Using Content-Aligned Assessment to Probe Middle School Students' Understanding of Ideas About Energy

Cari F. Herrmann-Abell and George E. DeBoer
AAAS Project 2061

NARST Annual Conference
Garden Grove, CA

April 17-21, 2009

We report the results of a pilot test of assessment items aligned to the middle school topic focused on forms of energy that was administered to 1728 sixth, seventh, and eighth grade students from 11 widely varying school districts across the country in the spring of 2008. This paper presents data from assessment items aligned to key ideas for motion energy, thermal energy, gravitational energy, and elastic energy, and describes how we use information gathered from the students to gain insight into students' thinking as well as insights into the quality of the items themselves. This work is part of a larger project funded by the National Science Foundation to develop items that are precisely aligned with national content standards. Each item is developed using a procedure designed to evaluate an item's match to important science ideas and its overall effectiveness as an accurate measure of what students do and do not know about those ideas.

Introduction

Now that states have begun the widespread testing in science mandated by the *No Child Left Behind* legislation (2002), it is more imperative than ever that the time spent in testing and preparing for tests be well used. The results of testing should be able to be used to improve science instruction, which means that the tests must be able to reveal what students do and do not know and the alternative ideas they have. Even before the state-mandated testing began, concerns were growing about the quality of science assessments and their alignment to state and national content standards (American Federation of Teachers, 2006). In response to these concerns, in 2004 we began to develop student assessment items in middle school science that are precisely aligned with established content standards, including those in *Benchmarks for Science Literacy* (AAAS, 1993) and the *National Science Education Standards* (NRC, 1996), and Web-based resources that support and promote the development and use of science assessment items aligned to science content standards. Each item is developed and analyzed using a procedure for evaluating an item's match to important science ideas and its overall effectiveness as an accurate measure of what students do and do not know about those ideas (see DeBoer et al., 2007, 2008a, 2008b for a more detailed discussion of the procedure). During item development, pilot testing is used to obtain feedback from students about the items. Then scientists and science education experts review the items using a set of criteria to ensure content alignment and construct validity. After revisions are made based on the reviews, the items are field tested on a large national sample to determine the psychometric properties of the items and clusters of items.

This paper describes results from the pilot testing of a set of assessment items aligned to the middle school topic of energy. It includes data on how well students performed on items that were designed to test each of the key ideas for this topic and a description of how pilot testing was used to gain insight into students' thinking about those ideas and insights into the quality of the items themselves.

Methodology

The pilot test reported on here included 1728 middle school students from 11 widely varying school districts across the country and was administered in the spring of 2008. Approximately 33% the students were in 6^{th} grade, 31% in 7^{th} grade, and 36% in 8^{th} grade. Approximately 6% of the students indicated that English was not their primary language and approximately half of the students were female and half were male (see Table 1).

Table 1
Demographic information

	Total	Female	Male	English	Non-English
Grade	% (N)	% (N)	% (N)	% (N)	% (N)
6th Grade	32.7% (565)	49.9% (282)	48.8% (276)	94.0% (531)	3.9% (22)
7th Grade	31.1% (538)	48.9% (263)	49.6% (267)	92.0% (495)	5.8% (31)
8th Grade	36.2% (625)	48.3% (302)	50.0% (311)	90.6% (566)	7.2% (45)
Total	100% (1728)	49.0% (847)	49.4% (854)	92.1% (1592)	5.7% (98)

During pilot testing, students responded in writing to the following questions about each of the items:

1. Where have you learned about this? School, Museum, Family, TV, Other, Never learned this
2. Circle any words on the test question you don't understand or aren't familiar with.
3. Is there anything about this test question that is confusing? Explain.
4. If there is a picture or table, is it helpful? If there no picture or table, would one be helpful?
5. Is answer choice A a correct answer to the test question? Yes, No, Not Sure
   Explain why or why not, or why you are unsure.
6. Is answer choice B a correct answer to the test question? Yes, No, Not Sure
   Explain why or why not, or why you are unsure.
7. Is answer choice C a correct answer to the test question? Yes, No, Not Sure
   Explain why or why not, or why you are unsure.
8. Is answer choice D a correct answer to the test question? Yes, No, Not Sure
   Explain why or why not, or why you are unsure.
9. Did you guess when picked your correct answer? Yes, No
10. Please suggest additional answer choices that could be used.

Examining the students' written responses to these questions enables us to determine: (1) whether students used the targeted learning goal to answer the question or if they used other knowledge instead, (2) whether the item was comprehensible to them, and (3) whether the answer choices were plausible to them. Mismatches between the answer choices that students select and the reasons they give provide us with information about whether the item is likely to yield false positive and false negative results. Written responses also provide information regarding misconceptions students may have.

During pilot testing, students answered questions related to five key ideas on the topic of energy. These five key ideas focus on the nature of four different forms or manifestations of energy. Ideas dealing with energy transformations, transfer, and conservation are dealt with in a separate topic. Those results will be reported at another time. Each student received four or five assessment items aligned to each of two of the targeted key ideas, for a total of eight to ten items. Two of the items on each test form were linking items common to each form. This enabled us to use Rasch modeling to compare item characteristics across forms. Each item was answered by approximately 100 students.

The ideas on which students were tested are based on Chapter 4, Section E of *Benchmarks for Science Literacy* (AAAS, 1993). The key ideas are:

Idea A: Motion energy (kinetic energy) is associated with the speed and the mass of an object.
Idea B: Thermal energy is associated with the temperature and the mass of an object and the material of which the object is made.
Idea C: Thermal energy of an object is associated with the disordered motions of its atoms or molecules and the number and types of atoms or molecules of which the object is made.
Idea D: Gravitational potential energy is associated with the distance an object is above a reference point, such as the center of the earth, and the mass of the object.
Idea E: Elastic potential energy is associated with the stretching or compressing of an elastic object and how easily the object can be stretched or compressed.

Each key idea was further clarified in order to state precisely what students would be expected to know (sub-ideas) and what they would not be expected to know (boundary statements). These clarification statements act as item writing specifications that ensure a close alignment between the items and the learning goals. For example, the clarification statement for Idea A includes the following sub-ideas and boundaries:

<u>Sub-ideas:</u>
1. Students should know that the motion energy of an object depends on both the speed and the mass of the object, and that motion energy depends *only* on these two factors. They should know that motion energy does not depend on other factors such as size, shape, material the object is made of, or direction of motion.
2. Students should know that any object that is moving has motion energy (kinetic energy). They should also know that the motion energy of an object that is not moving is zero.
3. They should know that objects that have the same mass and are traveling at the same speed have the same amount of motion energy.

4. They should know that increasing an object's speed will increase the motion energy of the object (regardless of how much the speed is increased) and decreasing an object's speed will decrease the motion energy of the object (assuming the mass of the object does not change).
5. Students should know that when the motion energy of an object increases, the speed of the object increases and that when the motion energy of an object decreases, the speed of the object decreases (assuming the mass of the object does not change).
6. They should know that, for objects that have the same mass, the object with the greatest speed will have the greatest motion energy and the object with the lowest speed will have the least motion energy.
7. They should know that, for objects that have the same mass, the object with the greatest motion energy has the greatest speed and the object with the least motion energy has the lowest speed.
8. Students should know that, for objects traveling at the same speed (greater than zero), the object with the greatest mass will have the greatest motion energy. They should also know that the object with the least mass will have the least motion energy.
9. Students should know that, for objects traveling at the same speed (greater than zero), the object with the greatest motion energy has the greatest mass and the object with the least motion energy has the least mass.
10. Students should know that, for objects traveling with the same amount of kinetic energy, the object with the greatest mass will have the lowest speed and the object with the least mass will have the greatest speed.
11. Students should know that, for objects traveling with the same amount of kinetic energy, the object with the greatest speed will have the least mass and the object with the lowest speed will have the greatest mass.

Boundaries:

1. Students are not expected to know or use the formula $½mv^2$. The sub-ideas above describe qualitative relationships.
2. Students are not expected to compare situations where both the mass and speed vary. In assessment items, either the mass or the speed will be held constant while the other varies so that both variables will not be changed at the same time. However, the case where one object is moving and the other is not (regardless of their masses) is valid.
3. The students are not expected to know the difference between "weight" and "mass." All of the contexts used in the assessment items will be ones where "mass" and "weight" are proportional to each other. When two objects are being compared, they will be in the same location.
4. This idea refers to motion with respect to the surface of the earth. An object is not moving if its position with respect to a point on the surface of the earth is not changing. Students are not expected to know Benchmark10A/M1*: Because every object is moving relative to some other object, no object has a unique claim to be at rest. Therefore, the idea of absolute motion or rest is misleading.
5. This idea is limited to translational kinetic energy. Students are not expected to know about other forms of kinetic energy such as vibrational kinetic energy and rotational kinetic energy.

We used Rasch modeling to analyze the pilot test data (Linacre, 2006). Rasch modeling was used to determine if the range of item difficulty was appropriate for the middle school students who were sampled, the extent to which each of the items was correlated with the entire set of items (point-measure correlation), and if the pattern of responses followed expectations such that the highest performing students on the entire set of items were more likely to select the correct answer for an item and the lowest performing students were less likely to select the correct answer. Any discrepancies prompted us to examine the items more carefully to determine the cause of the discrepancies.

Findings

The data had an adequate fit to the Rasch Model, suggesting a one-dimensional set of items all focused on the same basic theme. The results showed a separation index for the items of 3.98, which corresponds to a test reliability of 0.94. (The separation index represents the spread of the abilities or difficulties and indicates the approximate number of different levels of difficulty or ability. A separation index greater than 1.5 is considered acceptable.) The person separation index was 0.38, which corresponds to a person reliability of 0.13. The low person separation index and person reliability is due to the fact that each student responded to only 8 or 10 of the 79 items tested, so that there were not enough items on any given test form to reliably test the ability of the students at each ability level. In addition, the students were given the opportunity to say they were “not sure” about an item rather than to select one of the answer choices. This meant that some students may have known the correct answer to a question but selected “not sure” because they were uncertain about their knowledge. The “not sure” option could have added to the unreliability of the students’ answers. Students who responded “not sure” were marked incorrect. The point-measure correlation coefficients for the items ranged from 0.24 to 0.72 with a mean of 0.51. None of the items with low point-measure correlation coefficients had features that suggested that they were inappropriate for inclusion in this set of items.

Figure 1 shows the item-person map for the 79 items included on the pilot tests. The map shows the range of person abilities on the left side of a vertical line and item difficulties on the right side of the line. Low ability/difficulty is represented at the bottom of the map and high ability/difficulty at the top. As can be seen on the map, the mean item difficulty is slightly greater that the mean student ability, and there are items to measure the full range of student ability, including both ends of the student ability spectrum as well as the middle.

Figure 1: Item-person map showing the distribution of student abilities on the left and item difficulties on the right. When item difficulty and person ability match, the person has a 50% chance of answering the item correctly. Item difficulties are shown for the 79 items included on the pilot tests.

```
          PERSONS - MAP - ITEMS
               <more>|<rare>                          Idea A: Motion Energy (Kinetic energy)
 4              .##  +                                Idea B: Thermal Energy (substance level)
                     |                                Idea C: Thermal Energy (atomic level)
                     |                                Idea D: Gravitational Energy
                     |                                Idea E: Elastic Energy
                     |
                     |
                     |
                  .  |
 3                .  +
                     |
                     |
                  .  |  EG14-2
                  .  |  EG22-2
                  .  |
                  .  |  EG34-2
                  .  |T
 2                .  +
                  . T|  EG25-1
                  .  |
                  .  |  EG21-2  EG72-1
                 .#  |  EG32-2  EG55-3  EG6-4
                  .  |
                  .  |  EG45-2  EG59-1
                  #  |S EG24-1  EG56-3  EG60-1  EG67-1  EG72-2
 1                .  +  EG57-3
                ###  |  EG58-1
                 .# S|  EG27-2  EG66-2  EG70-1  EG9-2
                .##  |  EG50-5  EG74-1
                .##  |  EG47-2  EG52-3  EG61-2
                 .#  |  EG48-2  EG69-1
                .##  |  EG36-3  EG71-1
                .##  |  EG16-2  EG42-2
 0              .##  +M EG13-3  EG2-1   EG65-1
               .###  |  EG40-2  EG49-2  EG51-2  EG62-1  EG64-1
                 .#  |  EG20-2  EG35-3
               .### M|  EG11-1  EG17-2  EG17-3  EG29-2
                .##  |  EG1-4   EG19-2  EG19-3  EG3-1   EG31-2  EG38-2  EG53-2
                 .#  |  EG19-4  EG23-1  EG33-2  EG43-3  EG53-3
               .###  |  EG10-2  EG18-2  EG18-3
                 .#  |  EG36-2  EG37-2  EG54-2
-1             .###  +  EG15-2  EG41-2  EG7-2
                 .#  |S EG46-2  EG5-2
                 .#  |  EG30-2  EG44-2  EG52-2
                .## S|  EG26-2
                .##  |
                .##  |  EG28-2
                 .#  |  EG8-2
                  .  |
-2               .#  +  EG63-1
                  .  |T EG4-3
                  .  |
                  .  |
                  . T|  EG12-1
                  .  |
                  .  |                                M = mean ability/difficulty
                     |                                S = 1 standard deviation away from mean
-3 #############     +                                T = 2 standard deviations away from mean
               <less>|<frequ>                         Each '#' is 22 students.
```

The overall percentage of correct responses for the topic was 39.0%. Students had the most difficulty with ideas related to thermal energy and gravitational energy (Ideas B, C, and D), and they were most successful with questions testing the ideas of motion energy and elastic energy (Ideas A and E). The percent correct was 41.4% for items aligned to Idea A, 35.4% for Idea B, 31.3% for Idea C, 44.4% for Idea D, and 39.0% for Idea E. Table 1 shows the average Rasch item difficulty measure for each idea.

Table 2
Average Rasch Item Difficulty for the Energy Ideas

	Idea A	Idea B	Idea C	Idea D	Idea E
Average Difficulty	-0.41	0.00	0.20	0.24	-0.17

We also found that there was no significant difference in students' knowledge from grade to grade (see Table 3). This could be because different schools teach these ideas in different grades or not at all.

Table 3
Overall Percent Correct by Grade

Grade	6th Grade	7th Grade	8th Grade
% correct (N)	39.5% (565)	40.0% (538)	37.8% (624)

What we learned from pilot testing:

Idea A: Motion Energy
The items aligned to Idea A, that motion energy (kinetic energy) is associated with the speed and the mass of an object, tested the students' understanding of motion energy in a variety of ways. For example, in some items students were asked to determine the motion energy from information about the speed and the mass of the objects, and in other items the students were asked to determine the speed from information about the motion energy and the mass. Items were more or less difficult for students depending on how they were expected to reason with the knowledge. We found that students' success on these items varied from 11.9% correct to 72.4% correct (see Table 4).

Table 4
The results for the items aligned to the motion energy sub-ideas ordered from easiest to most difficult

Item Context (# of items in parentheses)	% Correct	Rasch Difficulty
Objects that are moving have motion energy, and objects that are not moving do not (2)	72.4%	-2.13
Direct statement of the relationship: motion energy depends on both the speed and the mass of the object (1)	61.9%	-1.01
Compare the speeds of two moving objects when given the amounts of motion energy; weight is held constant (1)	64.0%	-0.75
Compare the amounts of motion energy of two moving objects when given the speeds; weight is held constant (2)	43.6%	-0.41
Compare the amounts of motion energy of two falling acorns when given the weights; speed is held constant (1)	45.3%	-0.04
Compare the weights of two moving objects when given the amounts of motion energy; speed is held constant (1)	25.9%	0.71
The weights of two objects cannot be compared from the amounts of motion energy alone. You also need to know the objects' speeds. (1)	28.4%	1.16
The speeds of two objects cannot be compared from the amounts of motion energy alone. You also need to know the objects' weights/masses. (1)	11.9%	1.83

The easiest items were (1) ones that required the students to recognize that objects that are moving have motion energy and objects that are not moving do not have motion energy, and (2) ones that asked students to recognize the truth of a direct statement that the motion energy of an object depends on both the speed and mass of the object. It was more difficult for students to compare the speeds of two objects based on differences in amounts of motion energy but constant weight, and to compare the amounts of motion energy based on differences in speed but constant weight. More difficult yet was an item that required students to compare the amounts of motion energy from differences in weight but constant speed, and to compare the weights from differences in amounts of motion energy but constant speed. These are the first items in the progression in which mass/weight is not constant and, therefore, requires students to know that motion energy depends on mass as well as speed. Finally, the most difficult items were ones in which the students had to recognize that they needed more information than what was presented in the stem. Our analysis of this set of items revealed that we need additional items in which mass is not held constant to better test students' understanding that mass as well as speed contributes to motion energy. We have developed additional items and will further investigate this progression of understanding during the field testing being conducted in the spring of 2009.

The written comments that the students made during pilot testing led us to revise some of the motion energy items. For example, the students wrote that they were confused by the context for one item involving a boy and a girl sledding down a hill. The item stem stated that the boy started before the girl but that the girl was going faster than the boy, and the diagram showed the boy farther down the hill than the girl. The students didn't understand how the boy could be going slower yet be in front of the girl. Written comments included: "If the girl is going faster than the boy, and the boy got a head start they should be at equal length" and "If the boy started before the girl and the girl is going faster, then how come the picture is showing the boy going faster than the girl." In response, we revised the stem to explain that the reason why the girl was going faster was because she was pushed. We also added motion lines to the diagram to represent the differences in speed. Three motion lines were drawn behind the girl and two lines were drawn behind the boy.

Idea B: Thermal Energy (substance level)

Just as we found with Idea A, many students did not know that thermal energy depends on mass. They had no problem with the idea that it depends on temperature, but they were less certain that it depended on mass and the type of material. The results are shown in Table 5. The students had the least difficulty with items that asked them to compare the amounts of thermal energy of various objects when they were given the temperatures of the objects and told that the mass and type of material was the same. The students had the most difficulty with the items that asked them to compare the masses of two objects when they were given the amounts of thermal energy and told that the objects are made out of the same material and are at the same temperature.

Table 5

The results for the items aligned to the thermal energy sub-ideas ordered from easiest to most difficult

Item Context (# of items in parentheses)	% Correct	Rasch Difficulty
Compare the amounts of thermal energy of various objects when given the temperatures; mass and type of material are held constant (4)	54.5%	-1.16
Compare the temperature of two balls when given the amounts of thermal energy; mass and type of material are held constant (1)	54.6%	-0.67
Compare the amounts of thermal energy of various objects when given the masses; temperature and type of material are held constant (3)	22.7%	0.93
Compare the masses of two wooden blocks when given the amounts of thermal energy; temperature and type of material are held constant (1)	18.6%	2.26

The observed progression can be explained by the fact that the first two types of items shown in Table 5 do not require that students know that thermal energy depends on mass as well as temperature. Many of students know that thermal energy depends on temperature, but a smaller number know that it also depends on mass or the type of material involved. As with motion energy, this progression of understanding thermal energy will be further investigated during our field testing.

Idea C: Thermal Energy (atomic level)

From the items aligned to Idea C, that thermal energy is associated with the disordered motions of the atoms or molecules of an object, we learned that almost half of the students knew the thermal energy of an object depends on the speed of the molecules the object is made of, but fewer students knew that the thermal energy is also dependent on the number (21%) and type (28%) of molecules the object is made of. These results are analogous to the results from Idea B, which tests student understanding of thermal energy at the substance level. For an item that asked why all things have thermal energy, almost 20% of the students incorrectly thought that all things have thermal energy because all things are made up of atoms that are rubbing together. These students may be thinking about their experiences using friction to warm things, like rubbing their hands together to warm them.

When assessing students' understanding of Idea C, students were asked to link thermal energy to the *average speed of the molecules* in some items and to link thermal energy to the *motion energy of the molecules* in other items. Students scored consistently better on the items about the average speed of the molecules (45.3%) than on the items about the motion energy of the molecules (28.5%). It is possible that relating the motion energy of molecules to the thermal energy of the object is too abstract a task for middle school students. Because both of these are important concepts for students to understand, we will keep both kinds of items but report the results separately.

Idea D: Gravitational Energy

The results for the items aligned to Idea D, that gravitational potential energy is associated with the distance an object is above a reference point and the mass of the object, indicated that a number of students were interpreting the phrase "gravitational energy" to mean "the force of gravity." For example, one item, shown in Figure 2, asked how the gravitational energy of a rocket changed as it gets higher in the sky. About 34% of the students incorrectly selected answer choice B in which the gravitational energy is said to decrease as the rocket travels up into the sky (see Table 6). In fact, the force of gravity between objects does decrease with distance between the objects, but the gravitational potential energy *increases* (at least when the force of gravity is considered to be constant and the objects are near the reference surface). Written comments of students who selected this answer choice included: "the farther away you get from the earth, the less gravity" and "its going into space and space has no gravity." Similarly, another item asked how the gravitational energy of an object changed as it falls to the earth. Almost half of the students selected the answer choice that said that the gravitational energy increased as the object fell. Written comments included: "I think the pull of gravity gets stronger as the object gets closer to earth" and "The gravitational pull increases as the object falls." When asked what gravitational energy means, comments included: "to me it means gravity" and "it helps us stay on the ground."

Figure 2: An assessment item aligned to the idea of gravitational energy

---

A rocket is launched and is traveling up into the sky.

How does the gravitational energy of the rocket change as it gets higher in the sky?

A. The gravitational energy of the rocket increases as it gets higher.
B. The gravitational energy of the rocket decreases as it gets higher.
C. The gravitational energy of the rocket does not change as it gets higher.
D. The gravitational energy of the rocket depends on how fast the rocket is moving.

Table 6
Number and percentage of students who selected each answer choice

	A	B	C	D	Not Sure / No Response	Total
#	13	25	7	5	23	73
%	17.8%	34.2%	9.6%	6.8%	31.5%	100%

This finding is not surprising because it is very well documented that students often confuse energy and force (Duit, 1981, 1984; Kesidou & Duit, 1993; Kruger, 1990; Kruger, Palacino, & Summers, 1992; Papadouris, Constantinou, & Kyratsi, 2008; Stead, 1980; Summers & Kruger, 1993; Trumper, 1993, 1997a, 1997b, 1998; Trumper & Gorsky, 1993; Watts & Gilbert, 1983), and gravity is a specific instance of forces with which students are very familiar. Students' answer choices and written comments indicate that these assessment items are not a fair judge of students' understanding of gravitational potential energy. The items will be revised to address this issue. For example, students may need to be given more information in the item's stem to guide them to think about the potential energy that an object has with respect to a reference point and not the force of gravity. Perhaps just using the term "gravitational potential energy" instead of "gravitational energy" will be enough, but perhaps other revisions will be needed as well, such as telling them to assume that the force of gravity is constant near the surface of the earth. During field testing, we will be able to diagnose which students are thinking of the force of gravity by looking at the pattern of responses the students give to items aligned to gravitational potential energy. The students who are thinking about force should respond incorrectly to items that require the knowledge that gravitational potential energy increases as the object's distance from the center of the earth increases, and they should respond correctly to items that require the knowledge that gravitational potential energy increases as the object's mass increases.

Idea E: Elastic Energy

For Idea E, that elastic potential energy is associated with the stretching of an elastic object and with how easily the object can be stretched, students knew that, when comparing two identical stretched objects, the one that has more elastic energy is the one that is stretched more (67.2%) and the one that is stretched more has more elastic energy (58.2%), and when comparing two identical compressed objects, the one that is compressed more has more elastic energy (51.7%). Fewer students knew that when stretching two elastic objects, the one that is harder to stretch has more elastic energy (43.5%).

When comparing two springs that were not being stretched or compressed, 44.6% of the students chose the answer choice that stated that the longer spring had more elastic energy. Some students were thinking that the springs had elastic energy while they were not being stretched or compressed, in other words, that elastic energy is a property of an un-stretched object rather than of an object that is being stretched. Written comments included: "I think since the spring is bigger it has more energy" and "because the longer the spring the more elastic energy it has." Other students were thinking of the energy it would have if it were stretched or compressed. Written comments included: "Because the longer spring is able to be compressed more, therefore more elastic energy" and "I would think if it was longer it could stretch longer than a shorter one."

Our pilot test results also revealed that some students were unsure of the word "compress." A number of students circled "compress" as a word they were not familiar with, and an item that had the word "compress" but did not include a specific context or a picture had a lower percent correct and a higher Rasch difficulty (35.0% and .21, respectively) than an item that used the word and had a picture of a spring being compressed (47.8% and -.23). Without a real-world example, like a spring, some students were unsure if compressed objects had any elastic energy. Some students thought that only stretched objects had elastic energy. Written comments included: "elastic is to expand not get smaller" and "because elastic is how much you can stretch something." In the future, items that we develop that include the word "compress" will be set in a real-world context or utilize a diagram.

Conclusions

We have shown that the results of pilot testing can be used to provide information about how well a set of assessment items measures middle school students' understanding of ideas about energy. For this set of items, we found that many students knew that motion energy is associated with the speed of an object, thermal energy is associated with the temperature of an object, and elastic energy is associated with the stretching and compressing of an elastic object. But we also found that fewer students knew that motion energy and thermal energy are also associated with the mass of the object. Additionally, we learned that the difficulty of an item depends on the way in which the variables were presented in the item (i.e. expecting the students to determine the mass from the motion energy and speed versus expecting the students to determine the motion energy from the speed and mass).

The pilot test results suggested a number of revisions that could improve the validity of individual items as measures of students' understanding of the energy ideas. Students' responses to the items aligned to the idea about gravitational energy revealed that many students were confusing "gravitational energy" and "the force of gravity." Responses to the items aligned to the idea about elastic energy revealed that students may be confused by the word "compress" when the word is not used in a real-world context. Revisions to these items will be made to help avoid these confusions.

Acknowledgements

This work is funded by the National Science Foundation (Grants # ESI 0227557 and ESI 0352473).

References

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

American Federation of Teachers. (2006). Smart testing: Let's get it right (Policy Brief No. 19). Washington, DC: Author.

DeBoer, G.E., Herrmann Abell, C.F., and Gogos, A., (2007, March-April). *Assessment Linked to Science Learning Goals: Probing Student Thinking During Item Development*. Paper presented at the National Association for Research in Science Teaching Annual Conference, New Orleans, LA.

DeBoer, G.E., Herrmann Abell, C.F., Gogos, A., Michiels, A., Regan, T., & Wilson, P. (2008a). Assessment linked to science learning goals: Probing student thinking through assessment. In J. Coffey, R. Douglas, & C. Stearns (Eds.), *Assessing Student Learning: Perspectives from Research and Practice* (pp. 231-252). Arlington, VA: NSTA Press.

DeBoer, G.E., Lee, H.S., & Husic, F. (2008b). Assessing integrated understanding of science. In Y. Kali, M.C. Linn, & J.E. Roseman (Eds.), *Coherent Science Education: Implications for Curriculum, Instruction, and Policy* (pp. 153-182). New York, NY: Columbia University Teachers College Press.

Duit, R. (1981). Students' Notions about the Energy Concept -- Before and After Physics Instruction. Paper presented at the Conference on "Problems Concerning Students' Representation of Physics and Chemistry Knowledge", Germany.

Duit, R. (1984). Learning the energy concept in school - empirical results from the Philippines and West Germany. *Physics Education*, 19, 59-66.

Kesidou, S., & Duit, R. (1993). Students' conceptions of the second law of thermodynamics - An interpretive study. J*ournal of Research in Science Teaching*, 30(1), 85-106.
Kruger, C. (1990). Some Primary Teachers' Ideas about Energy. *Physics Education*, 25(2), 86-91.
Kruger, C., Palacino, D., & Summers, M. (1992). Surveys of English Primary Teachers' Conceptions of Force, Energy, and Materials. *Science Education*, 76(4), 339-351.
Linacre, J.M. (2006). WINSTEPS Rasch measurement computer program. Chicago: Winsteps.com.
No Child Left Behind Act of 2001, 20 U.S.C. § 6301 *et seq.* (2002).
Papadouris, N., Constantinou, C. P., & Kyratsi, T. (2008). Students' Use of the Energy Model to Account for Changes in Physical Systems. *Journal of Research in Science Teaching*, 45(4), 444-469.
Stead, B. (1980). Energy (Working Paper No. 17). In Learning in Science Project. Hamilton, New Zealand: Science Education Research Unit, University of Waikato.
Summers, M., & Kruger, C. (1993). Long term impact of a new approach to teacher education for primary science. Paper presented at the The Annual Meeting of the British Educational Research Association, Liverpool, England.
Trumper, R. (1993). Children's energy concepts: a cross-age study. *International Journal of Science Education*, 15(2), 139-148.
Trumper, R. (1997a). The Need for Change in Elementary School Teacher Training: The Case of the Energy Concept as an Example. *Educational Research*, 39(2), 157-174.
Trumper, R. (1997b). A Survey of Conceptions of Energy of Israeli Pre-Service High School Biology Teachers. *International Journal of Science Education*, 19(1), 31-46.
Trumper, R. (1998). A Longitudinal Study of Physics Students' Conceptions of Energy in Pre-Service Training for High School Teachers. *Journal of Science Education and Technology*, 7(4), 311-318.
Trumper, R., & Gorsky, P. (1993). Learning about Energy: The Influence of Alternative Frameworks, Cognitive Levels, and Closed-Mindedness. *Journal of Research in Science Teaching*, 30(7), 637-648.
Watts, D. M., & Gilbert, J. K. (1983). Enigmas in School Science: Students' Conceptions for Scientifically Associated Words. *Research in Science and Technological Education*, 1(2), 161-171.