Evaluating a Unit Aimed at Helping Students Understand Matter and Energy for Growth and Activity

Cari F. Herrmann-Abell
BSCS Science Learning
Joseph Hardcastle and Jo Ellen Roseman
AAAS Project 2061

Paper presented at the 2019 AERA Annual Conference
Division C: Learning and Instruction
Section 1d: Science
Toronto, Canada
April 5-9, 2019

Abstract

To support implementation of the *Next Generation Science Standards*, we designed a high school biology unit, *Matter and Energy for Growth and Activity* (MEGA), that engages students in explaining physical and life science phenomena using evidence, models, and science ideas about matter and energy changes within systems and transfers between systems. The unit's promise was evaluated using a randomized control trial (RCT) involving fifteen teachers from two schools. Teachers were randomly assigned to implement either the MEGA unit or district-developed activities that targeted the same learning goals. Pre- and post-tests were administered, and the data were analyzed using Rasch modeling and hierarchical linear modeling. Here we describe the unit and report on RCT results. Our data showed that, when controlling for pretest score, gender, language, and ethnicity, students in the treatment group performed better on the post-test than the students in the comparison group, indicating the MEGA unit has promise in improving students' understanding. We also discuss a number of challenges that arose when developing and evaluating the unit.

Introduction

The *Next Generation Science Standards* (NGSS) (NGSS Lead States, 2013) calls for instruction that fosters an integrated understanding of science and engineering practices, crosscutting concepts, and disciplinary core ideas. Currently there are few curriculum materials available to help students achieve this goal, especially at the high school level. To help fill this gap, we developed a twelve-week curriculum unit, *Matter and Energy for Growth and Activity* (MEGA), to help high school biology students gain a deeper and more integrated understanding of: (1) energy-releasing and energy-requiring chemical reactions in simple physical systems and complex biological systems and (2) how these reactions are coupled so that living organisms can carry out basic life functions. The curriculum intervention consists of instructional materials for students and additional resources and professional development for teachers.

Theoretical Framework

Extensive research has shown that ideas about the flow of matter and energy through living and nonliving systems are among the most difficult for students and are fraught with misconceptions (e.g. AAAS, n.d.). The concept of energy is particularly challenging and is often taught in isolation in different disciplines, using different terminology and representations that make it difficult for students to make connections across physical and life science and appreciate the usefulness of the same set of energy ideas in explaining phenomena (Becker & Cooper, 2014; Wolfson, Rowland, Lawrie, and Wright, 2014). The MEGA unit differs from existing materials in its development of a coherent story of energy across disciplines by engaging students in using a coherent set of science ideas and science practices to make sense of both physical and biological phenomena. The instructional approach involves:

- Engaging students in observing phenomena and identifying patterns in data involving changes in matter and energy.
- Helping students use models to make sense of matter changes during chemical reactions in terms of atom rearrangement and conservation, energy changes during chemical reactions in terms of bond breaking and forming, and energy transfer in terms of coupling energy-requiring to energy-releasing systems.
- Helping students generate science ideas to make sense of phenomena in simple physical systems and then apply them to phenomena in complex biological systems.
- Helping students explain novel phenomena using evidence, science ideas, and models.
- Facilitating integrative connections between physical and life science contexts through a coherent storyline that uses related phenomena, similar models, and a common visual and text language to represent and describe patterns in energy-related phenomena in physical and biological systems.

The instructional model for the unit is consistent with constructivist, conceptual change, and cognitive apprenticeship perspectives and on research that investigated aspects of these perspectives. For example, studies by Schmidt et al. (2005) and Ainsworth and Burcham (2007) lend support for having students experience a range of interesting phenomena that are explainable by an interrelated set of fundamental ideas and sequenced into a coherent content storyline; studies by Burke et al. (1998) and Tabak and Reiser (2008) lend support for engaging students in using models to make sense of the phenomena, and studies by Eaton et al. (1984), McDermott (1991), Minstrell (1984), and Osborne and Freyberg (1985) lend support to organizing instruction around a range of tasks that are structured to help students relate the phenomena, models, and representations to the science ideas, reconcile their own ideas with the science ideas, and use the science ideas to explain other relevant phenomena.

Research Question

This paper describes the results of a study conducted in Year 3 of the curriculum development project that used a randomized control trial to compare outcomes for students who used the MEGA unit with outcomes for a comparison group that used the business-as-usual district curriculum. The results of the study were used to answer the following research question:

- To what extent does the MEGA unit improve students' understanding of the targeted learning goals when compared with the business-as-usual curriculum?

Methodology

Learning Goals

NGSS lists performance expectations (PE) for students to master by the end of high school. These performance expectations were derived from the disciplinary core ideas, science and engineering practices, and crosscutting concepts detailed in *A Framework for K–12 Science Education* (National Research Council, 2012). The following PEs were selected as the target of the MEGA unit.

- Plan and conduct an investigation to provide evidence that feedback mechanisms maintain homeostasis. (HS-LS1-3)
- Develop and use a model to illustrate the hierarchical organization of interacting systems provide specific functions within multicellular organisms. (HS-LS1-2)
- Construct and revise an explanation based on evidence for how carbon, hydrogen, and oxygen from sugar molecules (glucose) may combine with other elements to form amino acids and/or other large carbon-based molecules (growth). (HS-LS1-6)
- Use a model to illustrate how photosynthesis transforms light energy into stored chemical energy. (HS-LS1-5)
- Use a model to illustrate that cellular respiration is a chemical process whereby the bonds of food molecules and oxygen molecules are broken and the bonds in new compounds are formed resulting in a net transfer of energy. (HS-LS1-7)
- Develop a model to illustrate that the release or absorption of energy from a chemical reaction system depends upon the changes in total bond energy. (HS-PS1-4)

The unit targets all of the disciplinary core ideas covered by these PEs including Structure and Function (LS1.A), Organization for Matter and Energy Flow in Organisms (LS1.C), Structure and Properties of Matter (PS1.A), and Chemical Reactions (PS1.B). The crosscutting concepts targeted in the unit include Energy and Matter and Systems and System Models. The science and engineering practices include Analyzing and Interpreting Data, Developing and Using Models, and Constructing Explanations.

About the MEGA Curriculum Unit

The first chapter of the unit develops ideas about matter changes and conservation in systems that are useful for explaining biological growth and repair. After students model components of the hierarchical structure of the human body, they then examine the polymers that make up body system organs and tissues and compare body polymers to polymers that make up their food. To review middle school ideas about atom rearrangement and conservation while developing high school ideas about systems and system models, students observe phenomena in which pure substances interact in simple systems and use models and systems thinking to make sense of how new substances form and why changes in mass do not violate ideas about matter conservation. The culminating lesson of Chapter 1 focuses on applying the ideas about atom rearrangement and conservation and practices developed in simple systems to complex biological systems involved in human growth and repair.

The second chapter develops ideas about energy changes within systems and energy transfer between systems that are useful for explaining the performance of an Olympic athlete. Students first observe energy changes in simple systems in which there is a clear indication that the energy of the system is changing and then represent the energy changes and transfers with bar graphs and flow diagrams. Next, students revisit the chemical reactions they observed in Chapter 1 from the perspective of energy changes within systems and transfers between systems and represent their observations with bar graphs and flow diagrams. Simple bond energy calculations and graphic representations are used to help students visualize and make sense of why some chemical reactions release energy, others require energy, and even energy-releasing chemical reactions nonetheless require an initial input of energy.

At this point, students are ready to examine how energy-releasing processes can be used to drive energy-requiring processes, first in simple physical systems and then in human body systems involved in playing tennis or running a marathon. Experimental data provides evidence that the reaction in muscle cells between $C_6H_{12}O_6$ and O_2 to form CO_2 and H_2O provides energy for motion of the body during exercise, allowing students to construct a simple energy flow diagram. Experimental data then provides evidence that cycles of ATP synthesis and breakdown couple cellular respiration to the sliding of proteins making up muscle fibers, which students use to revise their energy flow diagram so that it represents energy changes and transfers at the cellular level. And finally, students examine data from yogurt-producing bacteria and humans that provides evidence for an alternative energy-releasing chemical reaction that provides energy for both growth and motion in the absence of oxygen. Evidence comes from a mix of data students collect and data from published studies.

The unit's culminating lesson ties together the matter and energy storylines by engaging students in predicting changes in matter and energy during intense exercise, examining data that challenges their predictions, and gathering information to explain how various body systems attempt to maintain internal temperature and the concentrations of glucose, oxygen, and carbon dioxide within a narrow range (homeostasis).

Study Design

A randomized control trial was conducted in the fall and winter of 2017-2018 in one district in the Mid-Atlantic U.S. Fifteen teachers from two schools participated in the study. Treatment assignment of teacher was done randomly by a member of the external advisory board. In the treatment group, the MEGA unit replaced the students' usual curriculum materials for the basic biology class, and the unit's lessons were taught by the classroom teacher after the teacher participated in two days of face-to-face professional development. The students in the comparison group used a curriculum unit that targets the same science ideas and was developed by a team of teachers from the district.

Pre- and Post-tests

To determine the extent to which students' understanding of the targeted learning goals changed as a result of instruction, pre- and post-tests were developed by members of the research team. A total of 64 multiple-choice items and six constructed-response items were developed. The items targeted 11 key ideas as shown in Table 1. The items also required students to use specific science practices (constructing explanations, analyzing and interpreting data, and developing and using models) and crosscutting concepts (systems and system models and the flow of matter and energy). In addition, we included nine items that targeted science practices but did not require knowledge of any particular disciplinary core idea: four items assessed students' understanding of developing and using models, and six items assessed students' ability to evaluate experimental designs and draw conclusions from experimental data. These items were used to build three versions of the test with each version consisting of 33-35 items that were a mix of multiple-choice and constructed-response items. Seven items appeared on all three versions so that comparisons could be made across forms. Students were randomly assigned to take one version for their pre-test and then were randomly assigned to take another version for their post-test. Because the ideas targeted by this unit are challenging for high school students, we also administered the tests to a set of 420 biology students from a large public university in the northeast. This gave us a wider range of responses to both the multiple-choice and the constructed-response items, particularly at the higher end of the scale.

Table 1: *Topics Targeted by the Pre- and Post-tests*

Topic	Key Idea	# of items
Physical Science	Atom rearrangement & conservation	4
	Energy conservation	3
	Energy & chemical reactions	9
	Coupling chemical reactions	2
Life Science	Photosynthesis	6
	Cellular respiration	10
	Animal growth	7
	ATP	6
	Ecosystems	3
	Body systems	7
	Homeostasis	4
Practice	Developing & using models	4
	Designing investigations	6

In the analysis presented here, we decided to exclude the data from the three ecosystem items and from one of the animal growth items. These items are all constructed-response items that were either judged to be too difficult (perhaps because they required considerable knowledge transfer) or were too tangentially aligned to the learning goals discussed above.

Participants

Both high schools in the district volunteered to participate in the study and allowed us to randomly assign half of the teachers at each school to either treatment or comparison groups. A total of 735 students enrolled in basic biology classes participated in the trial, but the data reported here are from the 641 students who took both a pre-test and post-test. Demographic information is shown in Table 2.

Table 2: *Demographic Information for Students Included in the Study*

	Treatment	Comparison	Total

	(N = 404)	(N = 237)	(N = 641)
Grade			
9th	72%	70%	71%
10th	4%	3%	3%
11th	24%	25%	24%
12th	1%	1%	1%
Gender			
Male	47%	60%	52%
Female	53%	40%	48%
Ethnicity			
White	66%	63%	65%
Black	10%	8%	9%
Hispanic	9%	13%	11%
Asian	11%	12%	12%
Other	3%	4%	3%
Primary Language			
English	93%	93%	93%
Other	7%	7%	7%

Rasch Analysis

WINSTEPS (Linacre, 2016) was used to estimate Rasch student and item measures. The analyses done for this paper were based on the data from the dichotomously scored multiple-choice items and two of the constructed-response items, which were polytomously scored. Therefore, the "partial credit" model was used (Masters, 1982). The pre- and post-test data from the treatment and comparison groups were stacked, meaning that there are two rows for each student in the data file, one for pre-test responses and one for post-test responses (Wright, 2003). This results in two measures per student: a pretest measure and a post-test measure. The data file also included the responses from the 420 university students. The data's fit to the Rasch model was evaluated using the separation indices, infit and outfit mean-squares, and standard errors.

Hierarchical Linear Modeling

The Rasch post-test measures were modeled as outcome measures in a two-level hierarchical linear model (HLM) with students at level 1 and teachers at level 2. Student-level variables included pre-test measure, gender, ethnicity, and language. Whether or not the teacher implemented the MEGA unit was the only teacher-level variable. A fully unconditional model containing only the post-test outcome variable and no independent variables (other than an intercept) was estimated first. This was followed by a conditional model in which pre-test measure, gender, language, and ethnicity were included as controls and modeled as fixed effects. HLM 7 software was used in this study (Raudenbush et al., 2011). The method of estimation was restricted maximum likelihood. Effect sizes were calculated by dividing the coefficient by the square root of the pooled student-level unadjusted standard deviation.

Results

Rasch Fit

The Rasch fit statistics are presented in Table 3. The separation indices and corresponding reliabilities were 7.68 and 0.98 for the items and 1.35 and 0.65 for the students. The separation index for the items is considered acceptable—that is, greater than 2, according to Wright and Stone (2004). However, the separation index for the students was low. Additionally, the standard errors for the items and students were small (see Table 3). The infit and outfit mean-square values for the majority of the items and students were within the acceptable range of 0.7–1.3 for multiple-choice tests (Bond and Fox, 2007). Based on the fit statistics, we conclude that the data have an adequate fit to the Rasch model.

Table 3: *Rasch Fit Statistics*

	Items			Students		
	Minimum	Maximum	Median	Minimum	Maximum	Median
Standard error	0.04	0.12	0.09	0.32	2.18	0.38
Infit mean-square	0.86	1.22	0.99	0.59	1.91	0.99
Outfit mean-square	0.84	1.52	0.98	0.49	4.82	0.96
Separation index	7.68 (0.98)			1.35 (0.65)		

Fully Unconditional HLM

A fully unconditional HLM with no independent variables at either level was run to calculate the intraclass correlation coefficient. The results of the model are shown in Table 4. The intraclass correlation coefficient represents the proportion of variance in post-test measures that could be the result of teacher characteristics, such as the curriculum used or teachers' content knowledge. In this case, almost a quarter (23%) of the variance in post-test measures could be the function of teacher characteristics. Therefore, the proportion of the variance in post-test measures that exists at the individual level is 77%. A chi-square test indicated that post-test measures varied significantly between teachers ($\chi^2 = 197.16$, $p < 0.001$).

Table 4: *Fully unconditional HLM*

Variable	Value
Within-classroom variance (σ^2)	0.38
Between-classroom variance (τ)	0.11
Between-classroom SD	0.33
Reliability (λ)	0.90
Intraclass correlation (ρ)	0.23

Conditional HLM

The mixed-model for the conditional HLM is

$$\text{POSTTEST}_{ij} = \gamma_{00} + \gamma_{01} * \text{TREAT}_j + \gamma_{10} * \text{FEMALE} + \gamma_{20} * \text{BLACK} + \gamma_{30} * \text{ASIAN} + \gamma_{40} * \text{HISPANIC} + \gamma_{50} * \text{OTHER} + \gamma_{60} * \text{ENGLISH} + \gamma_{70} * \text{GRADE} + \gamma_{80} * \text{PRETEST} + u_{0j} + r_{ij}$$

where POSTTEST_{ij} and PRETEST_{ij} are the post- and pre-test scale scores for the student i within teacher j, respectively. TREAT is a dummy variable indicating whether or not the teacher implemented the MEGA unit. The comparison group, which was using the district curriculum, was used as the reference group. FEMALE is a dummy variable indicating the gender of student i in class j (female = 1; male = 0). Four dummy variables were created for ethnicity (BLACK, HISPANIC, ASIAN, and OTHER), and white was used as a reference group. ENGLISH is a dummy variable indicating whether or not English is the primary language of student i in class j (English = 1; other language = 0). All of the student-level variables were grand–mean centered and the teacher-level variable was uncentered. The terms u_{0j} and r_{ij} are the error terms associated with the teachers and students, respectively. The results of the conditional HLM are shown in Table 5.

Table 5: *Results from the conditional HLM*

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. $d.f.$	p-value
Teacher-level variables					
Intercept, γ_{00}	-0.77	0.11	-7.08	13	<0.001
Treatment, γ_{01}	0.35	0.15	2.33	13	0.04
Individual-level variables					
Female, γ_{10}	0.06	0.05	1.26	618	0.21
Black, γ_{20}	-0.16	0.08	-1.93	618	0.05
Asian, γ_{30}	0.12	0.08	1.47	618	0.14
Hispanic, γ_{40}	-0.03	0.08	-0.34	618	0.74
Other, γ_{50}	-0.05	0.13	-0.38	618	0.71
English, γ_{60}	0.04	0.10	0.41	618	0.68
Grade, γ_{70}	0.11	0.03	4.01	618	<0.001
Pretest, γ_{80}	0.25	0.04	5.70	618	<0.001
Random Effect	**Standard Deviation**	**Variance Component**	**d.f.**	χ^2	**p-value**
Intercept, u_0	0.27	0.07	13	159.71	<0.001
level-1, r	0.59	0.35			

The intercept indicates that the average post-test measure for the comparison group is -0.77 logits. The students in the treatment group scored, on average, 0.35 logits more than the comparison group so the average post-test measure for the treatment group is -0.42 logits. This corresponds to an effect size of 0.51, which is considered a medium effect size (Cohen, 1988). This suggests that, on average, the MEGA unit was more successful in improving students' understanding of the targeted ideas and practices than the comparison curriculum.

Performance by Topic

When we group the items by whether they assessed matter ideas, energy ideas, or integrated ideas about both, we find that the students in the treatment group make greater gains that the comparison group (see Table 7). However, the treatment group made smaller gains on items that require the integration of matter and energy than they did on the items that assess understanding of either matter or energy.

Table 7: *Percentage of Correct Responses by Matter and Energy Topics*

Topic	# of items	Group	Pre-test	Post-test	Gains
Matter	20	Comparison	34%	39%	5%
		Treatment	34%	45%	11%
Energy	3	Comparison	37%	41%	4%
		Treatment	38%	58%	20%
Matter and energy	32	Comparison	30%	32%	2%
		Treatment	30%	38%	8%

We also wanted to compare student performance on the different topics targeted in the unit, so we grouped the items into three categories: physical science, life science, and science practices. Overall, the treatment group outperformed the comparison group on the physical and life science items, with the treatment group making greater gains on the physical science items than the life science items (see Table 6). There was no difference in the gains made by the two groups on the items assessing science practices.

Table 6: *Percentage of Correct Responses by Topic*

Topic	Group	Pre-test	Post-test	Gains
Physical science	Comparison	34%	36%	2%
	Treatment	33%	46%	13%
Life science	Comparison	31%	35%	4%
	Treatment	32%	41%	9%
Science practices	Comparison	33%	41%	8%
	Treatment	37%	45%	8%

Discussion

Findings from our analysis of the data collected during the randomized control trial provide support for the instructional approach used to develop the MEGA unit. Students who experienced the unit gained a better understanding of the role that matter and energy play in chemical reactions occurring in simple physical systems and complex biological systems than students in the comparison group. Despite this finding, however, several issues emerged over the three years of our curriculum development project. Some of these issues we were able to address in successive revisions of the unit and teacher resources. Other issues are still unresolved and have implications not only for the MEGA unit but also for the wider implementation of NGSS.

Prerequisite Knowledge and Skills

Like most high school curriculum materials, the MEGA unit builds on essential middle school prerequisites. According to NGSS, students completing middle school are expected to (a) know that new substances are produced during chemical reactions because atoms making up molecules of reactants rearrange to form molecules of products without any atoms being created or destroyed and (b) be able to use an atom rearrangement and conservation model to make sense of physical and life science phenomena that involve chemical reactions. Regarding energy, NGSS expects students completing middle school to know that (a) some chemical reactions release energy and others require an input of energy, (b) the chemical reaction by which plants produce complex food molecules (sugars) requires an energy input (i.e., from sunlight) to occur, (c) cellular respiration in plants and animals involve chemical reactions with oxygen that release energy, and (d) when the motion energy of an object changes, there is inevitably some other change in energy at the same time.

The initial version of the MEGA unit was made up of a single chapter that focused primarily on the energy concept and the role of models and modeling as a tool for exploring phenomena. In developing the high school unit, we assumed that students had already mastered the middle school knowledge and skills related to matter as outlined above. During pilot testing of the initial version of the unit, it became clear that most high school students did not, in fact, have these prerequisite middle school knowledge or skills. In response, we revised the unit to include a new first chapter that revisited the middle school matter ideas and modeling skills while allowing for the introduction of high school science practices and crosscutting concepts. Because energy was the primary focus of the MEGA unit, we did not assume that high school students would already have an understanding of energy in the context of chemical reactions. Therefore, we designed the MEGA unit to build on elementary energy ideas by starting with a review of the indicators of energy changes and transfers and then introducing examples of energy-releasing and energy-requiring chemical reactions.

High Cognitive Complexity of NGSS Matter and Energy Ideas

The cognitive complexity of the complete set of high school matter and energy ideas described by NGSS is likely to be overwhelming for novice learners and beyond the scope of a single unit. To benefit fully from the MEGA unit, students should have a strong foundational understanding of ideas about changes to and conservation of matter before focusing on energy. When designing and testing *Toward High School Biology* (THSB), a middle school precursor unit that focuses on the role of matter in biological growth and repair, we had concluded that the matter story by itself was sufficiently ambitious. We had observed that even 8th grade honors students were inclined to explain losses in mass in terms of matter having been converted to energy and exhibited confusion about what a flame is. Given what could reasonably be accomplished in an eight-week unit, we limited the focus of the THSB unit to phenomena that could be explained by a coherent story of matter changes (e.g., focusing the plant growth story on making glucose and then reacting glucose to form polymers that become part of the plant's body, but not on the role of cellular respiration).

Because most high school students using the MEGA unit will not have had experience with our THSB middle school unit, we decided to include a new first chapter that would provide students with a "crash course" on the essential prerequisite ideas about matter. The revised MEGA unit develops the matter story that explains human growth in the first chapter and then integrates the energy story that explains motion and growth in the second chapter. Each of the chemical reactions included in the second chapter is introduced in the first chapter so that students would come into the second chapter with a firm grasp of the matter changes that occur during these reactions (e.g., glucose combustion, digestion of polymers from food to monomers, and the use of the monomers to build body polymers) and be ready to incorporate energy changes into their thinking. However, even with the addition of a first chapter on the matter story, the MEGA unit is not likely to provide sufficient support for students who enter high school without the middle school prerequisites. This is supported by our finding that students' scores on the post-test items that integrated matter and energy ideas were low (38% on average).

Teachers' Understanding of NGSS Ideas About Matter and Energy

It was clear during the professional development workshop that we conducted at the beginning of this study that many high school biology teachers have not incorporated correct ideas about energy associated with bond breaking and bond forming into a coherent story about chemical reactions in either physical or biological systems. Many think that ATP has "high-energy" bonds and that energy is released when the bonds are broken. Teachers have not learned the correct ideas that (a) breaking any bond requires an input of energy and (b) energy is released when any bond forms. These core ideas about energy associated with bond breaking and bond forming are essential to the understanding of energy that is expected of high school students. (The middle school story about energy is simply the empirical generalization that some chemical reactions give off energy and other chemical reactions require a net input of energy. No explanation/mechanism is expected of students nor are they expected to know why these phenomena are consistently observed.) Many biology teachers also do not know that the energy from an energy-releasing reaction can be used to "drive" an energy requiring reaction if two conditions exist: (a) the energy-releasing reaction releases more energy than is required by the energy-requiring reaction (more, that is, because some of the energy released is always transferred to the surroundings as heat) and (b) a suitable coupling mechanism exists (otherwise all the energy released will be transferred to the surroundings as heat). These core ideas can be used to explain how burning gasoline is used to make a car go and how glucose metabolism in cells is used to move muscles. Based on classroom observations, the two days of professional development teachers received before implementing the MEGA unit was not enough to make significant improvements in (1) their understanding of and (2) their comfort with teaching these cognitively complex ideas.

Biology Teachers' Resistance to Teaching Physical Science Ideas

Understandably, biology teachers' love of biology makes them not want to "waste" valuable class time teaching physical science ideas. They need to be convinced that focusing on physical science ideas is essential to their students' success in developing a coherent understanding of matter and energy for growth and activity in living things. To persuade teachers that this goal is important requires helping them to make explicit connections between physical science and biology and to see the value that physical science ideas have in making sense of biological phenomena. For NGSS-aligned curriculum materials to be successfully implemented, professional development will need to address biology teachers' perceptions of the role of physical science in biology. While we may not have convinced all of the teachers who participated in our study of the importance of physical science, the fact that students in the treatment group saw the greatest gains on the physical science items (13 percentage points) indicates that the MEGA unit shows promise for supporting the teaching of physical science ideas and representing energy changes and transfers in physical systems.

Challenge of Developing Valid NGSS Assessments

Few validated NGSS-aligned assessments currently exist, so curriculum developers are left to create their own assessments to study the promise of new NGSS-aligned units. The validation of assessments requires demonstrating that at least some students are able to successfully answer the most difficult items. As we developed items for the MEGA unit, we found that some of our assessments, the constructed-response items in particular, were too difficult for typical high school students. To calibrate the items with more precision, we recruited a sample of university biology students who would be expected to have had several opportunities to learn the ideas targeted by the items. Indeed, we found that the university students outperformed the high school students overall, but their performance on the constructed-response items was only marginally better.

Conclusions

This paper reports on data from a randomized control trial of MEGA, a new curricular unit designed to help students explain biological phenomena in terms of matter and energy changes and transfers within and between systems. Guided by a set of research-based design principles that align with the vision of NGSS, the unit was developed to improve upon currently available materials by helping students to construct a coherent story of matter and energy across the physical and life science disciplines.

The study was conducted in one school district and compared students who experienced the MEGA unit to students who experienced the district-developed curriculum. Rasch modeling was used to create student measures for both the pre- and post-tests. The post-test measures were then modeled as outcomes in a two-level HLM to investigate effects of the MEGA unit controlling for pre-test measure, gender, language, and ethnicity. The results of the model showed a positive correlation between using the MEGA unit and post-test measure with a medium effect size.

These results provide evidence of the promise of the MEGA unit for increasing students' understanding of the role of matter and energy changes during chemical reactions in making sense of biological growth and activity.

A number of challenges were encountered during the development and evaluation of the unit that may have limited its impact. These included the high cognitive complexity of the learning goals, students' lack of prerequisite understanding, and teacher's lack of understanding of, comfort with, and appreciation of physical science ideas. Additional research is needed to investigate ways of overcoming these challenges so that we can accomplish the vision of three-dimensional science understanding outlined by NGSS.

Acknowledgements

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150310 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

Ainsworth, S., & Burcham, S. (2007). The impact of text coherence on learning by self-explanation. *Learning and Instruction, 17*(3), 286-303.

American Association for the Advancement of Science. (n.d.) *Science Assessment Website*. Retrieved from http://assessment.aaas.org/

Becker, N. M., & Cooper, M. M. (2014). College chemistry students' understanding of potential energy in the context of atomic–molecular interactions. *Journal of Research in Science Teaching, 51*(6), 789-808.

Burke, K. A., Greenbowe, T. J., & Windschitl, M. A. (1998). Developing and using conceptual computer animations for chemistry instruction. *Journal of Chemical Education, 75*(12), 1658-1660.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd ed., New York: Academic.

Eaton, J. F., Anderson, C. W., & Smith, E. L. (1984). Student preconceptions interfere with learning: Case studies of fifth-grade students. *Elementary School Journal, 64*, 365-379.

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

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*, 149-174.

McDermott, L. (1991). Millican lecture 1990. What we teach and what is learned – closing the gap. *American Journal of Physics, 59*, 301-315.

Minstrell, J. (1984). Teaching for the understanding of ideas: Forces on moving objects. In C. W. Anderson (Ed.), *Observing science classrooms: Perspectives from research and practice.* 1984 Yearbook of the Association for the Education of Teachers in Science (pp. 55-73). Columbus, OH: ERIC Center for Science, Mathematics and Environmental Education. (ERIC Document Reproduction Service No. ED 255 355).

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

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

Osborne, R., & Freyberg, P. (1985). *Learning in science: The implications of children's science*. Auckland, New Zealand: Heinemann.

Raudenbush, S.W., Bryk, A.S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling*. Chicago, IL: Scientific Software International.

Schmidt, W. H., Wang H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies, 37*(5), 525-559.

Tabak, I., & Reiser, B. J. (2008). Software-realized inquiry support for cultivating a disciplinary stance. *Pragmatics and Cognition, 16*(2), 307-355.

Wolfson, A. J., Rowland, S. L., Lawrie, G. A., Wright, A. H. (2014). Student conceptions about energy transformations: Progression from general chemistry to biochemistry. *Chemistry Education Research and Practice, 15*, 168-183.

Wright, B. D. (2003) Rack and stack: Time 1 vs. Time 2. *Rasch measurement transactions*, *17*(1), 905-906. Retrieved March 12, 2012, from http://www.rasch.org/rmt/rmt171a.htm