DEVELOPMENT OF AN INTRODUCTORY PHYSICS PROBLEM-SOLVING ASSESSMENT TOOL

Timothy French and Karen Cummings
Rensselaer Polytechnic Institute, Troy, New York 12180

The physics education research group at Rensselaer is currently working to develop an assessment tool that will measure the problem-solving ability of introductory physics students. In its final form, the tool will consists of approximately 30-40 multiple-choice questions related to a limited number of classical mechanics topics. There are currently four types of questions included in the exam: attitudinal questions, quantitative problems that require students to identify the underlying principles used in solving the problem but not an explicit solution, questions that ask students to compare posed problems in terms of solution method, and quantitative problems requiring solution. Although the assessment is still under development, we have performed preliminary validation studies on questions requiring students to identify underlying principles. Specifically, both an ANOVA and a Fisher LSD test have been preformed. These evaluations showed (at the 98% and 95% confidence level, respectively) that wrong answers on assessment questions correlate to below average performance on the problem solving portion of the final course exam.

I. Introduction

Assessment tools play an important role in the use of educational research to improve physics education. Without some form of assessment, it is impossible to determine whether a change in pedagogy, curricular materials, or the use of technology has any impact on student learning or skill development. However, while much work regarding assessment has been done within the domain of conceptual learning, assessment tools which evaluate problem solving ability are not yet available.[1,2] This is despite the fact that development of problem solving ability is widely cited as an important goal in an introductory physics course. Hence, with support from the National Science Foundation, we are developing an assessment tool to measure the problem-solving ability of introductory physics students.

The exam under development is a multiple choice instrument, will ultimately contain 30-40 questions, and can be completed by students in less than an hour. While the selection of a multiple choice format makes the development of a valid assessment tool more difficult, it also allows grader-independent scoring of responses. Furthermore, because the assessment will be easy to give and grade, we increase the probability that it will be used.

Questions on the assessment tool address only problems within the regime of classical mechanics. Problem solving ability, like other skills and knowledge domains, is context dependent. Hence, success on this assessment will not necessarily be predictive of students' abilities in solving problems within other domains. Kinematics, application of Newton's second law, conservation of energy and conservation of linear momentum questions were chosen for inclusion because these topics are widely seen as being fundamental components of a first-semester undergraduate physics course.

II. Signatures of Expertise

We argue that there are several "signatures" of expertise in introductory physics problem solving. Experts, or advanced problem solvers, are not just more likely to get the correct answer to a given problem than are novices. They also possess certain characteristics and have certain attitudes toward aspects of problem solving that both facilitate their success *and* help to identify them as a group. These characteristics are the foundation of this assessment.

For example, experts are comfortable with trying an approach that is not certain to lead to a correct answer. They can ignore irrelevant information. They understand that problems can be solved in more than one way, and they readily adopt the simplest possible model of the situation (e.g. they will ignore friction when appropriate even if "ignore friction" is not explicitly stated in the problem). Experts reflect on and evaluate the reasonableness of their work at many points along the solution path. They have a rich set of actions they know that they can take if they cannot find a solution to the problem.

Additionally, expert problem solvers have (and know they have) a set procedure by which they solve certain types of problems. For example, they have a set procedure by which they approach the solution of a "conservation of energy" problem. Typically, novice problem solvers either do not have such a procedure or they fail to realize and exploit the fact that they do. We conjecture that this difference is the reason that experts categorize problems (i.e. group them as "similar") based on the conceptual, or "deep", features[3] of the problems (e.g. application of conservation of energy), while novices categorize based on "surface" features (e.g. problems containing rolling balls).

III Question Types

This assessment tool includes four types of questions. Examples of three types are shown in Figs. 1-3.

The first question type (Fig. 1) is designed to probe students' attitudes toward, and common behaviors during, the solution of introductory physics problems.

Figure 1: An example of a "Type 1" question that is used to probe student attitudes towards, and behaviors during, problem solving.

Imagine that you have just started to solve a physics problem. How likely are you to do the following? (Definitely, very likely, possibly, probably not, not at all likely).

1)Draw, redraw, or visualize a picture or graph
2)Make a list or table
3)Look for similar problems in the textbook and use them as a model
4)Look up an appropriate equation
5) Think about which physics concepts are important for the problem
6) Start with the solution (from the back of the book) and work backward
7)Make some assumptions about the situation
8) Ask yourself questions about what is going on in the problem and problem solution.

The second type of problem deals with the identification of the underlying principles used to solve a given problem (Fig. 2). Students are not required to find a numerical answer to these problems and are told not to solve them. This problem type provides us with information regarding a student's ability to identify nuances within the problem statement that clue experts into likely solution approaches.

Not requiring a complete solution to every question allows us to ask about a wider range of problems without making the test too time consuming. Furthermore, we assume that identifying a reasonable solution approach is a critical first step in the problem solution. Of course, even if a potentially productive approach is chosen, the correct solution of the problem is by no means guaranteed.

Figure 2: An example of a "Type 2" question that asks students to identify the best approach to use in solving the given problem.

Consider the following problem:
A ball is knocked off the edge of a cliff at 3 m/s and strikes the ground below. The cliff is 20 meters high. What is the velocity of the ball when it strikes the ground? Which of the following approaches would you use in solution of this problem.

A) Newton's Second Law (F=ma)
B) Either kinematics or conservation of momentum
C) Kinematics
D) Kinematics and then conservation of momentum
E) Conservation of momentum
F) None of the above

In the third type of question, two physics problems are presented. Some of the problems are standard "textbook" style questions, while others are "context-rich" problems gathered from the work of Heller and Heller[6]. As with the previous question type, students are not to find a numerical answer to the posed problems. Instead, they are told to determine whether or not the two problems are solved in a "similar" manner. This question type, shown in Fig. 3, focuses on the test-taker's problem categorization schemes. (Recall that experts categorize based on "deep" conceptual features)[4].

Figure 3: An example of a "Type 3" question used to probe student categorization schemes.

Consider the following two problems:
PROBLEM A) A 30 g wooden block rests on a frictionless surface. A bullet with mass 5 g, traveling with a speed of 275 m/s, passes through the block. The bullet's speed after passing through the block is 240 m/s. How fast is the block moving after the bullet leaves the block?
PROBLEM B) A 50 g lump of clay slides along a frictionless surface until it hits a 3 kg block. The block and clay, now united and moving together, have a final velocity of 0.5 m/s. What was the speed of the clay before the collision?

Are the problems solved in a similar manner?
A) Yes
B) No
C) I'm not sure.

The fourth question type (Not shown in a figure) requires the student to actually solve a quantitative problem. These problems probe whether a student can successfully pull all the required skills and abilities together. The distracters for this type of problem are generated through student interviews and by gathering data from free response problems given on exams in the introductory physics courses at Rensselaer.

IV Scoring the Assessment

Evaluation of student answers on the assessment is done through a comparison to the answers typically given by "expert" introductory physics problem solvers. Note that truly (i.e. non-subjective) correct answers only exist for the fourth type of question discussed above. For all the other question types, student responses are considered "correct" to the extent that they match the answer most commonly given by experts (e.g. professors, upper-level graduate students, practicing scientists). For example, if a student groups problems (in a Type 3 question) based on deep structure rather than surface features, this mirrors the thinking associated with an expert and so is judged to be "correct".

By giving the problem-solving assessment both pre- and post-instruction, we will ultimately be able to measure gains in problem solving ability (i.e. student motion towards the "expert" end of the continuum).

V Results of a Preliminary Validation Study

We have now completed a preliminary validation study on a subset of the Type 2 questions (i.e. "identify the approach used to solve" questions). Three questions of this type were included on the multiple choice portion of the Physics I final for the Spring 2001 semester at Rensselaer.[7] In order to test whether a student's score on these questions was positively correlated with their score on the free response, quantitative problem solving portion of the exam, we looked at a subset of 225 finals. (The questions on this portion of the exam are textbook, "show all your work to receive credit" style questions.)

The 225 exams were categorized based upon the Type 2 question score as follows: 0 wrong (N_0=148), 1 wrong (N_1=64), and 2 wrong (N_2=13). (No one got all 3 wrong). The average score out of 65 total points on the free response problem solving section for each group was 48.82, 43.46, and 39.54, respectively. Performing an ANOVA shows this to be a statistically significant correlation at the 98% confidence level.

In addition, a Fisher least significant difference (LSD) test was used to compare the three groups as pairs. This evaluation also showed (at a 95% confidence level) that incorrectly answering any one (or more) of the three Type 2 questions was predictive of a lower score on the free response, problem solving section of the exam. ("Lower" as used here is in comparison to the scores of students who answered all of the Type 2 questions correctly.)

While we still have a significant amount of work to do before this assessment tool is ready for dissemination, we feel the results of the ANOVA and Fisher LSD tests cited above indicate we are headed in a productive direction.

---

[1] D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," Phys. Teach. **30** (3), 141-158 (1992).

[2] R. Thornton and D. Sokoloff, "Assessing student learning of Newton's Laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula," Am. J. Phys. **66** (4), 338-352 (1998).

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

[6] K. Heller and P. Heller, *Cooperative Group Problem Solving in Physics*, University of Minnesota (1997).

[4] P.T. Hardiman, R. Dufresne, J. Mestre, "The relation between problem categorization and problem solving among experts and novices," Mem. & Cog. **17** (5), 627-638 (1989).

[7] K. Cummings, D. Kuhl, J. Marx, R. Thornton, "Evaluating innovation in studio physics," Phys. Educ. Res., Am. J. Phys. Suppl. **67** (7), S38-S44, (1999).