Using Rasch Modeling to Explore Students' Understanding of Elementary School Ideas About Energy

Cari F. Herrmann-Abell and George E. DeBoer

AAAS Project 2061

NARST Annual Conference
Chicago, IL

April 11-14, 2015

Abstract

Energy plays a central role in our society, so it is essential that all citizens understand what energy is and how it moves and changes form. However, research has shown that students of all ages have difficulty understanding these abstract concepts. This paper presents a summary of elementary, middle, and high school students' understanding of elementary school energy ideas. This work is part of a larger project to develop three vertically equated instruments to measure how students make progress in their understanding of the energy concept. The data presented here are from a field test of distractor-driven, multiple-choice items aligned to elementary-level ideas about the forms of energy and energy transfer. These items were tested with students in all three grade bands, even though they explicitly test elementary ideas. A total of 3,037 4th- through 12th-grade students in the U.S. participated, and Rasch modeling was used to analyze the data. Option probability curves were used to represent the distribution of correct answers and misconceptions across the range of student knowledge levels. The shapes of the curves and where they occur as a function of knowledge level suggest that specific misconceptions appear and disappear in sequence as students become more knowledgeable about energy.

Introduction

Energy is a critically important topic in the science curriculum, with many applications in the earth, physical, and life sciences, as well as in engineering and technology. As a result, it is essential that high quality assessments are available to determine what students do and do not know about energy and how those ideas develop throughout the grades. To this end, our research team at Project 2061, the science education reform initiative of the American Association for the Advancement of Science (AAAS), is developing and validating a set of three overlapping and vertically equated assessment instruments—one for the elementary grades, one for the middle grades, and one for high school—to monitor how students progress in their understanding of important ideas about energy. In addition to measuring what students know about energy, the instruments are designed to reveal their misconceptions.

In the *Framework for K-12 Science Education* (National Research Council [NRC], 2012) and the recently released *Next Generation Science Standards* (NGSS Lead States, 2013), "energy" serves as both a disciplinary core idea and a cross-cutting concept. These documents call for students to begin developing their understanding of energy in elementary school. By the end of grade five, students are expected to know that moving objects, sound, light, and heat are indicators of the presence of energy and that burning fuel or food releases energy. They are also expected to know that energy can be transferred from place to place by a variety of mechanisms such as electric currents.

Despite the central role that energy concepts play, they are highly abstract and often counterintuitive and can be challenging to students at all levels. Research on student learning about energy has revealed that students hold a number of misconceptions. For example, some students think that objects at rest have no energy and that energy is only associated with obvious activity or movement (e.g., Brook & Driver, 1984). Some students also think that living things have thermal energy but inanimate objects do not (e.g., Leggett, 2003). As a result, these students may think that it is "coldness" and not thermal energy that is transferred between two inanimate objects (e.g., Newell & Ross, 1996).

This paper focuses on the results of a field test of distractor-driven, multiple-choice assessment items that are aligned to elementary-level ideas about energy. The items were developed using a procedure for evaluating the items' match to specific science ideas and their overall effectiveness as an accurate measure of what students know about those ideas. Common student misconceptions are included as incorrect answer choices (distractors) so that the items can be used to diagnose why students are not selecting the correct answer. During item development, pilot testing was used to obtain written feedback from students on their ideas about energy and what they thought about the items. Then scientists and science education experts reviewed the items using a set of criteria to ensure content alignment and construct validity. After being revised, the items were then field tested on a large national sample to determine their psychometric properties.

Methodology

Field Tests
The data reported on here resulted from a field test of 31 items aligned to elementary-level ideas about energy. Because the ideas are considered to be foundational, they were tested at all three grade levels. Two forms of the test were created, and each form contained a subset of the 31 items. Each student received 22 items, and only students who responded to at least 30% of the items were included in the study. Each item was answered by an average of 1,622 students. Linking items allowed us to use Rasch modeling to compare item characteristics across forms. For each item, the students were asked to choose the one correct answer; students who chose more than one answer choice were marked incorrect.

Participants
The field test sample included 445 elementary school students (grades 4 and 5), 1,952 middle school students (grades 6 through 8), and 640 high school students (grades 9 through 12) from across the U.S. Table 1 summarizes the demographic information for each grade. About 10% of the students indicated that English was not their primary language, and approximately half of the students were female and half were male. The students were studying science but not necessarily physical science at the time of field testing.

Table 1
Demographic Information for Pilot Test Participants

Grade	Total	Female	Male	Primary Language is English	Primary Language is Not English
	% (N)	%	%	%	%
4th	7.8% (236)	46.2%	53.0%	85.2%	6.4%
5th	6.9% (209)	49.8%	45.9%	83.7%	15.8%
6th	20.3% (615)	51.5%	46.7%	89.4%	8.5%
7th	16.0% (487)	45.8%	49.9%	87.7%	7.6%
8th	28.0% (850)	51.4%	44.5%	89.2%	8.5%
9th	7.9% (239)	45.6%	49.0%	76.6%	15.9%
10th	4.6% (139)	61.9%	38.1%	82.7%	15.8%
11th	6.6% (199)	54.3%	43.2%	86.9%	11.1%
12th	2.1% (63)	46.0%	52.4%	77.8%	19.0%
Total	100.0% (3037)	50.1%	46.7%	86.6%	10.0%

Science Ideas
The field test items were aligned to science ideas derived from several documents that articulate and interpret national standards, including *Benchmarks for Science Literacy* (AAAS, 1993), *Atlas of Science Literacy* (AAAS, 2001; 2007), *A Framework for K-12 Science Education* (NRC, 2012), and the *Next Generation Science Standards* (NGSS Lead States, 2013). Note that each of the statements listed below describes elementary-level ideas that are less sophisticated than would normally be expected of middle or high school students. For example, students were tested only on the idea that kinetic energy (motion energy) is related to how fast an object is

moving, not on the idea that kinetic energy is also related to how much mass an object has, which is a middle school idea. The items in this study targeted the following ideas:

- *Kinetic energy*: The amount of energy an object has depends on how fast it is moving.
- *Thermal energy*: The amount of energy an object has depends on how warm it is.
- *Gravitational potential energy*: The amount of energy an object has depends on how high it is above the surface of the earth.
- *Elastic energy*: The amount of energy an elastic object has depends on how much the object is stretched, compressed, twisted, or bent.
- *Chemical energy*: Energy is released when fuel is burned. Energy is also released when food is used as fuel in animals.
- *Conduction*: When warmer things are touching cooler ones, the warmer things get cooler and the cooler things get warmer until they all are the same temperature.
- *Convection*: When air or water moves to another location, it can change the temperature of the air or water at that location.
- *Radiation*: When light shines on an object, the object typically gets warmer.
- *Dissipation*: Objects tend to get warmer when they are involved in energy transfers.

Because these are elementary-level ideas, the names for each form of energy (kinetic, thermal, gravitational potential, etc.) are not used in the items.

Rasch Modeling
The field test data were analyzed using Rasch modeling. 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 achievement level and the difficulty of the item, according to the following equation:

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

where P_{ni} is the probability that student *n* of achievement level B_n will respond correctly to item *i* with a difficulty of D_i (Bond & Fox, 2007; Liu & Boone, 2006). The person and item measures are expressed on the same interval scale, are measured in logits, and are mutually independent, which is not the case for percent correct statistics. In this study, the person measures are estimates of student achievement level for the topic, and item measures are estimates of item difficulty. WINSTEPS (Linacre, 2014) was used to estimate the student achievement levels and the item difficulties.

Distractor analysis. WINSTEPS was also used to generate option probability curves to track the progression of students' misconceptions within the sample. The curves are generated by plotting the probability that students will select each answer choice as a function of their overall level of understanding of the science topic. Typically, with the analysis of traditional multiple-choice items, probability curves are generated dichotomously, that is, for correct versus incorrect answers. The focus in that analysis is on whether or not the student selected the correct answer, and all of the incorrect answer choices are lumped together. The graph of those results shows two sigmoidal curves that cross, with the curve corresponding to the probability of selecting the correct answer typically increasing monotonically with increasing student understanding, and the

curve for the combined set of distractors typically decreasing monotonically with increasing student understanding (Haladyna, 1994).

For distractor-driven items, such as those included in our field test, past research has shown that the curves do not match the monotonic behavior of traditional items (Sadler, 1998). For that reason, we plotted a curve for each answer choice separately. The curves show the probability that students at a particular level of understanding of energy will choose each answer choice. The shape of the curve provides information about what types of students (in terms of the level of their overall understanding of energy) are more likely to select each answer choice, how persistent the misconception represented by the answer choice is, and how the popularity of the answer choice compares to the other answer choices. In other words, students who have a low overall level of understanding of the concept being measured tend to select particular misconceptions, while students with a higher level of understanding tend to select other misconceptions. Thus, the option probability curves provide additional information that is not available when the incorrect answers are lumped together. Identifying hierarchies of misconceptions allows us to diagnose more precisely how students' thinking develops.

Findings

Rasch Fit Statistics and Wright Map
The field test data showed a good fit to the Rasch Model, suggesting a unidimensional set of items targeting the same construct. A summary of the fit statistics is shown in Table 2. The item separation index was 12.80, with a corresponding reliability of 0.99. The person separation index was 1.66, with a corresponding reliability of 0.73. The lower reliability associated with the person separation index (.73) is due to the smaller number of items available to test students at the extreme ends of the scale, especially the higher end (see Figure 1). The standard errors for the field test items were low and ranged from 0.04 to 0.08. The infit mean-square values for all of the items and outfit mean-square values for all but four of the items were within the acceptable range of 0.7 to 1.3 for multiple-choice tests (Bond & Fox, 2007). The point-measure correlation coefficients for the items ranged from 0.17 to 0.54.

Table 2
Summary of Rasch Fit Statistics

	Min	Max	Median
Standard error	0.04	0.08	0.06
Infit mean-square	0.81	1.24	0.93
Outfit mean-square	0.62	1.72	0.86
Point-measure correlation coefficients	0.17	0.54	0.47
Item separation index (reliability)	12.80 (0.99)		
Person separation index (reliability)	1.66 (0.73)		

Figure 1 shows a Wright map comparing student achievement levels and item difficulties. Student achievement is shown on the left side of a vertical line from low achievement at the bottom to high achievement at the top. The spread of item difficulties are shown on the right side of the line ranging from easiest at the bottom of the map to most difficult at the top. The mean of the item difficulties was set at zero. When a student's achievement measure is at the same level on the map as an item's difficulty measure, the student has a 50% chance of answering that item correctly. The map in Figure 1 reveals that for these items and these students, the mean student achievement measure is higher than the mean item difficulty, indicating that the items were, on average, relatively easy for these students. Most of the items cluster around the lower range of student achievement, and there are very few items at the lowest and highest ends of the student distribution. The lack of items at the high-ability end of the scale is not surprising because the items were intended to test elementary school ideas about energy, and middle and high school students were included in the sample.

```
MEASURE   STUDENT - MAP - ITEM
              <more>|<rare>
    4             .  +
                     |
                     |
            .###    |
              .     |
                     |
                     |
    3             T+
            .##   |
            .###  |
              .    |
              .    |
            .####  |  RG19-3
            .####  |
    2             +
       .##########  S|
              .    |
           .######  |T
           .######  |
           .######  |
           .######  |
    1 .###########   +  RG48-2  RG50-2
           .######  |  RG79-2
           .###### M|S RG80-2
           .######  |  RG86-1
      .##########  |  RG58-2
            .####  |  RG21-3  RG29-3
            .#####  |  RG73-2
    0      .#######  +M RG11-3  RG62-2
            .####  |  RG11-4  RG59-2  RG60-2
            #### S|  RG16-2  RG52-2
           .######  |  RG55-2  RG56-2
            .###   |  RG12-2  RG47-2  RG54-2  RG63-2
            .###   |S RG13-2  RG14-3  RG15-2  RG17-2
            .#####  |  RG49-2
   -1         .#   +  RG51-2  RG53-2
              .#   |
              .#   |  RG61-2
              .# T|
               .  |T
                     |
               .  |
   -2         .  +
                     |
               .  |
               .  |
                     |
                     |
                     |
   -3             +
              <less>|<frequent>
```

Each "#" is 20 students. Each "." is 1 to 19 students.

Figure 1. Item – person map showing the distribution of student abilities on the left and item difficulties on the right. Student and item measures are expressed in logits. On the scale on the left, logits from -3 to +4, the zero point is set to match the mean item difficulty. Positive numbers indicate higher achievement/difficulty.

Progression of Difficulty by Idea
Table 3 summarizes the difficulty of the items that were used to test various energy ideas. The items aligned to the different *forms* of energy tended to be easier than the items aligned to ideas about energy *transfer*. This is consistent with past research on students' understanding of energy (Neumann, Viering, Boone, & Fischer, 2012; Liu & Collard, 2005; Liu & McKeough, 2005).

Table 3
Difficulty of Key Ideas as Measured by Field Test Items

Field Test Ideas	No. of Items	Difficulty			
		Min.	Max.	Mean	
Elastic Energy	4	-1.02	-0.29	-0.70	less
Chemical Energy	4	-0.69	-0.30	-0.59	difficult
Kinetic Energy	4	-1.22	0.04	-0.35	|
Dissipation	5	-0.74	0.22	-0.26	|
Thermal Energy	3	-0.47	0.36	-0.18	|
Radiation	2	-0.57	0.29	-0.14	|
Gravitational Potential Energy	4	-0.88	0.99	0.14	↓
Conduction	2	0.67	0.91	0.79	more
Convection	3	0.11	2.24	0.97	difficult

Grade-to-Grade Differences
Overall, students did well on the field test items (see Figure 1). The mean Rasch student measures were 0.42 for the elementary school students, 0.79 for the middle school students, and 0.92 for the high school students. The positive values indicate that, on average, the students' achievement level at all three grand bands is greater than the average item difficulty (which was set at zero). One-way ANOVA revealed statistically significant differences in the means across grade bands ($F = 27.44$, $p < .001$). A Bonferroni post hoc test showed that the difference between the elementary school level and the middle school level was statistically significant on the 0.001 level, and the difference between the middle school level and the high school level was statistically significant on the 0.05 level. There is improvement in understanding these fundamental ideas from elementary to middle to high school, with most of the growth occurring between elementary and middle school.

Distractor Analysis
Option probability curves that plot the probability of selecting an answer choice as a function of Rasch student measure were generated for each item. These curves present a visual image of the distribution of correct answers and misconceptions across the spectrum of student achievement (ranging from fourth to twelfth grade). When analyzing the option probability curves, we look for specific characteristics: (1) where the probability of selecting the answer choice peaks on the achievement spectrum, (2) the width of the peak, and (3) the height of the peak compared to other answer choices. Here we present option probability curves for five items.

Chemical energy. Figure 2 shows the option probability curves for an item assessing students' understanding that an animal gets the energy it needs to run from the food it eats. The correct answer choice (C) states that the animal gets energy by using food as fuel. Answer choice A corresponds to the misconception that animals make new energy while they sleep (Mann, 2010).

Answer choice B says that the animal gets energy by absorbing light from the sun and answer choice D says that the animal makes new energy while it runs.
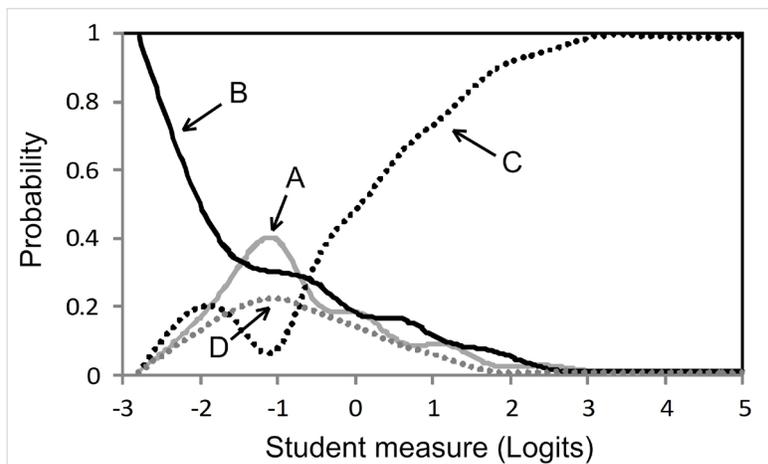
Figure 2. Option probability curves for an item aligned to chemical energy

The shape of the curve for answer choice B suggests that students with little understanding of energy (low student measure) think that animals get energy from the sun, but this misconception decreases rapidly as students' overall understanding of energy increases. The curves for answer choices A and D, both of which correspond to misconceptions about creating energy, also peak at low achievement levels and then gradually fall off as knowledge of energy increases. Students who select answer choices A (animals make new energy while they sleep) or D (animals make new energy when they run) may be thinking about their experiences of feeling more "energetic" after sleeping or while running, and associate this feeling with having more energy. As their understanding of energy progresses, the probability of selecting these misconceptions decreases, and the correct answer becomes the most probable answer choice selection.

Gravitational potential energy. Figure 3 shows the option probability curves for an item that asks students whether an object will have more energy when it is held three feet above the ground or when it is held six feet above the ground. The correct answer (B) states that the object will have more energy when it is six feet above the ground because the higher the object is, the more energy it has. Answer choice A (the object will have more energy at three feet above the ground) corresponds to the misconception that the closer an object is to the ground the more energy it has. Answer choice C (the object will have the same amount of energy when it is three feet or six feet above the ground) corresponds to the misconception that the amount of energy an object has does not depend on how high it is. Answer choice D (the object doesn't have any energy) corresponds to the well-documented misconception that energy is not associated with inanimate objects (e.g. Finegold & Trumper, 1989; Stead, 1980; Watts, 1983).

Students with the lowest level of understanding are most likely to select answer choice A. As the probability of selecting answer choice A decreases, the probability of selecting answer choice C increases and then plateaus, indicating that students at a wide range of achievement levels think that the amount of energy an object has does not depend on how high it is above the ground. At a Rasch student measure of about 0, the correct answer becomes the most likely answer choice to

9

be selected. The answer choice (D) stating that objects don't have any energy peaks around -2 but is less popular than the other answer choices, suggesting that most of students participating in the field test think that objects do have energy.
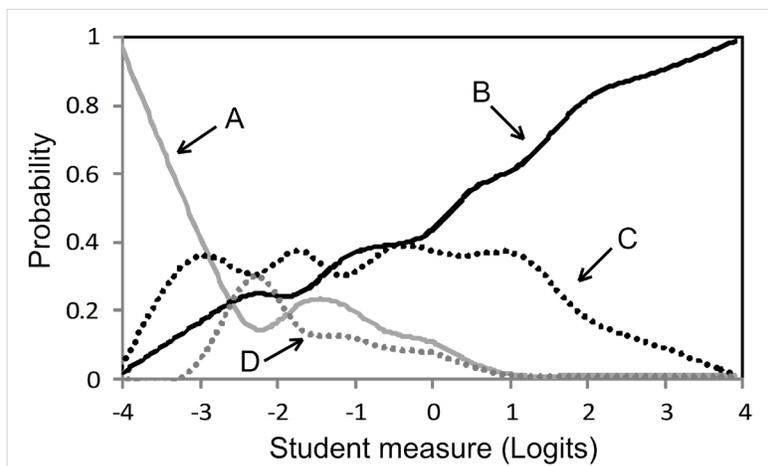
Figure 3. Option probability curves for an item aligned to gravitational potential energy

Conduction. The option probability curves shown in Figure 4 are for an item that asks about the temperature changes that will occur when a hot cookie is placed on a cool plate. The correct answer (A) states that the cookie will get cooler and the plate will get warmer until they are both at the same temperature. Answer choice B says that the cookie will get cooler and the plate will get warmer, but they will never be the same temperature. Answer choice C says that the plate will stay the same temperature and the cookie will get cooler until it is the same temperature as the plate. Answer choice D says that the plate will stay the same temperature and the cookie will get cooler, but they will never be the same temperature. Past research has shown that some students think that objects that are made of different materials and are at different temperatures will never reach the same temperature when they are brought in contact with each other (Erickson & Tiberghien, 1985; Kesidou & Duit, 1993; Wiser, 1986). It is likely that students think this because the objects feel like they are at different temperatures.

The option probability curves show that students with a low level of energy understanding (low student measure) are drawn to answer choices C and D which reflect the misconception that the plate would stay the same temperature and the cookie would get cooler. From the students' real world experience, they know that cookies cool when allowed to sit, but they are much less likely to recognize that the plate is also changing temperature. The intersection of the curves at the mid-range of achievement indicates that students with a mid-level understanding are equally likely to select any of the answer choices. This suggests that students at this level do not have strong conceptions about what happens when a warmer object is placed in contact with a cooler object. As students gain a better understanding of energy, they begin to understand what happens during conduction and are, therefore, most likely to select the correct answer that the plate will get warmer and the cookie will get cooler until they are both at the same temperature.
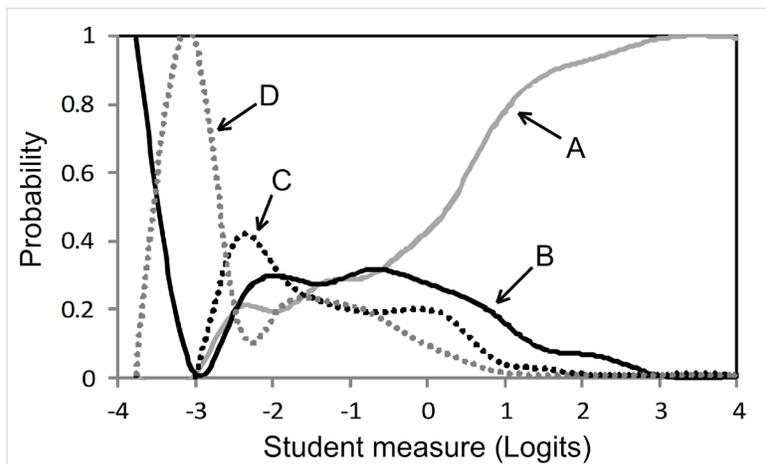
Figure 4. Option probability curves for an item aligned to conduction

Convection. The item for which option probability curves are shown in Figure 5 asks students to consider what would happen to the temperature at the other end of a room if a fan blows warmer air across the room. The correct answer (C) states that the warmer air being blown across the room makes the other end of the room warmer. Answer choice A says that fans can only cool a room, not warm a room, and answer choice B says that that the fan blowing warmer air will not change the temperature at the other end of the room. Answer choice D corresponds to the misconception that only very warm air can change the temperature of a location.
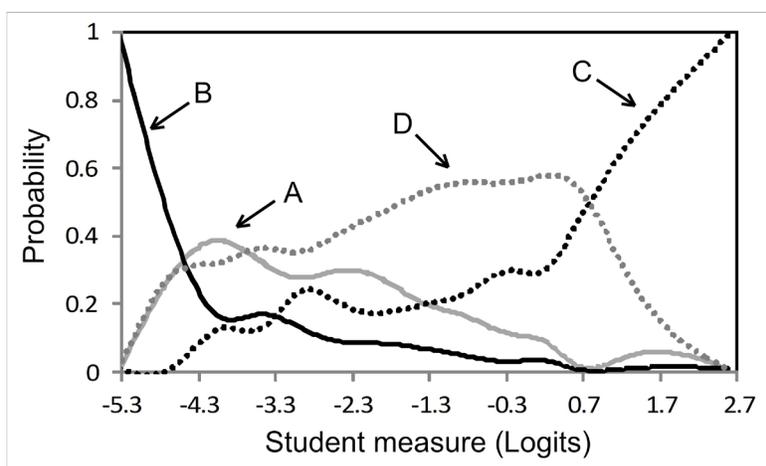
Figure 5. Option probability curves for an item aligned to convection

Students with a low overall level of understanding of energy (low student measure) are most likely to select answer choices A (fans can only cool) or B (the warmer air will not change the temperature of the room). The probability of selecting answer choice B is large for the least knowledgeable students but decreases quickly as overall understanding of energy increases. The probability of selecting answer choice A is higher for students with a Rasch student measure around -4 and decreases slowly as understanding increases, perhaps because students are more familiar with using fans to cool a room. As students gain an understanding of energy, they are more likely to select the answer choice that only very warm air will warm the other side of the

room (D), perhaps thinking that changes in temperature from the movement of slightly warmer air would be insignificant. This is a popular answer for a wide range of students and suggests that their everyday experiences have a significant impact on their development of a more formal understanding of science ideas. As students' understanding of energy improves, they realize that the temperature at the other side of the room has to increase at least a little bit even if the air that is blown across the room is just slightly warmer.

Dissipation. The option probability curves shown in Figure 6 are for an item that asks whether or not a remote control and the wheels of a remote controlled car get warmer after a child plays with them for an hour. The correct answer choice (C) states that the remote control will be warmer because it uses a battery, and the wheels will be warmer because they rubbed against the floor. Answer choice A says the remote control will be warmer but the temperature of the wheels will not change, and answer choice B says the wheels will get warmer but the temperature of the remote control will not change. Answer choice D states that neither will be warmer.
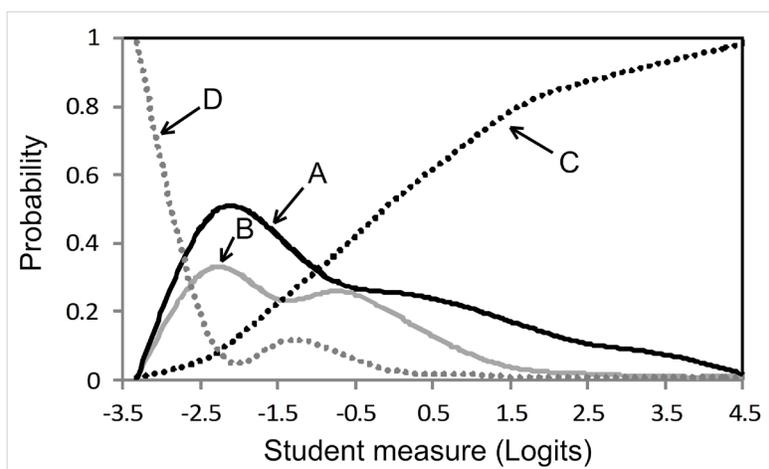
Figure 6. Option probability curves for an item aligned to dissipation

Answer choice D is the most probable answer choice selection for students with little understanding of energy (low student measure). Perhaps these students select this option because based on their real-world experiences, the temperature changes are often too small to be noticeable. The probability of selecting this answer choice decreases quickly as students become more knowledgeable about energy. In the mid-range of student achievement, students tend to focus on either the remote control or the wheels getting warmer, but not on both (answer choices A and B). At around a Rasch student measure of -1, the correct answer becomes the most probable answer choice selection. For this item, students' understanding progresses from thinking that neither object gets warmer, to one object gets warmer, to both objects get warmer.

Conclusions

Rasch modeling was used to analyze the results from a set of 31 multiple choice assessment items aligned to elementary level ideas about energy. Overall, we found that students' understanding of these energy ideas improves from elementary to high school, with the largest improvement occurring between elementary and middle school. The items testing ideas about the different forms of energy were, on average, easier than the items that tested ideas about energy transfer.

Option probability curves were generated for each item to investigate the order in which misconceptions appear and when they are likely to disappear as a function of overall student achievement with regard to understanding the energy concept. The curves for some items showed that some misconceptions rapidly decrease with increasing achievement levels while other misconceptions persist across a wide range of levels. The nature of the misconceptions and where they appear and disappear along the achievement continuum also provide insights into how students may be using everyday experiences as they form their ideas about energy. Classroom teachers and curriculum developers can use this information when selecting and sequencing instructional activities. Since instructional time is limited, educators need to select activities that support students' construction of the science ideas while contradicting the most popular and persistent misconceptions. The option probability curves can show which misconceptions are popular and persistent for which students.

Acknowledgements

References

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

American Association for the Advancement of Science. (2001). *Atlas of science literacy* (Vol. 1). Author.

American Association for the Advancement of Science. (2007). *Atlas of science literacy* (Vol. 2). Author.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Brook, A., & Driver, R. (1984). *Aspects of secondary students' understanding of energy: Full report*. Leeds, UK: The University of Leeds, Centre for Studies in Science Education and Mathematics Education.

Erickson, G., & Tiberghien, A. (1985). Heat and temperature. In R. Driver, E. Guesne, & A. Tiberghien (Eds.), *Children's ideas in science* (pp. 52–84). Philadelphia, PA: Open University Press.

Haladyna T. M. (1994). *Developing and validating multiple-choice test items*. Hillsdale, NJ: Erlbaum.

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.

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.

Liu, X., & Boone, W. J. (2006). Introduction to Rasch measurement in science education. In X. LiuW. J. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 1-22). Maple Grove, MN: 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.

Linacre, J. M. (2014). *WINSTEPS Rasch measurement computer program* (Version 3.81.0). Chicago: Winsteps.com.

Mann, M., & Treagust, D. F. (2010). Students' conceptions about energy and the human body. *Science Educational International, 21*(3), 144-159.

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.

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.

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.

NGSS Lead States. (2013). *Next generation science standards: For states, by state*s. Washington, DC: The National Academies 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.

Wiser, M. (1986). *The differentiation of heat and temperature: An evaluation of the effect of microcomputer teaching on students' misconceptions*. Cambridge, MA: Harvard Graduate School of Education.