

Symposium Proposal
2019 NARST Annual International Conference
Baltimore, MD
March 31 – April 3, 2019

Developing High School Biology Curriculum Materials that Support NGSS Teaching and Learning: Opportunities and Challenges

1. Jo Ellen Roseman
2. Louisa A. Stark
3. Cari F. Herrmann Abell
4. Kristin M. Bass
5. George DeBoer
6. Dina Drits-Esser
7. Joseph Hardcastle
8. Sheila Homburger
9. Molly Malone
10. Anat Yarden (discussant)

Abstract

Realizing the vision of the *Next Generation Science Standards* (NGSS) requires curriculum materials that integrate disciplinary core ideas and crosscutting concepts into a coherent storyline and engage students in using science ideas and practices to make sense of relevant and engaging phenomena. In this symposium, curriculum development groups who have collaborated in the design and testing of two high school biology units describe their iterative development processes and products, findings from pilot and field testing, and opportunities and challenges encountered in identifying appropriate phenomena, addressing prerequisites students may not have learned, and addressing teachers' lack of comfort with NGSS science practices and science ideas outside of biology. The symposium concludes with a synthesis of lessons learned that can inform the design of other NGSS-aligned curriculum materials for high school biology and beyond.

Symposium Overview

The most significant aspect of the reform recommendations laid out in the *Next Generation Science Standards* (NGSS) (NGSS Lead States, 2013) is the required integration of all three dimensions of *A Framework for K-12 Science Education* (National Research Council [NRC], 2012)—disciplinary core ideas, crosscutting concepts, and science and engineering practices—in a set of performance expectations. This three-dimensional approach aims to prepare all students to participate in science throughout their lives. While the vast majority of high school graduates will not pursue careers in research science, all will need to understand and use science ideas across disciplines to make sense of the natural and man-made world, to comprehend news coverage of scientific developments and issues, and to participate as citizens in policy decisions about the use of science to solve some of society’s most challenging problems of food, energy, health, and climate (NRC, 2009).

Integration of the three dimensions of science learning presents challenges to high school biology curriculum developers and teachers, especially for materials focused on matter and energy and evolution and heredity because these topics depend on physical science and mathematics prerequisites and because of the misconceptions that students bring. Today’s high school students may also lack experience in using science practices, and their teachers may be ill-equipped to address these deficiencies. Identifying interesting and appropriate phenomena to motivate and scaffold students’ learning of biology ideas and practices is also a challenge.

Two research and development teams collaborated in the design and testing of two high school biology units. One focuses on physical and life science ideas about matter and energy and systems as well as the practices of data analysis, modeling, and explanation of phenomena related to human growth and motion. The other focuses on ideas about heredity and evolution as well as the practices of data analysis and argumentation within the context of phenomena in which DNA-based variations in traits are acted upon by natural selection, leading to changes in populations over time and the diversity of life.

The proposed symposium will present case studies of the design and testing of these two high school biology units. Findings from the case studies provide important insights into the design of aligned curriculum and assessment materials that support NGSS teaching and learning. Each case study will (a) describe the unit and its instructional approach, (b) present findings from pilot and field tests, and (c) discuss challenges that arose during design and testing and ways these were addressed. The symposium will begin with 5 minutes of introductory remarks followed by two 30-minute case study presentations. Following the presentations, a discussant will comment on the case studies, synthesize the ideas that emerged, and consider lessons learned (10 minutes). There will be 15 minutes for interaction with the audience.

Defining the Problem

Despite the growing acceptance of NGSS in school districts across the country, few materials currently exist to support learning of disciplinary core ideas, crosscutting concepts, and science practices as called for in the new standards. For example, curriculum evaluations carried out by AAAS (2005) indicate that biology textbooks—both commercially developed and those funded by research grants—emphasized technical terminology over conceptual knowledge, provided few opportunities for students to draw connections between ideas, and did little to provide a meaningful narrative to weave ideas into a coherent story or to support teachers in their instruction. A more recent examination of energy concepts in biology textbooks showed that

current textbooks also fail, among other things, to help students develop a common language for talking about energy in non-living and living systems and to take advantage of simpler physical phenomena involving energy (e.g., batteries and combustion engines) to help students understand more complex biological phenomena (e.g., body building or competing in an Olympic event).

On the topic of evolution, researchers have called for a stronger genetics connection in students' study of evolution (Catley et al., 2010; Dougherty, 2009). However, few curriculum materials foster this integration, preventing students from easily making the essential conceptual connections between heredity and evolution (e.g., Biggs et al., 2009; Postlethwait & Hopson, 2009). Further, providing students opportunities to engage in the science practices of analyzing and interpreting data (Catley, Lehrer, & Reiser, 2004) and arguing from evidence (Asterhan & Schwarz, 2007) have been shown to impact evolution learning gains but such opportunities are rarely present in most curriculum materials.

Case Study 1: *Matter and Energy for Growth and Activity*

Matter and Energy for Growth and Activity (MEGA) is a twelve-week curriculum unit aimed at helping 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 accompanying materials and professional development for teachers.

Instructional approach. The MEGA unit differs from existing materials in its development of a coherent story of energy across disciplines by engaging students in using the same set of science ideas and science practices to make sense of both physical and biological phenomena. The instructional approach involves helping students to (1) observe phenomena and identify patterns in data involving changes in matter and energy; (2) 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; (3) generate science ideas in phenomena involving simple physical systems and then apply them to phenomena in more complex biological systems; and (4) explain novel phenomena using evidence, science ideas, and models; while also (5) facilitating integrative connections between physical and life science contexts through the use of related phenomena, similar models, and a common visual and text language. This instructional approach is consistent with constructivist, conceptual change, and cognitive apprenticeship perspectives (e.g. Ainsworth & Burcham, 2007; Tabak & Reiser, 2008; McDermott, 1991; Collins, Brown, & Newman, 1989).

The MEGA unit was designed to build on knowledge of middle school physical science ideas and practices that were developed in an 8th grade precursor unit *Toward High School Biology* (THSB) (AAAS/Project 2061, 2017), especially (a) the idea that new substances are produced during chemical reactions through atom rearrangement and (b) experience using data, science ideas, and models to explain phenomena involving changes in matter. This case study presents data to answer the following research questions: (1) To what extent does the MEGA unit show promise in improving students' learning of the targeted NGSS science ideas and practices? (2) Does students' use of THSB in 8th grade before experiencing the MEGA unit affect their performance on the MEGA pre- and post-tests?

Case Study 1: Methodology

Curriculum unit. The MEGA unit's first chapter develops ideas about matter changes and conservation in body systems that are useful for explaining biological growth and repair. After examining components of the hierarchical structure of the human body, students examine component body polymers and compare them to polymers that make up food. To review middle school ideas about atom rearrangement and conservation while developing high school ideas about systems and system models, activities engage students in observing phenomena in which pure substances interact in simple systems and in using models and systems thinking to make sense of why new substances form and why changes in mass do not violate ideas about conservation. The culminating lesson of Chapter 1 focuses on applying ideas about atom rearrangement and conservation and system inputs and outputs to make sense of the contributions of various human body systems to growth and repair.

The second chapter develops ideas about energy changes within systems and energy transfer between systems that are useful for explaining training and competing in athletic events. After observing energy changes in simple systems in which there is a clear indication that the energy of the system is changing and representing the energy changes and transfers with bar graphs and energy transfer models, 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, upward and downward arrows, and energy transfer models. Simple bond energy calculations and graphic representations are used to help students visualize and make sense of why some chemical reactions release energy and others require energy. Students then examine how energy-releasing processes can be used to drive energy-requiring processes, first in simple physical systems and then in biological systems. 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 homeostasis.

Study design. Seven teachers and their students from three schools in one school district in a mid-Atlantic state participated in the study. Two schools volunteered to implement the MEGA unit, and the third school implemented a district-developed unit that targeted the same ideas as the MEGA unit. In the two schools, 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.

Participants. A total of 449 students participated in the study, but the data reported here are from the 352 students who took both the pre- and post-tests. The students were about 58% male and 42% female, and approximately 21% of the students were white, 11% were Asian, 41% were African American, 23% were Hispanic, and 4% were of two or more ethnicities. Approximately 9% of the students were in 9th grade, 77% in 10th grade, 12% in 11th grade, and 2% in 12th grade.

Student pre- and post-tests. To determine how students' understanding of the targeted learning goals changed as a result of instruction, the research team developed three pre- and post-tests consisting of 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 another version for their post-test.

Teacher pre- and post-test. A teacher knowledge test was constructed to include eight multiple-choice items that targeted middle and high school ideas about matter and energy. In addition to

identifying the correct answer, the teachers were asked to identify the most popular incorrect answer (Sadler et al., 2013) and to explain why students selected that incorrect answer.

Rasch analysis. WINSTEPS (Linacre, 2017) was used to estimate Rasch student and item measures. The analyses were based solely on the data from the dichotomously scored multiple-choice items because the constructed-response items are currently being hand-scored. 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 a pre-test measure and a post-test measure for each student. The data showed a good fit to the Rasch model.

Case Study 1: Results

MEGA unit’s promise. Table 1 shows the average pre- and post-test scores by school. The scores are given in logits, which range from $-\infty$ to $+\infty$. The average item difficulty was set at 0 logits so negative scores mean that tests were difficult for the students. Overall, students in the comparison group who used the district-developed curriculum saw no gains from pre-test to post-test. Students at both schools in the treatment group who experienced the MEGA unit saw significant gains according to repeated measures t-tests (School #2: $t = -10.06$, $p < .001$; School #3: $t = -4.25$, $p < .001$). School #2 had a large effect size and School #3 had a medium effect size. This provides support for our instructional approach and indicates that the MEGA unit shows promise in improving students’ understanding of the targeted NGSS science ideas and practices (research question 1).

Table 1. *Summary of pre-test and post-test student measures*

School	Curriculum	N	Pre-test	Post-test	Gains	Effect size ^a
1	District	104	-0.84	-0.87	-0.03	-0.05
2	MEGA	100	-0.73	-0.09	0.64	1.42
3	MEGA	148	-1.00	-0.70	0.30	0.38

^aEffect size calculated by dividing the difference of the means by the pretest standard deviation (SD)

Prior use of THSB. Of the 248 students who used the MEGA unit, 33 had experienced the THSB unit in 8th grade. We predicted that students who experienced THSB in middle school would make larger gains when using MEGA because THSB provides the foundational knowledge upon which MEGA builds. Table 2 shows the average pre- and post-test scores for students at each treatment school who had used THSB versus those who did not. On average, students who experienced THSB gained 0.67 logits while students who did not experience THSB gained on average 0.40 logits. This suggests that students who experienced the THSB unit gain a better understanding of the high school ideas about matter and energy (research question 2).

Table 2: *MEGA pre- & post-test scores for students with & without THSB by school*

8 th grade unit	School	N	Pre-test	Post-test	Gains	Effect Size ^a
THSB	2	16	-0.68	0.23	0.91	1.78
	3	17	-1.26	-0.82	0.44	0.52
Not THSB	2	84	-0.74	-0.16	0.58	1.32
	3	131	-0.97	-0.69	0.28	0.36

^aEffect size calculated by dividing the difference of the means by the pretest standard deviation (SD)

Teacher knowledge. Observations during the two-day professional development workshop and results from the teacher pre-test showed that the teachers lacked knowledge of physical science

and systems ideas and experience using science practices to make sense of phenomena. Teachers performed poorly on the three items that targeted ideas about energy changes and chemical reactions. Similar results were found on the post-test indicating that the professional development and Teacher Edition were insufficient in improving teachers' understanding.

Case Study 2: Evolution: DNA and the Unity of Life

Evolution: DNA and the Unity of Life is an eight-week, five-module unit designed to build students' understanding of the underlying role of genetics in biological evolution using developmentally appropriate phenomena and applications of each dimension of NGSS. Developers of the curriculum unit theorize that students will better understand evolution when curriculum materials and instruction (1) integrate essential topics in heredity; (2) build students' abilities in the science practice of analyzing and interpreting skill-level-appropriate data about phenomena from published scientific research; (3) engage students in the construction of evidence-based arguments; and (4) frame ideas through the crosscutting concepts of patterns, systems and system models, and cause and effect. The curriculum consists of interactive multimedia and paper-based instructional materials for students, in-depth teacher guides, and just-in-time introductory videos for teachers.

Instructional approach. Each module is structured around a driving question and can be used on its own or with the other modules in sequence (see Table 3). When used in sequence, the modules establish an understanding of DNA as a blueprint for all living things, how DNA underlies variations in traits that are acted upon by natural selection, and how this leads to the diversity of life. A scaffolded claims-evidence-reasoning (CER) framework (Berland & McNeill, 2010) is integrated into the unit to incrementally build skills in constructing arguments from evidence over the course of the module sequence. Only non-human phenomena are used in order to increase acceptability of this evolution curriculum for the widest student and teacher population. Unit development was informed by several learning progressions (e.g., Catley et al., 2004) and drew on constructivist, conceptual change, and situated cognition theories of learning (e.g., Driver, 1995; Strike & Posner, 1992).

Table 3. *Driving questions and phenomena for each module.*

Module	Driving Question	Phenomena
Shared Biochemistry	<i>What shapes the characteristics of all living things?</i>	<ul style="list-style-type: none"> • Glowing fish • Firefly tails • Cognitive bias
Common Ancestry	<i>What is the evidence that living species evolved from common ancestral species?</i>	<ul style="list-style-type: none"> • Cetacean fossils, anatomy, embryos, and DNA • Number of shared genes between organisms
Heredity	<i>How do the differences in DNA that lead to differences in characteristics of organisms arise?</i>	<ul style="list-style-type: none"> • Canine trait similarities and differences
Natural Selection	<i>How do species change over time?</i>	<ul style="list-style-type: none"> • Highly variable pigeon traits • Rock pocket mice coloration • Variable lateral plate number in stickleback fish
Speciation	<i>How do new species arise?</i>	<ul style="list-style-type: none"> • Species is a human construct • Hawthorn fly speciation

Development and testing of the *Evolution* unit followed an iterative, multi-step, multi-year process in which student test data and feedback from teachers provided input to successive revisions of the unit. This case study focuses on the results of the final stage of testing, an efficacy trial comparing the unit with business-as-usual curricula in NGSS-adopting schools. Our guiding research questions were (1) Compared to a business-as-usual control, what is the effect of the *Evolution: DNA and the Unity of Life* curriculum materials on students' knowledge of evolution and ability to argue from evidence? (2) What are teachers' impressions of student learning and their own learning through engagement with the unit?

Case Study 2: Methodology

Study design. Ninth and tenth grade biology teachers were recruited via the curriculum developer's email list of over 20,000 educators nationwide and were selected to represent a broad range of geographic, socioeconomic, linguistic, and ethnic school-based contexts. Results from the pilot test of the curriculum unit indicated statistically significant increases in student scores from pre-test to post-test, setting the stage for a national efficacy field test that was designed to compare pre/post learning gains made by students whose teachers were randomly assigned to either the treatment (new unit) condition or control (business-as-usual) condition. The units used in both conditions targeted the same NGSS disciplinary core ideas for evolution. Teachers participated in a 1-1.5 hour training webinar before field testing.

Participants. Thirty-eight teachers (19 treatment, 19 control) and their students, representing 23 states and diverse teaching contexts, participated and are included in the analytic sample. Student were 50% female, 9% did not speak English as their primary language, 36% received free or reduced lunch, and 36% were from underrepresented ethnic or racial groups. No significant differences in the demographic categories were found between conditions.

Student pre- and post-tests. Assessments were developed by a project partner according to established test development procedures (Authors, 2014). The assessment tasks were aligned with the unit's learning goals and incorporated published scientific data but did not use the same phenomena as the unit lessons. Four test forms were created, each containing 26 multiple-choice and 2 constructed response items. Each test contained approximately the same number of items per topic and had the same average item difficulty. Students were randomly assigned to take one test version for their pre-test and another for their post-test. Students' constructed responses were scored using rubrics based on the CER framework. Three graders independently scored a subset of student responses, and clarifications were made to the rubric until sufficient inter-rater reliability was reached.

Teacher impressions and learning. Data on teachers' experience with and perceived learning from implementation of the unit in classrooms were collected through daily teacher logs, an end-of-implementation survey, an end-of-year survey, and post-participation interviews.

Data analysis. To analyze the significance of the treatment on students' pre- to post-test performance, percentage gains (post-test percentage correct minus pre-test percentage correct) were calculated for each student and compared using t-tests. Student gains were then fit to a mixed-effect hierarchical linear model (HLM) to estimate the influence of demographic and teacher effects.

Case Study 2: Results

Evolution unit's promise. The RCT efficacy trial results indicate that treatment and control students had statistically significant gains from pre- to post-test (see Table 4). Students in the treatment group had significantly higher gains than students in the control group ($t(1, 2,267) = 14.0, p < 0.001$). These findings provide support for the instructional approach used in the *Evolution* unit and suggest that the unit improves students' understanding of the targeted NGSS science ideas and practices (research question 1).

Table 4: Comparison of student pre-test & post-test scores for each curriculum unit

Unit	N	Pre-test (%)	Post-test (%)	Gain (%)	Effect Size ^a
Treatment (<i>Evolution</i>)	1165	42.2	60.4	18.3 ^b	1.14
Control (Business-as-usual)	1104	44.6	53.0	8.4 ^b	0.49

^aEffect size calculated by dividing the difference of the means by the pretest standard deviation

^bt-test $p < 0.001$

Additional results from HLM, module-level analysis, and the constructed response data (that assess argumentation skills) will be reported at the conference.

Teacher impressions and learning. Results show that teachers reported that (1) the unit carefully scaffolds the science practice of argumentation, with which students have little or no prior experience; (2) their practice was more student-centered when using the unit; (3) the unit was educative for them in terms of how to integrate the targeted NGSS practices and crosscutting concepts; and (4) the unit was educative for them in terms of how to integrate evolution and heredity for deeper student understanding (research question 2). Teachers also reported applying the unit's argumentation scaffold to other classes and shared what they had learned about NGSS with colleagues and administrators.

Discussion

In summarizing the challenges of NGSS implementation, a report for the Carnegie Corporation of New York (Bybee & Chopyak, 2017) articulates five characteristics that exemplify NGSS instructional materials. These same characteristics are used here to discuss challenges encountered in the two case studies.

Engage students in explaining phenomena. The development teams in both case studies tried to identify interesting and appropriate phenomena that students could make sense of using a set of coherent science ideas (both disciplinary core ideas and crosscutting concepts). For the MEGA unit, developers found data from a wealth of published studies on exercise physiology and human nutrition. For the *Evolution* unit, however, it was difficult to identify phenomena for which data are available that address key factors such as (a) variability, heritability and differential reproductive success as evidence for natural selection or (b) fossils, anatomy, embryonic, and DNA evidence for common ancestry.

Involve three-dimensional learning. Both development teams identified science practices (e.g., data analysis, modeling, explanation, and argumentation) that students would use, along with science ideas, to make sense of the phenomena. Because most students had no experience using science practices in middle school, both units needed to provide extensive scaffolding. For example, an activity in the MEGA unit first guides students in constructing a tentative model of how eggwhite proteins become muscle proteins, analyzing data from several experimental

studies, and using the evidence to revise their models prior to having student groups model how food starch becomes stored glycogen. Only then are individual students asked to model how triglycerides in butter can be used to produce human body fat or membrane phospholipids.

Likewise, the pilot test of the Natural Selection module for the *Evolution* unit revealed that students needed significantly more scaffolding for constructing arguments than the unit provided. The developers responded by incrementally building students' CER skills over the course of the 5-module sequence, using different phenomena for each level of practice. In the modules students are (1) introduced to the CER components of an argument, (2) identify evidence that supports a claim, (3) identify the appropriate reasoning that links evidence to claims, (4) organize evidence and write a supported argument, and (5) evaluate data and construct a written argument supported by an organizing worksheet.

In the MEGA unit, crosscutting concepts of both conservation and systems and system models were developed first in simple and/or more concrete systems (e.g., atom conservation modeled by counting LEGOs before computing with atomic masses) to scaffold development in more complex and/or more abstract systems (e.g., bar graphs and downward/upward arrows to model energy changes during chemical reactions and thick arrows to model energy transfers between systems). In the *Evolution* unit, crosscutting concepts were scaffolded using color codes to make areas of connection explicit and clear in the module outlines and in teacher notes.

Like their students, most teachers had not experienced three-dimensional learning nor did they have experience with implementing three-dimensional learning in their classrooms. The MEGA unit provided extensive scaffolding in the Student Edition itself as well as in the Teacher Edition (e.g., unit and chapter overviews, lesson guides, and teacher facilitation notes and prompts for responding to students' questions). For the *Evolution* unit, developers were challenged to provide supports that would fit into teachers' busy schedules. The unit's teacher materials were revised to provide easily accessible just-in-time supports for three-dimensional learning, including shorter, bullet points that required minimal reading .

Build K-12 progressions. The high school MEGA unit in Case Study 1 was originally designed to come after an 8th grade unit developed by the same team and targeting prerequisite ideas about atom rearrangement and conservation and several science practices. In reality, very few students in the study had experienced the 8th grade unit or anything like it, so developers added two lessons to the MEGA unit to help high school students make sense of the production of new substances during chemical reactions.

Developers of the *Evolution* unit in Case Study 2 faced similar problems in terms of students' prior knowledge and also different problems because of the modular nature of the unit. Because teachers often did not preview all of the lessons in a module or look ahead to later modules, they were unsure about how knowledge and skills were supposed to be built over the entire sequence. To address these issues, developers (a) added a short overview video at the beginning of each module, (b) clarified the lesson and module objectives in the teacher materials, and (c) added a "before you go on" list of student expectations for each module.

Align with Common Core English language arts and math. Both units engage students in writing evidence-based explanations and/or arguments that are consistent with the Common Core English Language Arts standards. In the MEGA unit, students use mathematics consistent with recommendations of 8th grade Common Core, such as analyzing and interpreting data from tables and graphs and using findings as evidence to support explanations about matter and energy

changes. Students also reason from correlational data to provide evidence of energy transfers from one system to another.

For the *Evolution* unit, developers found that some of the data they used for student data analysis in the first iteration of the Natural Selection module required mathematics that grade 9-10 biology students were not familiar with or found too difficult. Biology teachers who pilot tested this module indicated discomfort with teaching this math. The developers addressed this issue by either removing lessons that required this math or revising the lessons to exclude it. In developing the rest of the unit, they used middle-school-level math.

Link to sustained and continuous professional development. The different recruitment approaches of the two projects led to differences in their approaches to professional development. For the *Evolution* project, which recruited and provided online professional development to teachers nationwide, providing sustained and continuous professional development was not an option. As a result, developers built extensive teacher scaffolding into the materials themselves. While the MEGA unit also built extensive teacher scaffolding into the materials themselves, its long-term work with the participating school district makes sustained and continuous professional development an option.

Contributions

The proposed session will identify for the science education community, and for NARST members in particular, important issues that are central to the design and classroom implementation of NGSS-aligned curriculum. By describing how two development teams addressed the many challenges of realizing the vision of three-dimensional science learning and teaching in curriculum materials, the session provides other curriculum developers and researchers with important lessons learned. The two units can also serve as two different models of NGSS-aligned high school biology curriculum materials: the MEGA unit rooted in crosscutting concepts about matter and energy and the *Evolution* unit rooted in disciplinary core ideas about evolution and heredity. Together the units provide models for scaffolding students' engagement in the essential science practices of modeling and developing evidence-based arguments and explanations. Finally, the case studies provide examples of how to create materials that can be educative for teachers and serve as models for them as they strive to apply what they have learned to other areas of their classroom practice.

References

- Achieve, Inc. (2013). *Next generation science standards*. Retrieved from <http://www.nextgenscience.org/>
- 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. (2005). *High school biology textbooks: A benchmarks-based evaluation*. Retrieved from <http://www.project2061.org/publications/textbook/hsbio/report/default.htm>.
- American Association for the Advancement of Science /Project 2061. (2018). *Toward high school biology: Understanding growth in living things*. Arlington, VA: NSTA Press.
- Asterhan, C., & Schwarz, B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology, 99*(3), 626-639.
- Authors. (2014).

- Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, 94(5), 765-793.
- Bybee, R., & Chopyak, C. (2017). *Instructional materials and implementation of NGSS: Demand, supply, and strategic opportunities*. A report for the Carnegie Corporation of New York.
- Catley, K. M., Lehrer, R., & Reiser, B. (2004). *Tracing a prospective learning progression for developing understanding of evolution*. National Academies Committee on Test Design for K-12 Science Achievement.
- Catley, K. M., Novick, L. R., & Shade, C. K. (2010). Interpreting evolutionary diagrams: When topology and process conflict. *Journal of Research in Science Teaching*, 47(7), 861–882.
- Collins, A., Brown, J., & Newman, S. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum.
- Dougherty, M. J. (2009). Closing the gap: Inverting the genetics curriculum to ensure an informed public. *The American Journal of Human Genetics*, 85, 6-12.
- Driver, R. (1995). Constructivist approaches to science teaching. In L. P. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 385-500). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Genetic Science Learning Center, University of Utah Health Sciences. *Teach genetics: Evolution: DNA and the unity of life*. Retrieved from <https://teach.genetics.utah.edu/content/evolution/>.
- Linacre, J. M. (2017). WINSTEPS Rasch measurement computer program. Version 4.0.1. Beaverton, Oregon: Winsteps.com.
- McDermott, L. (1991). Millican lecture 1990. What we teach and what is learned – closing the gap. *American Journal of Physics*, 59, 301-315.
- Miles, M. A., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*, Sage, Thousand Oaks, CA.
- National Research Council. (2009). *A new biology for the 21st century*. Washington, DC: The National Academies Press.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.
- Postlethwait, J. H., & Hopson, J. L. (2009). *Modern biology (Teacher Wraparound ed.)*. Austin, TX: Holt, Rinehart, and Winston.
- Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal*. 50(5), 1020-1049.
- Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of science, cognitive psychology and educational theory and practice*. Albany, NY: State University of New York Press.
- Tabak, I., & Reiser, B. J. (2008). Software-realized inquiry support for cultivating a disciplinary stance. *Pragmatics and Cognition*, 16(2), 307-355.
- 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>.