Exploring Instructor Knowledge of Student Ideas with the Force Concept Inventory

Robert A. Haddard and Brian W. Frank

Department of Physics & Astronomy, Middle Tennessee State University, Murfreesboro, TN 37132

Abstract. Pedagogical content knowledge has been a useful construct for conceptualizing the knowledge-base that supports reform teaching practices. In physics education, we are still far from having established methods and instruments for assessing such knowledge. In an attempt to begin exploring possibilities for assessing instructor knowledge of student thinking, we asked a small sample of college physics instructors to take the Force Concept Inventory in two novel ways. Instructors were first asked to indicate the answer they think a typical novice student would choose prior to instruction and then to estimate the fraction of students answering correctly after instruction. We analyze how instructor responses compare with actual student data at our institution and discuss questions with significant mismatch—either ones that instructors overwhelmingly succeeded (or failed) at identifying student difficulties or questions where instructors were overwhelmingly pessimistic (or optimistic) about student performance. Implications for future work in this area of assessment are discussed.

Keywords: Pedagogical Content Knowledge

PACS: 01.40.Fk

INTRODUCTION

Pedagogical content knowledge, among other types of knowledge for teaching, has been shown to correlate with reform teaching practices [1,2]. Given such findings, pre-service physics teacher programs have been increasingly developing courses to help teachers develop this knowledge before entering the classroom [3]. Researchers have also recently placed increased attention to understanding the processes by which pre-service and in-service teachers interact with displays of student thinking in ways that help develop new knowledge [4,5]. The physics education research community, however, is still a long ways from having established methods and instruments for assessing pedagogical content knowledge, especially those that might afford comparisons and could be widely and readily used by a variety of stakeholders.

We report on some preliminary work whose aim is to take steps toward the eventual development of diagnostic instruments that would be capable of assessing instructor's knowledge of student thinking about force and motion. In our pilot study, five university physics instructors responded to questions from the Force Concept Inventory [6] (FCI) in two novel ways. Instructors first indicated the incorrect (multiple-choice) answer that they thought a typical novice student would choose prior to instruction and then estimated the fraction of students answering each question correctly after instruction. We were aiming to

gather information about instructor's awareness of (1) specific student difficulties and (2) their awareness of prevalence of those difficulties in their classrooms. We compare instructor responses to actual student data from our local institution, and discuss findings concerning the accuracy of instructor's predictions.

CONTEXT FOR RESEARCH

The Department of Physics and Astronomy at Middle Tennessee State University (MTSU) utilizes many reforms in their introductory algebra-based physics sequence, including Peer Instruction [7], and collaborative problem-solving. Students meet weekly for five hours in a setting that integrates laboratory investigations, problem-solving, and interactive demonstrations. Each section enrolls thirty-two students and is facilitated by a single university instructor. All sections collectively meet for an additional 1.5 hours of lecture taught by a lead instructor, also using interactive engagement methods.

Instructor's work with a common structure and pacing, presenting the same examples, guiding students during shared problem-solving and lab work, and facilitating similar class discussions. Instructors share grading responsibilities for common exams. Instructor's also administer the FCI at the start and end of the semester, having opportunities to familiarize themselves with the survey, but not the results. Thus

instructors have similar in-class opportunities to observe students working through the content.

DATA COLLECTION AND ANALYSIS

The survey was taken by five instructors who teach the course described above, whose teaching experience ranged from 5 to 27 years. The choice to use the FCI as a starting place for conducting preliminary work toward assessing instructor's knowledge of student thinking about force and motion has both positives and negatives. In terms of positives, the FCI consists of questions that are designed around research into student thinking, thereby emphasizing concepts and student difficulties that the PER community values and are known to be important for learning. This fact provides us potentially meaningful artifacts and contexts to give instructor's whose use with students have been validated. Additionally, since the FCI is administered to students at our institution, we are able to directly compare instructor perceptions to actual data in a local, contextualized manner.

There are also some negatives to using the FCI. For example, FCI distractors are useful when used with students. Those same distractors, however, may not be useful as distractors for instructors trying to anticipate what students will choose. In other words, instructors might spontaneously think students would say other things not options on the FCI. In addition, while the questions themselves are demonstrably useful for assessing student understanding of physics, we do not know the value they may for assessing instructor knowledge of student thinking.

Survey Structure and Analysis

We adapted the FCI for our purposes by asking instructors to respond to every FCI question in two novel ways. First, they were asked to indicate the "Most Common Incorrect" (MCI) answer, described as the incorrect answer that they think would be chosen most often by students *before* instruction. Second, instructors were asked to select a range that representing the percentage correct after instruction (PC), being forced to choose among 0-20%, 21-40%, 41-60%, 61-80%, and 81-100%. Instructors were also asked to select the correct answers, but all instructors answered every question correctly, except for one question which was answered wrong by one instructor.

In order to analyze the data, instructor responses were compared to actual student data pooled from two recent semesters of algebra-based physics ($N = 316$). To scrutinize instructor's performance in predicting PC, we examined their responses two ways. First, to provide an *absolute* sense of instructor's accuracy, we compared the percentage of students who actually

answered correctly to the range selected by the instructor. To get a *relative* sense of the instructor's accuracy, we sorted the FCI questions by difficulty and sorted the instructor's ranking of the questions as well. Both comparisons can be gleaned from Figure 1.

To analyze instructor's performance in predicting MCI answers, we used the following rule: If the most commonly-chosen distractor (based on data at MTSU) was chosen by students *more than twice as often* as the second most-commonly chosen distractor, than instructor's needed to choose that single distractor to be considered correct. Otherwise, instructor's could pick either the most common or second most common distractor and still be considered correct in their prediction. Four questions (7, 10, 13, and 21) were dropped from analysis, because students responses were too evenly distributed across all the distractors.

RESULTS

In this section, we briefly report of some general trends in the data. First, all five instructors were able to correctly identify the MCI answer(s) for a majority of questions, but there was significant variation in their performance, ranging from 54% at the low-end to 77% at the high-end.

Second, instructors struggled for the most part to select the exact bin representing the percentage of students answering correctly. The most successful instructor selected the right bin for 43% of the questions and the least successful instructor only 23%. However, instructors were quite successful at selecting a range within one neighboring bin, from 66% on the low end and 87% on the high end. Three instructors were, on average, neither consistently pessimistic or optimistic about student performance. One instructor was consistently overly optimistic (by ~ 0.7 bin) and one was consistently overly-pessimistic (by ~ 1.0 bins).

Questions Invoking Instructor Optimism

Instructors were consistently *over-optimistic* about student performance on Q17, Q25, Q26, and Q30, overestimating student success on average by 2.2, 1.2, 1.2, and 1.4 bins respectively. Figure 1 shows specific details for Q17, displaying both absolute and relative optimism. Instructor's over-optimism is perhaps not particularly surprising, because Q17, Q25, and Q26 represent the most difficult questions, and thus easiest to overestimated student performance. It is still, however, noteworthy that instructors do not seem to have a good sense of how difficult these specific questions are for students or that these are the most difficult. In particular, each question concerns the specific student tendency to associate *force as proportional the velocity*.

FIGURE 1. Instructors were consistently *overly-optimistic* on Q17. On a relative scale, this is shown by noting that all the boxed-in rectangles (representing the instructor’s ranking) are above the horizontal line (showing Q17 as the 2nd most difficult question). On an absolute scale, Q17 also falls in the darkest shaded region (meaning that < 20% of students answered it correctly), and instructors all predicted that students would perform higher (indicated by lighter shadings).

FIGURE 2. This graphs shows the relationship between student learning gains on each question and overall instructor success at predicting the MCI answer. With exception of Q16 (which has high gains due to false positives of students considering Newton’s 1st law), questions where instructors could not collectively identify the MCI answer(s) had very low gains. In contrast, questions where instructors were able to correctly identify the MCI answer(s) had a range of normalized gains.

For contrast, instructors did correctly identify Q9 (see Figure 3) as a challenging questions. It was the most accurate estimation the group made. Taken together it seems true that student performance is low on questions where students might be tempted to give answers consistent with “force being proportional to

velocity” and instructors seem largely unaware of how low their performance actually is on these questions.

Instructor Awareness of MCI Answers

Figure 2 combines information about instructor success at predicting MCI answers with student

learning gains. The density of data points on the right reflects the fact that instructors as a group were in general successful at selecting MCI answers. The graphs depicts that, for questions where instructors could identify the MCI answers, students sometimes had high gains, sometimes had medium gains, and sometimes had low gains.

In contrast, the left side of the graph depicts the relatively small number of questions where instructors could not collectively identify the MCI answers. One such question is Q9 with small negative gains, where none of the instructors could identify the MCI answer. With the exception of Q16, which is a Newton's 3rd Law question with high gains due to false positives, all questions where instructor's could not identify the MCI answers had low gains.

The speed of the puck just after it receives the kick is:
 (A) equal to the speed " v_o " it had before it received the kick.
 (B) equal to the speed " v_k " resulting from the kick and independent of the speed " v_o ".
 (C) equal to the arithmetic sum of the speeds " v_o " and " v_k ".
 (D) smaller than either of the speeds " v_o " or " v_k ".
 (E) greater than either of the speeds " v_o " or " v_k ", but less than the arithmetic sum of these two speeds.

Figure 3. Question 9. The MCI answer is (B), representing ~70% of incorrect answers. Instructors chose (C) as the MCI answer representing ~25% of incorrect answers.

DISCUSSION

Our pilot study to use the FCI as an exploratory tool to investigate instructor awareness of student difficulties leaves us with some useful groundwork for follow up. While all findings are certainly tentative and based on a limited sample, we believe two conditions warrant early sharing: (1) the immaturity of our field's advances in developing assessments of PCK and (2) the burgeoning use of courses and professional development opportunities that focus on developing PCK. Below, we discuss our tentative findings as exemplified in two questions, Q17 and Q9.

Q17 is an example where all the instructors correctly identified MCI answers, but instructors were *overly-optimistic* about student performance by a large margin. One interpretation could be that instructors have some knowledge of student difficulties concerning "force being proportional the velocity," but they underestimate just how difficult it is for students. Q25, Q26, and Q30 show similar patterns and are concern the same student difficulty.

Q9 represents the opposite situation—a question where the instructors accurately judge the difficulty, but fail to identify the MCI answer. One interpretation is that instructors know that students struggle with these *kinds* of questions, but instructors do not know why. The low normalized gains then might be explained in the following way: If instructors do not

have specific knowledge of student thinking, then it may be less likely to create high learning gains due to limitations in how enacted instruction addresses that thinking. Alternatively, low normalized gains on those topics could be explained at a curricular level. For example, since all the instructors teach in a common and shared curriculum, the curriculum may not adequately address the topics, leaving students with few meaningful learning opportunities and instructors with few opportunities to learn about their difficulties.

Q9 also raises questions for us concerning the kinds of distractors that maybe compelling for instructors. Specifically, in Q9 instructors picked (C), which is an answer that would be right in one dimension. Examining other questions where instructor's failed to identify the MCI answer reveals a potential bias toward picking answers that are "partially correct" (e.g., identifying some of the correct forces) rather than "completely wrong" (e.g., identifying a force of motion). That is to say, instructor's may at times be more inclined to choose answers that represent incomplete understanding or misapplications of right ideas rather than answers that are fully at odds with the physics.

Our future efforts will involve broadening the use of the FCI as preliminary tool as to include a broader range of instructors, both in terms of their experience and instructional settings, with continued efforts to examine potential correlations between instructor knowledge and student learning. Finally, this ongoing preliminary work is being used to inform the development of interviews and potential survey items that meaningfully probe at instructor's knowledge of student thinking about force and motion.

ACKNOWLEDGMENTS

We thank Corey Stockburger for help on this project.

REFERENCES

1. D.L. Ball, M.H. Thames, & G. Phelps, *J. Teach. Ed.* **59**(5),389-407 (2008).
2. H.C. Hill, D.L. Ball, & S.G Schilling, *J. Res. Math. Ed.* **105**(1), 11-30 (2008).
3. J.R. Thompson, W.M. Christensen, & M.C. Wittmann, *Phys. Rev. ST-PER* **7**(1), 010108 (2011).
4. D.B. Harlow, L.H. Swanson, & V.K Otero, *Journal of Science Teacher Education*, 1-21 (2012).
5. B.W. Frank & N.M Speer, 2012 *PERC Proc.*, 126-130 (2013).
6. D. Hestenes, M.Wells, & G.Schwachamer, *Phys. Teach.* **30**, 141-158 (1992).
7. E. Mazur. *Peer Instruction: A User's Manual*. Prentice Hall, Saddle River, NJ, 1997.