Using Rasch Modeling to Investigate a Learning Progression for Energy Ideas

Cari F. Herrmann-Abell and George E. DeBoer

AAAS Project 2061

Paper presented at the

2016 NARST Annual International Conference

Baltimore, MD

April 14-17, 2016

Abstract

Energy is a core concept in the teaching of science. Therefore, it is important to know how students' thinking about energy develops so that elementary, middle, and high school students can be appropriately supported in their understanding of energy. This study tests the validity of a proposed theoretical model of students' growth of understanding about energy that moves from a phenomenological understanding, to being able to explain phenomena using basic energy-related concepts, to being able to explain phenomena using more advanced energy-related concepts, often involving atomic/molecular explanations. The study examines results from the administration of 372 distractor-driven, multiple-choice test items aligned to a wide range of energy ideas from energy forms and transformations, to energy transfers, to energy dissipation and degradation, to energy conservation. Over 20,000 students from across the U.S. participated in the study. Rasch modeling provided linear measures of student performance and item difficulty. For most of the 14 targeted energy ideas, an analysis of the item difficulties validated the study's proposed theoretical model of the growth of understanding of the energy concept. Additionally, a cross-sectional analysis of student performance revealed that the high school students outperformed the middle school students and the middle school students outperformed the elementary school students.

Introduction

Energy is a critically important topic in the K-12 science curriculum, with many applications in the earth, physical, and life sciences and in engineering and technology. Therefore, it is essential that high quality assessments are available to determine what students do and do not know about energy and how their ideas develop throughout the grades. To this end, a team of assessment researchers is developing and validating a set of three 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 energy from late elementary school through high school and the misconceptions they may have.

A number of studies have investigated students' learning progressions for the energy concept (Liu & Mckeough, 2005; Liu & Collard, 2005; Lee & Liu, 2010). Liu and Mckeough (2005) used the responses from three populations of students from the U.S. (3rd & 4th graders, 7th & 8th graders, and 12th graders) to 27 multiple-choice and short-answer items from the Third International Mathematics and Science Study (TIMSS) database. In a follow up study, Liu and Collard (2005) administered three performance assessments to 67 students from one 4th grade class, one 8th grade class, and one high school physics class in the U.S. Lee and Liu (2010) selected 8 multiple-choice items and two explanation items from item sets released by TIMSS and the National Assessment of Educational Progress (NAEP) and tested them with 2,688 middle school students from across the U.S. These studies all support a progression in which students first perceive energy as activity, or the ability to do work. As students' understanding grows, they begin to distinguish different energy sources and forms of energy. Next comes an understanding of energy transfer, followed by an awareness of energy degradation. Finally, by the end of the progression, students are able to accept the highly abstract idea of conservation of energy.

Most previous research focuses on comparing the relative difficulties of these different conceptual categories about energy. More recently, researchers have been investigating students' growth of understanding within each of these conceptual categories (forms, transformations, transfer, conservation, etc.). For example, Neumann, Viering, Boone, & Fischer (2013) designed an assessment that not only tested the progression of the concepts but also tested a progression of complexity within each concept. The complexity progression starts with facts, then moves to simple connections, to qualified relationships, and finally to complex concepts. They administered this assessment to 1,856 German students in 6th through 10th grades. Their results did not support the proposed increase in complexity but rather that students typically make progress in understanding aspects of multiple energy ideas simultaneously, not mastering one idea before making progress on another. This model of an overlapping progression is supported by the *Next Generation Science Standards* (NGSS Lead States, 2013), which expects students to learn various aspects of the energy concept starting at grade four. The learning progressions that are assumed in the work being reported in this paper are consistent with this overlapping model, with basic ideas about forms, transfers, etc. being introduced in the elementary grades and increasing in sophistication throughout the grades.

Our work builds upon prior research in a number of ways. First, the learning progressions we propose and the 372 items we developed cover a more extensive and coherent set of important energy ideas including multiple forms of energy and multiple mechanisms of energy transfer. This enables us to get a more complete and detailed picture of students' progression of understanding of energy. Second, we collected data from a very large national sample of students in grades 4 through 12 (N = 21,061). This gives us a more comprehensive description than most studies do of how students' understanding of energy changes from grade-to-grade. Third, the items are designed not only to test for the correct scientific understanding but also to probe for common student misconceptions about energy. Incorporating these misconceptions into the distractors of the items gives students plausible answers to select from. This decreases the likelihood of guessing, thus giving us a more valid measure of what they actually know.

This paper presents a study of the validity of the overall proposed progression of understanding of energy ideas from forms of energy through transfer, dissipation and degradation, and finally conservation. The paper also presents data that tests students' progression of understanding within each of these categories.

Methodology

Energy Ideas. According to Duit (2014), there are four basic categories that comprise the energy concept: (1) energy forms and transformation, (2) energy transfer, (3) energy dissipation and degradation, and (4) energy conservation. For this study, the first two categories were partitioned into finer-grained ideas about the specific forms of energy and specific mechanisms of energy transfer (see Table 1). To guide the development of test items and to precisely describe the progression of understanding that was theorized for each idea, clarification statements were written to make the boundaries around the targeted knowledge explicit and to spell out the expectations for students at each level of the progression. These clarification statements were informed by several documents that present and interpret national science standards, including *Benchmarks for Science Literacy* (American Association for the Advancement of Science [AAAS], 1993), *Atlas of Science Literacy* (AAAS, 2001; 2007), *A Framework for K-12 Science Education* (NRC, 2012), and *Next Generation Science Standards* (NGSS Lead States, 2013).

Table 1
Energy Ideas Targeted by the Assessment Items

Ideas about the Forms of Energy	Ideas about Energy Transfer	Other Energy Ideas
Kinetic Energy	Conduction	Energy Conservation
Thermal Energy	Convection	Energy Dissipation & Degradation
Gravitational Potential Energy	Radiation	
Elastic Potential Energy	Transferring Energy by Forces	
Chemical Energy	Transferring Energy Electrically	
Energy Transformations	Transferring Energy by Sound	

Within each of the target ideas, three levels were defined, which in most cases correspond to the expectations at the three grade bands (elementary, middle, and high school). The levels progress from (1) a phenomenological understanding of energy-related events in the world, to (2) using basic energy concepts to explain phenomena, to (3) using more advanced energy concepts to explain phenomena, often requiring an atomic/molecular model of energy phenomena. For example, the learning progression for conduction starts with the expectation that students understand that when a warmer object is placed in contact with a cooler object, the warmer object will get cooler and the cooler object will get warmer. At the next level, students are expected to know that this phenomenon occurs because energy is transferred from the warmer object to the cooler one. At the highest level, students are expected to know that this energy is transferred by the random collisions of atoms and molecules that make up the objects. For gravitational potential energy, the progression starts with the expectation that students understand that the higher an object is above the earth, the more energy it has. At the next level, students are expected to know that gravitational potential energy of an object near the surface of the earth depends on the distance above the earth and the mass of the object. At the highest level, students are expected to know that gravitational potential energy is associated with the separation of mutually attracting masses. Descriptions of the progressions of understanding for each idea tested in this study are presented in Table 2. Note that for the transferring energy electrically idea, there are only two levels of understanding, and for the energy transformations idea, there is only one level.

Item Development. A total of 372 distractor-driven, multiple-choice items were tested with students from across the United States. Table 3 presents the number of items aligned to each energy idea by level. The table does not include the 13 items that were aligned to more than one energy transfer idea.

Item construction followed rigorous item development procedures that included (1) the identification of documented misconceptions, which were then used as distractors; (2) a careful evaluation of the items' alignment to the targeted ideas about energy; and (3) a close examination of the items for their overall psychometric effectiveness (DeBoer, Herrmann-Abell, & Gogos, 2007; DeBoer, Herrmann-Abell, Gogos, Michiels, Regan, & Wilson, 2008; DeBoer, Lee, Husic, 2008). The inclusion of misconceptions in the distractors increases the diagnostic power of the items by providing information about students' alternative ideas in addition to what science ideas they do and do not know (Sadler, 1998). Item alignment was determined using two criteria: *necessity* ensures that the targeted energy idea is needed to evaluate the answer choices, and *sufficiency* ensures that the targeted energy idea is enough by itself (Stern & Ahlgren, 2002). Careful alignment increases the validity of the inferences that can be made about what students know. Additionally, items were reviewed by a panel of scientists and science education experts to remove or reduce any construct irrelevant features such as issues with comprehensibility, test-wiseness, and inappropriate task contexts. We used Rasch modeling throughout the item development process to monitor the items' psychometric properties.

Table 2
Proposed progression of understanding for energy ideas

Energy Idea	Phenomenological level	Concept level	Advanced concept level
Kinetic Energy	The amount of energy an object has depends on how fast it is moving.	The kinetic energy (motion energy) of an object depends on the speed and the mass of the object.	Kinetic energy (motion energy) is proportional to the mass of a moving object and increases rapidly with increasing speed.
Thermal Energy	The amount of energy an object has depends on how warm it is.	The thermal energy of an object depends on the temperature and the mass of the object and the material of which the object is made.	The thermal energy of an object depends on the disordered motions of its atoms or molecules and the number and types of atoms or molecules of which the object is made.
Gravitational Potential Energy	The amount of energy an object has depends on how high it is above the surface of the earth.	The gravitational potential energy of an object near the surface of the earth depends on the distance the object is above the surface of the earth (or an alternate reference point), and the mass of the object.	Gravitational potential energy is associated with the separation of mutually attracting masses.
Elastic Potential Energy	The amount of energy an elastic object has depends on how much the object is stretched, compressed, twisted, or bent.	The elastic potential energy of an elastic object depends on the amount the object is stretched or compressed and how difficult it is to stretch or compress the object.	The amount of elastic potential energy stored in an elastic object increases when the object is stretched or compressed because stretching and compressing an object changes the distances between the atoms and molecules that make up the object.
Chemical Energy	Energy is release when fuel is burned. Energy is also released when food is used as fuel in animals	Some chemical reactions release energy into the surroundings, whereas other chemical reactions take in energy from the surroundings	Chemical energy is associated with the arrangement of atoms that make up the molecules of the reactants and products of a chemical reaction. Because the arrangement of atoms making up the molecules is different before and after the chemical reaction takes place, the amount of chemical energy in the system is also different.
Energy Transformations		Most of what goes on in the universe—from exploding stars and biological growth to the operation of machines and the motion of people—involves some form of energy being converted into one or more other forms of energy.	

Table 2 continued
Proposed progression of understanding for energy ideas

Energy Idea	Phenomenological level	Concept level	Advanced concept level
Transferring Energy by 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.	Conduction is the transfer of energy that occurs when a warmer object (or sample of matter) comes in contact with a cooler object (or sample of matter) without a transfer of matter.	Energy is transferred by conduction through a material by the random collisions of atoms and molecules that make up the material.
Transferring Energy by Convection	When air or water moves to another location, it can change the temperature at that location.	Temperature variations in fluids such as air and water lead to currents that circulate the fluid and transfers energy from place to place in the fluid	In a fluid, regions that have different temperatures have different densities. The differences in density lead to an imbalance between the downward gravitational force and upward (buoyant) forces exerted by the surrounding fluid, creating currents that contribute to the transfer of energy.
Transferring Energy by Radiation	When light shines on an object, the object typically gets warmer.	Light transfers energy from a light source to a receiver.	Energy can be transferred by electromagnetic radiation.
Transferring Energy by Forces	Pushes and pulls can transfer energy from one object to another resulting in a change in the objects' motion.	Energy is transferred mechanically whenever an object exerts a force, either by contact or at a distance, on another object that changes the objects' position or shape.	When two objects change relative position as a result of a gravitational, magnetic, or electric force, the potential and kinetic energies of the system change.
Transferring Energy Electrically	Energy can be transferred electrically when an electrical source is connected in a complete circuit to an electrical device.		Electrostatic potential energy can be stored in the separation of charged objects.
Transferring Energy by Sound	Sound can transfer energy from one location to another.	Energy can be transferred by sound when a vibrating object produces sound that travels through a medium to a receiver.	Energy is transferred by sound because of coordinated collisions between the atoms or molecules that make up the medium through which the sound travels.
Energy Dissipation & Degradation	Objects tend to get warmer when they are involved in energy transfers.	Transformations and transfers of energy within a system usually result in some energy being released into its surrounding environment causing an increase in the thermal energy of the environment.	Unless prevented from doing so, energy will become uniformly distributed.
Conservation of Energy	Everything has energy.	A decrease in energy in one object or set of objects always is accompanied by an increase in energy in another object or set of objects.	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.

Table 3
Item Count by Level of Progression for Each Idea

Energy Category	Energy Ideas	Number of Items Level 1	Level 2	Level 3
Forms of Energy	Kinetic Energy	5	27	8
	Thermal Energy	3	19	18
	Gravitational Potential Energy	6	23	6
	Elastic Potential Energy	4	11	3
	Chemical Energy	4	16	8
	Energy Transformations		29	
Energy Transfer	Conduction	4	18	4
	Convection	3	7	7
	Radiation	3	10	13
	Transferring Energy by Forces	4	13	6
	Transferring Energy Electrically	2		9
	Transferring Energy by Sound	2	3	7
Conservation of Energy		5	5	23
Energy Dissipation & Degradation		6	10	5

Field Tests. 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 items. The items were divided into 25 different test forms, ten for elementary students (Grades 4 and 5) and 15 for secondary students (Grades 6 through 12). The elementary forms of the field test included either 23 or 24 items, and the secondary forms included either 31 or 32 items. Linking items were used so that item characteristics could be compared across forms. Each item was answered by an average of 1,605 students. Items were field tested in May and June of 2015 in both online and paper-and-pencil formats. Students were given one class period to complete the test.

Participants. A total of 21,061 students participated in the field test but only students who responded to six or more items were included in the study (N=20,870). Students with highly unexpected responses were also excluded as described below in the Findings section. This made the final sample size 20,551 students. Table 4 provides a breakdown of the demographic information by grade level. The sample included students from schools in 42 different states across the U.S. and Puerto Rico. Elementary students (Grades 4 and 5) made up 14% of the sample, middle school students (Grades 6 through 8) 50%, and high school students (Grades 9 through 12) 36%. All of the students were studying science but not necessarily physical science at the time of testing.

Table 4
Demographic Information for Field Test Participants

	Elementary	Middle	High	Total
Grades	4-5	6-8	9-12	4-12
Number of Students	2967 (14%)	10207 (50%)	7377 (36%)	20551
Gender				
Male	48%	49%	46%	48%
Female	50%	48%	55%	50%
Ethnicity				
White	38%	48%	44%	45%
Asian	7%	4%	7%	5%
Black	17%	11%	10%	11%
Hispanic	17%	19%	22%	20%
Two or more ethnicities	10%	10%	11%	11%
Primary language				
English	87%	88%	85%	87%
Other	11%	9%	13%	11%

Rasch Modeling. WINSTEPS (Linacre, 2016) was used to estimate Rasch student and item measures. In the dichotomous Rasch model (Rasch, 1960/1980), the probability that a student will respond to an item correctly is determined by the difference in the student's performance level and the item's difficulty, 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 performance level B_n will respond correctly to item *i* with a difficulty of D_i (Bond & Fox, 2007; Liu & Boone, 2006; Boone, Staver, & Yale, 2014). When the data fit the Rasch model, the student performance level and item difficulties are (1) expressed on the same interval scale, (2) mutually independent, and (3) measured in log odds or logits, which can vary from $-\infty$ to $+\infty$. In our study, the average item difficulty was set at zero. Item difficulties above zero are more difficult, and item difficulties below zero are less difficult.

Findings

Rasch Fit. Initial analysis of the fit statistics showed that there were 10 items with outfit mean-square values outside of the acceptable range of 0.7 to 1.3 (Bond & Fox, 2007). The outfit statistic was used because it is unweighted and, therefore, sensitive to outliers. An investigation of the student response patterns for these items was conducted starting with the item with the highest outfit mean-square value. Data from students with highly unexpected responses, as indicated by a large Z-residual statistic greater than or equal to three, were removed from the data set. After removing these misfitting students, the total number of students was 20,551. The final fit analysis showed that all of the items were within the acceptable range for both infit and outfit indices. Table 5 summarizes the fit statistics for both the items and the students. The reliability of the item measures was 0.99 and the item separation index was 11.67. The reliability of the student measures was 0.66 and the item separation index was 1.40. This lower separation index and reliability for the student measures can be explained by the fact that each student responded to only a small percentage of the items in the item bank(about 7%), due to our use of matrix sampling during field testing. Therefore, there is less information available to estimate the student measures, which results in a lower reliability and higher standard errors. In contrast, differences in difficulty level of the items were easier to determine because such a large number of students responded to each item.

Table 5
Summary of Rasch Fit Statistics

	Item			Student		
	Min	Max	Median	Min	Max	Median
Standard error	0.02	0.11	0.06	0.37	1.93	0.40
Infit mean-square	0.84	1.27	0.99	0.44	2.17	0.99
Outfit mean-square	0.72	1.33	0.99	0.23	5.15	0.97
Point-measure correlation	0.00	0.53	0.34	-0.13	0.56	0.32
Separation index (reliability)	11.67 (0.99)			1.40 (0.66)		

Progression of Student Performance by Grade Band. ANCOVA was used to perform a cross-sectional analysis of the students' performance by grade band controlling for gender, ethnicity, and whether or not English was their primary language. To control for differences in instructional focus across the country, the researchers also controlled for the state students came from. Table 6 presents the F-ratios and degrees of freedom for grade band and each covariate.

Table 6
Results from the ANCOVA

Source	df	F	p
Grade band	2	395.54	<.001
Gender	1	11.00	<.01
Ethnicity	1	100.38	<.001
English as primary language	1	170.25	<.001
State	1	192.59	<.001
Error	19789		

The estimated marginal mean student performance was -0.54 for the elementary school students, -0.46 for the middle school students, and -0.17 for the high school students (see Table 7). Using the score-to-measure table generated by Winsteps, these measures correspond to a raw score of 142 out of 372 or 38% percent correct for elementary school students, 149 out of 372 or 40% for middle school students, and 172 out of 372 or 46% for high school students. A Bonferroni post hoc test showed that high school students performed significantly better than middle school students, and middle school students performed significantly better than elementary school students. Overall, the items were relatively difficult for this sample of students as indicated by the negative mean measures.

Table 7
Estimated Marginal Student Means by Grade Band

	Mean Student		95% Confidence Interval	
Grade band	Measure	Std. Error	Lower Bound	Upper Bound
Elementary	-0.54	.014	-.67	-.51
Middle	-0.46	.008	-.47	-.44
High	-0.17	.009	-.18	-.15

Progression of Difficulty by Energy Idea. To investigate the progression of item difficulty for the energy ideas, the average Rasch difficulties of items aligned to each idea were calculated (see Table 8). One-way ANOVA revealed statistically significant differences in the means of the 14 ideas, $F(13, 345) = 3.44$, $p < .001$. Bonferroni post hoc tests showed that the chemical energy items were significantly more difficult than the elastic potential energy items, the radiation items, and the kinetic energy items, and the conservation items were significantly more difficult than the items aligned to elastic potential energy, radiation, kinetic energy, and thermal energy.

Table 8
Difficulty of Energy Ideas as Measured by Field Test Items

	Energy	# of	Rasch Difficulty				
Energy Ideas	Category	Items	Min.	Max.	Mean	SD	
Elastic Potential Energy	Forms	18	-2.09	1.45	-0.45	0.88	less
Radiation	Transfer	26	-1.46	1.33	-0.32	0.62	difficult
Kinetic Energy	Forms	40	-1.49	2.07	-0.23	0.86	|
Thermal Energy	Forms	40	-1.31	0.76	-0.17	0.52	|
Energy Transformations	Forms	29	-0.80	0.72	-0.09	0.45	|
Gravitational Potential Energy	Forms	35	-1.48	1.20	0.00	0.62	|
Dissipation & Degradation	Diss/Deg	21	-1.52	1.44	0.00	0.81	|
Transferring Energy by Sound	Transfer	12	-0.57	0.79	0.01	0.45	|
Transferring Energy by Forces	Transfer	23	-1.26	1.20	0.04	0.80	|
Conduction	Transfer	26	-1.19	1.83	0.08	0.72	|
Convection	Transfer	17	-0.51	2.24	0.26	0.73	|
Transferring Energy Electrically	Transfer	11	-0.95	1.12	0.34	0.64	↓
Chemical Energy	Forms	28	-1.56	1.62	0.38	0.83	more
Conservation	Cons.	33	-1.01	1.97	0.50	0.82	difficult

When the items are grouped into the four basic conceptual categories that have been identified by others (see for example Duit, 2014), our analysis does not indicate a statistically significant progression of difficulty from energy forms and transformation, to energy transfer, to energy dissipation and degradation, to energy conservation as suggested by previous research (see Table 9). However, when the 14 energy ideas are rank ordered by difficulty (see Table 8), it is clear that items testing the forms of energy ideas tend to be easiest, items testing the transfer ideas next, and conservation items the most difficult. It is only the dissipation and degradation items that are easier than was predicted. This could be due to how this idea was defined. The first two levels of the progression for this idea deal with dissipation, progressing from an understanding that objects tend to get warmer when involved in energy transfers to an understanding that energy is released to the surroundings during energy transfers and is transformed into thermal energy. The third level introduces ideas about degradation (unless prevented from doing so, energy will become more uniformly distributed) but does not include an understanding of entropy. Neumann et al. (2013) argue that the dissipation ideas overlap in difficulty with transformation and transfer ideas, and that degradation and entropy ideas should be a separate level in the progression. If we separate our items into two groups (items targeting dissipation ideas and items targeting degradation ideas), we find that the dissipation items have an average difficulty of -0.15 and standard deviation of 0.82 and the degradation items have an average difficulty of 0.50 and standard deviation of 0.59. The dissipation items are in the same difficulty range as the energy transformation and transfer items, and the degradation items are in the same difficulty range as the conservation items. This supports the notion that degradation, but not dissipation, comes later in the progression of understanding energy.

Table 9
Item Difficulty by Energy Category

Energy Category	# of Items	Mean Rasch Difficulty	SD
Energy Forms and Transformation	190	-0.09	0.73
Energy Transfer	128	0.00	0.70
Energy Dissipation and Degradation	21	0.00	0.81
Energy Conservation	33	0.50	0.82

Progression of Difficulty by Level within Each Idea. To explore the validity of progressions of understanding within each idea, correlations between the item difficulty and level were calculated (see Table 10). Kendall's tau correlation coefficients showed a statistically significant correlation between difficulty and level for most ideas, which supports the validity of the progression as defined in Tables 2 and 3. As expected, items that required a phenomenological explanation were easiest, energy concept-based explanations next, and more advanced energy concept-based explanations hardest. For example, items that test the first level in the progression of understanding for conduction (a warmer object will get cooler when in contact with a cooler object) are, on average, easier than items that test the second level (energy is transferred from the warmer object to the cooler object). And the second-level items are, on average, easier than the items testing the third level (energy is transferred by random atomic collisions).

Table 10
Mean Item Difficulty by Level of Progression for Each Idea

	Mean Rasch Difficulty			Correlation	
Energy Ideas	Level 1	Level 2	Level 3	Kendall's τ	p
Elastic Potential Energy	-1.51	-0.16	-0.12	0.523	<.01
Radiation	-1.24	-0.51	0.05	0.579	<.001
Kinetic Energy	-0.99	-0.39	0.78	0.493	<.001
Thermal Energy	-0.99	-0.12	-0.07	0.167	n.s.
Gravitational Potential Energy	-0.46	-0.01	0.49	0.320	<.05
Dissipation & Degradation	-1.07	0.40	0.50	0.527	<.01
Transferring Energy by Sound	-0.51	-0.27	0.27	0.596	<.05
Conduction	-0.57	0.06	0.80	0.418	<.01
Transferring Energy by Forces	-0.93	0.07	0.41	0.298	n.s.
Convection	-0.06	-0.03	0.62	0.445	<.01
Transferring Energy Electrically	-0.72		0.58	0.572	<.05
Chemical Energy	-1.32	0.61	0.79	0.461	<.01
Conservation	-0.43	0.01	0.80	0.477	<.001

There were two ideas that did not follow the expected progression. For thermal energy, there was no significant difference between the second and third levels. For transferring energy by forces, the means followed the expected trends, but the correlation coefficient was not significant. An analysis of the items from these ideas was conducted to investigate the source of the deviation.

Thermal Energy. An analysis of the Wright map for the items aligned to the thermal energy ideas was performed to determine why the items did not fit the proposed progression. On the map (see Figure 1), students' performance level is on the left-hand side and item difficulties are on the right-hand side. Easier items and less knowledgeable students are toward the bottom, and harder items and more knowledgeable students are toward the top of the map. We noticed that items that targeted the idea that the thermal energy of an object also depends on the mass clustered toward the top of the map above the items that were part of the original level 3 (atomic/molecular ideas about thermal energy). In other words, although it is easy for students to think that thermal energy depends on the temperature of an object, it is very difficult for them to think about thermal energy as also depending on the mass of an object. In fact, this makes that idea even more difficult than the hypothesized level 3 idea. Therefore, the data supports a progression that starts with the idea that thermal energy depends on the temperature of an object, followed by the idea that thermal energy depends on the speed and number of atoms or molecules that make up the object, and ending with the idea that thermal energy depends on the mass of the object. The mean Rasch difficulties for the items aligned to the revised levels are shown in Table 11. Figure 1 presents the Wright map for the revised progression. Kendall's tau correlation coefficient was calculated using the revised progression and a large and significant value was found (see Table 11). We postulate that having a solid understanding of atomic/molecular ideas related to thermal energy is helpful for making sense of the idea that thermal energy depends on mass. If students understand that thermal energy increases as the number of atoms/molecules increases and that mass is a measure of the amount of matter/number of atoms/molecules that makes up the object, then they can reason that thermal energy increases as the mass increases.

Table 11
Mean Item Difficulty for the Revised Levels of Progression for Thermal Energy

Level	# of Items	Mean Rasch Difficulty
1) Thermal energy depends on temperature	8	-0.91
2) Thermal energy depends on the speed & number of atoms/molecules	14	-0.19
3) Thermal energy depends on mass	5	0.32
Kendall's τ	0.734	$p < .001$

```
           Students | Items
   2             .  +
                 .  |
                 .  |
                 .  |
                 .  |T                   Thermal Energy
                .#  |
                .# T|
   1           .##  +     Level 1    |       Level 2        |    Level 3
              .###  |                |                      |
              .###  |S               |                      |
             .####  |                |                      | EG0274 EG0323
            .##### S|                |                      |
            .#####  |                |                      | EG0344
         .########  |                | EG0484 RG2071 EG0743 | EG0614 EG0842
   0    .#########  +M               | EG0514 EG0506        |
       .##########  |                | EG0852 EG0565        |
     .############  |                | EG0463 EG0493 EG0474 |
     .############ M|                | EG0525 EG0652        |
     .############  |  EG0284 RG1801 |                      |
      .###########  |S EG0314 EG0334 | EG0536 EG0543        |
        .#########  |  EG0304        |                      |
  -1     .########  +                |                      |
          .####### S|  RG0563        |                      |
           .######  |  EG0263 RG0582 |                      |
              .###  |T               |                      |
              .###  |                |                      |
                .#  |
                .# T|
  -2             .  +
```

Figure 1. Wright map of the revised thermal energy progression. Each “#” is 136 students and each “.” is 1 to 135 students.

There were other thermal energy items that did not fit in this revised progression and are, therefore, not included in Figure 1 or Table 11. They are six items targeting the idea that all things have thermal energy and seven items that asked the students “the thermal energy of an object depends on which of the following.” The items that targeted the idea that all things have thermal energy are not testing any of the three levels of the progression. The “which of the following” items were not included because they were aligned to more than one level or tested the idea that thermal energy depends on the type of material/molecule the object is made up of, which was not part of the revised progression of understanding.

Transferring energy by forces. We looked also more closely at the Wright map for the items targeting ideas about transferring energy by forces to reevaluate our progression of understanding. The items aligned to the transferring energy by forces were clustered into three sub-ideas. At the lower end of the map are items that target the idea that energy can be transferred by contact forces. In the middle of the difficulty range are items that target the idea that energy can be transferred by noncontact forces. Finally, the items that cluster at the higher end target the idea that a change in position or shape is necessary in order for energy to be transferred by the force.

In our hypothesized progression, the upper level included ideas about the changes in kinetic and potential energy that take place when objects change relative position as a result of a gravitational, magnetic, or electric force, a very sophisticated idea. But what was most challenging for students is the idea that a change in position is needed for energy to be transferred, and energy will not be transferred if there is no change in position even if the force continues to act.

Two items appear to be outliers in their level groupings, which increases the range of scores for those groupings and their standard deviation. Item NG0504 in the level 1 group (contact forces) is more difficult than the other items at that level. This item, unlike the others, targets a very popular misconception that it is a force (not energy) that is transferred during a collision. Previous studies have shown that many students think that a force becomes part of a thrown or hit object (Fischbein, Stavy, & Ma-Naim, 1989; McCloskey, 1983; AAAS, n.d.). Because of this, the item was removed from this analysis. Item RG1102 in the level 2 group (noncontact forces) is an item asking students which magnet, a stronger or weaker one, will transfer more energy to a metal ball. This item is easier than the other items involving noncontact forces. It is possible that some students are simply associating the terms "stronger" and "more" rather than using their understanding of transferring energy by noncontact forces. However, the Rasch fit statistics for this item indicate that it is functioning properly so the item was retained in the analysis.

```
          Students | Items
  2              .  +
                 .  |         Transferring Energy by Forces
                 .  |
                 .  |     Level 1     |    Level 2     |   Level 3
                 .  |T                |                |
                .#  |                 |                |
                .# T|                 |                | RG1652
  1            .##  +                 | RG0773         | NG0523 RG1212
              .###  |                 | RG1132         |
              .###  |S NG0504         |                | RG1012
             .####  |                 | RG0742         | RG0763
            .##### S|                 | RG1172         |
            .#####  |                 | RG0753 RG1222  |
         .########  |                 | RG1162         |
  0     .#########  +M                |                |
       .##########  |                 | RG1602         |
     .############  |                 |                |
     .############ M|  RG0982         | RG1202         |
     .############  |  NG0512 RG0962  |                |
      .###########  |S RG0972         |                |
        .#########  |                 |                |
 -1      .########  +                 |                |
          .####### S|  RG0952         | RG1102         |
           .######  |  RG0992 RG1592  |                |
              .###  |T                |                |
              .###  |                 |                |
                .#  |
                .# T|
 -2              .  +
```

Figure 2. Wright map of the revised transferring energy by forces progression. Each "#" is 136 students and each "." is 1 to 135 students.

Table 12 shows the average Rasch difficulty for the items in the revised levels after removing the item discussed above. The Kendall's tau correlation coefficient supports this revised progression indicating that first students gain an understanding that energy can be transferred by contact forces, then progress to understanding that energy can be transferred by noncontact forces, and finally understanding that energy will only be transferred when the force acts over a distance.

Table 12
Mean Item Difficulty for the Revised Levels of Progression for the Transfer of Energy by Forces Idea

Level	# of Items	Mean Rasch Difficulty
1) Transferring energy by contact forces	7	-0.84
2) Transferring energy by noncontact forces	10	0.17
3) Forces must act over a distance in order to transfer energy	5	0.90
Kendall's τ	0.706	$p < .001$

Conclusions

This paper describes the use of Rasch modeling to analyze data from science assessment items aligned to a learning progression for energy. A cross-sectional analysis of the student measures showed that the high school students have a better understanding of the energy concept than the middle school students, and the middle school students have a better understanding than the elementary school students. For most of the energy ideas, an analysis of the item measures validated the study's description of the energy concept and how it progresses in conceptual complexity from a phenomenological understanding, to energy-concept explanations, to more advanced energy-concept explanations. For two energy ideas, the progressions were revised to better fit the data.

Given the wide application of these energy ideas, it is critical that students understand them and how to apply them in different contexts and that educators understand the difficulties that students may have. The results of this study can inform and improve science instruction on the topic of energy by providing information about how the energy ideas progress in difficulty. Because these items are designed to be carefully aligned with a progression of understanding for energy ideas in national content standards but not to any single curriculum or instructional approach, researchers and developers of curriculum materials will be able to compare the effectiveness of various materials and approaches with more precision and objectivity.

Acknowledgements

The research reported here was supported by 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 or the U.S. Department of Education.

References

AAAS Project 2061. (n.d.) AAAS Science assessment website. Retrieved from http://assessment.aaas.org.

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.

Boone, W. J., Staver, J. R., Yale, M. S. (2014). *Rasch Analysis in the Human Sciences.* Netherlands: Springer.

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

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. (2008). Assessment linked to science learning goals: Probing student thinking through assessment. In J. Coffey, R. Douglas, & C. Stearns (Eds.), *Assessing student learning: Perspectives from research and practice* (pp. 231-252). Arlington, VA: NSTA Press.

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

Duit, R. (2014). Teaching and learning the physics energy concept. In R. F. Chen, A. Eisenkraft, D. Fortus, J. Krajcik, K. Neumann, J. Nordine, & A. Scheff (Eds.). *Teaching and learning of energy in K-12 education* (pp. 67-85). New York: Springer.

Fischbein, E., Stavy, R., & Ma-Naim, H. (1989). The psychological structure of naive impetus conceptions. *International Journal of Science Education, 11*(3), 327-336.

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.

Linacre, J. M. (2016). *WINSTEPS Rasch measurement computer program*. Version 3.92.1. Beaverton, Oregon: Winsteps.com.

Liu, X., Boone, W. J. (2006). Introduction to Rasch measurement in science education. In X. Liu and W. 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.

McCloskey, M. (1983). Intuitive physics. *Scientific American, 248*(4), 122-130.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

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

Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching, 50*(2), 162–188.

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.

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.

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. Copenhagen , Denmark: Danish Institute for Educational Research. Expanded edition, Chicago: University of Chicago Press, 1980. (Original work published 1960).