Creating coupled-multiple response test items in physics and engineering for use in adaptive formative assessments

Laura Ríos*,†, Benjamin Lutz‡, Eileen Rossman†, Christina Yee*, Dominic Trageser†, Maggie Nevry†, Megan Philips†, Brian Self†

*Physics Department
California Polytechnic State University
San Luis Obispo, CA, USA
†Department of Mechanical Engineering
California Polytechnic State University
San Luis Obispo, CA, USA
‡Corresponding author: lrios02@calpoly.edu

Abstract—This Research Work-in-Progress paper presents ongoing work in engineering and physics courses to create online formative assessments. Mechanics courses are full of non-intuitive, conceptually challenging principles that are difficult to understand. In order to help students with conceptual growth, particularly if we want to develop individualized online learning modules, we first must determine student’s prior understanding. To do this, we are using coupled-multiple response (CMR) questions in introductory physics (a mechanics course), engineering statics, and engineering dynamics. CMR tests are assessments that use a nuanced rubric to examine underlying reasoning elements students may have for decisions on a multiple-choice test and, as such, are uniquely situated to pinpoint specific student alternate conceptions. In this paper, we describe our research efforts to create CMR test items for use in an adaptive learning technology for targeted intervention for a student’s unique learning needs. Our analysis was able to pinpoint some reasoning patterns that students employ to solve problems. We observed reasoning patterns that yield incorrect, partially correct, and correct answers. Our burgeoning process would aid instructors in tailoring their instruction or curricular materials for specific student needs.

Index Terms—formative assessment, coupled multiple responses, adaptive learning, statics, dynamics, mechanics, engineering education research, physics education research

I. INTRODUCTION

In introductory courses in physics and engineering, instructors are often met with persistent student misconceptions, or robust belief systems drawn from their visceral experience with physical phenomena. Because these experiences form a student’s mental model of physical phenomena, alternate conceptions of canonical mechanics are difficult to pinpoint and even more difficult to assess and address [1], [2], [3], [4].

One way to aid instructors in using class time productively (e.g., in active-learning endeavors involving peer-assisted learning) is to diagnose student alternate conceptions ahead of time via online formative learning tools, specifically adaptive technologies [5], [6], [7]. In this work-in-progress, we describe our research efforts which leverage existing online learning infrastructure (Concept Warehouse) [8]. The Concept Warehouse is an online instructional and assessment tool that helps students develop a deeper understanding of relevant concepts across a range of STEM courses and topics.

We begin by briefly reviewing student difficulties in statics and dynamics, in addition to our motivation for developing specifically coupled-multiple response (CMR) test items to tailor online learning tools. CMR items leverage a combination of student answers and explanations to identify specific ways of understanding, which can then be used to create targeted interventions to address robust misconceptions.

Because this is a work-in-progress paper, we then go on to describe the methodology of creating data-driven CMR test items to add a level of nuance in evaluating the students that is often missing from multiple-choice and online assessments. Our focus on methods is to communicate with interested researchers how our interdisciplinary team is undertaking the creation of CMR test items in different classroom contexts. We describe those contexts and provide examples from our ongoing work. We then discuss future work on validating the items. Ultimately, we expect CMR test items will guide how and when to deploy supplemental instruction or interventions that are tailored to a student’s particular reasoning.

II. BACKGROUND

A. Student difficulties in engineering and physics

Statics is often students’ first engineering mechanics course and typically poses a number of difficulties related to conceptual understanding [9]. As a result, a number of research efforts have examined both the relevant concepts within statics [10] and strategies for helping students overcome conceptual difficulties [11], [12]. One reason statics is challenging for students is the presence of threshold concepts [13], [14] and topics in the course that run counter to many students’
intuitive models of mechanics. For example, Newton’s Third Law is often noted as a threshold concept because many students to do not perceive the equal and opposite reaction forces in their experience of the world and this results in the incorrect application of forces in free body diagrams. These efforts at articulating and remediing statics concepts have been instrumental in the design of our concept questions as well as the ways we approached our analysis and coding.

Importantly, mechanics (physics) and dynamics (engineering) courses present similar types of difficulties. The literature in both physics and engineering education show that students encounter difficulties with foundational topics such as vector components [15], and also more sophisticated concepts such as rolling motion [16], which persist even with additional instruction. For example, studies with both physics and engineering students show similar challenges in constructing tension in massless strings [17]. Students tend to retain their ideas even with specific instruction on relevant topics. Altogether, the literature indicates that fostering a lasting understanding of foundational concepts remains an important aspect of instructional and research.

Finally, education researchers have shown the difficulty in mapping teaching goals onto learning assessments [2]. Thus, we propose that a learning assessment based on adaptive learning technology is well-suited for better alignment between teaching/learning goals and student assessment.

B. Adaptive learning

Adaptive learning is an educational method where computational algorithms are used to assess and deliver customized content [18]. Considering how numerous and deeply embedded student difficulties occur in introductory mechanics material, it is our hypothesis that deploying targeted interventions would improve learning gains. The interventions would be mediated by continuous formative assessments in an interactive, online format.

Previous work on adaptive tutorials (ATs) suggests that such learning experiences are generally positive and motivating. For example, [6], [7] developed ATs and showed that their students overcame common threshold concepts in 1st and 2nd year mechanics courses in engineering. The work is preliminary, but the findings point to increased course performance and positive student perceptions about the usefulness of the ATs for their conceptual understanding and engagement with the content. In a review of computerized ATs, [19] describes several instances of positive, significant gains in student performance. He also highlights the challenge of mapping student difficulties onto interactive, computerized content. With this in mind, we strive to address student misconceptions in a more nuanced way that multiple-choice online quizzes using coupled-multiple response items.

C. Coupled-multiple response tests

Coupled-multiple response (CMR) tests are assessments that use a nuanced rubric to examine underlying reasoning elements students may have for responses given on a multiple-choice test [20]. CMR assessments are time-intensive to create, but the outcome is well worth the effort; with a well-designed rubric, CMR assessments can pinpoint specific student conceptions and suggest areas for further intervention. Unlike multiple-choice questions that have one correct response and several incorrect distractors, CMR items consist of two parts: multiple choice (MC) and reasoning elements, a follow-on question that is coupled to their initial response. The MC options might indicate a general sense of alignment with canonical physics. In the reasoning element portion, students are given the option to give a justification for their initial answer. Students are given the option to select more than one reasoning element, which offers a more detailed window into their reasoning strategies. So, a student may get the right answer for the wrong reasons, which is distinct from getting the wrong answer but with proper reasoning. This type of rubric allows us to select more targeted interventions (e.g., supplemental instruction in the form of online videos) than is possible by assessing the correctness of the MC answer.

III. METHODOLOGY

To our knowledge, an explicit description of CMR item creation does not exist [21], [22]. Additionally, there is a general lack of discussion of the ways researchers have developed specific adaptive content and responses; we hope to address this gap by being explicit with both our methods and content development. In this section, we describe the analysis for three crucial courses: statics, dynamics, and mechanics.

A. Statics

We developed pilot concept questions to pose to students in class and collected qualitative responses from them via a free-response assessment. Questions were developed based off the author’s prior research in conceptual understanding and representation in various engineering mechanics contexts [23], [24], [25], [26]. In particular, this work focused on engineering statics concepts. Prior to the beginning of the quarter, the authors created concept questions across 7 content areas in statics where students are known to exhibit misconceptions. For the current paper, we offer the example of static friction.

Two objects (A and B) are different sizes, but have the same mass. If the two contact surfaces are made of the same materials, which will require a larger force to overcome friction and begin sliding? Explain your answer in 1-2 sentences.

![Static friction concept question](image)

Fig. 1. Statics concept question: static friction.

We coded the responses to this problem and 5 others according to existing literature related to student misconceptions in engineering statics [10]. In particular, we worked to identify the relevant concepts, skills, and resultant misconceptions that...
were present in students’ written explanations to a given problem.

As an example of our preliminary analysis, we presented students the problem shown in Figure 1. Students often noted that more force would be required to make the larger block slip because it had a larger contact area between the two surfaces. Such answers point to a preconception that static friction is influenced by the amount of surface area contact between two objects, when in reality the force of friction is related to (and, indeed, defined by) the normal force and the coefficient of static friction. This type of analysis is promising in helping identify specific interventions. In this case, an instructor may elect to revisit the parameters of the equation instead of focusing on restarting the lesson.

The responses were coded to capture the full range of student reasoning offered, and those reasonings will be used to populate the coupled multiple response questions developed within our adaptive learning modules.

B. Dynamics

MC questions with a free-response portion were used in a Dynamics course, as shown in Figure 2. We coded the “explain your answer” free-response portions of 70 students for common correct and incorrect reasoning. Although many students explained their answer by explicitly using the equation for polar motion, others used more conceptual reasoning in their responses. The team determined that two different reasoning questions were necessary to fully categorize the student responses – one to describe the acceleration along the slot (centripetal), and another to describe the acceleration perpendicular to the slot (Coriolis).

The circular platform rotates about the fixed pin, O, with constant angular velocity, \( \omega \), as shown in the figure below. A small box is moving in the slot towards O at a constant speed.

Future work will consist of turning the coded free-response questions to the reasoning elements in follow-on questions. The follow-on questions will be coded by the types of reasoning a student may engage to determine an answer. A preliminary, truncated result is shown below.

1) How did you determine if there was acceleration normal to the slot? Check all that apply.
   - (a) There is no acceleration in any direction because the speed along the slot is constant.
   - (f) The velocity in direction 1 changes because the radius changes

2) How did you determine if there was an acceleration in the direction of the slot? (Check all that apply)
   - (a) There is no acceleration in any direction because the speed along the slot is constant.
   - (f) The velocity in direction 1 changes direction toward 2

In the case that a student chooses (a) in both follow-up questions, we may infer that they are using reasoning related to the components of velocity in the radial and traverse direction and may have to revisit the relationship between the radial and transverse directions of acceleration with respect to the velocity vector. On the other hand, if a student selects (f) to both follow-up questions, we can infer that their reasoning includes the canonical answer. The combination of answers students give to follow-on questions 1 and 2 will inform the type of intervention the student would be directed to in the adaptive learning architecture.

C. Mechanics

The basis of the CMR items for mechanics were free response assessments used in physics studio courses based on tutorials created at the University of Washington [4] and the University of Maryland [27]. We coded 180 total responses for the example shown in Figure 3.

From the student numerical or algebraic answers, the most common correct and incorrect answers were determined. These populated the multiple choices. From the “explain your answer” and “show your work” free-response portions, we analyzed all the responses and through discussion, determined what constituted a meaningful error (e.g., equation was mis-interpreted) or a non-meaningful error (e.g., algebra mistake when rearranging a question). We cataloged common correct reasoning/work and common meaningful errors.

In ongoing work, we also catalog recurring skills or concepts that students applied to problems where those skills or concepts would not be helpful; this indicated to us that sometimes students pull from an existing “tool box” that should be addressed during class to refine when these should be applied productively. For example, the results for this
concept question show that while students often correctly identify that the radial acceleration is related to the radius of curvature of the trajectory, they sometimes conflate the radial acceleration with the magnitude of the acceleration given in the problem. This indicates that a productive targeted intervention would be centered on differentiating these two parameters. All these types of answers will populate the reasoning elements.

IV. NEXT STEPS

In addition to the ongoing work described for the three courses in the preceding sections, there are long-term next steps to take to integrate the CMR items into the existing adaptive learning architecture.

In line with recommendations for creating CMR assessment items [28], the next step on this project is to include expert validation. We will collect feedback from expert physicists and engineers for the CMR items on what other explanations/student reasoning they have observed in their own courses. We will also include distractors suggested from experts based on their instructional experience. To address the remaining challenge of creating a rubric for CMR items, the more detailed analysis of student responses will also generate an a priori codebook on what constitutes fully correct, partially correct, and incorrect reasoning patterns. After establishing interrater reliability, the rubric itself will be tested by seeing if it is statistically different from a free response grading using Student’s t-test.

Second, we will conduct think-aloud interviews with undergraduate students [29], [30]. The students will be asked to complete several of the CMR assessments and think aloud while taking them. The purpose is to make sure that students are interpreting the question correctly, that we address accessibility concerns, and that the interface is intuitive for a student to use. We do not want to conflate our measurement of student reasoning in mechanics problems with student confusion about directions or expectations. The last step is to collect data from a variety of instructional settings, instructors, and universities to conduct statistical validation tests based on classical test theory (e.g., Ferguson’s delta to measure whole-test discrimination) [31].

V. LIMITATIONS

Our data were student responses to free-response, in-person formative assessments that were both numerical and conceptual. Thus, our understanding of student resources is limited. Student resources are knowledge structures that students bring to the class to make sense of an initially confusing concept [32]. A student resources lens would offer us insight into the context of reasoning, which would aid us in developing more targeted interventions that are sensitive to the learning context. Further, education researchers in physics and engineering have advocated for productively building on student resources for lasting understanding instead of perceiving the resources as conceptions to be tossed away [33], [34]. A much more intensive, qualitative research effort, outside the scope of the present research question, would be required to make claims about student resources.

Another limitation to collecting data in the manner described above is the limited sample size in statics and dynamics. We collected approximately 70 student responses to populate the first iteration of the CMR questions in both statics and dynamics. As a result, in upcoming quarters, we will continue to administer this question and others to students to obtain a fuller range of responses.

VI. CONCLUSION

In this work-in-progress paper, we described how an interdisciplinary team of physics and engineering instructors and education researchers is using qualitative coding techniques to understand student reasoning pathways. The goal is to create targeted interventions for common difficulties in an online adaptive learning module. To achieve an adaptive learning scheme that addresses a student’s unique need, we are using coupled-multiple responses items, which offer a more nuanced look at student reasoning beyond correct versus incorrect answers on a MC question.

Our motivation is ultimately to create adaptive formative assessments to aid students in mastering important threshold concepts. Physics, statics, and dynamics are integral components of most subsequent engineering courses, and so mastering these concepts will help promote long term success in STEM. Further, the presence of these threshold concepts makes these courses “bottlenecks” within many engineering curricula, and also the point at which students are most likely to leave these programs. Therefore, developing research-based targeted interventions in these courses has the potential to increase both retention and student success.

ACKNOWLEDGMENT

We thank our colleagues Milo Koretsky, Dominic Dal Bello, Stamatis Vokos, Matthew Moelter, and John Chen for their efforts in the broader research and content development agenda. This research is supported by the California Governor’s Office of Planning and Research (http://opr.ca.gov/learninglab/).

REFERENCES


