# Developing an Energy Assessment for Elementary Education Majors

Tom Foster and Daniel Barnett

*Department of Physics, Southern Illinois University Edwardsville*
*Edwardsville, IL, 62026-1654*

**Abstract.** In support of an NSF-CCLI program, we developed a multiple-choice efficacy assessment for the energy concept. What makes this work novel amongst the sea of energy concept assessments is the intended audience: elementary and early childhood education majors. While these are smart and capable college students, their demographics require a different assessment than our engineering students. We will discuss the development of the assessment and our preliminary results.

**Keywords:** Energy Concept, Assessment Development, Elementary Education Pre-service Teachers
**PACS:** 01.40.Fk, 01.40.gf, 01.40.jc

## I: INTRODUCTION

Energy is the most fundamental concept and it underpins every science. How the concept is used and perceived varies by discipline, but energy is central. In physics, we have defined types of energy using the language of mathematics and hold the conservation of energy as one of our most sacred laws.

Ideally even young children should understand some of the basic concepts of energy and energy conservation. To achieve that goal, we need to make sure that future k-6 teachers also understand energy well enough to teach it. As college instructors in the post-FCI era, we understand that we need to assess our courses to ensure our students are learning.

## II: WHY ANOTHER ENERGY ASSESSMENT?

There are many energy assessments available today. Perhaps the most familiar to PERC readers would be Singh's Energy Concepts Survey [1] or Swackhammer's Energy Concept Inventory [2]. We also uncovered the ACER Physics Unit tests [3] from 1980. Each of these assessments generally looks at the energy concept through the traditional curriculum lens: every introductory physics course is a watered-down version of the calculus-based University Physics course. Therefore, a well-constructed conceptual test should be accessible to all physics students regardless of mathematical level.

However, these tests have limits and it is not mathematical sophistication. It is that these tests are imbedded in the culture of physics. They have face-validity to physics teachers, because they look like good physics questions. Furthermore, to allow for statistical analysis beyond averages, a concept test should have many questions on the same topic. While psychometrically an important principle, long exams might intimidate students who are already science-phobic. Finally, the coverage of an energy assessment does presume a standard, complete treatment of the energy concept. What if only a portion of the energy concept was taught? Students would get many questions wrong on the concept tests because they were not exposed to these detail of energy.

In short, every test has explicit and implicit goals when it is created. We felt that our setting did not match the goals of existing energy assessments.

### Foundations in Science

Since 1985, Southern Illinois University Edwardsville has been teaching various versions of Foundations in Science. This course began when faculty members from education and the science realized that elementary education majors needed an integrated science course tailored to their needs and career aspirations. In its current form, the course is a two-semester sequence with physics, chemistry, and a bit of astronomy as the content emphasis of the first semester and biology and geoscience in the second semester. One instructor (without additional support) teaches a semester of the course.

When possible, students learn the science by engaging with activities that follow the 5E [4] cycle. The activities are intended to be similar to activities that elementary students might do, so expensive equipment and "black boxes" are avoided in favor of cheaper, home-constructed devices.

The Illinois State Learning Goals [5] determines the specific content of the course. We teach only those topics which are identified by the state standards and descriptors for k-6 students. The energy concept appears in both physics and chemistry contexts allowing energy to occupy several days of instruction and to be a touchstone when discussing other topics. Therefore, testing students understanding of energy as presented in this limiting context, was our goal.

## Efficacy Versus Concept Inventory

Given the integrated nature of the course and the characteristics of our students, the authors decided to create an efficacy assessment and not a concept inventory. As Lindell [6] distinguishes assessments created for research purposes, we are not creating an exam to probe deeply the students understanding of energy (concept inventory). Our task was instead to develop a test to document the students' learning about energy during the course (efficacy test).

## Characteristics of Elementary Education Majors as Learners

While every course is populated with students that offer unique challenges to the instructor, teaching a course for elementary education majors is very different. Anecdotally, our Foundations course is "different" enough that during Tom Foster's sabbatical, none of the physics faculty members at SIUE wanted to teach the course, even with scheduling enticements. The Foundations is a very different course populated with unusual (by physics standards) students.

Elementary education majors are not dumb, but rather they face an incredible intellectual challenge. They need to master *all* the disciplines so they can teach each discipline to young learners (and pass the Illinois Content Test for licensure). They are trained to be Jacks-of-all-trades. When in-service elementary teachers are asked what they teach, the answer is nearly always "children," and not a specific subject [7]. Therefore, these learners' knowledge in any one content area is necessarily broad and thin, so they might come across as under-educated, but the opposite is true.

In spite of the need for a diverse background, we know that the elementary education majors at SIUE are generally neutral or negative about science. At the start of each Foundations course we ask the students to self-report their "affection" for science. We get very few "love science" comments. Rather, many of the students report a negative experience with science somewhere in their previous education. Therefore, when teaching or assessing these students it is important not to activate [8] these negative feelings.

An obvious, but important observation about the students in the Foundation Course is that the vast majority (>85%) of them are female. While it is not unusual for the population of the College Physics course for pre-meds to be mostly females, the pre-service elementary teachers tend to lack the same grade obsession that stereotypes the pre-med students. Couple that with the aforementioned attitudes toward science in the Foundations course and we conclude that not all groups of women in physics classes are the same.

Finally, it appears that most of the students in the Foundation course are in the *concrete* Piagetian level. While we have never tested this explicitly, we did notice a marked improvement in student understanding of conservation after adding a version of a Piagetian water conservation activity [9]. Therefore we have paid special attention to abstract concepts (like graphs) with this population of students, both in instruction and in assessment design.

# III: ENERGY ASSESSMENT DEVELOPMENT AND RESULTS

We began with the following goals for our energy assessment:

- Do not reinvent the wheel – borrow questions from previous assessments and inventories.
- The questions should not all be mathematical
- The objects in the questions should be concrete and gender neutral (e.g. no guns)
- The overall assessment should be short
- The assessment should include some graph interpretation.

What we created was a 15 item assessment which appear to be tied to the energy concepts taught in the course. Sometimes items were selected for inclusion on the assessment because they mirrored activities in class (i.e. energy bar charts) and other times item were selected because they pushed the students outside their comfort zones (i.e. ratios). We wanted a range of questions reflecting depth and breadth. Each scenario presented generally has several questions related to it providing the illusion of a shorter test. In addition, for the initial versions, we left spaces on the exams for the students to answer two questions to guide our development. Question 1 asked "what factors lead them to the answer they selected?" and question 2 asked "was the wording of this question unclear?" These two questions were at the end of each scenario inside a visually distinctive box.

The first version and associated paperwork was submitted to SIUE's IRB and the project was deemed exempt from further review by the IRB.

## Pilot Testing the Assessment

The development of the energy test was completed over the course of one semester using two different sections of the Foundations course. The test was modified based on the first section results prior to giving the test to the second section. Each section had about twenty-four students. It was administered to the students several weeks after they had studied energy. We used time in the classroom at the end of quicker lessons to administer the test. Participation was voluntary and the students were given a small amount of extra credit for their participation. We had a few students in each section choose not to participate.

The topics of each question are shown in table 1. Test items are grouped by scenario. Primary Concept is the concept the question was written to assess, while the Secondary Concept, when identified, represents other skills or ideas the students might need to succeed on the item.

**TABLE 1.** Items by concept.

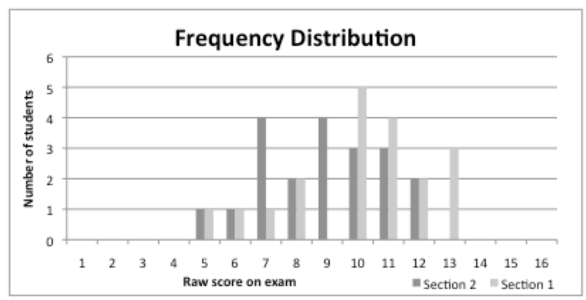
| Question Number | Primary Concept | Secondary Concept |
|---|---|---|
| Ball on Hill | | |
| Item 1 | Energy conservation | Thermal Energy |
| Two cars moving | | |
| Item 2 | KE defined | Ratios |
| Item 3 | Energy conservation | Ratios |
| Hammer on Moon | | |
| Item 4 | Mass defined | Value of "g" |
| Item 5 | GPE defined | Value of "g" |
| Child on Bike | | |
| Item 6 | Graph reading | KE used |
| Item 7 | Graph reading | KE used |
| Roller Coaster | | |
| Item 8 | GPE used | |
| Item 9 | GPE used | |
| Item 10 | Mass defined | |
| Ball dropped | | |
| Item 11 | Graph reading | Energy conservation |
| Item 12 | Graph reading | GPE used |
| Ball Bouncing | | |
| Item 13 | Bar Chart reading | Energy conservation |
| Item 14 | Energy conservation | Bar Chart reading |
| Item 15 | Energy conservation | Bar Chart reading |

With 15 questions and 20 students per trial, the authors are hesitant to use complex statistics, such as Item Response Theory or Factor Analysis, in analyzing the results of each revision of the test. Instead we relied heavily upon student feedback about the test items and student performance on each item and scenarios.

## Numerical Results

Part of the heartbreak of designing assessment for classes that one is teaching is the low scores. On the first cycle, the average on the test was 9.0 out of 15 (60% correct, N=19) and on the second test the average was 7.9 out of 15 questions (52%). The lower score by the second section may be due to attrition of knowledge about energy. The second section took the exam a month after the first section did. The reassuring side is that the test produces a good distribution of scores (see Figure 1).

**FIGURE 1.** Distribution of test scores for second section on energy exam.




The Cronbach Alpha was calculated for the test to measure the reliability of the instrument. The Cronbach Alpha of 0.6 for the first version and 0.45 for the second version. Since there are only 15 items and 20 students, interpreting these results is difficult; however, Cronbach Alphas greater than 0.8 are preferred.

## Behind the Numbers

It is clear from the item analysis (Table 2) that the students do not know the mathematical definitions of kinetic energy (KE) or gravitational potential energy (GPE). Both Items 2 and 5 had low success rates (Item difficulty is high), but they could generally use the principles (Items 6, 7, 8, 9, and 12) as types of energy.

There was much debate about the hammer on the moon question. The context is unfamiliar to the students (all of them are post-Apollo era) and the changing value of "g" might have been confusing. However, it is noteworthy that the students could locate the value of the mass given in the problem statement for the roller coaster (Item 10), but not for the moon context (Item 4).

**TABLE 2.** Percent of section 2 getting an item correct.

| Question Number | Primary Concept | Percent correct (N=20) |
|---|---|---|
| Ball on Hill | | |
| Item 1 | Energy conservation | 85 ± 8 |
| Two cars moving | | |
| Item 2 | KE defined | 0.0 |
| Item 3 | Energy conservation | 5 ± 5 |
| Hammer on Moon | | |
| Item 4 | Mass defined | 10 ± 6 |
| Item 5 | GPE defined | 30 ± 10 |
| Child on Bike | | |
| Item 6 | Graph reading | 85 ± 6 |
| Item 7 | Graph reading | 25 ± 10 |
| Roller Coaster | | |
| Item 8 | GPE used | 90 ± 7 |
| Item 9 | GPE used | 75 ± 10 |
| Item 10 | Mass defined | 85 ± 8 |
| Ball dropped | | |
| Item 11 | Graph reading | 85 ± 8 |
| Item 12 | Graph reading | 60 ± 11 |
| Ball Bouncing | | |
| Item 13 | Bar Chart reading | 80 ± 9 |
| Item 14 | Energy conservation | 25 ± 10 |
| Item 15 | Energy conservation | 50 ± 10 |

For the two cars moving, the item difficulties for Item 2 and 3 would suggest removing the question. However, if Item 3 is scored by checking for consistency between the item answers (in other words, did the students use the conservation of energy to answer Item 3 based upon their answer to Item 2) then nearly every student gets the item correct (80%)

Items 14 and 15 also have low score and it is unclear if the students are confused by the energy concept or the energy bar chart representation or even bar charts in general.

Finally, a holistic view of table 2 reveals that there are essentially two items types: those in which the students score well (>80%) and those in which the students do poorly (<25%). There are only about 25% items in between. When we look at the Item Discrimination we see only one item (7) that the high scoring students generally got wrong, but that correlation is weak (R = –0.20). With small numbers it is difficult to interpret this results as well.

# IV: FUTURE WORK

The next step is to use this energy efficacy test for what it was designed for - pre and post testing. Ideally we will put the test in electronic form and assign it as homework. Even while the exam is being taken, it might be interesting to probe the students with a survey or interviews specifically about the moon context and their confidence with using energy bar charts. Finally, we would be interested in other pre-service courses for elementary education majors using test to see if results are generalizable and allow us to increase the sample size. Contact Tom Foster at tfoster@siue.edu for an electronic version of the test.

# ACKNOWLEDGMENTS

The authors would like to thank the NSF for partially funding this work under the CCLI program, grant DUE-0837417. In addition, this work was supported by SIUE under the Undergraduate Research Associate program. Finally, we'd like to thank the anonymous reviewers for their insightful comments.

# REFERENCES

1. C. Singh and D. Rosengrant, *Am. J. Phys. 71, 607 (2003)*.
2. G. Swackhamer, & D. Hestenes, An energy concept inventory. Unpublished manuscript, Arizona State University. (2005).
3. Australian Council for Educational Research, Hawthorn. 1980. ACER Physics Unit Tests: Unit Tests, Diagnostic Aids, [and] Teachers Handbook. (1980). ERIC, EBSCOhost (accessed July 7, 2011)
4. R. Bybee,. Teaching secondary school science: strategies for developing scientific literacy. Upper Saddle River, N.J: Pearson/Merrill/Prentice Hall, (2008).
5. http://www.isbe.state.il.us/ILS/ (accessed July 7, 2011)
6 R. Lindell, E. Peak, T. Foster. "Are they all created equal? A comparison of different concept inventory development methodologies." Invited paper; Proceedings of the 2006 Physics Education Research Conference, ed. By P.Heron, L Hsu and L. McCullough. (2006)
7 K. Appleton Elementary Science Teaching. In Abell, S., & Lederman, N. (Ed.) Handbook of Research on Science Education (pp. 493-536). Hillsdale: Lawrence Erlbaum. (2007).
8 C.M. Steele, J. Aronson. *Journal of Personality and Social Psychology* 69 (5): 797–811. (1995).
9 http://www.devpsy.org/teaching/theory/piaget2info_processing.html (accessed July 7, 2011)