., Stewart J., Stewart G., Papak A., and Lindell R. Gender fairness within the force concept inventory. *Phys. Rev. PER.*

[125] Marchand G. and Taasoobshirazi G. Stereotype threat and women's performance in physics. *Int. J. Sci. Educ.*, 35:3050, 2013.

[126] Bian L., Leslie S., and Cimpian A. Gender stereotypes about intellectual ability emerge early and influence childrens interests. *Science*, 355:6323, 2017.

[127] Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review.*

[128] Madsen A., Mckagan S., and Sayre E. How physics instruction impacts students' beliefs about learning physics: A meta-analysis of 24 studies. *Phys. Rev. ST PER*, 11:010115, 2014.

[129] Nokes-Malach T., Marshman E., Kalender Z.Y., Schunn C., and Singh C. Investigation of male and female students' motivational characteristics throughout an introductory

physics course sequence. *2017 Physics Education Research Conference Proceedings*, pages 276–279, 2018.

[130] Lomax R. G. *Statistical concepts: A second course (3rd ed.)*. Lawrence Erlbaum Associates, 2007.

[131] Meltzer D. Relation between students problem solving performance and representational mode. *Am. J. Phys.*, 73:463, 2005.

[132] Harper K. Student problem solving behaviors. *Phys. Educ.*, 44:250, 2006.

[133] Gladding G, Gutmann B., Schroeder N., and Stelzer T. Clinical study of student learning using mastery style versus immediate feedback online activities. *Phys. Rev. ST Phys. Educ. Res.*, 11:010114, 2015.

[134] Tobias S. Research Corporation, 1990.

[135] Lin S. Using isomorphic problems to learn introductory physics. *Phys. Rev. STPER*, 7:020104, 2011.

[136] Good D. J. Master's thesis, University of Pittsburgh, 1982.

[137] Elby A. Another reason that physics students learn by rote. *American Journal of Physics*, 67:S52–S57, 1999.

[138] Mason A. and Singh C. Reflection and Self-Monitoring in Quantum Mechanics. In Sabella M., Henderson C., and Singh C., editors, *American Institute of Physics Conference Series*, volume 1179 of *American Institute of Physics Conference Series*, pages 197–200, 2009.

APPENDIX A

AAPS SURVEY

AAPS survey given to physics students

To what extent do you agree with each of the following statements when you solve physics problems?

Answer with a single letter as follows:

A. Strongly agree

B. Agree somewhat

C. Neutral or do not know

D. Disagree somewhat

E. Strongly disagree

1. If I am not sure about the right way to start a problem, I am stuck unless I go see the teacher/TA or someone else for help.

2. When solving physics problems, I often make approximations about the physical world.

3. In solving problems in physics, being able to handle the mathematics is the most important part of the process.

4. In solving problems in physics, I always identify the physics principles involved in the problem first before looking for corresponding equations.

5. "Problem solving" in physics basically means matching problems with the correct equations and then substituting values to get a number.

6. In solving problems in physics, I can often tell when my work and/or answer is wrong, even without looking at the answer in the back of the book or talking to someone else about it.

7. To be able to use an equation to solve a problem (particularly in a problem that I have not seen before), I think about what each term in the equation represents and how it matches the problem situation.

8. There is usually only one correct way to solve a given problem in physics.

9. I use a similar approach to solving all problems involving conservation of linear momentum even if the physical situations given in the problems are very different.

10. If I am not sure about the correct approach to solving a problem, I will reflect upon physics principles that may apply and see if they yield a reasonable solution.

11. Equations are not things that one needs to understand in an intuitive sense; I routinely use equations to calculate numerical answers even if they are non-intuitive.

12. Physics involves many equations each of which applies primarily to a specific situation.

13. If I used two different approaches to solve a physics problem and they gave different answers, I would spend considerable time thinking about which approach is more reasonable.

14. When I solve physics problems, I always explicitly think about the concepts that underlie the problem.

15. When solving physics problems, I often find it useful to first draw a picture or a diagram of the situations described in the problems.

16. When answering conceptual physics questions, I mostly use my "gut" feeling rather than using the physics principles I usually think about when solving quantitative problems.

17. I am equally likely to draw pictures and/or diagrams when answering a multiple-choice question or a corresponding free-response (essay) question.

18. I usually draw pictures and/or diagrams even if there is no partial credit for drawing them.

19. I am equally likely to do scratch work when answering a multiple-choice question or a corresponding free-response (essay) question.

20. After I solve each physics homework problem, I take the time to reflect and learn from the problem solution.

21. After I have solved several physics problems in which the same principle is applied in different contexts, I should be able to apply the same principle in other situations.

22. If I obtain an answer to a physics problem that does not seem reasonable, I spend considerable time thinking about what may be wrong with the problem solution.

23. If I cannot solve a physics problem in 10 min, I give up on that problem.

24. When I have difficulty solving a physics homework problem, I like to think through the problem with a peer.

25. When I do not get a question correct on a test or homework, I always make sure I learn from my mistakes and do not make the same mistakes again.

26. It is more useful for me to solve a few difficult problems using a systematic approach and learn from them rather than solving many similar easy problems one after another.

27. I enjoy solving physics problems even though it can be challenging at times.

28. I try different approaches if one approach does not work.

29. If I realise that my answer to a physics problem is not reasonable, I trace back my solution to see where I went wrong.

30. It is much more difficult to solve a physics problem with symbols than solving an identical problem with a numerical answer.

31. While solving a physics problem with a numerical answer, I prefer to solve the problem symbolically first and only plug in the numbers at the very end.

32. Suppose you are given two problems. One problem is about a block sliding down an inclined plane with no friction present. The other problem is about a person swinging on a rope. Air resistance is negligible. You are told that both problems can be solved using the concept of conservation of mechanical energy of the system. Which one of the following statements do you MOST agree with? (Choose only one answer.)

A. The two problems can be solved using very similar methods.

B. The two problems can be solved using somewhat similar methods.

C. The two problems must be solved using somewhat different methods.

D. The two problems must be solved using very different methods.

E. There is not enough information given to know how the problems will be solved.

33. Suppose you are given two problems. One problem is about a block sliding down an inclined plane. There is friction between the block and the incline. The other problem is about a person swinging on a rope. There is air resistance between the person and air molecules. You are told that both problems can be solved using the concept of conservation of total (not just mechanical) energy. Which one of the following statements do you MOST agree with? (Choose only one answer.)

A. The two problems can be solved using very similar methods.

B. The two problems can be solved using somewhat similar methods.

C. The two problems must be solved using somewhat different methods.

D. The two problems must be solved using very different methods.

E. There is not enough information given to know how the problems will be solved.

AAPS survey given to astronomy students

To what extent do you agree with each of the following statements when you solve astronomy problems?

Answer with a single letter as follows:

A. Strongly agree

B. Agree somewhat

C. Neutral or do not know

D. Disagree somewhat

E. Strongly disagree

1. If I am not sure about the right way to start a problem, I am stuck unless I go see the teacher/TA or someone else for help.

2. When solving astronomy problems, I often make approximations about the physical world.

3. In solving problems in astronomy, being able to handle the mathematics is the most important part of the process.

4. In solving problems in astronomy, I always identify the astronomy principles involved in the problem first before looking for corresponding equations.

5. "Problem solving" in astronomy basically means matching problems with the correct equations and then substituting values to get a number.

6. In solving problems in astronomy, I can often tell when my work and/or answer is wrong, even without looking at the answer in the back of the book or talking to someone else about it.

7. To be able to use an equation to solve a problem (particularly in a problem that I have not seen before), I think about what each term in the equation represents and how it matches the problem situation.

8. There is usually only one correct way to solve a given problem in astronomy.

9. I use a similar approach to solving all problems involving conservation of linear momentum even if the physical situations given in the problems are very different.

10. If I am not sure about the correct approach to solving a problem, I will reflect upon astronomy principles that may apply and see if they yield a reasonable solution.

11. Equations are not things that one needs to understand in an intuitive sense; I routinely use equations to calculate numerical answers even if they are non-intuitive.

12. Astronomy involves many equations each of which applies primarily to a specific situation.

13. If I used two different approaches to solve an astronomy problem and they gave different answers, I would spend considerable time thinking about which approach is more reasonable.

14. When I solve astronomy problems, I always explicitly think about the concepts that underlie the problem.

15. When solving astronomy problems, I often find it useful to first draw a picture or a diagram of the situations described in the problems.

16. When answering conceptual astronomy questions, I mostly use my "gut" feeling rather than using the astronomy principles I usually think about when solving quantitative problems.

17. I am equally likely to draw pictures and/or diagrams when answering a multiple-choice question or a corresponding free-response (essay) question.

18. I usually draw pictures and/or diagrams even if there is no partial credit for drawing them.

19. I am equally likely to do scratch work when answering a multiple-choice question or a corresponding free-response (essay) question.

20. After I solve each astronomy homework problem, I take the time to reflect and learn from the problem solution.

21. After I have solved several astronomy problems in which the same principle is applied in different contexts, I should be able to apply the same principle in other situations.

22. If I obtain an answer to an astronomy problem that does not seem reasonable, I spend considerable time thinking about what may be wrong with the problem solution.

23. If I cannot solve an astronomy problem in 10 min, I give up on that problem.

24. When I have difficulty solving an astronomy homework problem, I like to think through the problem with a peer.

25. When I do not get a question correct on a test or homework, I always make sure I learn from my mistakes and do not make the same mistakes again.

26. It is more useful for me to solve a few difficult problems using a systematic approach and learn from them rather than solving many similar easy problems one after another.

27. I enjoy solving astronomy problems even though it can be challenging at times.

28. I try different approaches if one approach does not work.

29. If I realise that my answer to an astronomy problem is not reasonable, I trace back my solution to see where I went wrong.

30. It is much more difficult to solve an astronomy problem with symbols than solving an identical problem with a numerical answer.

31. While solving an astronomy problem with a numerical answer, I prefer to solve the problem symbolically first and only plug in the numbers at the very end.

32. Suppose you are given two problems. One problem is about a block sliding down an inclined plane with no friction present. The other problem is about a person swinging on a rope. Air resistance is negligible. You are told that both problems can be solved using the concept of conservation of mechanical energy of the system. Which one of the following statements do you MOST agree with? (Choose only one answer.)

A. The two problems can be solved using very similar methods.

B. The two problems can be solved using somewhat similar methods.

C. The two problems must be solved using somewhat different methods.

D. The two problems must be solved using very different methods.

E. There is not enough information given to know how the problems will be solved.

33. Suppose you are given two problems. One problem is about a block sliding down an inclined plane. There is friction between the block and the incline. The other problem is about a person swinging on a rope. There is air resistance between the person and air molecules. You are told that both problems can be solved using the concept of conservation of total (not just mechanical) energy. Which one of the following statements do you MOST agree with? (Choose only one answer.)

A. The two problems can be solved using very similar methods.

B. The two problems can be solved using somewhat similar methods.

C. The two problems must be solved using somewhat different methods.

D. The two problems must be solved using very different methods.

E. There is not enough information given to know how the problems will be solved.

Favorable responses to AAPS survey

1. D/E
2. A/B
3. D/E
4. A/B
5. D/E
6. A/B
7. A/B
8. D/E

9. A/B
10. A/B
11. D/E
12. D/E
13. A/B
14. A/B
15. A/B
16. D/E
17. A/B
18. A/B
19. A/B
20. A/B
21. A/B
22. A/B
23. D/E
24. A/B
25. A/B
26. A/B
27. A/B
28. A/B
29. A/B
30. D/E
31. A/B
32. A/B
33. A/B

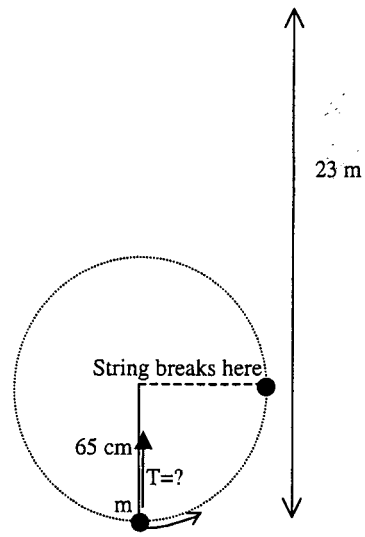
APPENDIX B

PROBLEM VARIATIONS

Below are example problems given in the TA professional development class related to the problem variations activities. These served as concrete examples of a problem "type" to help guide discussions in class and during interviews. However, such discussions and interviews were aimed at probing TAs' views about the problem types in a general sense.

Problem A

A 1.8 kg mass is attached to a frictionless pivot point and is moving in a circle at the end of a 65 cm string. The string breaks when the mass is moving directly upward and the mass rises to a maximum height of 23.0 m. What is the tension in the string one-quarter turn before the string breaks? Assume that air resistance can be neglected.



A) What velocity, v_1, must the stone have when released in order to rise to 23 meters above the lowest point in the circle?

B) What velocity, v_o, must the stone have when it is at its lowest point in order to have a velocity v_1 when released?

C) What force will you have to exert on the string at its lowest point in order for the stone to have a velocity v_o?

Problem B

You are whirling a stone tied to the end of a string around in a vertical circle having a radius of 65 cm. You wish to whirl the stone fast enough so that when it is released at the point where the stone is moving directly upward it will rise to a maximum height of 23 meters above the lowest point in the circle. In order to do this, what force will you have to exert on the string when the stone passes through its lowest point one-quarter turn before release? Assume that by the time that you have gotten the stone going and it makes its final turn around the circle, you are holding the end of the string at a fixed position. Assume also that air resistance can be neglected. The stone weighs 18 N.

A) 1292 N
B) 1258 N
C) 1248 N
D) 1210 N
E) None of the Above

Note: The choices are based on common student problems.

Problem C

You are working at a construction site and need to get a 3 lb. bag of nails to your co-worker standing on the top of the building (60 ft. from the ground). You don't want to climb all the way up and then back down again, so you try to throw the bag of nails up. Unfortunately, you're not strong enough to throw the bag of nails all the way up so you try another method. You tie the bag of nails to the end of a 2 ft. string and whirl the string around in a vertical circle. You try this, and after a little while of moving your hand back and forth to get the bag going in a circle you notice that you no longer have to move your hand to keep the bag moving in a circle. You think that if you release the bag of nails when the string is horizontal to the ground that the bag will go up to your co-worker. As you whirl the bag of nails around, however, you begin to worry that the string might break, so you stop and attempt to decide before continuing. According to the string manufacturer, the string is designed to hold up to 100 lbs. You know from experience that the string is most likely to break when the bag of nails is at its lowest point.

Problem D

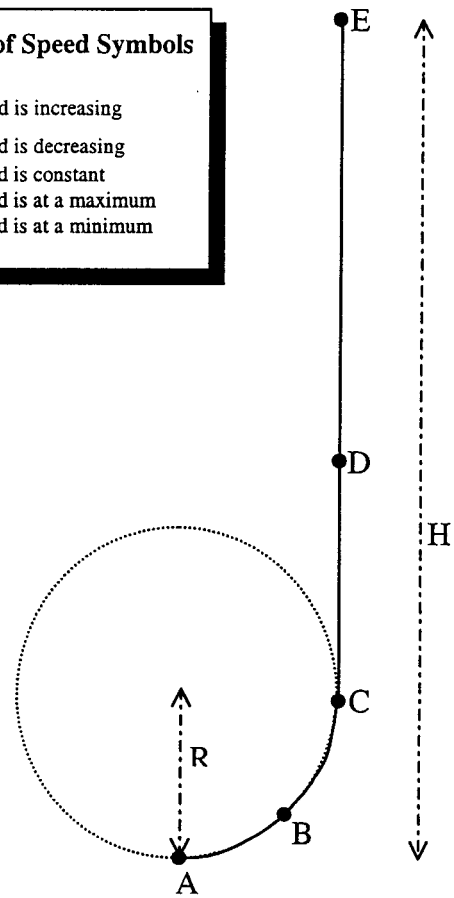
You are whirling a stone tied to the end of a string around in a vertical circle of radius R. You wish to whirl the stone fast enough so that when it is released at the point where the stone is moving directly upward it will rise to a maximum height, H, above the lowest point in the circle. In order to do this, what force will you have to exert on the string when the stone passes through its lowest point one-quarter turn before release? Assume that by the time that you have gotten the stone going and it makes its final turn around the circle, you are holding the end of the string at a fixed position. Assume also that air resistance can be neglected.

A) For each point labeled in the diagram, circle the symbol(s) that describe how the speed of the stone is changing.

Point	Change in Speed
A	↑ ↓ = max min
B	↑ ↓ = max min
C	↑ ↓ = max min
D	↑ ↓ = max min
E	↑ ↓ = max min

Change of Speed Symbols

↑	Speed is increasing
↓	Speed is decreasing
=	Speed is constant
max	Speed is at a maximum
min	Speed is at a minimum



B) At each point on the diagram, draw and label a vector representing the acceleration of the stone.

C) At each point, draw and label vectors to represent all of the forces acting on the stone.

Problem E

You are whirling a stone tied to the end of a string around in a vertical circle having a radius of 65 cm. You wish to whirl the stone fast enough so that when it is released at the point where the stone is moving directly upward it will rise to a maximum height of 23 meters above the lowest point in the circle. In order to do this, what force will you have to exert on the string when the stone passes through its lowest point one-quarter turn before release? Assume that by the time that you have gotten the stone going and it makes its final turn around the circle, you are holding the end of the string at a fixed position. Assume also that air resistance can be neglected. The stone weighs 18 N.

The correct answer is 1292 N

APPENDIX C

PROBLEM VARIATIONS WORKSHEETS

Below are the final versions of worksheets given in the TA professional development class related to the problem variations activities. Not all worksheets were used in all years, but each year, some iteration of the following worksheets were used.

Name

<u>"Different Problem Type for the same physics scenario" Activity</u>

Activity goals: To reflect on different ways of posing physics problems that focus on the same physical situation.

1. Write a typed, ½ page essay about each problem type (examples illustrating each problem type are problems A-E) that answers the following questions:

a. Have you encountered this type of problem before? If yes, in what instructional situations have you seen this type of problem before?
b. What do you like about this type of problem?
c. What would you change about the problem in order to make it an even better problem (make it more instructionally beneficial)?
d. In what ways do you think this type of problem is beneficial for introductory physics students' learning?
e. Do you think this type of problem is helpful in teaching students problem solving strategies that can help them in future problem solving? Explain why or why not.
f. If you had full control of teaching an introductory physics course, in what way(s) or instructional settings would you use this type of problem to achieve an instructional goal (e.g., as a homework problem, quiz problem, exam problem, in small group problem solving, etc.)? Why?
g. If you had full control of teaching an advanced physics course, would you use this type of problem to achieve an instructional goal in an advanced physics course? Why or why not?

2. Now compare the 5 problem types and rate each problem on a scale of 1 to 5:
Rate the problem types based upon how **instructionally beneficial** they are for an introductory student (1 being least beneficial and 5 being the most beneficial) and explain your reasons for your rating:

Rate	Explain your reasons for the rating
Problem type A Rating (circle one): 1 2 3 4 5	
Problem type B Rating (circle one): 1 2 3 4 5	
Problem type C Rating (circle one): 1 2 3 4 5	
Problem type D Rating (circle one): 1 2 3 4 5	
Problem type E Rating (circle one): 1 2 3 4 5	

Rate the problem types based upon the **level of challenge** for introductory students (1 being the least challenging and 5 being the most challenging) and explain your reasons for your rating:

Rate	Explain your reasons for the rating
Problem type A Rating (circle one): 1 2 3 4 5	
Problem type B Rating (circle one): 1 2 3 4 5	
Problem type C Rating (circle one): 1 2 3 4 5	
Problem type D Rating (circle one): 1 2 3 4 5	
Problem type E Rating (circle one): 1 2 3 4 5	

Rate the problem types based upon your likelihood to **use** for introductory students if you had complete control of teaching a class (1 being the least likely to use and 5 being the most likely to use) and explain your reasons for your rating:

Rate	Explain your reasons for the rating
Problem type A Rating (circle one): 1 2 3 4 5	
Problem type B Rating (circle one): 1 2 3 4 5	
Problem type C Rating (circle one): 1 2 3 4 5	
Problem type D Rating (circle one): 1 2 3 4 5	
Problem type E Rating (circle one): 1 2 3 4 5	

Rate the problem types based upon how much you **like** the problems (1 being the least liked and 5 being the most liked) and explain your reasons for your rating:

Rate	Explain your reasons for the rating
Problem type A Rating (circle one): 1 2 3 4 5	
Problem type B Rating (circle one): 1 2 3 4 5	
Problem type C Rating (circle one): 1 2 3 4 5	
Problem type D Rating (circle one): 1 2 3 4 5	
Problem type E Rating (circle one): 1 2 3 4 5	

Stage 1 – "Different Problem Type for the same physics scenario" Activity guidelines

Activity goals: to look at alternative ways of asking problems about the same physical situation and to clarify our considerations in designing different problem types for our students to meet different instructional goals in different instructional settings.

Instructions:

1. Look at the five problem types and note the different **features** in each problem type. Note that these are example problems to illustrate each problem type and you should be thinking about the features of well-designed problems of a particular problem type in general (for multiple-choice problem type, consider both conceptual and quantitative multiple-choice problems even though the example problem illustrating this problem type is quantitative)

2. Fill in the column in the table called **"Problem type features"**: describe the main features of each problem type.

3. Fill in the column in the table called **"Problem type features"**: mark "✓" for the problem types that contain these features and "✗" for the problem types that do not contain these features.

4. Fill in the column in the table called **"Requirements"**: Different problem types or ways of asking problems require different things from students. Explain what requirements the various features of a problem type pose for students. Do the features increase or decrease the amount of work and thinking required of students to solve the problem.

5. Fill in the column in the table called **"Preference"**: For each problem type features, please provide your preference - would you like/not like to include this feature in the problems you give to your students in a particular situation?

6. Fill in the column in the table called "**Pros/Cons of problem type feature**": Explain the pros and cons of using this feature on a homework or quiz problem. Please try to suggest at least one "pro" and "con" for each problem type feature. This column should explain your preferences for including/excluding the feature on a HW or quiz problem.

Problem type features	**Sorting problem types by features**					**Requirements** **If this feature is present in the problem statement, does it increase or decrease requirements of the student and how?**	**Preference**		**PROS/CONS of problem type feature (Please try to suggest at least one PRO and one CON for each problem type feature, even if you are not likely to make use of it)**
	A	B	C	D	E		HW	QUIZ	
e.g. Drawing	✓	✗	✗	✓	✗	If a drawing is present in the problem statement, it decreases the requirements of the students. They don't have to go through the steps of visualizing and drawing a representation of the physical situation.	✗	✓	PRO: Visualizing the problem is an important problem solving strategy. On HW students have time to visualize and draw, so I will not provide a drawing for them. CON: Visualizing and drawing the problem takes time for the student. On a quiz or final exam, I would provide a drawing which would help students when they are stressed.

Problem type features	Sorting problem types by features					Requirements If this feature is present in the problem statement, does it increase or decrease requirements of the student and how?	Preference		PROS/CONS of problem type feature (Please try to suggest one PRO and one CON for each problem type feature, even if you are not likely to make use of it)
	A	B	C	D	E		HW	QUIZ	

Problem type features	A	B	C	D	E	**Requirements** **If this feature is present in the problem statement, does it increase or decrease requirements of the student and how?**	**Preference**		**PROS/CONS of problem type feature** **(Please try to suggest at least one PRO and one CON for each problem type feature, even if you are not likely to make use of it)**
							HW	QUIZ	

Stage 2 - Group Worksheet, "Different Problem Types for the same physics scenario" activity

In case there is no agreement explain what you disagreed upon and why.

Problem type features	Sorting problem types by features					Requirements If this feature is present in the problem statement, does it increase or decrease requirements of the student and how?	Preference		PROS/CONS of problem type feature (Please try to suggest at least one PRO and one CON for each problem type feature, even if you are not likely to make use of it)
	A	B	C	D	E		HW	QUIZ	

Problem type features	A	B	C	D	E	**Requirements** **If this feature is present in the problem statement, does it increase or decrease requirements of the student and how?**	**Preference**		**PROS/CONS of problem type feature** **(Please try to suggest at least one PRO and one CON for each problem type feature, even if you are not likely to make use of it)**
							HW	QUIZ	

Stage 3 - Post-discussion, "Different Problem Types for the same physics scenario" Activity

Instructions: Attached is a set of problem type features that another group of instructors/TAs noticed in the five problem types for the same physics scenario. Please think about all the features of a problem type (including the ones that appear on the list attached, the ones that you originally noticed, and the ones that were discussed in class discussion) again.

On the worksheet,

1. List ALL the problem type features that you NOW believe have pedagogical implications in some instructional setting to meet some instructional goal in the problem features column.
2. For each of the problem type features, if you noticed it originally in Stage 1 – "Different Problem Types for the same physics scenario" Activity, please write down how you originally named this feature in the problem type feature column. If you did not originally notice it, please choose a name yourself.
3. Please go over the attached list of problem type features and select the one that describes the same problem type feature and write down its number in the first column. If none of the problem type features in the list correspond to the one you describe, please write down "other" in the first column.

Number	Problem type Feature
1	Qualitative
2	Multiple-choice
3	Broken into parts
4	Real-world context
5	Wordy
6	Diagram/Drawing given
7	Complex or multi-step

Number	Problem type features	Sorting problem types by features					Requirements If this feature is present in the problem statement, does it increase or decrease requirements of the student and how?	Preference		PROS/CONS of problem type feature (Please try to suggest at least one PRO and one CON for each problem type feature, even if you are not likely to make use of it)
		A	B	C	D	E		HW	QUIZ	
6	e.g., Drawing	✓	✗	✗	✓	✗	If a drawing is present in the problem statement, it decreases the requirements of the students. They don't have to go through the steps of visualizing and drawing a representation of the physical situation.	✗	✓	PRO: Visualizing the problem is an important problem solving strategy. On HW students have time to visualize and draw, so I will not provide a drawing for them. CON: Visualizing and drawing the problem takes time for the student. On a quiz or final exam, I would provide a drawing which would help students when they are stressed.

Number	Problem type features	Sorting problem types by features					Requirements If this feature is present in the problem statement, does it increase or decrease requirements of the student and how?	Preference		PROS/CONS of problem type feature (Please try to suggest at least one PRO and one CON for each problem type feature, even if you are not likely to make use of it)
		A	B	C	D	E		HW	QUIZ	

Number	Problem type features	A	B	C	D	E	Requirements If this feature is present in the problem statement, does it increase or decrease requirements of the student and how?	Preference		PROS/CONS of problem feature (Please try to suggest one PRO and one CON for each problem feature, even if you are not likely to make use of these features)
								HW	QUIZ	

Number	**Problem type Features**	A	B	C	D	E	**Requirements** **If this feature is present in the problem statement, does it increase of decrease requirements of the student and how?**	**HW**	**QUIZ**	**PROS/CONS of problem type features (Please try to suggest at least one PRO and one CON for each problem type feature, even if you are not likely to make use of it)**

Compare the problem types (see examples to illustrate each problem type - A-E) and answer the following questions in the tables below (**Important**: make sure your rankings are for a particular class of problem type since the example problem for each problem type is just one specific example to illustrate the problem type):

Rank the problem types based upon how **challenging** they are for an introductory student (1 being least challenging and 5 being the most challenging) and explain your reasons for your ranking:

Rank	Explain your reasons for the ranking.
1. Problem type ___	
2. Problem type ___	
3. Problem type ___	
4. Problem type ___	
5. Problem type ___	

Rank the problem types based upon the **instructional benefits** for introductory students (1 being the least instructionally beneficial and 5 being the most instructionally beneficial) and explain your reasons for the ranking you gave:

Rank	Explain your reasons for the ranking.
1. Problem type ___	
2. Problem type ___	
3. Problem type ___	
4. Problem type ___	
5. Problem type ___	

Follow-Up Homework Assignment: Design of Different Types of Introductory Physics Problems

Using the quantitative problem you created, re-write your problem in the following ways:

1). Pose the problem as two separate multiple-choice (MC) problems: a quantitative MC problem and a conceptual MC problem. Carefully consider what you include in your choices. Will an incorrect answer tell you anything about what a student did not understand and provide feedback to you about how to improve student learning?

2). Pose the problem as a "context-rich" problem, in which you create a realistic narrative of the scenario and students must interpret what is being asked of them. As you write the context-rich problem, think about a way in which such a problem could be used to help student's development of problem-solving skills. Is it possible for the students to use a "plug and chug" approach for the problem you created?

3). Pose the problem as a conceptual open-ended problem which does not require explicit calculation. It may be helpful to identify and write down concepts which are central to your quantitative problem in order to construct an analogous problem which is entirely conceptual (qualitative). Will preceding this problem with the corresponding quantitative problem you created earlier help your students think about how they would approach the quantitative version of the problem better?