

Chapter 7

Developing and Using Distractor-Driven Multiple-Choice Assessments Aligned to Ideas About Energy Forms, Transformation, Transfer, and Conservation

Cari F. Herrmann-Abell and George E. DeBoer

7.1 Introduction

In today's society, people are asked to use their knowledge about energy to make decisions about what kinds of light bulbs to use, to evaluate arguments related to global climate change, and to think about national energy resource policy. Students learn about energy in their physical science classes and also in life and earth science classes when they are studying topics such as photosynthesis and respiration or weather and climate. But, research has shown that in spite of the instruction they receive, students (and adults) hold a wide range of misconceptions and alternative ideas about energy. Because energy is such an important concept, it must be taught well, and part of teaching it well is understanding what students know and do not know about energy before, during, and after instruction. Having assessments designed specifically to pinpoint students' conceptual problems and their causes is essential for teachers. To respond to this need for high quality diagnostic assessments, in 2004 we began to develop a bank of assessment items that are precisely aligned to middle school science ideas in the life, physical, and earth sciences and that can be used to diagnose common misconceptions and students' difficulties with energy concepts (DeBoer et al. 2008a).

Our work builds on and extends the existing knowledge base about student understanding of energy in a number of ways. First, the assessment results are based on a very large national sample of students in grades 6–12 ($N = 23,744$). This gives us a more broad-based description than most studies do of how students' understanding of energy changes from grade-to-grade. We can also make comparisons across grades and across concepts because all students in grades 6–12 were tested from a common pool of linked test items. Second, our assessment items are precisely aligned to specific energy concepts including energy forms,

C.F. Herrmann-Abell (✉) • G.E. DeBoer
American Association for the Advancement of Science/Project 2061, Washington, DC, USA
e-mail: cabell@aaas.org

transformation, transfer, and conservation. This enables us to explore problems that students are having understanding particular concepts, not just energy in general. Third, the items are designed not only to test for the correct scientific understanding but also to probe for common misconceptions. This provides the opportunity to make a detailed analysis of the alternative ideas students hold that may be giving them difficulty. Incorporating common misconceptions into answer choices also gives students plausible answers to select from so they are less likely to guess, thus giving us a more valid measure of what they actually know.

In this chapter, we describe the development of assessment items for the energy topic, and then we review the results of national field testing of those items focusing on what students know, what they do not know, and the misconceptions they have.

7.2 Item Development

7.2.1 Selecting a Set of Target Learning Goals

The first step in our item development process is selecting a set of learning goals that will serve as the targets for the assessment items. Energy is a very broad topic with applications in many fields of science, so the topic has to be narrowed down to a coherent set of learning goals before beginning item development. Our goal was to choose a set of ideas about energy that can be connected together and reinforce each other in a coherent storyline, that target particularly problematic ideas about energy that students have, that are appropriate for middle school students and reflect the progression of their understanding of energy based on empirical research reported in the literature, and that are consistent with the recommendations in national standards documents.

7.2.2 Consulting the Research Literature on Students' Understanding of Energy

Before beginning item development, we conducted a thorough review of the research literature to determine which ideas about energy students were most likely to be learning in school and, therefore, appropriate for assessment. One body of research on students' understanding of energy has focused on energy as a single unified concept (Watts 1983; Trumper 1990, 1993; Nicholls and Ogborn 1993), not on the different "forms" or manifestations of energy such as thermal, chemical, and elastic energy. For example, Watts (1983) classified students' ideas about energy into seven general types: anthropocentric, depository, ingredient, activity, product, functional, and flow-transfer. Trumper (1990) later expanded on Watts' work by splitting the depository framework into two, the original passive

“depository” framework and an “active” deposit or “cause” framework, and adding the transformation framework, which is the accepted scientific view. Although that work demonstrates the variety of ways that students may think about energy, we chose to focus on the ideas students have about the different ways that energy is manifested in their everyday lives. Thus, forms of energy (e.g., motion, thermal, and gravitational potential energy) and form conversions are key aspects of our assessments.

Research on students’ ideas about energy revealed a number of misconceptions. For example, some students associate energy only with obvious activity or movement (Brook and Driver 1984; Finegold and Trumper 1989; Kruger 1990; Kruger et al. 1992; Stead 1980; Summers and Kruger 1993; Trumper 1990, 1998; Trumper and Gorsky 1993; Watts 1983). To them, objects at rest have no energy at all. Regarding thermal energy, students often think that living things have thermal energy but inanimate objects do not (Stead 1980; Solomon 1983; Watts 1983; Finegold and Trumper 1989; Kruger 1990; Trumper 1990, 1993; Kruger et al. 1992; Trumper and Gorsky 1993; Leggett 2003). Ideas about gravitational potential energy frequently include the idea that “potential” energy is the potential to have energy in the future (Stead 1980; Summers and Kruger 1993). Students may also confuse force and energy, especially thinking that objects in motion have a force within them (Fischbein et al. 1989; McCloskey 1983). Students also report that it is “coldness” that is transferred between two objects at different temperatures (Brook et al. 1984; Clough and Driver 1985; Newell and Ross 1996), not thermal energy. Finally, it is common for students to think that energy can be created or destroyed (Brook and Driver 1984; Kesidou and Duit 1993; Kruger 1990; Loverude 2004; Papadouris et al. 2008; Stead 1980; Trumper 1998). These alternative ideas were built into the assessment items as distractors (Sadler 1998).

Regarding the progression of students’ understanding of energy over time, a number of recent empirical studies have shown that students develop ideas about forms of energy first, followed by transformations and transfers, and finally conservation. Liu and Mckeough (2005) used data from selected items from the Third International Mathematics and Science Study (TIMSS) database. Using the partial credit Rasch model, they demonstrated support for their hypothesized sequence of development of the energy concept. In their proposed progression of understanding, students first perceive energy as activity, or the ability to do work. As students’ understanding progresses, they begin to distinguish different energy sources and forms of energy. Next comes an understanding of energy transfer, followed by an appreciation of energy degradation, and finally an acceptance of the concept of conservation of energy. Liu and Collard (2005) validated those results in a follow-up study on students in grades 4, 8, 10, 11, and 12 using performance assessments and Many-Facet Rasch measurement. Lee and Liu (2010) found further support for the conclusion that energy conservation was the most difficult concept for students using 10 two-tiered items based on released TIMSS items addressing energy sources, energy transformation, and conservation of energy. The items were administered to a large sample of middle school students, and the results showed that the conservation of energy items required the highest level of knowledge integration compared to the

other two concepts. This same trend was also found in a study of German students' understanding of energy (Neumann et al. 2012). Neumann and his colleagues found that 6th grade students had an understanding of forms and sources of energy, 8th grade students had an understanding of energy transformations and transfer, and only some 10th grade students had reached an understanding of conservation of energy.

7.2.3 Clarification of the Target Learning Goals

To choose the set of ideas for testing, we relied primarily on AAAS Project 2061's previous work developing the strand maps published in the *Atlas of Science Literacy* (Association for the Advancement of Science [AAAS] 2001, 2007). These maps show how student understanding might progress from grade to grade and across concepts to create a complex mental network of interconnected knowledge about the world. The Energy Transformations map in *Atlas of Science Literacy, Volume 2* (AAAS 2007) and Chapter 4 (The Physical Setting), Section E (Energy Transformations) of *Benchmarks for Science Literacy* (AAAS 1993) were used to guide our choice of energy ideas as targets for assessment:

- *Motion Energy (kinetic energy)* is associated with the speed and the mass of an object.
- *Thermal Energy (substance level)* is associated with the temperature and the mass of an object and the material of which the object is made.
- *Thermal Energy (atomic level)* is associated with the disordered motions of an object's atoms or molecules and the number and types of atoms or molecules of which the object is made.
- *Gravitational Potential Energy* is associated with the mass of an object and the distance the object is above a reference point, such as the center of the earth.
- *Elastic Energy* is associated with the stretching or compressing of an elastic object and how easily the object can be stretched or compressed.
- Energy can be *transformed* within a system (e.g. motion to thermal, gravitational potential to motion, etc.).
- Energy can be *transferred* from one object or system to another in different ways: by conduction, mechanically, electrically, or by electromagnetic radiation.
- Energy is *conserved*. Regardless of what happens within a system, the total amount of energy in the system remains the same unless energy is added to or released from the system.

Each of these ideas was further clarified into sub-ideas to state precisely what students would be expected to know and boundary statements to indicate what they would not be expected to know. 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 conservation of energy includes the following sub-ideas and boundary statements:

Students should know the following sub-ideas:

1. Regardless of what happens within a system, the total amount of energy in the system remains the same unless energy is added to or released from the system, even though the forms of energy present may change.
2. If the total amount of energy in a system seems to decrease or increase, energy must have gone somewhere or come from somewhere outside the system.
3. If no energy enters or leaves a system, a decrease of one form of energy by a certain amount within the system must be balanced by an increase of another form of energy by that same amount within the system (or a net increase of multiple forms of energy by that same amount). Similarly, an increase of one form of energy by a certain amount within a system must be balanced by a decrease of another form of energy by that same amount within the system (or a net decrease of multiple forms of energy by that same amount).
4. Energy can neither be created nor destroyed, but it can be transferred from one object or system to another and/or be changed from one form to another.
5. If energy is transferred to or from a very large system (or a very complex system), increases or decreases of energy may be difficult to detect and, therefore, it may appear that energy was not conserved.

Boundaries:

1. Students are not expected to quantitatively keep track of changes of energy in a system.
2. Assessment items will avoid using the phrase “energy conservation” or “conservation of energy” because of the misconceptions associated with them.
3. Students are not expected to know about energy-mass conversions such as nuclear reactions or other subatomic interactions.

(Note: Clarifications of the other targeted ideas can be found at <http://assessment.aaas.org>.)

7.2.4 Efforts to Ensure Validity

Each item was developed using a procedure designed to ensure its match to the targeted idea and its overall effectiveness as an accurate measure of what students do and do not know about the idea. We used a set of criteria developed by AAAS Project 2061 to evaluate the content alignment of assessment items and to minimize construct-irrelevant factors that make it difficult to interpret a student’s response to an item. The full description of the analysis criteria used during item development can be found at: http://www.project2061.org/research/assessment/assessment_form.htm. Additional details of the analysis procedure can be found in DeBoer et al. (2007, 2008a, b).

Content Alignment Two content alignment criteria were used: necessity and sufficiency (Stern and Ahlgren 2002). The necessity criterion addresses whether the knowledge described in the learning goal is needed to answer correctly. For a multiple choice question, meeting the necessity criterion means that the knowledge in the learning goal is needed to evaluate all of the answer choices, including incorrect answer choices targeting known misconceptions (Sadler 1998). In other words, the knowledge that was targeted by the item had to be necessary to evaluate each answer choice, both correct answers and distractors. For example, an item aligned to conservation of energy would meet the necessity criterion only if it required students to use the knowledge that the total amount of energy in a system remains the same unless energy is added to or released from the system to select the correct answer and eliminate all of the distractors. The second content alignment criterion is sufficiency, which addresses whether the knowledge described in the learning goal is enough by itself to successfully complete the item. The student should not be expected to use additional knowledge not covered by the learning goal to evaluate the answer choices and select a correct answer. In the conservation of energy example above, the item would *not* meet the sufficiency criterion if it required students to use knowledge of chemical reactions to analyze the answer choices. That would be going beyond the knowledge targeted by the item and, therefore, would not a fair measure of students' understanding of the learning goal. In a few cases, we decided to target more than one learning goal in an item, but that was done intentionally, and the target learning goals were specified prior to item development. In general, however, targeting more than one learning goal in the same item makes it more difficult to pinpoint exactly where students have gaps in their knowledge, so we avoided as much as possible alignment to multiple learning goals in what were intended to be highly diagnostic items.

Construct Validity The set of criteria that we used to ensure construct validity and minimize construct-irrelevant features included comprehensibility, appropriateness of task context, and test-wiseness. (1) To meet the comprehensibility criterion, the item had to make it clear what question is being asked, avoid unfamiliar general vocabulary or unnecessarily complex sentence structure, use words or phrases that did not have ambiguous meanings, and present diagrams, pictures, graphs, and tables that could be easily understood. (2) To ensure that the task context was appropriate and fair, it had to be familiar to most students, so that one group of students was not advantaged or disadvantaged because of their familiarity with the context, be clear and easy to understand, use information and quantities that are reasonable and believable, and accurately represent scientific or mathematical realities or make clear when idealizations are involved. (3) When analyzing for test-wiseness, the plausibility of the distractors was considered, along with whether one answer choice differed in length or detail, whether one answer choice was qualified differently, whether one answer choice contained vocabulary at a different level of difficulty, whether a pair of answer choices contained logical opposites that may lead students to eliminate answer choices, and whether the language in one answer choice mirrored the language in the stem.

7.2.5 Pilot Testing

After items were drafted, they were pilot tested with students in middle and high school to obtain feedback from them about the items. The pilot test included follow-up questions for each item to give us insight into how well the item was meeting the assessment criteria described above. Questions asked students to describe anything they found confusing about an item, to circle words they were unfamiliar with, and to comment on the helpfulness of diagrams and tables. Student responses to these questions provided information about how well the item was meeting the criteria related to construct irrelevant factors. Students were also asked to write explanations for why they selected or rejected each answer choice, to indicate if they guessed, and to indicate where they had learned about the topic. This allowed us to determine what knowledge students were using to answer the items, which helped us to evaluate the content alignment of the items. After the pilot test, a panel of scientists and science education and assessment experts was convened to review the items using the same criteria that were used in item development. After revisions were made based on the pilot testing and expert reviews, the items were field tested on a large national sample to determine the psychometric properties of the items. A subset of the items used in the field test is available at <http://assessment.aaas.org/topics/EG#/>.

7.2.6 Field Testing and Data Collection

Field testing of the items took place in the spring of 2009 and the spring of 2010. Because we were testing more items than students could finish in a typical class period, we created multiple test forms that contained subsets of the available items. Linking items were included so that we could use Rasch modeling to compare item characteristics across forms and between years. During the field test, students were asked to choose the single correct answer for each item. Items for which students chose more than one answer choice were marked incorrect.

We sent invitations to participate in the field testing directly to teachers or to school and district administrators who then recruited teachers to participate. Teachers were selected to participate on a first-come first-served basis but, when necessary, we adjusted our selections to achieve representation from urban, rural, and suburban schools from different parts of the US. The field tests included 14,484 middle school students and 9,260 high school students from 48 states across the country. (See Table 7.1 for demographic information.) The teachers received the testing materials by mail and administered the field tests to whatever science classes they were teaching (either life, physical, or earth science). Item sampling was used such that each student received 30–44 assessment items, and each item was answered by an average of 2,530 middle school students and 1,665 high school students.

Table 7.1 Demographic information for the students who participated in the field test^a

| Grade | Total | Female | Male | English | Non-English |
|------------|----------------|--------|--------|---------|-------------|
| | % (N) | % | % | % | % |
| 6th Grade | 18 % (4,330) | 49.6 % | 48.7 % | 87.1 % | 10.3 % |
| 7th Grade | 22 % (5,177) | 49.1 % | 48.6 % | 88.0 % | 9.0 % |
| 8th Grade | 21 % (4,977) | 48.9 % | 49.3 % | 88.2 % | 8.7 % |
| 9th Grade | 11 % (2,673) | 49.5 % | 48.0 % | 88.7 % | 8.9 % |
| 10th Grade | 12 % (2,822) | 49.3 % | 48.6 % | 88.8 % | 8.1 % |
| 11th Grade | 10 % (2,452) | 51.6 % | 46.7 % | 90.6 % | 6.9 % |
| 12th Grade | 6 % (1,313) | 53.5 % | 45.1 % | 88.7 % | 8.6 % |
| Total | 100 % (23,817) | 49.6 % | 48.3 % | 88.2 % | 8.9 % |

^aGender and language columns do not total to 100 % because not all students specified their gender or primary language

7.3 Rasch Modeling

We used Rasch modeling to analyze the field test data. In the dichotomous Rasch model, the probability that a student will respond to an item correctly is determined by the difference in the student’s ability and the difficulty of the item, according to Eq. 7.1:

$$\ln \left(\frac{P_{ni}}{1 - P_{ni}} \right) = B_n - D_i \quad (7.1)$$

where P_{ni} is the probability that student n of ability B_n will respond correctly to item i with a difficulty of D_i (Liu and Boone 2006; Bond and Fox 2007). Student ability and item difficulty are expressed in logits, which can range from $-\infty$ to ∞ . It is important to note that the ability and difficulty measures are expressed on the same interval scale and that the Rasch model assumes that these two measures are mutually independent, which is not the case for classical percent correct measures. (Note: Rasch modeling uses the term “ability” to refer to the students’ understanding of the ideas being targeted by the items. It should not be interpreted as an underlying, innate quality of the student, but more narrowly as the students’ understanding of the topic at the time of the field test.)

WINSTEPS (Linacre 2012) was used to estimate student abilities and item difficulties for all of the students and all of the items from both field tests. From these parameters, we were able to determine how well the range of item difficulties matched the range of student abilities and the extent to which each of the items correlated with the entire set of items (point-measure correlation). We also looked to see if the pattern of student responses followed expectations such that the most able students were most likely to answer the most difficult questions correctly. According Rasch (1960):

a person having a greater ability than another should have the greater probability of solving any item of the type in question, and similarly, one item being more difficult than another one means that for any person the probability of solving the second item correctly is the greater one (cited in Wright and Stone 1999).

Table 7.2 Summary of Rasch fit statistics

| | Min. | Max. | Median |
|----------------------------------------|------|--------------|--------|
| Standard error | 0.02 | 0.08 | 0.04 |
| Infit mean-square | 0.85 | 1.19 | 0.99 |
| Point-measure correlation coefficients | 0.13 | 0.51 | 0.38 |
| Item separation index (reliability) | | 14.29 (1.00) | |
| Person separation index (reliability) | | 1.80 (0.76) | |

7.3.1 *Model Fit*

The field test data had a good fit to the Rasch model, suggesting that the items were measuring a unidimensional energy construct. A summary of the fit statistics is shown in Table 7.2. The separation index (14.29) and corresponding reliability (1.0) were high for the item data. The separation index indicates the approximate number of different levels of item difficulty or person ability that can be discriminated. A separation index greater than 2 is considered acceptable according to Wright and Stone (2004). The lower person separation index (1.80) for our data is due to the relatively small number of items available to test students at the extreme ends of the scale, especially the lower end (see Figs. 7.1 and 7.2). This means that there is less information available to measure the ability of very low or very high ability students. In contrast, because there were so many students responding to each item, differences in difficulty level of the items is easier to determine, which can be seen in the very high item separation index and reliability estimate. Additionally, the standard errors for the items were small. The infit mean-square values for all of the items fell within the acceptable range of 0.7–1.3 for multiple-choice tests (Bond and Fox 2007). Infit statistics are reported here because they give more weight to the responses of students with abilities closer to the item difficulty, whereas outfit statistics are unweighted and, therefore, are more sensitive to outlying scores.

7.3.2 *Wright Maps*

Figure 7.1 shows the Wright map for 91 field test items aligned to ideas about five different forms of energy. Figure 7.2 shows the Wright map for 95 items aligned to ideas about energy transformation, transfer, and conservation. The maps show the range of student abilities on the left side of a vertical line and the range of item difficulties on the right side of the line. The scale runs from low ability/difficulty at the bottom of the maps to high ability/difficulty at the top. The mean of the item difficulties is set at zero. When a student's ability matches an item's difficulty, the student has a 50 % chance of answering the item correctly. The maps show a good match between mean item difficulty and mean student ability, with mean item difficulty being just slightly higher than mean student ability. The maps also

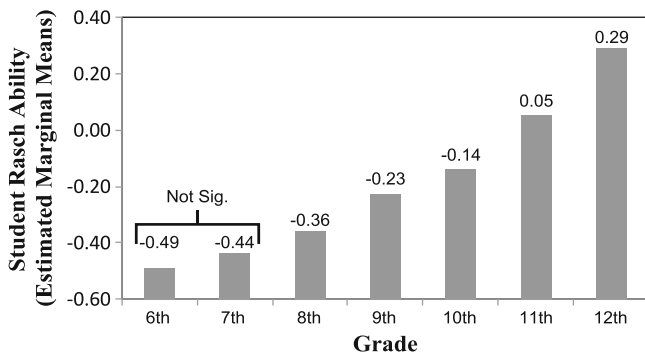


Fig. 7.3 Average student ability by grade for the field test

covariance was performed controlling for gender and whether the students identified English as their primary language. Both gender [$F(1, 22698) = 51.1$] and English as the primary language [$F(1, 22698) = 302$] were significantly correlated with student ability measures ($p < .001$). ANCOVA showed that differences in average ability by grade were significant at the .001 level of significance [$F(6, 22698) = 247$]. The estimated marginal means for the student ability measures are reported in Fig. 7.3. A Bonferroni post hoc test showed that the differences in mean ability for all grades were statistically significant on the .001 level, except for the difference between sixth and seventh grades. Knowledge of the energy ideas increased steadily from seventh to twelfth grades (see Fig. 7.3). This trend of increasing ability can be attributed to more students having the opportunity to learn these energy ideas as they progress through the grades and the greater maturity of students in the higher grades that made them more likely to understand the energy ideas covered by the items, which are often abstract and counterintuitive.

7.5 Students' Knowledge and Misconceptions

In this section, we report on the results of our field test that provided insight into what students know and do not know about energy and what misconceptions they hold. Results for each key idea are presented below.

7.5.1 Motion Energy

Students were tested for their understanding of the idea that motion energy (kinetic energy) is associated with the speed and mass of an object. Although all the motion energy items focused on some aspect of that basic idea, the items also varied in the

Table 7.3 Item difficulties for items aligned to the motion energy sub-ideas from easiest to most difficult

| Item context | No. of items | Difficulty | | | |
|--------------------------------------------------------------------------------------------------------------------|--------------|------------|-------|--------|-------|
| | | Min. | Max. | Median | Mean |
| Items in which motion energy and speed vary and weight is held constant | 10 | -1.00 | -0.36 | -0.82 | -0.78 |
| Items in which motion energy and weight vary and speed is held constant | 4 | 0.08 | 1.09 | 0.22 | 0.40 |
| Items in which the stem does not provide enough information to be able to compare the motion energy of two objects | 4 | 0.69 | 1.12 | 0.94 | 0.92 |

mental processing required of the students. In other words, students were asked to use their knowledge in a variety of ways. For example, in some items, students were asked to determine the motion energy from information about the speed and mass of the objects, and in other items the students were asked to determine the speed from information about the motion energy and mass. In other items, students were asked to compare the motion energy of two objects given information about the speed of the objects but were expected to recognize that they could not make such a comparison because they were not given information about the mass of the objects. Items were more or less difficult for students depending on how they had to reason with the knowledge.

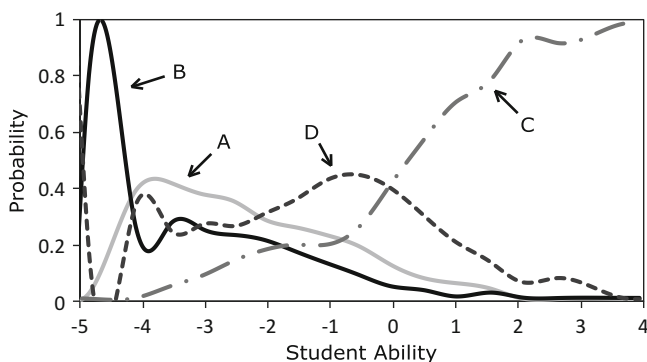
To quantify the effect that different ways of asking about motion energy had on student performance, we divided the items into three categories: (1) items in which the motion energy and the speed of the objects varied and the weight/mass was held constant, (2) items in which the motion energy and the weight/mass varied and the speed was held constant, and (3) items in which the stem did not provide enough information to be able to compare the motion energy of two objects. Table 7.3 summarizes the difficulty of the items in each of these categories. One-way ANOVA revealed statistically significant differences in the means for the three categories [$F(2, 15) = 65.55, p < .001$]. Using a Bonferroni post hoc test, the items in which the weight/mass was held constant were significantly easier than the items in the other categories ($p < .001$). Because the weight/mass was held constant, these items only required students to use their knowledge of the relationship between motion energy and speed. Items in the other two categories, which required students to know that motion energy depends on mass as well as speed, were more difficult.

Table 7.4 shows the percentage of correct responses for each of the categories by grade level. As expected, the high school students performed better than the middle school students on the items in all three categories. The high school students were more aware of the dependence of motion energy on both speed and mass and better able to handle the more sophisticated mental operations required for answering correctly.

These results from items aligned to the idea of motion energy indicate that knowing that motion energy depends on speed comes before knowing that motion

Table 7.4 Percentage of correct responses and grade level differences for items aligned to the motion energy sub-ideas

| Item context | Middle school | High school | χ^2 | Sig. |
|--------------------------------------------------------------------------------------------------------------------|---------------|-------------|----------|-------|
| Items in which motion energy and speed vary and weight is held constant | 57.8 % | 61.4 % | 59 | <.001 |
| Items in which motion energy and weight vary and speed is held constant | 25.0 % | 36.3 % | 357 | <.001 |
| Items in which the stem does not provide enough information to be able to compare the motion energy of two objects | 22.7 % | 29.7 % | 115 | <.001 |

**Fig. 7.4** Option probability curves for an item aligned to motion energy

energy depends on mass, in other words, knowing that energy is an extensive property of objects. This is supported by an analysis of the option probability curves for an item shown in Fig. 7.4. Option probability curves show the probability of selecting each answer choice as a function of student ability. For this item, answer choice A corresponds to the misconception that an object has energy because a person gives energy to the object. Answer choice B corresponds to the misconception that inanimate objects do not have energy but people do. Answer choice C is the correct answer, corresponding to the idea that motion energy depends on mass. Answer choice D corresponds to the misconception that motion energy depends only on speed, not on mass.

The misconceptions in answer choices A and B were popular among students at the lower end of the ability spectrum. It is reasonable to assume that as students receive formal instruction on the topic they let go of these human-centered views of energy. Students with abilities between -2 and 0 were more likely to select answer choice D (motion energy depends only on speed) possibly because when motion energy is first introduced, instruction focuses mainly on the speed of objects, with mass held constant. Finally, in the progression of understanding, students with ability level above zero (above the mean) were increasingly more likely to select the correct answer.

We also found from our analysis of items in this set that students held additional misconceptions about motion energy. One item revealed that about a quarter of both the middle and high school students thought that a ball that is thrown has motion energy while it is moving, but a ball that is dropped does not have motion energy while it is moving. Results from another item indicated that 35 % of the students thought that the motion energy of an object depends on both the speed and the direction of motion. These misconceptions may be related to the belief that energy is associated with how difficult it is to move an object or how hard a person has to pull or push it (Brook and Driver 1984). Students who have this misconception would think, for example, that a person walking uphill would have a different amount of motion energy than when walking downhill even if the person was walking at the same speed in both cases. For the item that compared the thrown ball and the dropped ball, students' responses may reflect their belief that the dropped ball has no motion energy because no effort was put into dropping the ball.

7.5.2 Thermal Energy (Substance Level)

Similar to our findings related to students' understanding of motion energy, we found that students were less likely to associate thermal energy with the mass of the object than with its temperature. To quantify the difference, we divided the items into two categories: (1) items in which the thermal energy and the temperature of the object or objects varied and the weight/mass and the type of material were held constant, and (2) items in which the thermal energy and the weight/mass varied and the temperature and the type of material were held constant. Table 7.5 presents the average difficulty of the items in each of these categories. A t-test confirmed that the items in which the weight/mass varied were more difficult than items in which the temperature varied [$t(8) = -9.27, p < .001$].

Table 7.6 shows the percentage of correct responses for each of the categories by grade level. These results are similar to those found for the motion energy items, showing that a number of students are unaware that the amount of energy an object has depends on how much of the object there is. As with the motion energy items, the high school students performed better than the middle school students on the items in both categories.

We also identified a number of misconceptions that students held about thermal energy. About 43 % of the middle school students and 31 % of the high school students thought that living things, including humans and plants, have thermal energy but dead things, such as dead plants, and inanimate objects, such as pennies, do not. It has been well reported in the literature that students often associate energy with living things, not inanimate objects (Stead 1980; Solomon 1983; Watts 1983; Finegold and Trumper 1989; Kruger 1990; Trumper 1990, 1993; Kruger et al. 1992; Trumper and Gorsky 1993; Leggett 2003). Also, some students associate thermal energy with warmth. For example, on an item involving two pieces of metal, about

Table 7.5 Item difficulties by for items aligned to the thermal energy sub-ideas

| Item context | No. of items | Difficulty | | | |
|------------------------------------------------------------------------------------------------------|--------------|------------|-------|--------|-------|
| | | Min. | Max. | Median | Mean |
| Items in which thermal energy and temperature vary and weight and type of material are held constant | 5 | -1.02 | -0.42 | -0.67 | -0.71 |
| Items in which thermal energy and weight vary and temperature and type of material are held constant | 5 | 0.18 | 0.46 | 0.40 | 0.36 |

Table 7.6 Percentage of correct responses and grade level differences for items aligned to the thermal energy sub-ideas

| Item context | Middle school | High school | χ^2 | Sig. |
|------------------------------------------------------------------------------------------------------|---------------|-------------|----------|-------|
| Items in which thermal energy and temperature vary and weight and type of material are held constant | 56.6 % | 59.1 % | 18.5 | <.001 |
| Items in which thermal energy and weight vary and temperature and type of material are held constant | 35.4 % | 36.8 % | 5.7 | <.01 |

31 % of the middle school students and 26 % of the high school students thought that the piece of metal that feels warm has thermal energy but the piece of metal that feels cold does not.

7.5.3 Thermal Energy (Atomic Level)

From the items aligned to the idea that thermal energy is associated with the disordered motions of the atoms or molecules of an object, we learned that most of the students knew that the thermal energy of an object depends on the speed of the molecules the object is made of (67 % of middle school students and 71 % of high school students). Far fewer students knew that the thermal energy is also dependent on the number and type of molecules the object is made of (34 % of middle school students and 35 % of high school students). These results are analogous to the results described above from testing students' understanding of thermal energy at the substance level. Other misconceptions that were revealed in student answer selections included the idea that thermal energy is due to atoms rubbing together, that only living things have thermal energy, and, more generally, that only warm things have thermal energy.

On one item, students were asked if all things have thermal energy and then asked why or why not. The option probability curves in Fig. 7.5 show the probability of selecting each answer choice as a function of student ability. In this item, answer choice A corresponds to the misconception that all things have thermal energy because all things are made up of atoms that are rubbing together (Wiser 1986; Kesidou and Duit 1993). Answer choice B is the correct answer corresponding to the idea that thermal energy is the result of atoms in constant motion. Answer choice C

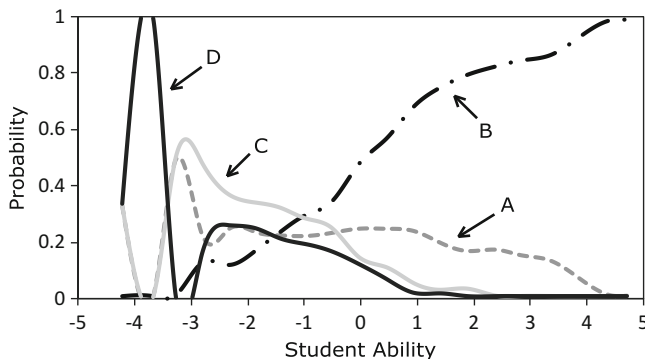


Fig. 7.5 Option probability curves for an item aligned to atomic-level thermal energy

corresponds to the misconception that only warm or hot things have thermal energy, and answer choice D corresponds to the misconception that only living things have thermal energy.

The misconceptions in C and D were popular at the lower end of the ability spectrum, and the probability of selecting these choices decreased to zero at a student ability of approximately 2. The probability of selecting answer choice A was rather constant from ability -3 to 2 and then decreased slowly. The shape of this curve indicates that the misconception that thermal energy is a result of atoms rubbing together is present over a very wide range of ability levels. This misconception most likely results from students thinking about their experiences using friction to warm things, like rubbing their hands together to warm them.

7.5.4 Gravitational Potential Energy

Results of pilot testing the gravitational potential energy items revealed that middle school students confused the phrase “gravitational energy” with the force of gravity (Herrmann-Abell and DeBoer 2009). For an item asking how the gravitational energy of a rocket changes as it gets higher in the sky, about one third of the students chose the answer stating that the gravitational energy decreases as the rocket gets higher. Their written responses confirmed that they were thinking about the force of gravity. For example, students wrote “the farther away you get from the earth, the less gravity” and “its going into space and space has no gravity.” In response to a follow-up question asking what “gravitational energy” meant to them, students responded “to me it means gravity” and “it helps us stay on the ground.” Students’ answer choices and written comments indicated that assessment items using “gravitational energy” may not be a fair judge of middle school students’ understanding of gravitational potential energy because the items do not do enough to cue the students away from “gravity” toward an energy context.

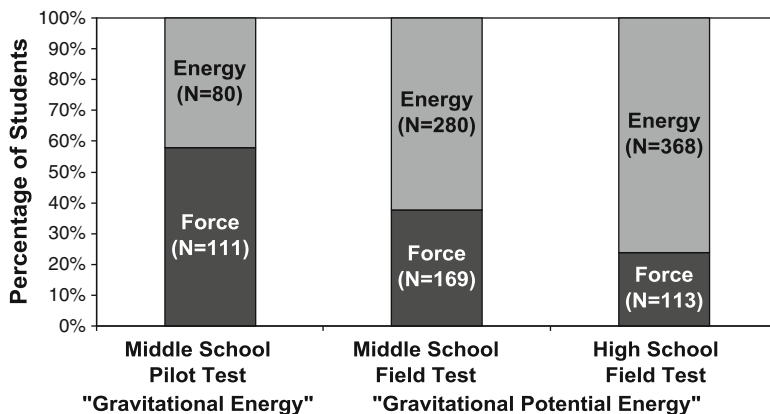


Fig. 7.6 The percentage of students coded “Force” or “Energy” for the pilot and field tests. ($\chi^2 = 74.2$, $p < .001$)

To help reduce this confusion and improve validity, the items were revised after pilot testing and the phrase “gravitational potential energy” was used in place of “gravitational energy.” To test the effect of this change, we examined the pattern of responses to estimate how many students were still responding to the items in terms of the force of gravity. We expected that students who were thinking about the force of gravity would select the answer choice that said the gravitational potential energy decreases as the distance increases (thinking that the gravitational force between two objects decreases as the distance between the objects increases) or the answer choice that said that gravitational potential energy remains the same as the distance increases (thinking that the force of gravity is constant near the surface of the earth). Additionally, we expected that students who were thinking about the force of gravity would respond correctly to items that required the knowledge that gravitational potential energy increases as the object’s mass increases because the force of gravity does increase as mass increases.

To compare response patterns from the pilot and field tests, we coded students by their answer choice selections on the set of items. Those students who chose both (1) answer choices that said the gravitational potential energy decreases or stays the same as the distance increases and (2) answer choices that said the gravitational potential energy increases as mass increases were coded “Force.” Students who responded to the majority of the questions correctly (i.e., those students who were answering the items using a correct understanding of gravitational potential energy) were coded “Energy.” Students who did not follow either of these patterns were not coded.

The bar graph in Fig. 7.6 shows the percentage of students who were coded either “Force” or “Energy” for both the pilot test and field test. The field test data were separated by grade level. In the pilot test, where the phrase “gravitational energy” was used in the items, more students’ answers were coded “Force” than were coded

Table 7.7 Item difficulties for items aligned to the gravitational potential energy sub-ideas

| Item context | No. of items | Difficulty | | | |
|-------------------------------------------------------------------------------------------|--------------|------------|-------|--------|-------|
| | | Min. | Max. | Median | Mean |
| Items in which gravitational potential energy and height vary and weight is held constant | 10 | -0.74 | 0.52 | 0.14 | -0.04 |
| Items in which gravitational potential energy and weight vary and height is held constant | 5 | -0.32 | -0.15 | -0.19 | -0.23 |

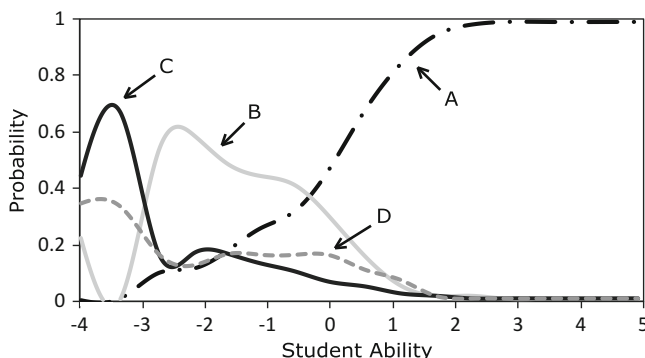
“Energy.” In the field test, where the phrase “gravitational potential energy” was used in the items, more middle school students’ answers were coded “Energy” than were coded “Force.” This suggests that the phrase “gravitational potential energy” did indeed help more students correctly think in terms of energy. With regard to the high school students who participated in the field test, we expected that they would be less likely to confuse gravitational potential energy with the force of gravity because they would have had more instruction on gravitational potential energy than the middle school students. The results shown in Fig. 7.6 support this, as indicated by the fact that more high school students’ answers were coded “Energy” than middle school students. All of these results give us confidence that the field test items provide a better measure of students’ knowledge of gravitational potential energy than the pilot test items do.

As we noted earlier, for the motion energy and thermal energy ideas, students tended to be less familiar with the effect of mass on energy than the effect of motion or temperature on energy. This was not the case for items related to gravitational potential energy (see Table 7.7). Items in which the gravitational potential energy and weight/mass varied and the height above the ground was held constant were not more difficult than items in which the gravitational potential energy and height above the ground varied and the weight/mass was held constant [$t(13) = 0.875$, $p > .05$]. This is probably because it is just as reasonable to think that the energy of an object held above the ground is a function of its weight as it is to think that it is a function of its height, especially if the student is visualizing the object’s energy in terms of the damage it can do when it is dropped. Apparently, it is less obvious to students that thermal energy and motion energy are also associated with mass.

As with thermal energy and motion energy, the most difficult item in the set of items aligned to gravitational potential energy was one in which the gravitational potential energy was held constant and both the height and weight/mass varied (difficulty = 0.99) but, as before, this is as much a function of the sophistication of the mental processing required as it is about the science content being assessed. It is difficult for both middle and high school students to make the mental manipulation that requires them to determine that an object that is higher above the ground must weigh less than an object closer to the ground if the two objects have the same gravitational potential energy. Perhaps it is not surprising that reasoning about an inverse relationship, such as the relationship between height and weight when gravitational potential energy is constant, is more difficult than reasoning about a

Table 7.8 Percentage of correct responses and grade level differences for items aligned to the gravitational potential energy sub-ideas

| Item context | Middle school | High school | χ^2 | Sig. |
|-------------------------------------------------------------------------------------------|---------------|-------------|----------|-------|
| Items in which gravitational potential energy and height vary and weight is held constant | 39.9 % | 49.2 % | 487 | <.001 |
| Items in which gravitational potential energy and weight vary and height is held constant | 46.9 % | 50.7 % | 31 | <.001 |
| Item in which height and weight vary and gravitational potential energy is held constant | 23.2 % | 29.5 % | 38 | <.001 |

**Fig. 7.7** Option probability curves for an item aligned to gravitational potential energy

direct relationship, such as the relationship between gravitational potential energy and height when weight is constant. This also points to the fact that in addition to the science knowledge that is being targeted, each item asks students to reason with that knowledge, which also affects the likelihood of a correct answer.

Table 7.8 shows the percentage of correct responses for each of the categories by grade level. As with motion energy and thermal energy, the high school students outperformed the middle school students in all of the categories.

The field test revealed a number of misconceptions that students held about gravitational potential energy. About 32 % of the middle school students and 26 % of the high school students thought that gravitational potential energy depends on the speed of a moving object. The trend this misconception follows across ability levels is shown in the option probability curves in Fig. 7.7. In the item analyzed there, the students were shown a diagram of a coconut falling from a palm tree and landing on the ground. The students were asked when the coconut has the most gravitational potential energy and why. Answer choice A is the correct answer and explains that the coconut has the most gravitational potential energy before it falls off the tree because that is when the coconut is at the highest point. Answer choice B corresponds to the misconception that the coconut has the most gravitational potential energy while it is falling because gravitational potential energy depends on the speed. Answer choice C says that the gravitational potential energy increases as

the coconut gets closer to the ground and answer choice D says that the gravitational potential energy of the coconut is the same all the time because gravitational potential energy depends only on the mass.

Students of low ability (-4 to -3) were most likely to choose answer choice C (gravitational potential energy increases as distance decreases) most likely because they were thinking about the force of gravity as discussed above. Students with abilities ranging from -3 to a little less than 0 had a high probability of selecting the misconception that gravitational potential energy depends on the speed (answer choice B). The probability of this misconception decreased steadily to 0 at an ability level of 2 . The correct answer A became the most probable answer choice selected at an ability of 0 . The misconception that gravitational potential energy depends only on mass was the second most probable answer choice for students of low ability (less than -3) but it was not the most probable at any ability level.

Another misconception the students had was that gravitational potential energy depends on how likely an object is to fall (16% of middle school students, 12% of high school students). These students thought that the gravitational potential energy of a rock resting on a flat surface at the top of a cliff depends on how close it is to the edge of the cliff. This misconception could be related to the misconception that potential energy is the potential to start moving, which was present in 17% of the middle school students and 16% of the high school students. Similar misconceptions have been previously cited in the literature. For example, one study found that some students believe that an object on a table has much less gravitational potential energy than an identical object at the same height but not supported by the table because the object not supported by the table has the potential to fall (Loverude 2004). Other studies have shown that students believe that potential energy is the potential to have energy in the future, not now (Stead 1980; Summers and Kruger 1993). For example, a student in Stead's study said "you could use all that water as for hydro-electricity, you know, make hydroelectricity, so it could be used for energy. Potential energy is not energy at all but it could be converted to energy – you could get energy out of it."

7.5.5 *Elastic Energy*

Students were also tested on the idea that elastic potential energy is associated with how much an elastic object is stretched or compressed and with how easily it can be stretched or compressed. Items that tested the idea that when comparing two identical stretched objects the one that is stretched more has more elastic energy (difficulty = -0.87), and the idea that when comparing two identical compressed objects the one that is compressed more has more elastic energy (difficulty = -0.65) were easier than an item that addressed the knowledge that when stretching two elastic objects the one that is harder to stretch has more elastic energy (difficulty = -0.25). When asked directly what elastic energy depends on, 38% of

the middle school students and 32 % of the high school students thought that elastic energy does not depend on how difficult it is to stretch or compress an object.

The most difficult item in this set was one that compared the elastic energy of two springs that were not being stretched or compressed at all (difficulty = 1.43). About 42 % of the middle school students and 38 % of the high school students chose the answer that stated that the longer spring had more elastic energy. These students may be thinking that elastic potential energy is a property of an un-stretched object rather than of an object that has been stretched. Written comments from an earlier pilot test indicated that this was the case (Herrmann-Abell and DeBoer 2009).

7.5.6 Energy Transformation

Items aligned to the idea of energy transformation revealed that a little under half of the students knew that energy can be transformed within a system. The percentage of correct responses to the items aligned to this idea was 42 % for the middle school students and 52 % for the high school students, and the mean Rasch difficulty for the 23 items targeting this idea was -0.06 . Only a few students selected answer choices that explicitly stated that energy cannot be transformed (14 % middle school and 12 % high school).

Student feedback obtained during the pilot test stage of item development provided evidence that some of the difficulties students had with the energy transformation items can be attributed to a lack of knowledge about the individual forms of energy (Herrmann-Abell and DeBoer 2010). In order to be successful on items testing the transformation of one form of energy into another, students must have an understanding of the individual forms of energy and be able to detect changes in those forms of energy. For example, in order to describe the energy transformations involved when a ball falls to the floor, a student has to know that the motion energy of the ball is increasing because motion energy depends on speed and the speed of the ball increases as it falls. The student also has to know that the gravitational potential energy of the ball is decreasing as it falls because gravitational potential energy depends on the height above the earth, and the height of the ball is decreasing as it falls. Misunderstanding motion energy and gravitational potential energy would cause a student to respond incorrectly to this energy transformation item. For example, one student chose the incorrect answer that said the motion energy is transformed into gravitational potential energy while the ball *falls* and wrote:

It makes sense the motion energy will decrease & gravitational will increase because it's falling. It will change from one type of energy to the next. The gravitational energy wouldn't decrease until point 3 to 4 [after the ball bounces and is traveling back up into the air].

This student thought that gravitational potential energy decreases when the distance between the floor and the ball increases. This incorrect idea about a form of energy led him to incorrectly identify the energy transformation involved when the ball falls.

Table 7.9 Percentage of correct responses and grade level differences for items aligned to ideas about energy transfer

| Type of energy transfer | Middle school | High school | χ^2 | Sig. |
|-------------------------|---------------|-------------|----------|-------|
| Electrical | 52 % | 56 % | 30.9 | <.001 |
| Radiation | 43 % | 49 % | 174.3 | <.001 |
| Conduction | 39 % | 50 % | 918.7 | <.001 |
| Mechanical | 39 % | 42 % | 18.7 | <.001 |

7.5.7 Energy Transfer

Students' knowledge of four different ways energy can be transferred from one system to another was also tested. The mean Rasch difficulty for the 43 items targeting the idea of energy transfer was -0.17 . Table 7.9 compares the percentage of correct responses from the field test for middle and high school students broken down by type of energy transfer. As with the other topics, the high school students performed better than the middle school students. The largest difference between the grade levels was on items testing students' understanding of conduction (11 percentage points). The smallest gain between grade levels was on items testing ideas about mechanical energy transfer (3 percentage points).

Energy Transferred Mechanically A closer look at the items aligned to ideas about mechanical energy transfer revealed two difficulties that many students had. First, many students did not know that in order for energy to be transferred mechanically there must be a change in position (the push or pull that is required for a mechanical transfer must act over a distance). About 32 % of middle school students and 38 % of high school students selected the answer that states that energy is transferred mechanically whenever one object pushes or pulls on another object even if the objects do not move.

Second, the misconception that both energy and a force are transferred during a mechanical interaction was widespread at all of the grade levels tested (32 % middle school and 43 % high school). Support for this is shown in the option probability curves for one of the items targeting this misconception (see Fig. 7.8). In this item, a black marble rolls across the floor and hits a white marble. After the collision, the black marble stops rolling and the white marble starts to roll. The students were asked what is transferred during this interaction. Answer choice A says that a force is transferred. Answer choice B says that motion energy (kinetic energy) is transferred. Answer choice C says that both a force and motion energy are transferred, and answer choice D says that neither are transferred. As shown in Fig. 7.8, students of low ability (less than -1) were most likely to choose that a force is transferred. The probability of selecting that both a force and energy are transferred (answer choice C) is significant over a wide range of abilities (from -2.5 to 3.5), although students with abilities of approximately 1 and higher were more likely to select the correct answer than this misconception.

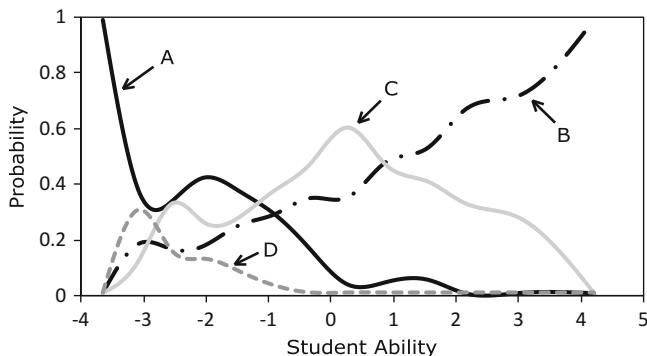


Fig. 7.8 Option probability curves for an item aligned to energy transferred mechanically

The misconception that an object has a force within it, or that a force becomes part of an object when it is thrown or hit, has been documented in previous studies (Fischbein et al. 1989; McCloskey 1983). In our assessment work on the topic of force and motion, we found that middle school students chose this impetus misconception 67 % of the time (AAAS 2013). A similar misconception was revealed by an energy transformation item, where 47 % of the middle school students and 45 % of the high school students thought that the motion energy of a book that has been shoved across a table is transformed into both a force and thermal energy.

Transferring Energy by Radiation Students also had difficulty with the idea that all objects transfer energy by means of electromagnetic radiation, whether the object is in contact with another object or not. Distractors that were aligned to the misconception that only objects that are glowing radiate energy were selected 23 % of the time by the middle school students and 21 % of the time by the high school students. Perhaps this is due to students' linking radiation to visible light and not considering radiation that cannot be detected by the eye. Additionally, some students thought that the objects must be in contact in order to transfer energy by radiation. For one item that asked students to recognize a statement of the general principle of energy transfer by electromagnetic radiation, 29 % of the middle school students and 24 % of the high school students selected the incorrect answer that stated that objects give off energy in the form of electromagnetic radiation at any temperature, but they must be touching each other in order for the energy to be transferred between objects. The idea that energy can be transferred by radiation to objects not in contact was especially difficult for students in an item that asked them to describe what happens to energy in the context of hot food cooling on a counter. The most popular answer choice for this item was that energy was transferred only to the things the food is touching, like the air and the counter, and not to things the food is not touching, like the kitchen walls (40 % of middle school students and 45 % of high school students). In this real-world example of hot food, students seemed to be focused more on conduction than on radiation.

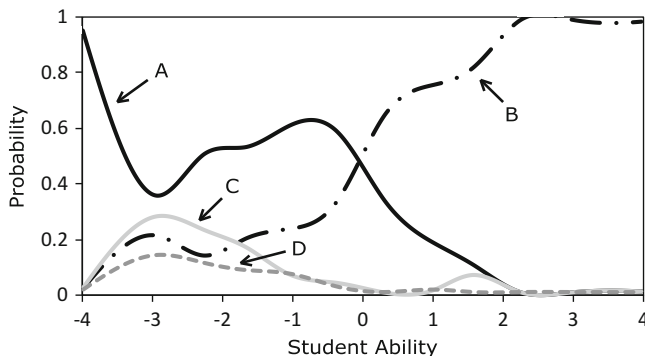


Fig. 7.9 Option probability curves for an item aligned to energy transferred by conduction

Transferring Energy by Conduction With respect to conduction, students performed well and showed the most growth from middle school to high school of all of the energy transfer ideas tested (see Table 7.9). Nevertheless, students at both levels still held several misconceptions. Prior research has shown that one of the most common misconceptions students hold about conduction is that when a cold and a warm object are placed in contact with each other, the warm object gets colder and the cold object gets warmer because “coldness” is transferred from one object to the other (Brook et al. 1984; Clough and Driver 1985; Newell and Ross 1996). We tested the prevalence of this misconception in several items. Overall, the misconception was chosen 31 % of the time by middle school students and 23 % of the time by high school students. The misconception was particularly strong in situations involving frozen objects. For example, Fig. 7.9 shows the option probability curves for an item involving an ice pack and a warm can of juice in a lunch bag. Answer choice A, which corresponds to the “coldness” misconception, was the most popular distractor at all ability levels. Students with an ability level over zero were more likely to choose the correct answer that stated that thermal energy was transferred from the can of juice to the ice pack, but the probability of selecting answer choice A didn’t reach zero until an ability level of around 2. Middle school students selected answer choice A 59 % of the time, and high school students selected the misconception 46 % of the time. The other distractors C and D stated that the can of juice got cold because lunch bags are used to keep food cold and that no energy was transferred, respectively.

Transferring Energy Electrically Over half of the students knew that energy is transferred electrically from an electrical source to an electrical device only when the electrical circuit is complete. The most popular distractors were ones that said that energy could still be transferred when the circuit was not complete. For example, 26 % of the middle school students and 27 % of the high school students thought that a power plant would transfer some energy to a lamp in a house even when the lamp is off and the circuit is not complete. Results from the pilot test

indicated that these students thought that the lamp, like most modern-day electrical devices, uses some energy at all times. Students wrote “a little bit of energy is used while any electronic is plugged in,” “although the lamp uses very little power, as long as it is connected to the socket, power flows,” and “that is why you must unplug your electrical units when you don’t need them.”

7.5.8 Conservation of Energy

Items aligned to the conservation of energy idea had a mean Rasch difficulty of 0.70 and were the most difficult items in the set. The percentage of correct responses to the items aligned to this idea was 28 % for middle school students and 37 % for high school students. These results are consistent with previous research that has indicated that learners do not fully grasp conservation of energy until very late in the developmental progression (Liu and Collard 2005; Liu and McKeough 2005; Neumann et al. 2012). However, on an item that involved identifying a statement of the general principle of conservation of energy, 42 % of middle school and 56 % of high school students answered correctly. These results suggest that the problem lies primarily in not being able to apply the principle of conservation to real-world events. Although many students can recognize a correct statement of the principle of conservation of energy, they are less able to draw upon that basic idea to analyze specific situations where it applies. For example, the most difficult conservation items required students to use the idea of conservation of energy to predict the speed of objects. On one of these items, students were asked to predict (in an idealized environment) the speed of a ball after it goes over a hill on a track in which there is no energy transferred between the track and the ball or between the ball and the air around it. The item required students to recognize that because energy is conserved, and because the heights of the track before and after the hill are equal, the ball must be traveling at the same speed. On three items of this kind, the percent correct was 13 % for middle school students and 19 % for high school students. The additional cognitive load involved in interpreting the scenario, realizing that the conservation principle needs to be used, accepting a situation in which there is no friction, and drawing logical inferences made these items much more difficult than simply recognizing the truth of a general statement about conservation of energy. Similar results have been reported by Chabalengula et al. (2012). Most of the students in their study knew the principle of conservation of energy but were unable to apply it to biological contexts.

Past research has also shown that it is common for students to think that energy can be created or destroyed (Brook and Driver 1984; Kesidou and Duit 1993; Kruger 1990; Loverude 2004; Papadouris et al. 2008; Stead 1980; Trumper 1998). In our sample, distractors involving the creation of energy were chosen 30 % of the time by middle school students and 25 % of the time by high school students. Distractors involving the destruction of energy were chosen 23 % of the time by middle school students and 20 % of the time by high school students.

7.6 Implications for Instruction

The findings from this work do more than simply confirm much of the existing research on students' understanding of energy, or provide a set of assessment items precisely aligned to core ideas about energy. It is our expectation and hope that these findings can be used to inform changes to instruction about energy. To cite just a few examples, our results show how important it is to make clear to students, particularly at the middle school level, the differences between gravitational potential energy and gravity as a force. This can also serve as an opportunity to discuss the differences between forces and energy in general. We also suggest that to help avoid confusion when talking about energy, use of the full phrase "gravitational potential energy" may help to cue students to the idea of the potential energy that a raised object has rather than to the force between two objects. In our study, we found convincing evidence that making this distinction decreased the likelihood that students would think about the force of gravity when asked about energy.

We also found that many students may know general energy principles but not be able to apply the principles to new contexts. This disconnect between teaching scientific principles separate from their application to real-world phenomena is perhaps one of the most significant failings of science education today. To correct this problem, we recommend that energy instruction make the application of general principles to real-world phenomena explicit. Instruction about energy should include a variety of phenomena from different science disciplines so that students have experience applying their energy knowledge in physical, life, and earth science contexts. And students should be asked to explain those specific events of the world in terms of the general principles that apply.

Our work also shows that many students are not aware of the factors that the different forms of energy depend on. This is particularly important in the cases of motion energy and thermal energy, where students are not aware of the mass component. We suggest that instruction include activities that allow the students to examine the effect that changing different variables has on the amounts of the different forms of energy. This is a particularly good opportunity to be explicit about the fact that the energy of an object depends on how much mass the object has, whether it is thermal energy, motion energy, or gravitational potential energy. Some example curriculum units that focus students on discovering the factors or indicators of the different forms of energy are *How can I use trash to power my stereo?* and *Why do some things stop and others continue going?* (Fortus et al. 2005, 2012).

7.7 Moving Forward

Although our work has focused on learning goals from the AAAS *Benchmarks for Science Literacy* (AAAS 1993) and to some extent from the *National Science Education Standards* (National Research Council [NRC] 1996), there is also considerable overlap with the National Research Council's *A Framework for K-12*

Science Education (2012) and the *Energy Literacy Framework* (U.S. Department of Energy 2012). For example, both frameworks include ideas about energy transfer, conservation of energy, and energy degradation. (We have recently started developing assessments aligned to ideas about energy dissipation and degradation.) Places where the overlap is not as strong include ideas about force fields, power, and engineering. As we move forward with our own work on the development of an instrument to test students' understanding of energy concepts from elementary grades through high school, we will consider incorporating a number of these additional ideas recommended in the *Framework for K-12 Science Education* and in the *Energy Literacy Framework*.

7.8 Conclusions

Even after years of instruction, many students have a very limited and unsophisticated understanding of the formal conventions for thinking and talking about energy. For example, students may know that motion energy is related to motion and that thermal energy is related to temperature, but they are unaware that both are also related to mass. They often think that potential energy is related to the potential to have energy; they confuse force and energy, as in the case of the force of gravity and gravitational potential energy; and they say that it is "coldness" that is transferred between two objects at different temperatures rather than thermal energy.

Given the widespread significance of these fundamental ideas about energy, it is critical that students understand them and are able to apply that knowledge in a variety of different contexts. It is also critical that we are able to effectively assess students' difficulties with these energy concepts as well as find out what they do understand. This chapter details our use of precisely aligned, distractor-driven assessment items to investigate students' understanding of energy and the misconceptions and alternative ideas they have that impede their learning. Overall, we found that students' understanding of energy ideas improves steadily from 7th grade to 12th grade. This continuous improvement is most likely attributable to a combination of the formal instruction they are receiving about energy in school, the out-of-school experiences they are having, and their increased ability to interpret the scenarios they are presented with based on their overall cognitive maturation. Certainly, it is encouraging to see growth, but the growth is modest, and there is no point where the improvement begins to accelerate, where students finally "get it." In fact, we found that the high school students held many of the same misconceptions that the middle school students held. Option probability curves revealed that some misconceptions quickly decrease with increasing ability levels while others were popular over wide ranges of abilities. This knowledge of which misconceptions are the most prevalent for which groups of students can be valuable for improving energy education. Additionally, because the items developed for this study were designed to be aligned to key energy ideas but not to any single curriculum or instructional strategy, researchers and curriculum developers will be able to use

the items, along with the field test results described in this chapter, to inform the development and test the effectiveness of various interventions that they use. The items, clarification statements, misconceptions, and field test results are available on the AAAS Project 2061 Science Assessment website (<http://assessment.aaas.org/>).

Acknowledgements The research reported here was supported by the National Science Foundation, through Grants # ESI 0227557 and ESI 0352473, and the Institute of Education Sciences, U.S. Department of Education, through Grant R305A120138 to the American Association for the Advancement of Science. The opinions expressed are those of the authors and do not represent views of the Institute, the U.S. Department of Education, or the National Science Foundation.

References

- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.
- American Association for the Advancement of Science. (2001). *Atlas of science literacy, volume 1*. Washington, DC: AAAS/NSTA.
- American Association for the Advancement of Science. (2007). *Atlas of science literacy, volume 2*. Washington, DC: AAAS/NSTA.
- American Association for the Advancement of Science. (2013). AAAS Project 2061 Science Assessment website. Retrieved October 31, 2012. Available from <http://assessment.aaas.org/>
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah: Lawrence Erlbaum Associates.
- Brook, A., & Driver, R. (1984). *Aspects of secondary students' understanding of energy: Full report*. Leeds: The University of Leeds, Centre for Studies in Science Education and Mathematics Education.
- Brook, A., Briggs, H., Bell, B., & Driver, R. (1984). *Aspects of secondary students' understanding of heat: Full report*. Leeds: The University of Leeds, Centre for Studies in Science Education and Mathematics Education.
- Chabalengula, V. M., Sanders, M., & Frackson, M. (2012). Diagnosing students' understanding of energy and its related concepts in biological context. *International Journal of Science and Mathematics Education, 10*(2), 241–266.
- Clough, E., & Driver, R. (1985). Secondary students' conceptions of the conduction of heat: Bringing together scientific and personal views. *Physics Education, 20*(4), 176–182.
- DeBoer, G. E., Herrmann-Abell, C. F., & 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: 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: Columbia University Teachers College Press.
- Finegold, M., & Trumper, R. (1989). Categorizing pupils' explanatory frameworks in energy as a means to the development of a teaching approach. *Research in Science Education, 19*(1), 97–110.
- Fischbein, E., Stavy, R., & Ma-Naim, H. (1989). The psychological structure of naive impetus conceptions. *International Journal of Science Education, 11*(3), 327–336.

- Fortus, D., Krajcik, J. S., Nordine, J. C., Plummer, J., Rogat, A., & Switzer, A. C. (2005). *How can I use trash to power my stereo?* Ann Arbor: University of Michigan Center for Highly Interactive Classrooms, Curricula, & Computing in Education.
- Fortus, D., Weizman, A., Nordine, J., Jin, H., & Abdel-Kareem, H. (2012). Why do some things stop and others keep going? In J. Krajcik, B. Reiser, D. Fortus, L. Sutherland, & D. Fortus (Eds.), *Investigating and questioning our world through science and technology (IQWST)*. Norwalk: Sangari Active Science.
- Herrmann-Abell, C. F., & DeBoer, G. E. (2009, April). *Using content-aligned assessment to probe middle school students' understanding of ideas about energy*. Paper presented at the National Association for Research in Science Teaching annual conference, Garden Grove, CA.
- Herrmann-Abell, C. F., & DeBoer, G. E. (2010, March). *Probing middle and high school students' understanding of energy transformation, energy transfer, and conservation of energy using content-aligned assessment items*. Paper presented at the National Association for Research in Science Teaching annual conference, Philadelphia, PA.
- Kesidou, S., & Duit, R. (1993). Students' conceptions of the second law of thermodynamics – An interpretive study. *Journal 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.
- Lee, H. S., & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education*, 94(4), 665–688.
- Leggett, M. (2003). Lessons that non-scientists can teach us about the concept of energy: A human-centered approach. *Physics Education*, 38(2), 130–134.
- Linacre, J. M. (2012). Winsteps® (Version 3.75.0) [Computer Software]. Beaverton: Winsteps.com. Retrieved from <http://www.winsteps.com/>
- Liu, X., & Boone, W. J. (2006). Introduction to Rasch measurement in science education. In X. Liu & W. J. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 1–22). Maple Grove: JAM Press.
- Liu, X., & Collard, S. (2005). Using the Rasch model to validate stages of understanding the energy concept. *Journal of Applied Measurement*, 6(2), 224–241.
- Liu, X., & McKeough, A. (2005). Developmental growth in students' concept of energy: Analysis of selected items from the TIMSS database. *Journal of Research in Science Teaching*, 42(5), 493–517.
- Loverude, M. E. (2004, August). *Student understanding of gravitational potential energy and the motion of bodies in a gravitational field*. Paper presented at the Physics Education Research conference, California State University, Sacramento, CA.
- McCloskey, M. (1983). Intuitive physics. *Scientific American*, 248(4), 122–130.
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2012). Towards a learning progression of energy. *Journal of Research in Science Teaching*, 50(2), 162–188.
- Newell, A., & Ross, K. (1996). Children's conception of thermal conduction – Or the story of a woolen hat. *School Science Review*, 78(282), 33–38.
- Nicholls, G., & Ogborn, J. (1993). Dimensions of children's conceptions of energy. *International Journal of Science Education*, 15(1), 73–81.
- 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.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

- Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching*, 35(3), 265–296.
- Solomon, J. (1983). Messy, contradictory and obstinately persistent: A study of children's out-of-school ideas about energy. *School Science Review*, 65(231), 225–229.
- Stead, B. (1980). Energy (Working Paper No. 17). In *Learning in science project*. Hamilton: Science Education Research Unit, University of Waikato.
- Stern, L., & Ahlgren, A. (2002). Analysis of students' assessments in middle school curriculum materials: Aiming precisely at benchmarks and standards. *Journal of Research in Science Teaching*, 39(9), 889–910.
- Summers, M., & Kruger, C. (1993). *Long term impact of a new approach to teacher education for primary science*. Paper presented at the annual meeting of the British Educational Research Association, Liverpool, England.
- Trumper, R. (1990). Being constructive: An alternative approach to the teaching of the energy concept – Part one. *International Journal of Science Education*, 12(4), 343–354.
- Trumper, R. (1993). Children's energy concepts: A cross-age study. *International Journal of Science Education*, 15(2), 139–148.
- Trumper, R. (1998). A longitudinal study of physics students' conceptions of energy in preservice 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.
- U.S. Department of Energy. (2012). *Energy literacy: Essential principles and fundamental concepts for energy education*. Washington, DC: Author.
- Watts, D. M. (1983). Some alternative views of energy. *Physics Education*, 18(5), 213–217.
- Wiser, M. (1986). *The differentiation of heat and temperature: An evaluation of the effect of microcomputer teaching on students' misconceptions* (Technical Report). Cambridge, MA: Harvard Graduate School of Education.
- Wright, B. D., & Stone, M. H. (1999). *Measurement essentials*. Wilmington: Wide Range.
- Wright, B. D., & Stone, M. H. (2004). *Making measures*. Chicago: Phanon Press.