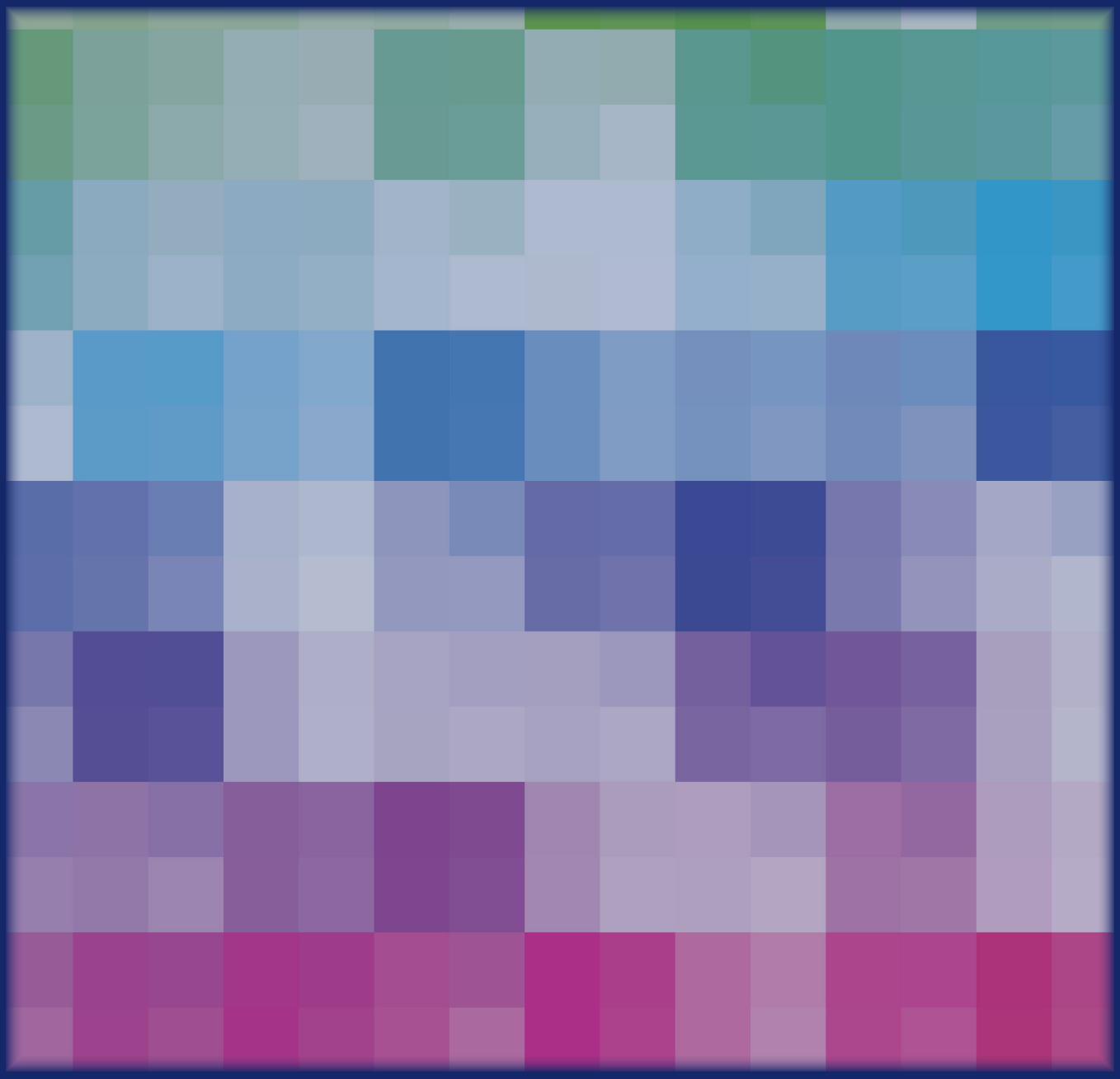
Opening the Door to Physics Through Formative Assessment



Edited by Cary Sneider, PhD
Associate Research Professor
Portland State University, Portland, Oregon
and
Brenda Wojnowski, EdD
President and CEO
WAI Education Solutions

Opening the Door to Physics Through Formative Assessment

Edited by Cary Sneider, PhD
Associate Research Professor
Portland State University, Portland, Oregon
and
Brenda Wojnowski, EdD
President and CEO
WAI Education Solutions

© 2013 Center for Science Education, Portland State University

ISBN 978-0-9828243-2-0

Key Words: Formative Assessment, Physics Instruction, Science Curriculum, High School, Research

Permission is hereby granted to print copies of this work or to circulate electronic copies for educational purposes, provided the above copyright is included in each copied document.

Reference as:

> Sneider, C., & Wojnowski, B. (Eds.) (2013). *Opening the door to physics through formative assessment* [monograph]. Portland, OR: Portland State University. Retrieved from National Science Education Leadership website: www.nsela.org/publications

Each of the articles in this monograph was partially supported by a National Science Foundation grant to Portland State University (NSF grant # 1020385 Bridging the Gap Between High School and College Physics: An Exploratory Study). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Please address all correspondence to Cary Sneider, csneider@pdx.edu, 503-788-6424

Design by Amy Layton, Sanger, Texas

Table of Contents

Acknowledgements iv

Preface vi
Cary Sneider, Portland State University, Portland, OR

Chapter 1. Course Structure and Learning Model of *Energizing Physics* 1
Jennifer Wells and Cary Sneider, Portland State University, Portland, OR

Chapter 2. Formative Assessment: Attending to the Substance of Student Thinking. 9
Lauren Yeiser and Cary Sneider, Portland State University, Portland, OR

Chapter 3. Shifting the Focus From Summative to Formative Assessment 18
Aaron Osowiecki, Boston Latin School, Boston, MA

Chapter 4. GUIDE-ing Students to a Better Understanding of Physics 24
Stephen Scannell, Gresham High School, Gresham, OR

Chapter 5. Summative Evaluation of *Energizing Physics* *32*
William Walker, High Tech High, San Diego, CA,
Stephen Scannell, Gresham High School, Gresham, OR,
Aaron Osowiecki and Jesse Southwick, Boston Latin School, Boston, MA,
Cary Sneider, Portland State University, Portland, OR

Chapter 6. Depth vs. Breadth in a New Inquiry Curriculum. 41
Jesse Southwick, Boston Latin School, Boston, MA

Chapter 7. The Road Ahead 48
Cary Sneider, Portland State University, Portland, OR

Acknowledgments

Project Staff

Project Director, Cary Sneider, Portland State University, Portland, OR
Administrative Assistant, Jennifer Wells, Portland State University, Portland, OR
Energizing Physics Authors, Aaron Osowiecki and Jesse Southwick, Boston Latin School, Boston, MA
Physics Teachers, Stephen Scannell, Gresham High School, Gresham, OR and William Walker, Westview High School, Portland, OR
Editor, Brenda Wojnowski, WAI Education Solutions (Wojnowski and Associates, Inc.)

We wish to acknowledge the following individuals who have contributed to this project:

Project Advisors

J. Myron Atkin, Emeritus Professor of Science Education, Stanford University
Marilyn Decker, Director of Science, Milton Public Schools, Milton, MA
Karen Draney, Senior Education Researcher, University of California Berkeley

2008-2009 Pilot Teachers

Megan Maresh, Westwood High School, Westwood, MA
Ellen Russell, Westwood High School, Westwood, MA

2009-2010 Pilot Teachers

Chris Doss, Boston Latin School, Boston, MA
Raquel Jacobson-Peregrino, Boston Latin School, Boston, MA

2010-2011 Pilot Teachers

Chris Doss, Boston Latin School, Boston, MA
Nan Ketpura-Ching, La Pietra Hawaii School for Girls, Honolulu, HI
Caitlin Santos, Cantwell School, Cantwell, AK

2011-2012 Pilot Teachers

Shu-Yee Chen, Newton North High School, Newton, MA
Chris Doss, Boston Latin School, Boston, MA
Michael Hazeltine, Newton North High School, Newton, MA
Jeffrey Lockhart, Westview High School, Beaverton, OR
Brian Morrison, Malden High School, Malden, MA
Alex Njoku, Boston Latin School, Boston, MA
Aaron Osowiecki, Boston Latin School, Boston, MA
Stephen Scannell, Gresham High School, Gresham, OR
Jesse Southwick, Boston Latin School, Boston, MA
Justin Strom, Westview High School, Beaverton, OR
William Walker, Westview High School, Portland, OR

We wish to express our appreciation to the following individuals who encouraged our efforts at our home institutions: William Becker, Director of the Center for Science Education at Portland State University, and Lynne Mooney Teta, Headmaster, Boston Latin School. Thanks also to Steve Rasmussen, founder of Key Curriculum Press, who introduced Aaron Osowiecki and Jesse Southwick to Cary Sneider and urged us to work together. Additionally, we very much appreciate the support of Joseph Reed, our program officer at NSF. We also wish to thank Elizabeth Allan, past president of the National Science Education Leadership Association (NSELA), Janey Kaufmann, chair of the NSELA Professional Development Committee, and all of the members of that committee for volunteering the time and expertise needed to provide excellent and helpful peer reviews for this manuscript.

Finally, we acknowledge and appreciate the efforts of more than 1,000 students who enrolled in *Energizing Physics* and helped to test the curriculum and various methods of formative and summative assessment developed for the course between 2008 and 2012.

Opening the Door to Physics Through Formative Assessment

Preface

In August 2010, my colleagues and I were awarded an exploratory grant from the National Science Foundation[1] to develop a formative assessment system for a new physics course called *Energizing Physics*. The course was developed by Aaron Osowiecki and Jesse Southwick, two physics teachers in Boston, and was tested by a dozen more physics teachers during the 2011-2012 school year. Nearly 1,000 students participated in the pilot, which took place on both sides of the continent, in Oregon and Massachusetts. This monograph reports the results of our work. Although very much a work-in-progress, we believe that our findings will be helpful to others doing similar work, both in physics and in other science domains. Before providing an overview of this body of work, which is the main subject of this preface, we provide a very brief sketch of the history of physics education so as to situate our work within a broader landscape.

Background

The improvement of physics education became a high priority in the U.S. after the launching of Sputnik in 1957. Development of PSSC Physics in the 1960s was a major break from the past, which established a high standard for laboratory-based science that continues to influence high school science today (Habershaim, 2006). The subsequent development of Harvard Project Physics made physics accessible to a wider spectrum of students by taking a historical approach and maintaining a strong inquiry element (Holton, 2003). Nonetheless, for many years enrollment in high school physics classes remained low, dropping to about 18% of students in the 1970s and 80s.

Interest in physics began to pick up again in the 1990s, due to an increasing sense of urgency about our education system. Many state departments of education increased high school graduation requirements to include more high school science and mathematics, thereby encouraging more students to enroll in all science courses, including physics. To accommodate the many additional students who may not have been as well-prepared academically as typical physics students a decade earlier, many teachers chose to focus on conceptual understanding of physics rather than mathematical problem solving, both to reduce cognitive mathematical preparatory demand and increase student interest. This approach is evident in the popularity of *Conceptual Physics* (Hewitt, 2009), an introductory high school textbook first published in 1987, which has continued to grow with increased enrollments in physics (Popkin, 2013). A complementary effort has been to offer physics as a first year course (Popkin, 2013), to provide a firm physical foundation for chemistry and biology. As a result of these efforts the number of students enrolled in high school physics has grown substantially since the 1990s, so that today physics enrollment stands at more than a third of all high school students (Neuschatz et al., 2005; Tesfaye & White, 2010). However, the field of physics education continues to be plagued with a dearth of well-trained physics teachers, with fewer than 50% having completed a college degree with a physics major (Neuschatz et al., 2005).

1 Bridging the Gap Between High School and College Physics: An Exploratory Study, Grant #1020385.

As we describe in **Chapter 1: Course Structure and Learning Model of *Energizing Physics***, our efforts have taken a different path. We began with the development of a new physics course by two teachers who wanted to reach a broad spectrum of students, but without giving up mathematical problem solving, which they viewed as essential to preparing students for college and careers (Osowiecki & Southwick, in press). These developers modified the course each year, seeking new ways to help students who struggled to acquire core ideas and skills. The work that we report in this monograph began with the recognition that the addition of a formative assessment system could provide an exceptionally valuable set of tools that would enable teachers to accomplish the primary goal of the course—to enable all students to succeed in high school physics while acquiring the interests, predispositions, and skills that would open the door to further studies of physics and related science and engineering courses in high school and beyond.

Formative Assessment Today

As we were writing this monograph, we came across Elliot Bennett's critical, but constructive, review of formative assessment literature (Bennett, 2011). We found the themes in his paper to parallel what we have learned from our own experiences over the past three years as we developed and tested our formative assessment system. In the following paragraphs, we will reference the key findings from Bennett's paper to introduce the seven chapters in this monograph.

Bennett (2011) makes a strong argument that to be maximally effective, a formative assessment system should be grounded in a discipline. His argument includes both the role of the teacher and intellectual tools and instruments. Regarding the teacher, he notes that "a teacher who has weak cognitive-domain understanding is less likely to know what questions to ask of students, what to look for in their performance, what inferences to make from that performance about student knowledge, and what actions to take to adjust instruction" (p. 15). Regarding intellectual tools and instruments, he claims that a cognitive-domain model is necessary to specify the progression of concepts and skills that students are expected to acquire, as well as tasks that enable teachers to determine where students stand along that progression.

The formative assessment system that we have developed for *Energizing Physics* supports Bennett's argument that formative assessment needs to be grounded in a discipline. In **Chapter 1: Course Structure and Learning Model of *Energizing Physics*** we describe the theory of action that undergirds the pedagogical approach, as well as the big ideas of physics and nature of science that lend coherence to the course content. The chapter also explains how detailed learning targets were developed to provide the cognitive-domain model as called for in Bennett's paper, and describes a means for assigning students to a level of accomplishment on each learning target. Observations and teacher reports during the pilot study showed that the specificity of learning targets and a means for judging levels of accomplishment were very valuable in determining which students needed additional help and in adjusting instruction accordingly.

A second key finding in Bennett's paper was a dichotomy in how formative assessment is defined. On the one hand some educators think of formative assessment in terms of instruments or quizzes that enable a teacher to gather evidence of student accomplishments so as to determine the future course of instruction. On the other side are those who think of formative assessment as a process that enables perceptive teachers to gain insight into student understanding. We were surprised at the vehemence with which individuals on both sides of this divide defended one point of view or another. In the end, we came to the decision that both are important, and we were gratified to find that Bennett came to the same conclusion:

> Arguably, each position is an oversimplification. It is an oversimplification to define formative assessment as an instrument because even the most carefully constructed, scientifically supported instrument is unlikely to be effective instructionally if the process surrounding its use is flawed. Similarly, it is an oversimplification to define formative assessment as a process since even the most carefully constructed process is unlikely to work if the 'instrumentation,' or methodology being used in that process is not well-suited for the intended purpose. 'Process' cannot somehow rescue unsuitable instrumentation, nor can instrumentation save an unsuitable process. A strong conceptualization needs to give careful attention to each component, as well as to how the two components work together to provide useful feedback. (Bennett, 2011, p. 7)

Chapter 2: Formative Assessment: Attending to the Substance of Student Thinking focuses on the process side of the discussion and addresses what Bennett and others have called a major flaw in the seminal work by earlier advocates of formative assessment—paying too much attention to whether or not the students "get" a specific idea rather than listening closely to gain insight into their thinking processes. This chapter identifies three different pedagogical approaches: *evaluative listening* in which the teacher listens for a specific vocabulary term, assuming that students understand the idea if they use the right words; *interpretive listening* in which the teacher listens to the substance of students' thinking; and *hermeneutic listening* in which the teacher is open to learning along with students during classroom conversations.

Chapter 3: Shifting the Focus From Summative to Formative Assessment concerns the instrumentation side of the discussion by describing the formative assessment tools that have been developed and tested during the *Engineering Physics* pilot study. The chapter explains how the tools are embedded within an instructional sequence that uses a variation of the 5E learning cycle: Engage, Explore, Explain, Elaborate, Evaluate (Bybee et al., 1989; Bybee, 2002). The chapter also provides a theory of action to describe how each tool is expected to operate, and describes how one of the tools was modified during the pilot study to make it more effective.

One of the major themes of Bennett's 2011 paper is the essential need for formative assessment to make valid inferences about the sorts of difficulties that students are experiencing:

> The centrality of *inference* in formative assessment becomes quite clear when we consider the distinctions among errors, slips, misconceptions, and lack of understanding. An error is what we observe students to make—some difference between a desired response and what a student provides. The error we observe may have one of several underlying causes. Among other things, it could be a slip—that is, a careless procedural mistake; or a misconception, some persistent conceptual or procedural confusion (or naïve view); or a lack of understanding in the form of a missing bit of conceptual or procedural knowledge, without any persistent misconception. Each of these causes implies a different instructional action, from minimal feedback (for the slip), to re-teaching (for the lack of understanding), to the significant investment required to engineer a deeper cognitive shift (for the misconception). (Bennett, 2011, p. 17)

A recent study of 181 teachers and nearly 10,000 middle school students (Sadler et al., 2013) provides additional support to Bennett's argument. Both students and their teachers were given a multiple-choice test of physical science concepts in which distractors were common misconceptions. As expected, students' and teachers' scores were correlated, suggesting that teachers who were more knowledgeable about science

were more effective teachers. In addition, teachers were asked to indicate how they thought their students were likely to respond to the questions. Analysis of the data showed that teachers who were aware of their students' misconceptions had much larger classroom gains than teachers who knew only the correct answer, but were not aware of their students' misconceptions.

The shift from summative to formative assessment described in Chapter 3 also discusses how the *Teacher's Guide* to *Energizing Physics* helps teachers anticipate the misconceptions that students may have, how best to use formative assessment tools to surface those ideas, and how to address them.

Chapter 4: GUIDE-ing Students to a Better Understanding of Physics describes a template developed for use in the course aimed at helping students and teachers sort out the different kinds of errors that Bennett describes. The paper is written from the perspective of a physics teacher, who was not one of the course authors, explaining how the template functions both to sort out different types of errors, and also as an instructional tool to help students approach physics problems systematically. The chapter also describes the results of a survey of other teachers and students in the pilot study concerning their perceptions of the effectiveness of the tool.

Although Bennett's paper is focused on formative assessment it also addresses the value of summative assessment. He points out that the common statement that formative assessment is assessment *for* learning, whereas summative assessment is assessment *of* learning, is overly simplistic, since a good summative assessment can also help to support learning. That is, students who study to prepare for a summative assessment learn as they are reviewing material in preparation for the test. The test itself can help students integrate prior learning, help them retain what they have learned, and sometimes help them learn a certain amount of new material. Summative assessments that are well aligned with course content and formative assessments can carry out their primary purpose of documenting what students know and can do, and thereby inform modifications of the course materials prior to being used again.

Chapter 5: Summative Assessment of *Energizing Physics* describes the development of a pre-post- assessment instrument and summarizes the findings of the pilot test year in response to the following research questions:

1) To what extent do students who enroll in *Energizing Physics* improve in their conceptual understanding of physics?

2) Are some physics concepts more challenging to learn than others? If so, which concepts are most challenging?

As a pilot, the purpose of the summative assessment was to inform the improvement of the course and assessment system. It was not intended to test the value of formative assessment. However, as we contemplated how a full or quasi-experiment could be designed to do so, we realized that while it would be possible to conduct a future study to compare the efficacy of *Energizing Physics* with a different course, it might not be possible to teach the same course with and without formative assessment, since such assessment is deeply embedded in the fabric of daily instruction for the course.

Chapter 6: Depth vs. Breadth in a New Inquiry Curriculum addresses the conundrum raised by our decision to emphasize depth over breadth. Taking the time for students to build conceptual and mathematical models, to conduct formative assessment, and to modify instruction appropriately forced us to limit the amount of content, which has implications for preparing students to take standardized high-stakes tests. Given the increasing number of students who take physics that would not have done so a decade ago, we believe the depth vs. coverage tradeoff is warranted. Nonetheless, it is an important issue for discussion.

Chapter 7: The Road Ahead is an epilogue to the series of papers, in which we describe a second pilot year that will be undertaken with the support of a commercial publisher that plans to publish *Energizing Physics*. It also addresses an issue that we have only just begun to consider and implement—professional development. Again, we find ourselves in agreement with Bennett's (2011) assertion that effective professional development takes time, especially in light of the knowledge that more than half of all physics teachers did not major in physics:

> Even if we can find a practical way to help teachers build pedagogical skill, deep domain understanding, and a sense of the measurement fundamentals, teachers need significant time. They need time to put that knowledge, skills, and understanding to practice, for example, to learn to use or adapt purposefully constructed, domain-based, formative-assessment materials. Such materials might include items, integrated task sets, projects, diagnostic tests, and observational and interpretive guides. Teachers also need time to reflect upon their experiences with these materials. If we can get teachers to engage in iterative cycles of use, reflection, adaptation, and eventual creation—all firmly rooted in meaningful cognitive-domain models—we may have a potential mechanism for helping teachers better integrate the process and methodology of formative assessment with deep domain understanding. (Bennett, 2011, p. 19)

It is fitting to close with a reminder to the reader that the formative assessment system for *Energizing Physics* is a work-in-progress. This monograph describes some of our challenges as well as our successes, and raises more questions than answers. We address these issues in **Chapter 7: The Road Ahead** where we discuss our own next steps and offer to help others pursuing similar ends so that they may avoid blind alleys and choose productive avenues to advance research in physics education and formative assessment.

References

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5-25.

Bybee, R. (2002). *Learning Science and the Science of Learning*. Arlington, VA: NSTA Press.

Bybee, R., Buchwalk, S., Crissman, D., Heil, D., Kuerbis, P., Matsumoto, P., & McInerney, J. (1989). *Science and technology education for the elementary years: Frameworks for curriculum and instruction.* Washington, D.C.: The National Center for Improving Instruction.

Habershaim, U. (2006). PSSC Physics, a personal perspective. *American Association of Physics Teachers.* Retrieved from http://www.compadre.org/portal/pssc/docs/Haber-Schaim.pdf

Hewitt, P. (2009). *Conceptual Physics: The High School Program.* Upper Saddle River, NJ: Prentice Hall.

Holton, G. (2003). The Project Physics course: Then and now. *Science & Education, 12*, 779-786.

Neuschatz, M., McFarling, M., & White, S. (2005). Highlights from the 2005 High School Physics Teachers Survey. *American Institute of Physics*. Retrieved from http://www.aip.org/statistics/trends/highlite/hs2/hshigh.pdf

Osowiecki, A., & Southwick, J. (in press). *Energizing Physics.* New York, NY: Freeman. (The course is also described at energizingphysics.com)

Popkin, G. (2013). "Physics First" battles for acceptance. American Physical Society. Retrieved from http://www.aps.org/publications/apsnews/200907/physicsfirst.cfm

Sadler, P. M., Sonnert, G. P., 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*(2).

Tesfaye, C. L., & White, S. (2010). High school physics courses and enrollments. *AIP Focus On*. Retrieved from http://www.aip.org/statistics/trends/reports/highschool3.pdf

Chapter 1

Course structure and learning model of *Energizing Physics*

Jennifer Wells and Cary Sneider
Portland State University, Portland, OR

Editors' Note: Chapter 1 describes a learning model based on Lev Vygotsky's social development theory and asks the question: To what extent does this theory provide a useful theory of action to explain how students acquire the knowledge and skills of high school physics? This chapter also describes the development of detailed learning targets and discusses how students can be assessed on their level of understanding of each learning target on a four-point scale. How might the theory suggest actions that a teacher could take to help students who are functioning at level 1 or 2 on a four-point scale?

Course structure and learning model of *Energizing Physics*

First year college physics is a gateway course for the vast majority of tomorrow's scientists and engineers, and for many other professions as well. The failure and dropout rate in these courses is high, severely limiting the pipeline of students who will make up tomorrow's technical workforce, and narrowing the opportunities for thousands of potential scientists, engineers, and technicians (Gainen, 1995).

The source of the problem may well be at the high school level, where students' interest in science tends to decline. Osborne's review of about 150 studies on students' attitudes towards science (Osborne, 2003) found that boys' and girls' initially positive interests in science tend to decline from age 11, and that the decline continues as students encounter high school science courses, especially physics, which they find to be difficult, boring, and disconnected from society.

One approach to the lack of student interest in high school physics has been to reduce the mathematical requirements and emphasize physics concepts and applications. High school curricula, such as *Conceptual Physics* (Hewitt, 2009), support this approach and have become very popular. There has also been a movement to change the traditional order of the school curriculum by starting with physics in 9th grade, where it can serve as a foundation for the other science disciplines (Lederman, 2001, 2005), and provide a foundational physics course to a much higher percentage of high school students (AAPT, 2006). A conceptual physics approach is well suited to this new sequence since our experience leads us to believe that few 9th grade students are sufficiently confident with algebra to handle a mathematically rigorous course.

Mathematical rigor is not the only challenge that physics students encounter. Educational researcher David Hammer (1994, 1996, 1997) reported than when he taught high school physics he found that many of his students believed that physics meant memorizing information and procedures supplied by the

professor or textbook and had little to do with the "real world." Hammer wanted his students to develop a sense of the discipline's underlying principles and coherence, to recognize that these principles apply to the every-day physical world, and to realize that learning is a challenging process of revising one's understanding of how the world works, both through laboratory experiences and lively and sometimes contentious discussions with other students.

Genesis of *Energizing Physics*[2]

Like David Hammer, high school physics teachers Aaron Osowieki and Jesse Southwick struggled to provide the kind of experiences that would help their students develop knowledge and skills in physics, as well as an understanding of physics as a means for conceptualizing how the world functions. They were dissatisfied with existing textbooks which rewarded students for memorizing facts and learning to use an algorithmic approach to problem solving, rather than engaging students in "doing physics" by applying logical thinking and an appropriate level of mathematics. Moreover, they perceived existing textbooks as lacking in coherence, with chapters on different topics that were not always related to the laboratory activities.

Aaron and Jesse began by producing individual lessons to help students develop a coherent understanding of physics and physics problem solution. Eventually they developed an entire first-year high school physics course that engaged students in mathematical problem solving through the means described by Hammer (1994, 1996, 1997). They decided to focus the course on the concept of energy based on their beliefs that: 1) students have an intuitive understanding of energy on which they could build; 2) energy is a scalar quantity so students do not need to learn to work with vectors right away; and 3) energy literacy is the most important concept in physics in aiding students to become productive citizens.

Learning Model

In *Energizing Physics* (Osowiecki & Southwick, in press) there is no distinction between lecture and laboratory activities. Each day the students work in small groups to explore physical phenomena, applying ideas and skills learned earlier, so the course scaffolds the gradual development of ever more powerful models that students can apply to the physical world. Instruction takes place within the context of laboratory activities, discussions, and short lectures relevant to the task at hand. Activities include engineering design projects that require the students to apply physics concepts. For example, in the *Bungee Egg Drop Challenge*, students apply their knowledge of energy conservation and Hooke's Law to provide a thrilling, yet safe, bungee drop for a raw egg. The course materials include a teacher guide and student reader; but the most important resource is the student workbook which structures daily activities, such as lab work, concept applications, discussion prompt questions, and embedded assessments. The class often engages in a large group discussion led by the teacher, who presents ideas "just in time" to support the next stage of small group work. In brief, teamwork in conducting meaningful activities and solving problems is the essence of the *Energizing Physics* experience.

An early champion of this approach was Lev Vygotsky, the Russian psychologist who was especially interested in how children learn science. In his famous book, *Thought and Language* (1934, revised 1962, revised 1986), Vygotsky pointed out that words are more than just vocabulary. Words are concepts that represent a class of objects or events that are similar in some important way. A person who has acquired the

2 Information on *Energizing Physics* can be found at energizingphysics.com

concept of "tree," for example, can identify many different kinds of objects as trees, including ones they had never seen before.

Vygotsky had little use for classroom lectures. "Practical experience shows that direct teaching of concepts is impossible and fruitless. A teacher who tries to do this usually accomplishes nothing but empty verbalism, a parrot-like repetition of words by the child, simulating knowledge of the corresponding concepts but actually covering up a vacuum" (Vygotsky, 1962, p. 83). Instead, Vygotsky's experiments with teaching and learning led him to the conclusion that children learn best in a social environment, where they are guided in their learning through engaging questions and discussions with peers. Although most of his work was with elementary-age children, he claimed that similar principles applied to older students as well.

Formative assessment aligns with Vygotsky's theory of learning in a way that traditional assessment does not. Traditionally, a student's learning is measured by giving a test at the end of a lesson or unit to see what a student can do when working alone. In formative assessment the teacher observes students' capabilities and understandings during the lesson, as they interact with other students. The process of formative assessment is similar to Vygotsky's concept of a *zone of proximal development*, meaning the range of problems that a student is able to solve with the assistance of others. When students learn to solve problems in their zone they not only gain new capabilities, they are also challenged at just the right level, so that learning is engaging and enjoyable. If the task is too easy the student will be bored; if it is too hard the student can become discouraged. It's the teacher's job to make sure that the level of challenge meets each student's needs, which is why formative assessment is so important. If the teacher doesn't understand what the student can do or cannot do, it is impossible to set an appropriate level of challenge so the student can integrate the new ideas into their existing knowledge structures.

During the summer before the pilot, we developed a matrix to describe the level of accomplishment that we wanted our students to achieve on both the conceptual and mathematical aspects of the course. The matrix is shown in Table 1. Levels 0, 1, and 2 are steps towards a solid understanding as represented by level 3, while level 4 is advanced.

Table 1. Levels of understanding the two major constructs

Level	**Conceptual Understanding** Students can…	**Mathematical Modeling** Students can…
4 Recognizing Limitations or Modifying Models	(4a) Describe how to adjust the models for different assumptions. (4b) Understand the limitation/assumptions of the model.	(4a) Describe how to adjust the models for different assumptions. (4b) Understand the limitation/ assumptions of the model.
3 Application	Apply model (diagram/words) to accurately predict/explain/solve.	Apply mathematical model to accurately predict/explain/solve.
2 Representation	Use diagrams/words to model the system under investigation.	Adjust/Combine the appropriate tools to represent the specific situation.
1 Identification	Identify the relevant physical features of a situation.	Identify variables and appropriate tools related to a system under investigation.
0 No Connection	Unable to represent a physical situation with an appropriate model.	Unable to connect variables to the situation.

We tested the reliability of the matrix to characterize students' thinking during a summer professional development program for the pilot teachers. We provided the teachers with student papers in response

to some of the short assessment quizzes, and, in nearly all cases, the teachers agreed on the level of performance or differed by no more than one level. The use of the matrix throughout the year provided a useful way to discuss student achievement with the students and with each other.

During this year's pilot test of *Energizing Physics* teachers used the feedback from student work in various ways to help them decide how to proceed in class. For example, one teacher drew upon the students who earned four points on the short quizzes to work with other students in small groups, while the teacher circulated among the groups helping where needed. In a proficiency-based education program a team of two teachers worked together to divide the students into two skill-groups. Students who worked at level 1 or 2 worked with one of the teachers to revisit the concepts they misunderstood or missed, while students at level 3 or 4 worked with the other teacher to apply their understanding to more challenging problems.

The 5E Learning Cycle

As an example of how these pieces fit together, imagine a lesson about the transfer of potential to kinetic energy that begins with a series of wooden ramps with different heights and slopes. Grooves on the slanted edge of each ramp allow students to roll marbles smoothly down the sloped surfaces, much as Galileo did 400 years ago. The teacher ***engages*** students' prior knowledge and interest by asking them to predict which ramps will yield the fastest moving marble as it leaves the bottom of the ramp. The students are then free to ***explore*** the apparatus and refine or modify their predictions as they use meter sticks to measure the height of the ramps and length of the sloped surfaces.

At this stage of the activity most of the students think the steepness of the slope is the most important factor, while others think it's the length of the slope or height of the top of the ramp. The students are then given stopwatches and assigned to teams to design experiments to test their predictions. They collect data and ***explain*** their results. Most of the teams conclude that height of the ramp is the most important factor, but don't understand why that factor is so important. At this point the teacher ***extends*** the learning by introducing the concepts of gravitational potential energy and kinetic energy, to provide a more satisfying explanation for why height of the ramp is the most important factor.

Students ***evaluate*** their work throughout the activity, commenting on the accuracy of their predictions, the quality of their experiment design and observations, and their understanding of the connection between their findings—that height is the important factor—and the concepts of potential and kinetic energy presented by their teacher.

As illustrated by the bold font in the paragraph above, the curriculum materials engage students in qualitative and quantitative thinking facilitated through a slightly modified 5E learning model (Bybee, 2002; Bybee et al., 1989), which plays out in nearly every lesson or sequence of lessons as follows:

Engage. The teacher engages students' interest and probes their current ideas.

Explore. Students explore a natural phenomenon by engaging in a meaningful task.

Explain. Students explain the phenomenon, and listen to the explanations of others.

Extend. The teacher helps students deepen their understanding.

Evaluate. Students assess their science practices and understanding of key concepts.

Recall Vygotsky's notion that a word is really a concept, in light of this lesson about potential energy. Vygotsky observed that students first recognize differences before they perceive the underlying similarities. Shooting arrows from a bow and bungee jumping from a bridge at first appear to be very different kinds of events. But in both cases potential energy is transferred to kinetic energy. These events undergo a similar transformation to that of water plunging over a waterfall. Essential parts of the learning process involve

becoming consciously aware of the concept and developing the ability to express the concept in common language, which is why small group and class discussions as well as student writing are so valuable. In the case of physics, it is also important for the students to express these relationships mathematically, and to use their growing skill in mathematics to predict what will happen in a given instance—an ability that is not only important in science, but also in engineering and many other fields.

First Steps Toward a Formative Assessment System

The first major activity under our grant from the National Science Foundation[3] was to attend a one-week workshop at the Berkeley Evaluation and Assessment Research (BEAR) Center at the University of California, conducted by Director Mark Wilson and Senior Researcher Karen Draney. The objectives of the workshop were for the Berkeley team to share a vision for producing valid and reliable assessment instruments, and to help our team develop a clear description of "constructs" that represent what students are expected to learn. With the help of the Berkeley researchers we arrived at the following broad constructs for the abilities that we wanted *Energizing Physics* students to acquire:

Conceptual understanding of physics, which is a student's ability to apply a conceptual model of a physical principle to a situation in the physical world; and

Mathematical modeling ability, which is the ability to represent and manipulate an appropriate model mathematically to make predictions and solve physics problems.

We constructed a matrix like the one shown in Table 1 to represent succeeding levels of knowledge/capability for each core idea in the course. Table 2 illustrates the conceptual and mathematical levels of achievement with regard to the concept of energy conservation, as it is presented in Chapter 2. Notice how closely the mathematics table mirrors the conceptual table. Essentially these are two sides of the same coin, which is the way that we want students to come to see these complementary ways of visualizing the world.

Table 2. *Energizing Physics* Chapter 1 Constructs—Conservation of Energy

Level	Conceptual Understanding Students can…	Mathematical Modeling Students can…
4 Recognizing Limitations or Modifying Models	Recognize the limitations of the conservation of energy principle for a particular situation.	Adjust the model to compensate for the limitations of the model in the particular situation.
3 Application	Utilize the conservation of energy principle to describe changes in the behavior of a system.	Apply the mathematical model to accurately predict/explain/solve for situations involving changes in energy for a system.
2 Representation	Sketch the situation, determine the types of energy at different locations, and identify if work is done on the system.	Able to build a mathematical model for a situation using the conservation of energy principle.
1 Identification	Identifies the variables related to the energy of a system.	Able to connect values associated with the energy of a system.
0 No Connection	Unable to identify the variables associated with a situation.	Unable to connect variables to the situation.

3 Bridging the Gap Between High School and College Physics: An Exploratory Study, Grant #1020385

Coming out of our work at Berkeley, we realized that the Teacher Guide needed to very clearly illustrate what levels 0-4 looked like for each of the core physics concepts. To do that we had to specify what we wanted students to be able to do. Will Walker, a physics and chemistry teacher who had recently joined the staff, proposed that we develop "learning targets," a means that he found useful in specifying what capabilities he wanted his students to develop. Written from the student's perspective, these numbered targets allowed teachers and students to track the accomplishment of each objective through formative and summative assessments, identifying areas of difficulty and success. Table 3 shows some of the learning targets associated with conservation of energy.

Table 3. Learning Targets for Conservation of Energy in Chapter 2 of *Energizing Physics*

LT 1.3 (I can) Quantify kinetic energy, gravitational potential energy, and total energy.
LT 1.4 (I can) Use energy conservation and energy models to quantify variables in simulations involving two types of energy.
LT 1.7 (I can) Use the relationship between force and stretch to quantify spring force, stretch, and/or stiffness.
LT 1.9 (I can) Use energy conservation and energy models to quantify variables in simulations involving three sources of energy.

Overall Structure of the Course

In light of research showing that high school science students are more likely to succeed in college physics courses if they spend more time focusing on fewer concepts in greater depth (Schwartz et al., 2008) our research team decided to focus efforts on chapters 1-5 of the *Energizing Physics* text, which presented the most important concepts of the course.

As illustrated in Figure 1, once a concept is introduced, it is used not only in that chapter but also later in the course to reinforce learning. For example, in Chapter 1 students begin by designing a device to measure average speed. The concept of average speed and the process of engineering design are then used throughout the course. In Chapter 2 students measure the average speed of marbles rolling off ramps to "discover" conservation of energy. The chapter goes on to introduce different types of energy (elastic and gravitational potential energy and kinetic energy) and the law of energy conservation. In Chapter 3 students apply their understanding of average speed and different types of energy to more complex cases involving work and various forces, including friction. Newton's second law of motion is the primary focus of Chapter 4 in which students consider balanced and unbalanced forces that cause acceleration. Chapter 5, which concerns energy and power in electrical systems, incorporates concepts from the prior chapters, and then has students extend their learning by analyzing the electrical system in their homes and recommend ways that their families can save energy and money.

The rationale for engaging students in applying the same concepts several times throughout the course is based on decades of research showing that even after "correcting" their misconceptions in physics, few students will give up their prior beliefs until they see that their new learnings are useful in solving other problems. (See, for example, Minstrell and Kraus, 2005).

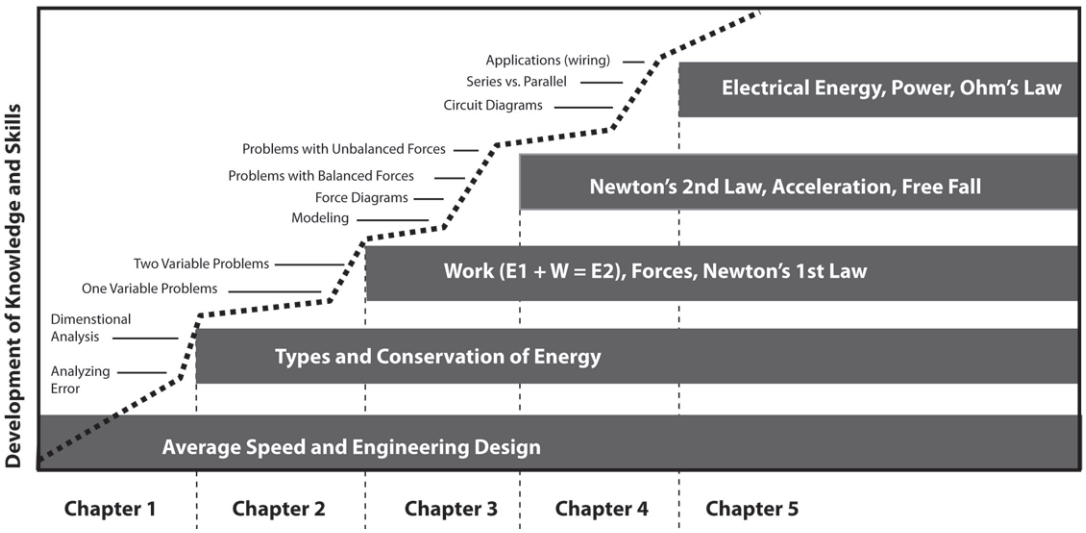


Figure 1. Overview of chapters 1-5. Major concepts in each chapter are shown in shaded bars. The dotted line is a qualitative view of student progress, and challenging new skills are listed to the left of the dotted line.

The dotted line in Figure 1 is a qualitative view of student progress, with a steep slope indicating where students are learning challenging new skills (listed to the left of the dotted line) and a shallower slope indicating smoother progress as students consolidate their understanding and apply concepts and skills to new situations.

Chapters not covered by this pilot study, but which some of the pilot teachers found time toward the end of the academic year to teach, were Chapter 6: Waves, Chapter 7: Thermal Energy, Chapter 8: Multiple Objects and Multiple Dimensions, and Chapter 9: Optics. As a general guideline, however, teachers were not expected to cover all nine chapters, but rather for all of the teachers to teach at least the first five chapters, plus those that each teacher thought were most relevant to their students' needs and interests.

Discussion

In summary, with *Energizing Physics,* students learn the conceptual and mathematical sides of physics by approaching the discipline in the ways that scientists do—through observation, measurement, analysis, prediction, testing, and discussion— to make sense of what they have learned.

In order to accommodate formative assessment strategies embedded directly within the curriculum, we have identified learning targets as clearly as possible, so that students and teachers can determine what skills the students have acquired, and where gaps exist. Our hypothesis is that as more students receive the assistance they need to make sense of physics, they will be less fearful of the subject, and that, over time, the reputation of physics as a "tough subject" will change, not because it is made easier by eliminating the most challenging concepts or leaving out the math, but because students will have more opportunities to make sense of their learning experiences and see physics as a way to make sense of the world around them.

Acknowledgements

We are indebted to the teachers and students who participated in this study. This work was partially supported by a National Science Foundation grant to Portland State University (NSF grant # 1020385 Bridging the Gap Between High School and College Physics: An Exploratory Study). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

AAPT, (2006). *Physics first: An informational guide for teachers, school administrators, parents, scientists, and the public.* High School Committee of the American Association of Physics Teachers. Retrieved from http://www.aapt.org/aboutaapt/ennouncer/upload/physicsfirst.pdf

Bybee, R. (2002). *Learning science and the science of learning.* Arlington, VA: NSTA Press.

Bybee, R., Buchwalk, S., Crissman, D., Heil, D., Kuerbis, P., Matsumoto, P., & McInerney, J. (1989). *Science and technology education for the elementary years: Frameworks for curriculum and instruction.* Washington, D.C.: The National Center for Improving Instruction.

Gainen, J. (1995). Barriers to success in quantitative gatekeeper courses. In J. Gainen & E. W. Williamson (Eds.), *Fostering Student Success in Quantitative Gatekeeper Courses* (61, 5-14). San Francisco: Jossey-Bass.

Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction, 12*(2), 151-183.

Hammer, D. (1996). More than misconceptions: Multiple perspectives on student knowledge and reasoning, and an appropriate role for education research. *American Journal of Physics, 64*(10), 1316-1325.

Hammer, D. (1997). Discovery learning and discovery teaching. *Cognition and Instruction, 15*(4), 485-529. Retrieved from http://dx.doi.org?10.1207/s1532690xci1504_2

Hewitt, P. (2009). *Conceptual physics: The high school program.* Upper Saddle River, NJ: Prentice Hall.

Lederman, L. M. (2001). Revolution in science education: Put physics first. *Physics Today, 54*(9), 11-12.

Lederman, L. M. (2005). Physics First? *The Physics Teacher, 43*(1), 6.

Minstrell, J., & Kraus, P. (2005). Guided inquiry in the science classroom. In M. Donovan, & J. Bransford (Eds.), *How Students Learn Science in the Classroom* (pp. 475-512). Washington, D.C.: National Academy Press.

Osborne, J. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049-1079.

Osowiecki, A., & Southwick, J. (in press). *Energizing physics.* New York: Freeman. (The course is also described at energizingphysics.com)

Schwartz, M. S., Sadler, P. M., Sonnert, G, & Tai, R. H. (2008). Depth Versus Breadth: How Content Coverage in High School Science Courses Relates to Later Success in College Science Coursework. *Science Education, 93*, 798-826.

Vygotsky, L. S. (revised 1962). *Thought and Language.* (E. Hanfman & G. Vakar, Trans., Ed.). Cambridge, MA: The MIT Press. (Original work published 1934.)

Vygotsky, L. S. (revised 1986). *Thought and Language.* (A. Kozulin, Trans., Ed.). Cambridge, MA: The MIT Press. (Original work published 1934.)

Chapter 2

Formative assessment: Attending to the substance of student thinking

Lauren Yeiser and Cary Sneider
Center for Science Education
Portland State University, Portland, OR

Editors' Note: This chapter reviews a number of recent papers that take issue with earlier reports of formative assessment studies, questioning whether or not examples in those studies showed teachers attending to the substance of their students' thinking, and pointing out other concerns. These ideas have profound implications for professional development, training of new teachers, and evaluation of current teachers. They also have important implications for curriculum development, as we will show in subsequent chapters about formative assessment in the *Energizing Physics* curriculum. How might these ideas affect your own professional work?

Formative assessment: Attending to the substance of student thinking

The purpose of formative assessment is to facilitate the learning process and not to assess what students have learned in order to award a final grade. Formative assessment has three components: identifying learning goals, assessing where students are with respect to those goals, and using that assessment to inform adaptive instruction. Over the past two decades, formative assessment has become a widely accepted means to enhance student performance, and has become a routine element in national discussions on education reform (Furtak, 2012).

Despite the general acceptance of formative assessment, several researchers have recently taken a critical view of the most widely cited studies (e.g., Coffey et al., 2011; Bennett, 2011). One concern is that many prominent publications have ignored the substance of student thought. Another is that prior studies have failed to distinguish between students' knowledge of an idea and their understanding of it. For the purpose of this paper, "knowledge" refers to content knowledge, while "understanding" includes the additional ability to synthesize, apply, and extrapolate from that knowledge. For formative assessment to be maximally effective, teachers must be able to gauge students understanding, not just students' content knowledge.

The purpose of this review is to briefly summarize a few key papers on the educational strategy of formative assessment, and to draw conclusions about the importance of both curriculum and professional development in maximizing the value of formative assessment for the science classroom.

Seminal Studies

In 1998, Paul Black and Dylan Wiliam (1998a) published what would become a seminal research review about formative assessment based on their analysis of 580 research studies and their own research. Their influential article in *The Phi Delta Kappan* (Black & Wiliam, 1998b) positioned formative assessment as a critical component of education reform. The authors claimed that, "innovations that include strengthening the practice of formative assessment produce significant and often substantial learning gains" (Black & Wiliam, 1998b, p. 141). Consequently, changes that foster formative assessment would be pivotal in raising academic performance. They argued that national standards, which emphasize summative assessments, neglect "the process of teaching and learning in the classroom" (Black & Wiliam, 1998b, p.139), and that it is what teachers and students do in the classroom that drives learning.

The studies that Black and Wiliam reviewed covered a large diversity of topics, from peer assessment to teacher questioning behavior, but all are assumed to have one thing in common: the information gained during assessment is used as feedback to adapt instruction and, therefore, directly impacts what teachers and students do in the classroom.

Since their 1998 publication, Black and Wiliam have made multiple contributions to the body of formative assessment literature. In 2009, the authors produced an extensive review on the theory of formative assessment. Importantly, this article provides a comprehensive definition of formative assessment, and clearly articulates the accepted strategies of formative assessment. Black and Wiliam (2009) concluded:

> Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited. (p. 9)

In accordance with this definition, the authors proposed five strategies that instructors can implement to support an atmosphere of formative assessment:

(1) Clearly delineate and articulate learning intentions and criteria for success.
(2) Support classroom activities that elicit evidence of student understanding.
(3) Provide feedback that stimulates the learning process.
(4) Foster productive peer evaluation and instructions.
(5) Promote student investment in learning. (Black & Wiliam, 2009)

Black and Wiliam's research reviews and compelling popular articles have led to a number of books and articles for teachers about how best to conduct formative assessment. For example, *Everyday Assessment in the Science Classroom* (Atkin & Coffey, 2003) is a collection of essays about the various methods and uses of formative assessment. Black and Wiliam's publications are among the most cited works on formative assessment and collectively the two authors have shaped the discussion of formative assessment.

Attending to the Substance of Student Understanding

In an insightful critique of prior work on formative assessment Randy Bennett (2011) has emphasized the importance of distinguishing between eliciting evidence of student knowledge and making inferences from that evidence. If a teacher makes incorrect inferences, subsequent changes in instruction will be without valid justification and learning may be hindered. Bennett posits that eliciting evidence demonstrates

content knowledge, whereas inference allows the distinctions among "slips –that is, a careless procedural mistake; or a misconception, some persistent conceptual or procedural confusion (or naïve view); or a lack of understanding in the form of a missing bit of conceptual or procedural knowledge, without any persistent misconception" (Bennett, 2011, p. 17).

It follows that a critical component of a well-designed formative assessment activity is that "different observers [can] draw similar inferences about a student's skills from the same evidence; that the inferences drawn are consistent with other, more in-depth methods of characterizing what a student knows and can do" (Bennett, 2011, p. 14). Consequently, Bennett argues that to realize the considerable promise of formative assessment, we must not only discuss formative assessment strategies, but also empirically demonstrate how implementation of those strategies fosters an understanding of the substance of student thought.

Bennett (2011) was not the first author to criticize formative assessment literature for neglecting the substance of student reasoning. In the "Missing Disciplinary Substance of Formative Assessment" Coffey et al. (2011), commented that "in its concentration on strategies for the teacher, the literature [on formative assessment] overlooks the disciplinary substance of what teachers and students assess" (p. 1). To support this argument, the authors analyzed four prominent formative assessment studies, and contrasted these studies with examples from their own research. The authors found that in focusing on general teacher strategies, there was minimal discussion of how those strategies reveal the substance of student thinking. As a result, there was tacit acceptance of the traditional notions of content as a body of predetermined correct information (Coffey et al., 2011). To illustrate this point, compare the following examples. The first is taken from Black et al. (2003), and the second from the research of Coffey et al. In the first example, the teacher asked the students to explain why plant growth is, in part, dependent on location.

> Monica: That one's grown bigger because it was on the window [pointing].
> Teacher: On the window? Mmm. What do you think Jamie?
> Jamie: We thought that…
> Teacher: You thought…?
> Jamie: That the big 'un had eaten up more light
> Teacher: I think I know what Monica and Jamie are getting at, but can anyone put the ideas together? Window-Light-Plants? (Black et al., as cited by Coffey et al., 2011, p. 1110)

The teacher continued in a similar manner until one student mentioned the word photosynthesis, at which point the teacher wrote photosynthesis on the board and asked "can anyone put Plant, Light, Window and Photosynthesis together and tell me why these two plants have grown differently?" (Black et al., as cited by Coffey et al., 2011, p. 1110)

As Coffey et al. (2011) explained, when the teacher asked the students to put the ideas together, (s)he did not have enough information to know what the students meant by those ideas. For example did Jamie think that plants eat light in the same way that humans eat food? Did Monica think that the plant grew faster on the window because of light, heat, or better view? The teacher neglected the substance of student thought and instead focused on eliciting "correct" terminology. The class discussion was intended to introduce the concept of photosynthesis and connect photosynthesis with plant growth, not to dissect student understanding of the relevant concepts. As a result, the teacher "tacitly treated subject matter as

information, and the assessment of student thinking as a check on its correctness against that information" (Coffey et al., 2011, p. 1122).

Compare the above example to one from data collected by one of Coffey et al. authors during a three-year study of how high school teachers interact with students, and how those interactions altered instruction. In this example, the teacher, Terry, intended to briefly review the definition of matter before beginning a new chemistry unit. When he asked his students to define matter, they mentioned that it could be seen; it could be touched; it takes up space; and it had mass. As opposed to simply elaborating on the terminology mentioned by students, Terry asked leading questions to determine what students meant in each case. These questions led to a discussion of whether or not air is matter. The class initially concluded that air was not matter, but through a series of discussions and examples with balloons, the class came to the consensus that air was matter because it took up space and had weight. The key difference between the discussion led by Terry and that presented by Black et al. (2003) was that, in the second case, Terry attempted to understand the substance behind students' utterances. His discussion did not end when students had listed the correct terminology, but instead examined the students' understanding of the words they were using (Coffey et al., 2011). As these examples demonstrate, it is not what teachers do, but rather what they notice that makes assessment formative. In both cases the teachers engaged the class in discussions, but, unlike the first teacher, Terry probed to understand the meaning behind the students' words before deciding how to adapt instruction to fit his students' needs.

Different Types of Listening

The contrast between the discussion lead by Terry and that described in Black et al. (2003) highlights the need to characterize different types of listening. Such a characterization is described in "Listening for differences: An evolving conception of mathematics teaching" (Davis, 1997). The study summarized a yearlong collaboration between the author (Davis) and a middle school math teacher (Wendy). The research explored how Wendy's listening evolved when working collaboratively with Davis to prepare units, review relevant literature, and team teach. Davis identified three types of listening: evaluative, interpretive, and hermeneutic. Evaluative listening is the most restrictive and is characterized by listening *for* something, not *to* the speaker. When engaging in evaluative listening, Wendy framed her questions with a preconceived answer in mind, and would guide students to that answer or answer the question herself. Interpretive listening diverges from evaluative listening in that the listener pays attention to the reasoning and substance expressed by the speaker. During interpretive listening, questions are response seeking, and the listening attempts to "access the subjective sense being made" (Davis, 1997, p. 365). That said, during interpretive listening, the listener is still constrained by the expectation of a right and wrong answer. Hermeneutic listening, on the other hand, is a genuine dialogue in which both participants' ideas may change. As explained by Davis:

> Instead of seeking to prod learners toward particular predetermined understandings [during hermeneutic learning] Wendy seems to have engaged, along with her students, in the process of revising her own knowledge of mathematics…what seems to have been abandoned is the belief that teaching is a matter of *causing* learners to acquire, master, or construct particular understandings through some pre-established (and often learner-independent) instructional sequences. In [the case of hermeneutic learning] learning was a social process, and the teacher's role was one of participating, of interpreting, of transforming, of interrogating – in short, of *listening*. (p. 371)

It is important to clarify that in discussing hermeneutic listening, Davis (1997) was not proposing that teachers abandon content knowledge and allow students to reinvent the wheel. Rather, Davis argued that teachers should be open to new ways of thinking about and applying content knowledge, and should recognize that students' understanding of content knowledge is affected by social and historical circumstances.

Applying Davis' (1997) characterization of learning to the concerns raised by Coffey et al. (2011) and Bennet (2011), it becomes clear that for assessment to truly be formative, teachers must abandon evaluative listening, and at least strive for interpretive, and, if possible, hermeneutic listening. Although Davis focused his analysis on verbal discussions, these same categories can be applied to written communications. In the next section "listening" will refer to listening during both written and verbal transactions.

Teaching Teachers to Listen

A powerful example of the pedagogical changes that can occur when teachers abandon evaluative listening and learn to understand the substance of student thinking, is demonstrated in "Teacher learning in mathematics: Using student work to promote collective inquiry" (Kazemi & Franke, 2004). The study was designed to describe what teachers learn through a collaborative discussion of student work. The authors collected data during seven workgroups including ten teachers from the same elementary school. The data included transcripts of audio recordings, student work, written teacher reflections, and a final teacher interview. The workgroups met once a month throughout the school year and were guided by a research facilitator. Prior to the workgroups, each teacher had his/her classes solve the same mathematical problem. Student responses to these problems were the basis of each meeting. In each workgroup the facilitator explicitly asked the teachers to discuss the details of problem solving strategies employed by students while answering the common problem.

Kazemi and Franke (2004) found that there were two shifts in teachers' workgroup participation. The first shift demonstrated an enhanced ability to attend to the substance of student thought. Over the course of the study, the teachers went from being unable to discuss the specifics of how students solved the problems, to having an in-depth understanding of the various strategies utilized by students. An important paradigm shift was the recognition that insight into student thought requires more than simply asking students to solve a problem. Teachers began asking students to explain, demonstrate, model, and/or defend their answers. As the teachers learned to listen to their students, they found that students employed a diversity of creative and intelligent strategies. Importantly, "The student work also allowed the teachers to begin to see themselves as mathematical thinkers when they were willing to struggle through student strategies they did not understand" (Kazemi & Franke, 2004, p. 230). Thus learning to attend to the details of student strategies, not only provided the opportunity to recognize that students have powerful mathematical ideas, but also positioned the teacher as a learner. This is hermeneutic listening.

The insight into student thought gained by the teachers over the course of the study facilitated the second shift—the development of instructional trajectories. As teachers acquired an understanding of student strategies, they also gained important insight into how they could design instruction to build on students' current mathematical thinking. For example, one teacher taught her students to work with a tens bar by having her students solve problems involving multiples of ten. This adaptive change came from the recognition that her students understood the concept of the tens bar but did not comprehend its utility (Kazemi & Franke, 2004). Over the course of the year, the participating teachers learned how to elicit and interpret evidence of student thought, and used their insights to inform instruction. Critically,

the participants went beyond determining if their students understood a predefined body of content knowledge, and learned how to listen for the substance behind their students' words.

Kazemi and Franke (2004) demonstrated that teachers could learn to engage in the substance of student thinking. This next study illustrates that achieving that goal can be very challenging. In "Linking a learning progression for natural selection to teachers' enactment of formative assessment," Furtak (2012) explored the utility of using learning progressions of natural selection as a scaffold for formative assessment. The study took place over a two-year period at an ethnically and socioeconomically diverse high school. Data from six biology classes were analyzed and included teacher interviews, classroom videotapes, and stimulated teacher recall interviews. The first year of the study was dedicated to professional development. The teachers participated in discussions of how to elicit and analyze student reasoning, how to address naive conceptions, and how to address the critical components of formative assessment. Teachers were also required to review videos of each other implementing formative assessment activities and analyze student thinking and teacher responses. At the end of the first year, the participants came together and compiled a short list of formative assessment activities to implement the following year.

Furtak (2012) found that while teachers were able to pick out students' ideas as they related to the learning progression, four of the six teachers simply corrected misconceptions without adapting instruction or considering the substance of student thought. These teachers primarily engaged in evaluative listening and "drew upon the learning progression to more quickly identify and do away with ideas they viewed as wrong" (Furtak, 2012, p. 1206). The inability of these four teachers to use learning progressions as an aid in eliciting students' thinking highlights the difficulty in supporting hermeneutic listening.

Talenquer et al. (2013) reported similar findings in an analysis of what prospective teachers noticed when evaluating student understanding during an inquiry-based unit. The exploratory study took place at a public university and included 43 participants enrolled in an undergraduate secondary science teacher preparation program. All participants had completed courses which emphasized assessment practices and focused on comprehending student understanding of scientific phenomena. Participants were given a video of another teacher's lesson and a set of associated artifacts ranging from that teacher's plans to samples of student work, and were asked to "select two forms of evidence to analyze and evaluate the extent to which student understanding is demonstrated in each form" (Talenquer et al., 2013, p. 194). The authors (Talenquer et al.) analyzed participants' responses using a constant comparison method and identified emerging patterns. The observed patterns were used to construct a coding system that allowed the authors to organize and compare results. Overall Talenquer et al. found that the participants' assessments of student understanding focused on the presence or absence of expected concepts, and in most cases failed to evaluate the ideas expressed by the students.

Effective Formative Assessment Strategies

Although the challenge of making every teacher aware of student understanding may seem formidable, there are many promising tactics that may make it more approachable. One avenue of research that has yet to receive significant attention is the relative effectiveness of different formative assessment prompts in exposing the substance of student thought. In "Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts," Furtak and Ruiz-Primo (2005, 2008) examined the effectiveness of four different formative assessment prompts at eliciting middle school students' conception of density. The authors defined assessment prompts as something that "specifies what the student will say, do, or make to provide the necessary evidence about the knowledge to be tapped. An assessment prompt

is linked necessarily with the format in which the student response will be captured and with a strategy to judge the quality of the students' response" (Furtak & Ruiz-Primo, 2008, p. 801). To answer their research question, the authors analyzed group discussions and written responses to four different assessment prompts. The four prompts were: 1) graph, 2) predict an outcome based on an experimental setting [Predict, Observe, Explain (POE)], 3) respond to an open-ended question that was central to the unit (CR), and 4) predict an outcome based on experimental conditions that were a step beyond the current content (PO). The authors found that when using prompts that feature fewer constraints and more novel experimental settings, CR and PO are more likely to elicit a range of student thinking, whereas more conventional prompts, POE and graph, demonstrated student knowledge. In addition, the diversity of student responses was greater when the prompts were presented as a written activity rather than as part of classroom discussion (Furtak & Ruiz-Primo, 2008). This study demonstrated that the design of formative assessment prompts can directly affect the nature of student responses, and that some prompts are better than others at exposing the substance of student thought.

The same authors present similar findings in "Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry" (Ruiz-Primo & Furtak, 2007). In this study the authors presented an in-depth analysis of three teachers' use of informal formative assessment. The data included videotapes of classrooms during all ten sessions implementing a physical science unit of the *Foundational Approaches to Science Teaching* (FAST) curriculum. Each video was transcribed and coded. The authors' coding was nested within the context of ESRU cycles. In an ESRU cycle, a "teacher Elicits a question; the Student responds; the teacher Recognizes the student's response; and then Uses the information collected to support student learning" (Ruiz-Primo & Furtak, 2007, p. 57). Student achievement during the unit was quantified by comparing the results of a 38-item multiple-choice pre-test with student performance during the embedded assessment. Across all three teachers, the most common eliciting question focused on the application of a known hypothesis, procedure, or observation. Very few assessment conversations involved "formulation of explanations, evaluation of quality of the evidence, comparing or contrasting other's ideas" (Ruiz-Primo & Furtak, 2007, p. 69). One of the teachers, who had the most complete ESRU cycles, frequently asked students to further clarify their understanding and was the only teacher to include questions asking students to compare and contrast different responses and evaluate the quality of evidence. That teacher's students significantly outperformed those of the other two teachers during the embedded assessments. This finding supports the principle that formative assessment is most effective when assessment prompts go beyond eliciting content knowledge and encourage student reasoning (Ruiz-Primo & Furtak, 2007).

Discussion

Black and Wiliam's 1998 publications pushed formative assessment to the forefront of education reform. Since then, formative assessment has become a generally accepted means to enhance student learning (Black & Wiliam, 2009). Despite the wide acceptance of the value of formative assessment some researchers have begun to criticize the seminal studies. Both Coffey et al. (2011) and Bennett (2011) have argued that prominent assessment studies have ignored the substance of student thinking and instead illustrate instances in which teachers gauge learning by the students' abilities to repeat pre-determined scientific facts. Davis (1997) categorized this type of listening as evaluative. During interpretive listening the teacher listens carefully to understand the students' meaning behind the words they are using; and during hermeneutic listening the teacher enters a the discussion as a co-learner, open to changing his own

ideas in response to the students' comments (Davis 1997). According to these studies, effective formative assessment requires at least interpretive and, if possible, hermeneutic listening.

Kazemi and Franke (2004) demonstrated that teachers could learn to engage in the substance of student thinking by examining student work. However, developing the skill can take considerable time and not all efforts are successful. In a subsequent study Furtak (2012) reported that a full year of professional development with a group of six biology teachers on use of learning progressions to elicit and analyze student reasoning met with limited success. Talenquer et al. (2013) reported similar discouraging results after a course in formative assessment strategies for 43 undergraduate students in a secondary science teacher preparation program.

Despite these discouraging findings, the use of certain prompts to stimulate class discussions and written work has been found to elicit higher level student thinking. Furtak and Ruiz-Primo (2005, 2008) examined a number of different strategies for using prompts to elicit student thinking, including variations of the familiar Predict, Observe, Explain (POE), and found that some strategies were more effective than others. In a different study (Ruiz-Primo & Furtak, 2007) the authors reported that student achievement was linked to the use of strategies such as asking students to compare and contrast different ideas and to evaluate the quality of evidence.

The remaining papers in this series build on the body of literature on formative assessment by describing a number of different formative assessment methods that were developed and pilot tested as part of an overall effort to develop a formative assessment system for *Energizing Physics* (Osowiecki & Southwick, in press), a high school course developed at Boston Latin School. Although it was not possible to tease out the relative effectiveness of these methods during the pilot study of this course, a summative evaluation (described in chapter 5 of this monograph) demonstrated the overall effectiveness of the course to significantly increase students' understanding of physics, including those who were taking it as high school freshmen.

Acknowledgements

This work was partially supported by a National Science Foundation grant to Portland State University (NSF grant # 1020385 Bridging the Gap Between High School and College Physics: An Exploratory Study). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Atkin, J. M., & Coffey, J. (2003). *Everyday assessment in the science classroom.* Arlington, VA: National Academy Press.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy and Practice*, *18*(1), 5-25.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice.* London: Open University Press.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-71.

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *The Phi Delta Kappan*, *80*(2), 139-148.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 1-31.

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching, 48*(10), 1109-1136.

Davis, B. (1997). Listening for differences: An evolving conception of Mathematics Teaching. *Journal for Research in Mathematics Education, 28*(3), 355-376.

Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, *49*(9),1181-1210.

Furtak, E. M., & Ruiz-Primo, M.A. (2005). Questioning cycle: Making students' thinking explicit during scientific inquiry. *Science Scope*, *28*, 22-25.

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education*, *92*(5), 799-824.

Kazemi, E., & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education*, *7*(3), 203-235.

Osowiecki, A., & Southwick, J. (in press). *Energizing Physics*. New York, NY: Freeman. (The course is also described at energizingphysics.com)

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, *44*(1), 57-84.

Talanquer, V., Tomanek, T., & Novodvorsky, I. (2013). Assessing students' understanding of inquiry: What do prospective science teachers notice?. *Journal of Research in Science Teaching*, *50*(2), 198-208.

Chapter 3

Shifting the focus from summative to formative assessment

Aaron Osowiecki, Boston Latin School, Boston, MA

Editors' Note: This chapter describes formative assessment instruments built into the structure of *Energizing Physics* to help teachers determine whether or not their students have developed the intended knowledge and skills. The focus is on one of the instruments to illustrate how it was improved part way through the pilot to improve its usefulness and reduce the time it takes to administer. As you read through the chapter consider how a teacher's capabilities are likely to interact with use of the instrument. What sort of preparation would a teacher need to use the instrument effectively?

Shifting the focus from summative to formative assessment

To most people assessment means exams, standardized tests, and end of unit projects. Although these summative assessments provide important information on what students learned, they rarely help students deepen their understanding or provide useful feedback for teachers in time to help students who may be struggling. Unlike summative assessment, teachers conduct formative assessment during the lesson or unit, allowing students and teachers to make adjustments so as to maximize student achievement during the lesson.

The most important distinction between formative and summative assessments is how they are used. Summative assessment is often thought of as assessment *of* instruction—to determine student grades, or to evaluate the curriculum or the teacher, while formative assessment is generally assessment *for* instruction—so that both students and teachers can find out how well students are learning the lesson and make adjustments if needed.

Most curricula leave it up to the teacher to figure out how to incorporate formative assessment into classroom practice. Convinced of its importance, my fellow physics teacher, Jesse Southwick, and I developed *Energizing Physics* (Osowiecki & Southwick, in press), an introductory physics course with formative assessment tools embedded into each lesson. Although we still administer end-of-unit tests, mid-terms, and final exams, and use these summative assessments to award grades, we have shifted the center of gravity of our course from summative to formative assessments.

Support from the National Science Foundation[4] allowed us to evaluate and refine our formative assessment tools during the 2011-2012 school year with 13 teachers and almost 1,000 students. In this article I'll describe how we've embedded formative assessment into daily lessons, touch on a number of the tools

[4] Bridging the Gap Between High School and College Physics: An Exploratory Study, Grant #1020385).

that we've embedded in the course, and focus on how we have modified one of the tools, which we call *DYGIT (Did You Get IT?)*

The Formative Assessment Tools

Rather than center the course on the teacher, we built the course lessons in *Energizing Physics* using the student-centered 5E learning model developed by Roger Bybee (2002) and colleagues (Bybee et al., 1989) at BSCS. Our commitment to the 5E model means there is no separation between "lecture" and "lab." Any given class may involve starter questions, laboratory activities, discussion within small groups, facilitated all-class discussions, and/or direct instruction by the teacher.

Since the 5E model provides continuous opportunities to observe student thinking, it forms the foundation of our embedded formative assessment system, summarized in Table 1. That is, teachers pay attention to the substance of students' thinking throughout each of the 5E phases. Students also participate in formative assessment by monitoring their own learning during interactions with their peers and the instructor as they progress through each stage.

Table 1. Formative Assessment Tools in *Energizing Physics* (EP) Linked to the 5E Model

The 5E Model[4A]	Formative Assessment Tools in EP
Engage lessons provide the opportunity for science teachers to identify students' current concepts and misconceptions. Although provided by a teacher or structured by curriculum materials, these activities introduce major ideas of science in problem situations. The theme here might be—*how do I explain this situation?*	***Energizing Questions*** are designed to engage students' interests and probe their prior knowledge. Energizing questions set the stage for introducing a new idea.
Explore lessons provide a common set of experiences for students and opportunities for them to "test" their ideas with their own experiences and those of peers and the science teacher. The theme for this phase is—*how do my exploration and explanation of experiences compare with others?* Students have the opportunity to compare ideas that identify inadequacies of current concepts. Here, the theme is—*how does one challenge misconceptions?*	***Activity Questions*** guide student investigations of phenomena relevant to the concept that is being introduced. These questions, in the context of students' activities, help to reveal their current understanding of physics, including any misconceptions they may have. It is an especially good opportunity for students to compare their own understanding with the ideas of fellow students.
Explain lessons provide opportunities for students to use their previous experiences to recognize misconceptions and begin making conceptual sense of the activities through the construction of new ideas and understandings. This stage also allows for the introduction of formal language, scientific terms, and content information that makes students' previous experiences easier to describe and explain. The theme is—*this is a scientific explanation*.	Students are challenged to demonstrate their understanding by solving physics problems using a template called ***GUIDE***. The GUIDE format provides a structure to help students approach the problem systematically. It also helps the teacher determine where the students may be having difficulty, for example by applying the wrong concept, setting up the problem, or making computational errors.

continued on next page

[4A] From Bybee, Rodger (2002). Learning Science and the Science of Learning. Arlington, VA: NSTA Press. p. 32. The last two Es (evaluate and elaborate) have changed places in the EP course, since we want students to first evaluate their understanding of a core idea before extending the application of the idea to other situations. Students' evaluation of their own work also occurs at other points in a lesson. Table 1 displays the order of the assessment tools associated with the 5Es in the course.

Table 1, continued

Evaluate lessons can serve as a summative assessment of what students know and can do at this point. Students confront a new activity that requires the understandings and abilities developed in previous activities. The final theme is—*how do students understand and apply scientific concepts and abilities?*	***DYGIT (Did You Get IT?)*** is a short quiz that provides an opportunity for students to assess their own understanding, and for the teacher to determine which students are struggling and how to help them. Though formative assessment happens informally throughout the lesson *DYGIT* places it explicitly in the lesson model.
Elaborate lessons apply or extend the student's developing concepts in new activities and relate their experiences to the current activities. Now the theme is—how does the new explanation work in a different situation?	This phase, which we call ***Extend***, introduces an additional formative assessment activity called ***What's the Big Idea?*** that challenges students to summarize the main point of the lesson. It also introduces ***Practice Problems*** in which students apply their newly acquired concept understandings to more sophisticated problems and situations.

For more *formal* formative assessment data, we developed quick (5-10 minute) quizzes for the end of each lesson, which we call ***DYGIT (Did You Get IT?).*** To ensure timely feedback, we focused the quizzes on the specific learning targets of each lesson while making them quick to grade. (We give examples of learning targets in Chapter 1 of this monograph).

Formative Assessment in Action

Support from the National Science Foundation allowed Jesse and me to observe the pilot teachers at schools in three neighboring cities. These observations, and our own experiences, indicated that formative assessment takes time. For example, the pilot teachers reported that the pace of the course slowed during the "extend" phase of the lesson when the teacher introduces new ideas to help students deepen their understanding of the phenomena. New to the concepts, students often made mistakes and required time to determine solutions to a problem. Often, the teacher needed to bring the class together to review possible solutions and guide students. While this process resulted in a slower pace, we found it to be an essential piece for students to internalize concepts or skills before moving on to the next lesson.

Engaging students and allowing them to explore, evaluate and extend their skills and concepts while making instructional adjustments gives students the opportunity to develop a solid understanding and ability to *use* the concepts and skills of physics. In fact, at Boston Latin School our students performed much better on the quizzes and exams after we implemented these assessment tools than in previous years, providing evidence in favor of the depth side of the depth vs. breadth debate. (Jesse will take up that issue in Chapter 6). In other words, it takes time to develop deep understanding.

While observing other teachers, Jesse and I were surprised by the difficulty of identifying specific moments of formative assessment. We attribute some of this difficulty to the global use of formative assessment in the 5E learning cycle. Teachers constantly gathered information, monitored the pace, and made minor adjustments throughout the lessons. Our observations agree with Black and Wiliam's comment that "formative assessment and instruction are indivisible" (1998, p. 7).

We also attribute the difficulty of identifying formative moments to the way that formative assessment tools are woven into the *Energizing Physics* curriculum. We designed the questions to build in complexity and to challenge common misconceptions. Pilot teachers receive a comprehensive *Teacher's Guide*, which describes the lessons, highlights the challenges and recommends methods to help students recognize accurate and inaccurate reasoning. This design and the periodic checks between students and teachers build in adjustments in timing as part of the lesson flow.

Fine-tuning the DYGIT Tool

As the explicit evaluation stage of the 5E model, DYGIT questions provide an opportunity for students to test their grasp of the lesson's learning target. One pilot teacher commented that the DYGIT questions highlighted the difficulty that some of the students encountered when connecting new concepts to the mathematical models that they needed to solve physics problems. In that case, the problem was due to lack of conceptual understanding, rather than limited mathematical abilities.

In their regular feedback interviews, several of the pilot teachers highlighted the effectiveness of the DYGIT quizzes used at the end of many activities. Students also appreciated these low stakes opportunities for additional targeted practice in a test-like environment. Although valuable as a formative assessment tool, our observations identified two problems with DYGIT as we initially implemented the course design. First, the questions were used to spark large group discussion and problem solving activity. While the questions we selected enabled us to assess the substance of students' understanding, they did not give us information on *all* of the students. The less confident students tended to wait as their more confident peers responded to the questions. Without full participation, teachers could not accurately gauge all students' understanding. Second, the time needed to solve the question and the required class discussion added a significant number of minutes to each lesson, slowing the pace of the entire course.

So, based on the results of the 2011-2012 pilot, we converted the DYGITs from open-ended problems to multiple choice questions modeled after the well-known series *Uncovering Students' Ideas in Science* (Keeley et al., 2005). Each probe consists of two parts: (1) a multiple choice selection with research-based distractors, and (2) student justification of the answer choice. The following example from our energy conservation unit illustrates this transition:

Learning Target: (I can) Use energy conservation and energy models to quantify variables in situations involving TWO types of energy.
Old DYGIT You push a 50 g marble so that it is moving at 4 m/s at the top of a 1-meter high ramp. What will be the marble's speed when it reaches the bottom of the ramp?
New DYGIT Released at the top of a 1-meter high ramp, a 50 g marble will reach 4.5 m/s at the bottom. Which of the following choices equals the marble's speed if it was pushed at 4.0 m/s at the top? . a. 4.0 m/s b. 4.5 m/s c. 6.0 m/s d. 8.5 m/s Justify your choice.

Figure 1. Old and new versions of DYGIT formative assessment tool.

While the calculation required in the old DYGIT intimidated some students, the new DYGIT can be approached without calculation and student choices highlight their conceptual understanding of energy and its conservation and provide feedback. For example, in the sample DYGIT above:

a. This incorrect choice indicates that the student believes the marble will roll down the ramp at constant speed.

b. This incorrect choice indicates that the marble's speed at the bottom does not depend on the marble's initial speed (in the lesson students learn that without an initial speed the final speed depends on the strength of gravity and the ramp's height).
c. This choice is correct, indicating that the student understood the lesson.
d. This incorrect choice indicates that students believe they can just add the top speed to the original bottom speed, not factoring in that kinetic energy involves the square of speed.

Teachers can use the new DYGIT to quickly make a visual assessment of the class using one of the following techniques:

- Student Voting - students raise hands to vote for a certain choice.
- Four Corners - have students move to the corner corresponding with their choice.
- Stick Bars - Students record their choice on sticky notes, which are then used to make a bar graph (also known as an "affinity graph").

Students then justify their choices as the teacher guides the class towards the preferred answer with the reasoning behind it. With this data in hand, teachers will be able determine the most appropriate way to proceed with the lesson.

The Extend phase, which aimed at helping students deepen their understanding and sometimes involved presentation of a new concept, often included two assessment tasks: 1) *What's the Big Idea?,* in which the students summarized what they perceived as the major point of the lesson, and 2) *Practice Problems*, in which the students applied what they learned in a new setting. Although these assessment tools were very effective in finding out if the students really did get the big idea of the lesson, some of the Practice Problems were unnecessarily complicated, consisting of many parts. We revised some of the problems by limiting the number of parts and complications. These changes allowed students to extend their understanding and provided the teacher with feedback on students' progress without extending the length of the course.

Discussion

As Black and Wiliam pointed out in their seminal paper, *Inside the black box* (1998), external or "high stakes" testing has come to dominate K-12 assessment, although it has little impact on instruction. In contrast, formative assessment, which has been shown to have a substantial effect on student achievement, has had very little support. Our observations during the pilot test have convinced us that Black and Wiliam are correct in their judgment. By building formative assessment into the *Energizing Physics* curriculum, we have provided tools that future teachers of the course will be able to use to assess the progress of their own students, and to adjust instruction accordingly.

Although we have done our best to embed formative assessment in everyday instruction through the development and modification of instructional materials, we close with a comment on the importance of professional development. As summarized in Chapter 2, a number of researchers (e.g., Coffey et. al., 2011; Bennett, 2011; Davis, 1997; Kazemi & Franke, 2004, Furtak, 2012; and Talenquer, 2013) and the advisors on this project have pointed out that effective formative assessment depends on a teacher's desire and ability to listen to the substance of student thought—not simply to determine if the students have the "right idea," but rather to hear the meaning behind their words. Because, without attending closely to the substance of students' thinking *during class*, it is not possible to help them build on the ideas that they bring to the problem at hand, and unravel any misconceptions they may have. Although, on the whole, our pilot teachers were up to the task, thanks in part to a one-week professional development workshop and to their status as

experienced physics teachers who are willing to try something new, we realize that professional development will be the lynchpin to future success.

Acknowledgements

We are indebted to the teachers and students who participated in this study. This work was partially supported by a National Science Foundation grant to Portland State University (NSF grant # 1020385 Bridging the Gap Between High School and College Physics: An Exploratory Study). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5-25.

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappa,* 80(2), 1-17.

Bybee, R. (2002). *Learning science and the science of learning.* Arlington, VA: NSTA Press.

Bybee, R., Buchwalk, S., Crissman, D., Heil, D., Kuerbis, P., Matsumoto, P., & McInerney, J. (1989). *Science and technology education for the elementary years: Frameworks for curriculum and instruction.* Washington, D.C.: The National Center for Improving Instruction.

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching, 48*(10), 1109-1136.

Davis, B. (1997). Listening for differences: An evolving conception of Mathematics Teaching. *Journal for Research in Mathematics Education, 28*(3), 355-376.

Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, *49*(9), 1181-1210.

Kazemi, E., & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education*, *7*(3), 203-235.

Keeley, P., Erberle, F., & Farrin, L. (2005). *Uncovering students' ideas in science: 25 formative assessment probes.* Arlington, VA: NSTA Press.

Osowiecki, A., & Southwick, J. (in press). *Energizing Physics.* New York, NY: Freeman. (The course is also described at energizingphysics.com)

Talanquer, V., Tomanek, T., & Novodvorsky, I. (2013). Assessing students' understanding of inquiry: What do prospective science teachers notice?. *Journal of Research in Science Teaching*, *50*(2), 198-208.

Chapter 4

GUIDE-ing students to a better understanding of physics

Stephen Scannell, Gresham High School, Gresham, OR

Editors' Note: Bennett (2011) pointed out the importance of sorting out the specific errors that students make to determine how best to help them. This chapter describes a template developed by the *Energizing Physics* team to help teachers pinpoint students' difficulties. The chapter also describes comments from students and teachers about what they liked or didn't like about using the template. Based on this data, are there ways that teachers could maximize their use of the template while avoiding some of the drawbacks?

GUIDE-ing students to a better understanding of physics

Problem solving is an integral part of many introductory physics classes, yet many students struggle with solving physics problems. Students may say, "Where do I start?"; "What equation do I use?"; "How do I do the math?"; "Is this right?"; or "I understand the concept, but I just can't do the problems. What am I doing wrong?" If their questions are specific, I can identify where my students are having trouble and provide the help they need. However, in many cases my students don't know why they are having difficulties, and cannot frame a useful question, which makes it difficult for me to identify what they do know, where their knowledge is lacking, and how to help them overcome their specific challenges. To complicate matters, research shows that even students who can solve traditional quantitative problems correctly may not have a functional understanding of the physics concepts (Kim & Pak, 2001).

Energizing Physics (EP), written by Aaron Osowiecki and Jesse Southwick (in press), two high school teachers at Boston Latin School, is the culmination of ten years of work developing a physics course with "energy" as the overarching theme. The course uses best practices identified in physics education research and is designed to develop students' conceptual and quantitative understandings of physics. To address student difficulties with problem solving, the authors developed a template to lead students through each major component of the problem solving process (Heller, Keith, & Anderson, 1991; Gok, 2010; Reif, Larkin, & Brackett, 1976). The result of their work is the **GUIDE** template, in which each of the five steps forms a mnemonic that is easy for students to remember.

Given: Assign variables to represent the given information. Sketch the situation.

Unknown: Assign a variable to the unknown information.

Identify Tools: What concepts/equations will you use to solve the problem?

Do the Math: Apply the tools and solve the problem.

Evaluate: Review your work. Does your result make sense?

The GUIDE template was embedded into many aspects of the course, being used for most problem solving activities, including homework, lab activities, formative and summative assessments, and, in course resources, answer keys and Pencasts[5] of selected problems. The template is shown with a sample problem in Figure 1.

Physics Problem Solving with GUIDE	
Given	Assign variables to represent the given information. Sketch the situation.
Unknown	Assign a variable to the unknown information.
Identify the Tools	What concepts and equations will you use to solve the problem?
Do the Math	Apply the math tools and solve the equation.
Evaluate Your Answer	Review your work. Does the result make sense?

Makayla measured the ramp to be 45 arms long. She timed a ball rolling down the ramp to be 15 swings of a pendulum. What was the average speed of the ball?

G	U	I	D	E
Δx = 45 arms Δt = 15 swings	v	$v = \frac{\Delta x}{\Delta t}$	$v = \frac{45 arms}{15 swings}$ $v = 3.0 \frac{arms}{swing}$	Units came out with distance on top and time on the bottom just like speed limits.

Figure 1. GUIDE Problem Solving Handout/Poster.

This article describes what I learned from using GUIDE as I piloted *Energizing Physics* with two classes of physics students in 2011-2012, and from the results of a survey taken by six other pilot teachers and 141 of their students.

The Need for a Problem Solving Strategy

Researchers have found that students benefit by having an explicit strategy when approaching problem solving (Maloney, 1994; Larkin & Reif, 1979; Reif, 1981). Explicit problem solving strategies address both the qualitative and quantitative aspects of a problem, while a traditional strategy, sometimes called a "textbook strategy," tends to focus only on the quantitative aspect of a problem (Huffman, 1997). Without an explicit structure, most students do not approach problem solving in a strategic manner, but instead may focus on a "plug and chug" approach—by guessing which equation to use and plugging in numbers—or they may use no systematic approach at all (Walsh, Howard, & Bowe, 2007). In contrast, expert problem solvers will look at a problem qualitatively, focusing on the physics concepts before developing equations. Expert problem solvers also evaluate their work, making revisions as necessary (Reif, 1981). GUIDE supports students in using more expert-like problem solving strategies and provides two opportunities to move away from textbook problem solving by adding two steps: **I** (identify tools) and **E** (evaluate). In the

5 Information on Pencasts can be retrieved from http://www.livescribe.com/en-us/pencasts/

I step, students identify concepts first and then identify relevant equations. For the **E** step, students are asked to evaluate their work. Evaluation helps the students reinforce their understanding and provides a place for them to ask questions. As their instructor, I find that the **I** and **E** steps provide insight into the substance of my students' thinking with respect to both the physics concepts and ways to go about setting up and solving problems.

What Distinguishes GUIDE from Other Strategies

GUIDE is not an entirely new idea. There are other problem solving strategies, such as WISE (What's happening, Isolate the unknown, Substitute, Evaluate) (Wright & Williams, 1986), and GUESS (Given, Unknown, Equation, Substitute Solution) (Beall, 2012), among others (Gok, 2010). What I like about GUIDE is that the word is easy to remember and it "guides" students towards solving the problem. It also provides for both qualitative and quantitative analysis of student thinking. The goal of GUIDE is to engage students in thinking deeply about the problem, including evaluating their work, not just leading them to a quantitative answer.

GUIDE: A Common Language

One of the unexpected benefits of GUIDE is that it provides a common language for teachers and students to use in discussing problems and the problem solving process. It is especially helpful with the student who says, "I don't get it." The series of questions implicit in GUIDE helps the student and teacher find out what part the student doesn't get. Does the student have questions about the givens (**G)** or the unknowns (**U**)? If the student is aware of the givens and unknowns does he or she have questions identifying (**I**) the physics concepts and relationships involved? Is the student challenged by doing (**D**) the math? If the student has accomplished all of the above, did he run into a problem when he tried to evaluate (**E**) his work by checking to see if the solution has the correct units and a reasonable numerical value, and seems to make sense in relation to the original problem?

An example of an in-class assessment is shown in Figure 2. It was given after students had learned how to calculate kinetic energy (KE) and gravitational potential energy (GPE), and to determine the elastic potential energy (EPE) and total energy (TE) using the principle of conservation of energy. It asks students to determine the type and amount of each type of energy at three points along the path of a bungee jumper, both quantitatively and graphically. In the *Energizing Physics* course, this problem assesses students' abilities to accomplish the following learning target for the lesson:

> **Learning Target C2.9:** (I can) Use energy conservation to quantify kinetic energy, gravitational potential energy, and elastic potential energy at various points in a bungee drop.

The GUIDE format provides a structure for students to organize their work for this multi-step problem while also providing the teacher with several points of reference to determine student understanding. The student whose work is shown in Figure 2 does all of the calculations correctly, but in the **E** section, the student demonstrates a possible misunderstanding of the problem. The student begins with a condition for the answer, "If energy is conserved, there are 14760 J with gravity= 9.8m/s^2," indicating that the student understands that total amount of energy in the system is constant. However, the following statement, "As Jordan jumps his GPE decreases and his KE increases, when he gains EPE, his KE decreases to nothing," indicates that the student either omitted the initial condition in the problem, that Jordan initially jumps upward at 10 m/s, or had a misunderstanding of the problem. This statement provides an opportunity for

the teacher to follow-up with the student to clarify whether the issue was one of communication (omitting important information) or a misunderstanding of the physics. Assessment scores are based on a scoring rubric that focuses on students' understanding of physics concepts, and their ability to use their knowledge (Figure 3).

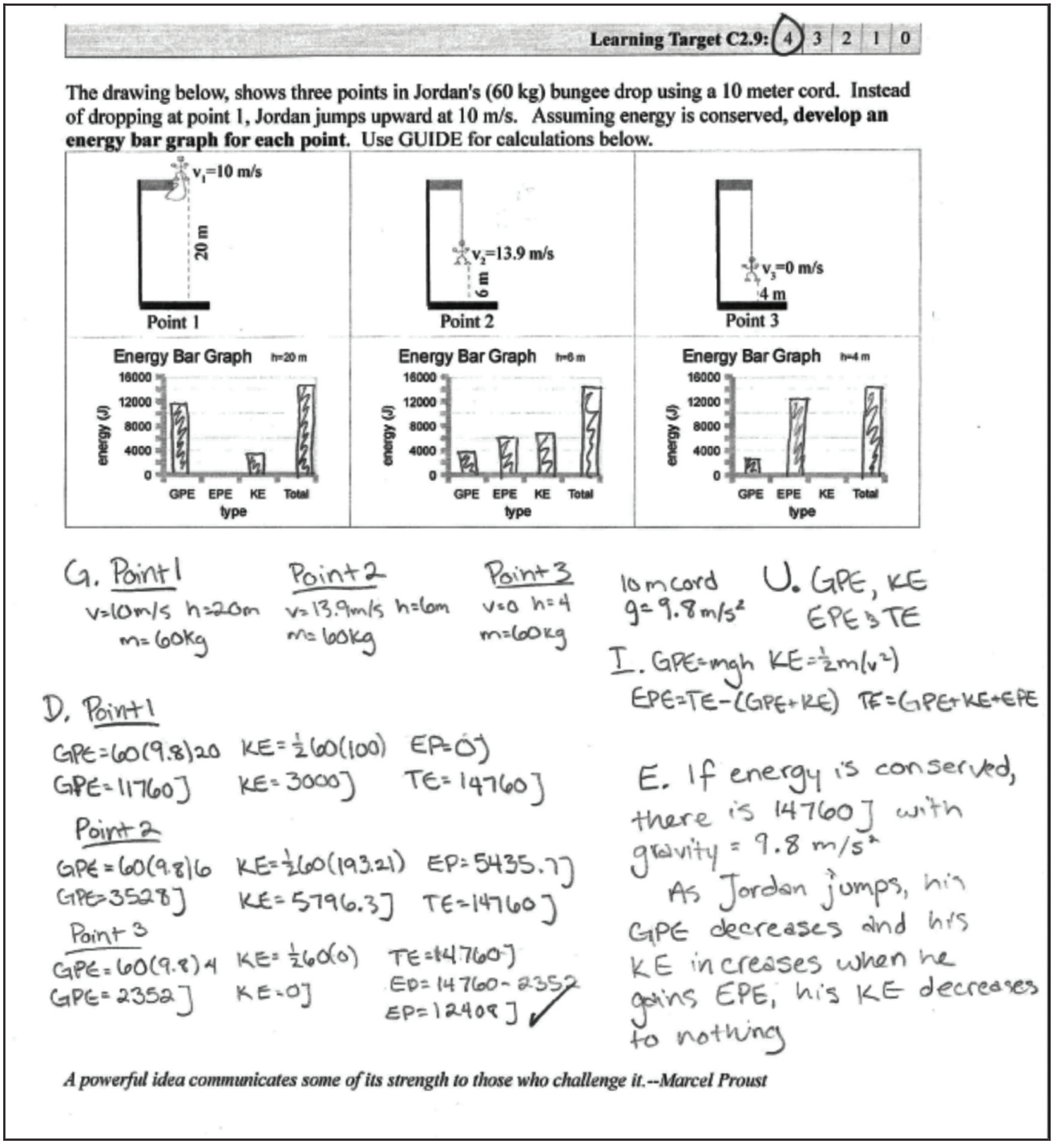


Figure 2. Example of Survey using GUIDE.

Scoring Guide for DYGIT Quizzes		
Rating	**Performance**	**Description**
4	Complete understanding	Problem solved with all parts of GUIDE complete, units and significant figures correct
3	Solid understanding	Problem solved but with minor omissions or errors
2	Some understanding but major gaps	Problem not solved correctly, one or more major concepts included
1	Limited understanding	Some relevant information given
0	No information	No answer or not relevant

Figure 3. Four Point Scoring Guide for DYGIT Quizzes.

Teacher Perspective of GUIDE

The GUIDE format was well received by pilot teachers. Survey results from six pilot teachers showed that all teachers found GUIDE to be helpful for students in solving physics problems, while also improving students' problem solving skills. Additional teacher comments from the survey included:

"Clear path to solve problems."
"Gives students who are sloppy a template to follow and makes grading easier."
"I liked that I could easily follow their work —see what tools they were using and then how they used them."
"Guide gives a standard approach that helps get students started, and organizes their thinking. It has also given EP a signature format for all of its solutions. For students and authors and teachers it gives a common reference point."

Student Perspective of GUIDE

Students found GUIDE to be helpful as well. In a survey completed by 141 students (Figure 4), a majority of students were able to identify aspects of GUIDE that they found helpful, with only 15% of students saying they didn't gain any benefit from using GUIDE and only 4% of students saying the GUIDE did not help them. 73% of students found GUIDE to be "helpful" or "very helpful." One of the things we were hoping for was that students might find the strategy helpful in another class (typically science or math). This turned out to be the case for 35% of students (score of 3 or greater on a 5-point Likert Scale from 1 = did not help to 5 = was a great help), while 43% indicated it was not helpful in other classes. In addition to these survey results, selected comments from the survey indicated that using the GUIDE template did have a positive impact for students. Below is a selection of student comments taken from the student survey:

- "I have to admit that GUIDE has grown on me. It helps the problem solver to organize all the given information into one area, and it reminds the problem solver of what equation to use to solve."
- "Showed me what part of the problem I didn't understand."
- "If I make a mistake its easy to find."
- "It provides a structured method in case I'm not sure how to begin solving a problem."

- "I like how it organizes my work which makes it easier for me to get the correct answer and easier to go back and analyze my work later for studying etc. I also like how it makes finding what the problem is asking for easier."
- "The other thing I liked about GUIDE is that you could use it in any situation, not necessarily just math."
- "I think it was very helpful in my learning of physics."
- "GUIDE made my notes clearer so I could study better, organized my thoughts before I solved the problem, and made me less confused when I had to solve the problem because I had everything written out."
- "I really like GUIDE because it helps me completely understand how to answer problems. It may take longer but I like that I can see step by step, my train of thought."

When responding to, "List two things you don't like about GUIDE" the following quotes were typical:

- "Sometimes it could get tedious, and I didn't understand what we were supposed to do for the E step."
- "Sometimes it may take up a lot of space on my paper."
- "Spending too much time organizing the problem."
- "The E part is silly and unnecessary."
- "GUIDE took much longer than just solving the problem."

What benefits did you gain by using GUIDE? Check all that apply. (N=141)

	N	%
GUIDE provided a method for solving a problem.	89	63%
In using GUIDE I could figure out what information was important.	74	52%
In using GUIDE I knew what calculations to perform.	64	45%
In using GUIDE I could figure out what equation I was supposed to use.	63	45%
In using GUIDE I could figure out what I was supposed to find out.	62	44%
In using GUIDE I was able to check my work and verify my answer.	31	22%
I didn't gain any benefit from using GUIDE.	21	15%

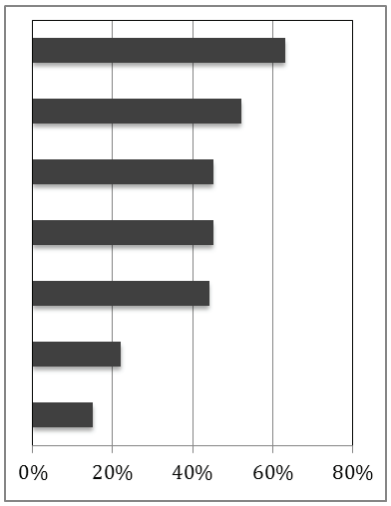
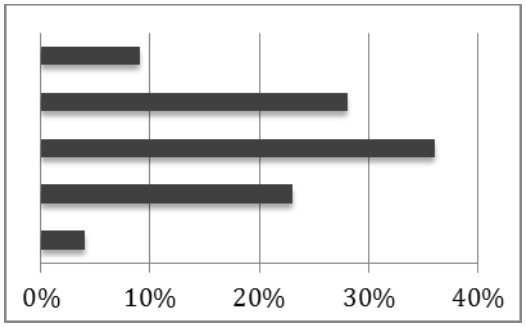


Figure 4. *continued on next page*

On a scale of 1 to 5, how would you rate GUIDE with regard to helping you solve physics problems? (N=141)

	N	%
1 Did not help	5	4%
2	33	23%
3	51	36%
4	40	28%
5 Was a great help	12	9%



On a scale of 1 to 5, how would you rate GUIDE with regard to helping you solve problems in other classes? (N=141)

	N	%
1 Did not help	60	43%
2	31	22%
3	30	21%
4	13	9%
5 Was a great help	7	5%

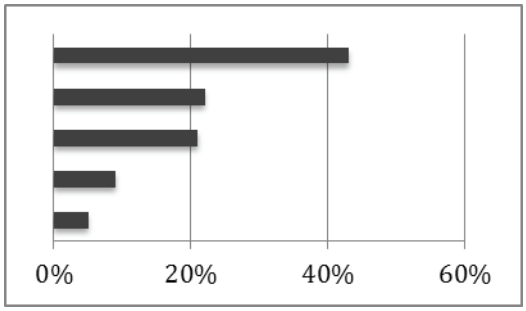


Figure 4. Student responses from GUIDE survey.

Discussion

My experience in using GUIDE with my own physics students during the 2011-2012 school year was very positive. The GUIDE format provided us with a common language for discussing how to approach and solve problems. It helped my students become more expert problem solvers and it helped me understand the nature of my students' difficulties and, therefore, how I could provide the assistance they needed.

The GUIDE strategy was also seen as an effective instructional strategy by the other pilot teachers and a majority of students surveyed. Student complaints about the GUIDE strategy were typically that using GUIDE took too much time, required too much writing, and contained repetitive steps. Some of the negative student comments indicated that students often fail to understand the importance of evaluating their own work to see if it makes sense. Despite these negative comments, we feel that the GUIDE problem solving strategy was helpful in facilitating instruction of physics concepts and problem solving and moving students towards being more "expert-like" in their approach to problem-solving.

While both teachers and students found value in using GUIDE, we did not measure how GUIDE impacted student performance. A study investigating the effectiveness of the WISE strategy found that

students who used the strategy most of the time did well in the course, while those who did not use the strategy, including those students identified with strong math skills before the start of the course, did poorly (Wright & Williams, 1986). Since GUIDE has a similar purpose to the WISE strategy, that finding lends some support to the value of GUIDE. However, we will still need to test GUIDE through an experimental or quasi-experimental study to determine its strengths and weaknesses. Meanwhile, we invite our fellow physics teachers to consider using this strategy. We would look forward to any feedback from instructors who use it.

Acknowledgements

We are indebted to the teachers and students who participated in this study. This work was partially supported by a National Science Foundation grant to Portland State University (NSF grant # 1020385 Bridging the Gap Between High School and College Physics: An Exploratory Study). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Beall, L. (2012). Metagognition: The effects of cognitive strategy instruction on the problem solving skills of high school algebra 2 students. Masters thesis. Bozeman, MT: University of Montana.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5-25.

Gok, T. (2010). The general assessment of problem solving processes in physics education. *Eurasian Journal Of Physics And Chemistry Education, 2*(2). Retrieved from http://www.eurasianjournals.com/index.php/ejpce/article/view/492/225

Heller, P., Keith, R., & Anderson, S. (1991). Teaching problem solving through cooperative grouping. Part 1: Group vs. individual problem solving. *American Journal of Physics*, *60*(7), 627-636.

Huffman, D. (1997). Effect of explicit problem solving instruction on high school students' problem-solving performance and conceptual understanding of physics. *Journal of Research in Science Teaching*, *34*(6), 551-570.

Kim, E., & Pak, S. J. (2001). Students do not overcome conceptual difficulties even after solving 1000 traditional problems. *American Journal of Physics, 70*(7), 759-765.

Larkin, J. H., & Reif, F. (1979). Understanding and teaching problem solving in physics. *European Journal of Science Education, 1*(2), 191-203

Maloney, D. P. (1994). Research on problem solving: Physics. In D. Gabel (Ed.), *Handbook of research on science teaching and learning*. Washington D.C: National Science Teachers Association.

Osowiecki, A., & Southwick, J. (in press). *Energizing Physics*. New York, NY: Freeman. (The course is also described at energizingphysics.com)

Reif, F. (1981). Teaching problem solving—a scientific approach. *The Physics Teacher, 19*(5), 310.

Reif, F., Larkin, J. H., & Brackett, G. C. (1976). Teaching general learning and problem solving skills. *American Journal of Physics*, *44*(3), 212-217

Walsh, L. N., Howard, R. G., & Bowe, B. (2007). Phenomenographic study of students' problem solving approaches in physics. *Physical Review Special Topics-Physics Education Research*, *3*(2) 020108, 1-12.

Wright, D. S., & Williams, C. D. (1986). A wise strategy for introductory physics. *The Physics Teacher, 24*(4), 211-216.

Chapter 5

Summative evaluation[6] of *Energizing Physics*

William Walker, High Tech High, San Diego, CA,
Stephen Scannell, Gresham High School, Gresham, OR,
Aaron Osowiecki and Jesse Southwick, Boston Latin School, Boston, MA,
Cary Sneider, Portland State University, Portland, OR

Editors' Note: As a pilot study, these findings are limited in scope. The authors conclude that it is probably not possible to isolate the formative assessment elements to determine their contributions to student learning. Yet research on alternative methods of formative assessment is needed. Is there a study design that could provide such information?

Summative evaluation of *Energizing Physics*

Abstract

This paper reports the results of a summative evaluation of *Energizing Physics*, a new full-year physics curriculum involving 12 pilot teachers and 597 students who completed pre- and post-tests. Findings were that students made significant gains from pre-test to post-test in solving problems related to Average Speed, Conservation of Energy, Work and Energy, Force and Motion, and Electrical Energy. Effect sizes ranged from moderate to high. A further analysis showed that 9th graders made significant gains, although not as much as students in grades 10, 11, and 12, indicating that *Energizing Physics* could be used as a Physics First course.

Introduction

The 2011-2012 academic year was intended as a pilot study year in order to improve instructional materials and develop a formative assessment system for *Energizing Physics* (Osowiecki & Southwick, in press), a new high school physics course. However, we also wanted to find out if the current version of the curriculum helped our students learn physics. For this purpose, the authors developed the *Energizing Physics* Concept Inventory (EPCI), a 40-item multiple-choice test to be used before and after a year of study. We used the EPCI as a pre-post-test to answer the questions:

1. To what extent do students who enroll in *Energizing Physics* improve in their conceptual understanding of physics?

[6] In this monograph the term *assessment* refers to a process for determining a student's knowledge or capabilities, while *evaluation* refers to a process for determining the value of instruction.

2. Are some physics concepts more challenging to learn than others? If so, which are more challenging?

Since some students were taking the course as 9th graders, we also wanted to find out how well they did in contrast to students who took the course in grades 10, 11, and 12 and who had received additional instruction in mathematics. So our third research question was:

3. Are 9th graders as capable of learning physics as students in grades 10-12?

As a pilot study, we expected to make improvements in *Energizing Physics* both by listening to suggestions from pilot teachers and students and by comparing results of the pre- and post-tests in order to determine which parts of the course needed strengthening. Consequently, we were not yet ready to undertake an experimental study to compare *Energizing Physics* with other physics curricula. This pilot year's goal was not to find out if *Energizing Physics* was better than other courses, but rather to determine if *Energizing Physics* was effective and if it, therefore, warranted further study to compare its effectiveness with curricula currently on the market.

Method

Design

All treatment students received *Energizing Physics*, the experimental treatment intervention, and were given the EPCI to complete before and after the course. The quasi-experimental design is briefly summarized in Figure 1.

N_{9th}	O_{EPCI}	X	O_{EPCI}
$N_{10th\ 11th,\ 12th}$	O_{EPCI}	X	O_{EPCI}

N — non-randomized groups
X — *Energizing Physics* for one academic year
O_{EPCI} — Energizing Physics Concept Inventory

Figure 1. Study design.

Subjects

597 students (of nearly 1,000 in the pilot) who completed both the EPCI as both a pre-test and post-test were included in the study. The two contrast groups included 241 students in grade nine and 356 students in grades ten through twelve. The students came from six school districts in the greater metropolitan areas of Boston, Massachusetts and Portland, Oregon. Although one of the schools had only 4% of its students on free and reduced lunch status, the other schools had free and reduced lunch percentages ranging from 38% to 50% per grade level.

The twelve experienced physics teachers whose students participated in the study had attended a week-long institute in *Energizing Physics* during the summer prior to the study. The teachers were given preliminary printed and electronic drafts of all curriculum materials, any equipment that they needed to teach the course

that they did not already have, and, in some cases, a scanner to copy student papers for analysis by project staff.

Intervention

Energizing Physics is a new introductory high school course designed to help students develop a broad spectrum of capabilities in "doing physics." Although it emphasizes conceptual understanding of physics principles, *Energizing Physics* engages students in using moderately sophisticated mathematics to solve problems so they will be prepared to enroll in an upper level physics course for students who expect to major in science or engineering in college. Integrated into the course are a number of embedded formative assessment activities aimed at providing both the teacher and students with feedback to inform instruction. The course structure, instructional model, and learning targets are described in Chapter 1 of this monograph. Formative assessment elements of the course are described in Chapters 3 and 4.

Instruments

The primary instrument used to collect data on student achievement was the *Energizing Physics* Concept Inventory (EPCI), which consisted of 40 multiple-choice items. The test focused on conceptual understanding of the major physics concepts, along with basic math skills such as reading graphs and doing simple calculations accessible to 9th graders. Physics jargon was also replaced with common language so that capable students who had taken physical science and middle school math, would be able to engage in the problems on the pre-test.

Additional data sources included observations and interviews of the pilot teachers by the course authors in the Boston area and two other staff members in the Portland, Oregon area. All four observers were also physics teachers (peer observers) who were teaching the course, and whose students were included in the study.

A valuable partner in this project was the Berkeley Evaluation and Assessment Research (BEAR) Center at the University of California, led by Director Mark Wilson and Senior Researcher Karen Draney. The Berkeley group had developed methods for assessment and an approach for helping developers create assessment systems for their science courses (Wilson & Sloan, 2000; Wilson & Scalise, 2003; Wilson & Draney, 2004; Wilson, 2005; Wilson & Scalese, 2006). Prior to beginning this project core staff attended a one-week workshop conducted by the BEAR team. One outcome of the workshop was to develop a clear description of "constructs" that represent what students are expected to learn. The team developed two overarching constructs: 1) conceptual understanding of physics—that is, a student's ability to apply a conceptual model of a physical principle to a situation in the physical world; and 2) mathematical modeling, which consists of the ability to represent and manipulate the model mathematically to make predictions and solve physics problems. A formative assessment tool called GUIDE, which is described in chapter 4 enabled us to separate students' difficulties with mathematics from difficulties in understanding physics concepts, enabling the teacher to differentiate instruction and help students with the particular problems they were encountering.

The EPCI primarily assessed the conceptual understanding of a physics construct. An important step in developing the EPCI was to use item response modeling to determine if the test consisted of a sufficiently wide range of difficulty levels to measure initial understanding as well as changes in understanding of physics as a result of the course. The result of that analysis, expressed as a Wright Map, is shown in Figure 2.

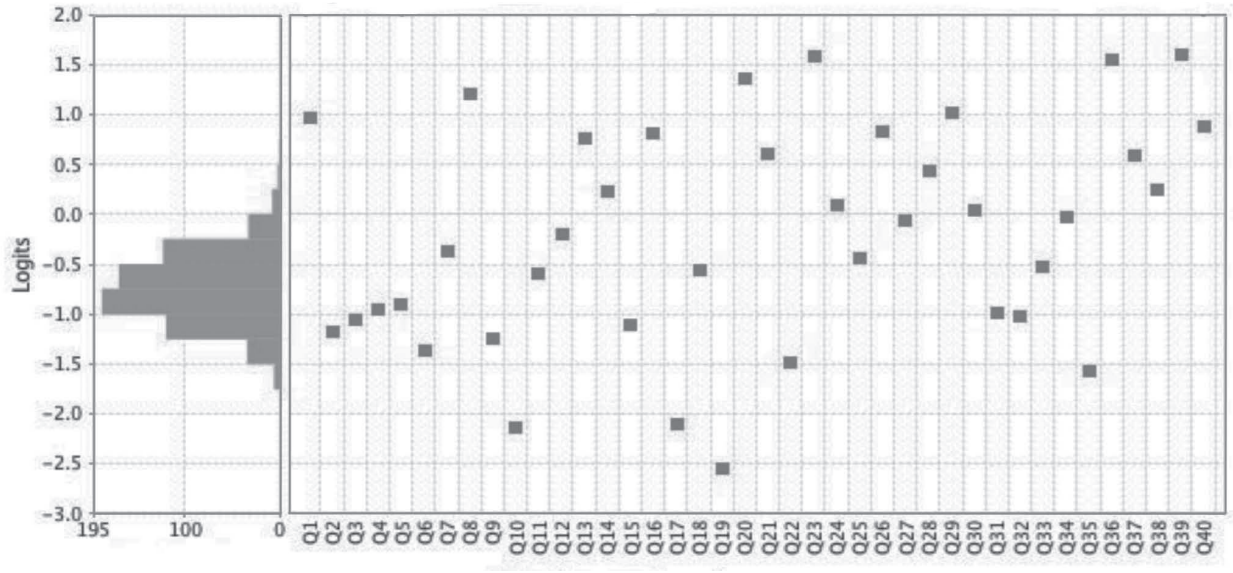


Figure 2. Wright Map of EPCI. A Wright map, based on item response theory (IRT), was built using software created by the BEAR Center with data from the EPCI pre-test. The left side shows a distribution of student ability levels on the pre-test. Each square on the right side represents the difficulty of an item such that students at that level have a 50% chance of responding correctly.

As an example of the kinds of questions included on the EPCI and on the Wright Map, consider the following three questions in which students are asked to apply their understanding of gravitational potential energy.

6. Attached to a bungee cord, Julia drops from point 1 to point 2 as shown in the drawing. She then bobs up and down before coming to a stop. When she comes to a stop, she will be at:

a. Point 1 c. Between points 1 and 2
b. Point 2 d. Lower than point 2

7. In the situation described in #6, at which point does Julia have the most energy?

a. Point 1 c. Point 1 = Point 2
b. Point 2 d. Can't tell.

8. Which choice correctly identifies the source of energy for Julia's bungee jump?

a. Julia climbed up to the dropping platform.
b. Bungee cords naturally have energy.
c. Julia stepped off of the platform.
d. Julia connected to the bungee cord.
e. All of the above.

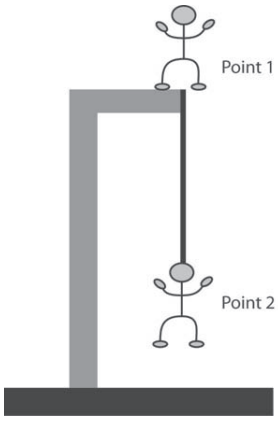


Figure 3. Questions about gravitational potential energy from the EPCI. As shown on the Wright map, above, on the pre-test nearly all of the students answered question 6 correctly, only a few answered question 7 correctly, and none of the students could answer question 8. (Correct answers are 6c, 7a, and 8a.)

Three of the EPCI questions were not used in the present analysis. Question 38, concerning thermal energy, and questions 39 and 40, concerning energy and society, were not taught by many of the pilot teachers. The other questions were grouped into five categories, representing core ideas in the first five chapters:

- Average Speed (questions 1-4)
- Conservation of Energy (questions 5-14)
- Work and Energy (questions 15-24)
- Force and Motion (questions 25-34)
- Electrical Energy (questions 35-37)

Findings

The findings are organized according to the three questions that motivated this assessment:

1) To what extent do students who enroll in *Energizing Physics* improve in their conceptual understanding of physics?
2) Are some physics concepts more challenging to learn than others? If so, which are more challenging?
3) Are 9th graders as capable of learning physics as students in grades 10-12?

1) To what extent do students who enroll in *Energizing Physics* improve in their conceptual understanding of physics?

Based on all students (N=597), the mean score on the EPCI given prior to instruction was 14.49 (39% correct, SD = 4.694) and the mean score after completing five chapters was 20.52 (55% correct, standard deviation SD = 6.359). A Student's t-test of the two means showed a significant difference ($t = 28.335$, $p < .001$). A further calculation showed a large effect size (Cohen's $d = 1.08$).

In summary, the statistical analyses show that the students made significant gains in their understanding of physics. However, as physics teachers, a success rate of 55% of answers correct is not acceptable to us, even though we acknowledge that at several of the schools the test was given at the end of the year, after grades were submitted, so the students may not have given the assessment their best effort. Nonetheless, we gathered a lot of information during the pilot year, and we have already made many improvements in the course.

2) Are some physics concepts more challenging to learn than others? If so, which are more challenging?

The results are shown in Table 1 and Figure 4. Since this analysis involved six significance tests, we used a more stringent test for significance ($p < .05/6 = p < .008$). Students gained significantly for the total test and for each of the five areas.

Calculation of Cohen's d revealed that effect sizes were moderate with respect to Average Speed (d=.49) and Electrical Energy (d=.48), and large with respect to Conservation of Energy (d=.77), Work and Energy (d=.71), and Force and Motion (d=1.03).

In summary, students gained in their understanding of physics in all five major concepts presented in the first five chapters of *Energizing Physics*. Their weakest areas concerned Average Speed and Electrical Energy. While it is possible that these questions on the EPCI were relatively challenging, we are paying special attention to how these parts of the course might be strengthened.

Table 1. Pre-Post Scores for all students (N = 597)

N=597	Pre-Test			Post-Test			t	p <	d
	Mean	SD	%	Mean	SD	%			
Average Speed (1-4)	1.66	1.095	41%	2.21	1.137	55%	12.107	.001	0.49
Conservation of Energy (5-14)	4.07	1.637	41%	5.41	1.857	54%	15.640	.001	0.77
Work and Energy (15-24)	4.22	1.869	42%	5.64	2.106	56%	15.471	.001	0.71
Force and Motion (25-34)	3.41	1.929	34%	5.74	2.540	57%	23.238	.001	1.03
Electrical Energy (35-37)	1.13	0.726	38%	1.52	0.885	51%	8.813	.001	0.48
Total	14.49	4.694	39%	20.52	6.359	55%	28.335	.001	1.08

% = Percent Correct SD = Standard Deviation d = Cohen's d, a measure of effect size

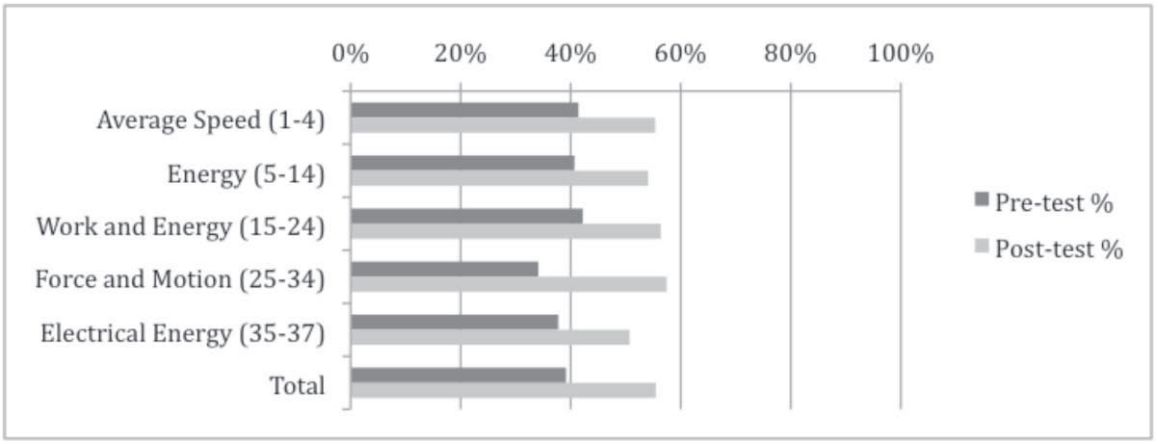


Figure 4. Pre-Post Scores for all students.

3) Are 9th graders as capable of learning physics as students in grades 10-12?

We performed similar calculations as above with 9th graders (N = 241) and with 10th, 11th, and 12th graders as a group (N = 356). The results are shown in Tables 2 and 3. Since this analysis involved twelve significance tests, we used an even more stringent test for significance ($p < .05/12 = p < .004$). As shown below, students in both groups gained significantly in all five areas. Furthermore, the effect sizes are moderate to high for both groups, although they tend to be somewhat higher for the older students.

Figure 5 compares post-test scores on the EPCI for students in grade 9 with post-test scores of students in grades 10, 11, and 12. Consistent with Table 2, the figure reveals that ninth graders learned physics in all five areas. However, they did not do as well as the 10th - 12th grade students in any category, but especially with respect to Force and Motion, where the older students excelled.

With regard to the question of whether or not *Energizing Physics* can serve as a "physics first" course, the answer is a qualified yes. The 2011-2012 pilot was the first year that the course was used with ninth graders. Their performance on the EPCI post-tests indicates that ninth graders can indeed grapple with physics concepts and learn a significant amount through this course that would prepare them for a more advanced high school course. However, ninth graders are not likely to achieve as much as older students, especially with respect to such areas as interpreting motion graphs, which require practice in reading graphs of complex functions.

Table 2. Pre-Post Scores for students in 9th grade (N = 241)

N=241	Pre-Test			Post-Test			t	p <	d
	Mean	SD	%	Mean	SD	%			
Average Speed (1-4)	1.31	1.055	33	1.71	1.114	43	5.228	.001	0.37
Energy (5-14)	3.73	1.597	37	5.17	1.787	52	10.679	.001	0.85
Work and Energy (15-24)	3.91	1.823	39	5.16	2.111	52	8.793	.001	0.63
Force and Motion (25-34)	2.97	1.878	30	4.32	2.333	43	9.507	.001	0.64
Electrical Energy (35-37)	1.03	0.752	34	1.41	0.932	47	5.293	.001	0.45
Total	12.95	4.354	35	17.77	5.956	49	16.056	.001	0.92

% = Percent Correct SD = Standard Deviation d = Cohen's d, a measure of effect size

Table 3. Pre-Post Scores for students in 10th 11th and 12th grade (N = 356)

N=356	Pre-Test			Post-Test			t	p <	d
	Mean	SD	%	Mean	SD	%			
Average Speed (1-4)	1.89	1.059	47	2.56	1.018	64	11.699	.001	0.65
Energy (5-14)	4.30	1.624	43	5.57	1.888	56	11.464	.001	0.72
Work and Energy (15-24)	4.43	1.873	44	5.96	2.043	60	12.804	.001	0.78
Force and Motion (25-34)	3.71	1.909	37	6.71	2.204	67	23.662	.001	1.46
Electrical Energy (35-37)	1.20	0.702	40	1.59	0.846	53	7.071	.001	0.50
Total	15.53	4.633	42	22.39	5.940	61	23.978	.001	1.29

% = Percent Correct SD = Standard Deviation d = Cohen's d, a measure of effect size

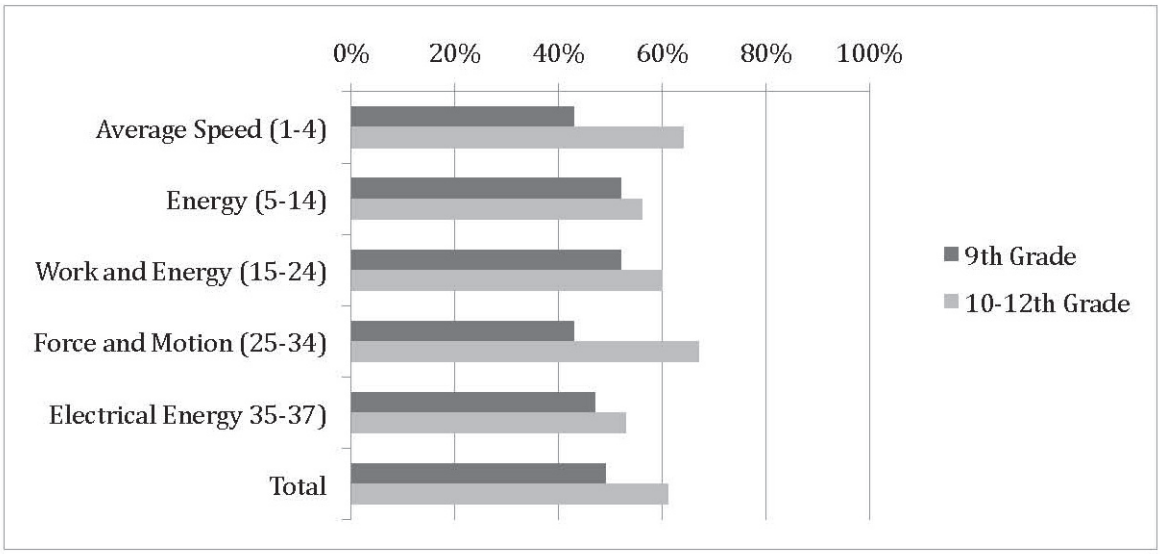


Figure 5. Post-test scores for 9th graders vs. post-tests scores for 10th 11th and 12th graders.

Discussion

In answer to our three research questions we found that *Energizing Physics* students do indeed improve their conceptual understanding of physics fundamentals. Although the effect size is large, as one would expect from a full-year physics course, we would like to see further improvement in student skills. Given that the treatment was a pilot test of the course, we can expect performance to improve as we fine-tune the course content and teacher guide.

By breaking down the EPCI into five topics we were able to determine that students significantly improved their scores in all areas. Effect sizes were medium to large with the greatest gains in Force and Motion, Conservation of Energy, and Work and Energy and the least gains in Average Speed and Electrical Energy.

Finally, we found that 9th grade students improved their scores significantly in all five content areas. However, they did not do as well on the post-test as more mature students in grades 10, 11, and 12. Consequently, *Energizing Physics* may be used as a "physics first" course with the recognition that 9th graders will find it more challenging than older students.

We also want to call attention to the *Energizing Physics Concept Inventory* (EPCI) as a tool that might be used by others to measure students' growing understanding of physics fundamentals. Since this study aimed at not only measuring the impact of the course, but also developing an assessment instrument, a revised version of the EPCI (Version 4.0) will be provided on the *Energizing Physics* website for teachers and researchers who register and obtain a password.

We conclude by pointing out an important limitation of these findings. Our study involved no comparison group that used an alternative physics course. Consequently our findings are limited to answering the questions discussed above. However, as a preliminary study it does answer the question of whether or not further research studies that compare *Energizing Physics* with other physics curricula are warranted. As we have shown, it is indeed worth further study, and we encourage other researchers who were not on the development team, and therefore can be unbiased, to undertake such a study.

This study was also limited in that it did not compare instruction with and without formative assessment. Given that formative assessment tools are woven into the fabric of the course it is probably not possible to isolate the formative assessment elements to determine their contributions to student learning. However, the present study, along with the other chapters in this monograph, provide what Elliot Bennett has called a "concrete instantiation" that illustrates what theory-based formative assessment looks like and how it might work in a real setting (Bennett, 2011, p. 8).

Acknowledgements

We are indebted to the teachers and students who participated in this study. This work was partially supported by a National Science Foundation grant to Portland State University (NSF grant # 1020385 Bridging the Gap Between High School and College Physics: An Exploratory Study). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), p. 5-25.

Osowiecki, A., & Southwick, J. (in press). *Energizing Physics*. New York, NY: Freeman. (The course is also described at energizingphysics.com)

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Erlbaum.

Wilson M., & Draney, K. (2004). Some links between large-scale and classroom assessments: The case of the BEAR Assessment System. In M. Wilson, (Ed.): *Towards coherence between classroom assessment and accountability. 103rd Yearbook of the National Society for the Study of Education, Part II*. Chicago, IL: University of Chicago Press.

Wilson, M., & Scalise, K. (2003). Reporting progress to parents and others: Beyond grades. In J. Atkin, & J. Coffey, (Eds.). *Everyday Assessment in the Science Classroom*. Arlington, VA: NSTA Press.

Wilson, M., & Scalise, K. (2006). Assessment to improve learning in higher education: The BEAR Assessment System. *Higher Education, 52*, 635-663.

Wilson, M., & Sloan, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, *13*(2), 181-208.

Chapter 6

Depth vs. breadth in a new inquiry curriculum

Jesse Southwick, Boston Latin School, Boston, MA

Editors' Note: Depth vs. breadth is a trade-off issue that never seems to end. This paper by one of the authors of *Energizing Physics* describes the conundrum: although research studies indicate that depth is more important for learning, high stakes tests have difficulty measuring depth, and more easily reveal gaps in breadth. What are some possible pathways that might lead science educators to find a way to a more satisfactory resolution for students both in terms of learning outcomes and standardized test scores?

Depth vs. breadth in a new inquiry curriculum

Teachers everywhere struggle with the age-old question of "depth vs. breadth" in planning their year. Adding more depth in one topic, in the form of additional activities, projects, or deeper conceptual questions, involves a cost to the number of topics covered in the year. In our decade-long development of our new curriculum, *Energizing Physic* (Osowiecki & Southwick , in press), my co-author and I have embraced depth. In this paper we will share our rationale, our experience and some implications.

Rationale for Depth over Breadth

1) Depth gives better preparation for college. The basic question of depth vs. breadth has been debated and researched for years, but we have found the most convincing piece of evidence supporting the choice of depth over breadth to be Schwartz et al.'s (2008) dramatic finding that students who had studied at least one topic for a month or more in their high school science class got significantly better grades in college science classes than those who reported no coverage in depth. Judging from college grades, the advantage to covering at least one high school physics concept in depth was equivalent to an additional two-thirds of a year preparation in contrast to students whose teachers covered all of the major physics subjects. Learning some topics deeply, allowing time for the students to see connections, extending ideas to deeper levels, or connecting concepts to experiments or projects, seems to prepare students for success in college. Their conclusion is clear and well justified: "teachers should use their judgment to reduce coverage in high school science courses and aim for mastery by extending at least one topic in depth over an extended period of time" (Schwartz, 2008, p. 798).

2) Depth means repeatedly connecting ideas for understanding. The depth we are embracing involves taking physics beyond a collection of facts and formulas to memorize, or simple algebra problems to "plug and chug." Physics ideas must be repeatedly connected to each other and to data in the real world in

order for students to see physics as a coherent system that applies to the world around them. The traditional coverage model of physics teaching does not lead to this deeper level of understanding. This contrast has been articulated by other science education researchers as well:

> Students who have difficulties often view physics knowledge as a collection of facts, formulas, and problem solving methods, mostly disconnected from everyday thinking, and they view learning as primarily a matter of memorization. By contrast, successful learners tend to see physics as a coherent system of ideas, the formalism as a means for expressing and working with those ideas, and learning as a matter of reconstructing and refining one's current understanding. (Hammer & Elbe, 2003, p. 54)

Clearly a different physics course is needed from the traditional coverage course—a course where students have repeated chances to connect the models of physics to experiments, to new situations, to other models, to engineering design, and to everyday life.

3) Depth is embraced by science standards. This preference for depth over breadth has been a key idea expressed in every science standards document that has come out over the past two decades, beginning with *Science for All Americans* (AAAS, 1989), the *Benchmarks for Science Literacy* (AAAS, 1993), and *National Science Education Standards* (NRC, 1996), and culminating with *A Framework for K-12 Science Education* (NRC, 2012). The most recent standards document, the *Next Generation Science Standards* (NGSS/Achieve, 2013), references the most important ideas in the *Framework*, including the following statement about the value of depth over breadth:

> Second, the framework focuses on a limited number of core ideas in science and engineering both within and across the disciplines. The committee made this choice in order to avoid the shallow coverage of a large number of topics and to allow more time for teachers and students to explore each idea in greater depth. Reduction of the sheer sum of details to be mastered is intended to give time for students to engage in scientific investigations and argumentation and to achieve depth of understanding of the core ideas presented. Delimiting what is to be learned about each core idea within each grade band also helps clarify what is most important to spend time on and avoid the proliferation of detail to be learned with no conceptual grounding. (NGSS/Achieve, Appendix E, p. 1)

The *Framework for K12 Science Education* also supports our vision for what we should do with the "depth" gained by teaching fewer topics in greater detail. The "Disciplinary Core Ideas" dimension, in other words the "content," is only one of three dimensions we should be working to teach. The other two dimensions, "Scientific and Engineering Practices" and "Crosscutting Concepts" indicate to us that significant time and energy should be spent in each science class working on doing science (as specified by the practices) and connecting concepts together (with crosscutting concepts). The eight scientific and engineering practices in Dimension 1 of the *Framework for K12 Science Education* are listed below:

Dimension 1: Science and Engineering Practices

1. Asking questions (for science) and defining problems (for engineering)
2. Developing and using models
3. Planning and carrying out investigations
4. Analyzing and interpreting data

5. Using mathematics and computational thinking
6. Constructing explanations (for science) and designing solutions (for engineering)
7. Engaging in argument from evidence
8. Obtaining, evaluating, and communicating information (NRC, 2012, p. 41)

Looking at Dimension 1 more closely, we see the *Framework*'s vision of a different type of science classroom where students spend significant time and energy doing science. If all eight of these practices are important and worth doing, we must acknowledge they can't be done in a single lab period. Building each of these practices into students' experience of science takes time. In our reading of the *Framework*, science students should be spending days or weeks (at least) of their year conducting experiments, grappling with the data and conclusions. Students should also have some engineering projects in their physics class.

Energizing Physics

Since 2004, Aaron Osowiecki and I have been developing *Energizing Physics* in the classrooms of Boston Latin School in Boston, MA. Boston Latin is an urban public exam school with a diverse and talented student body. Our introductory physics course has a mix of students in grades 10, 11 and 12. We were informed and inspired by the *National Science Education Standards* (1996) and relevant research in the field to create a new type of physics class experience. The course is an inquiry curriculum, built around a progression of class activities during which students answer scaffolded questions through experiments, small group work, class discussions, and problem solving. We have been informed by research into misconceptions, the 5E learning cycle (Bybee, 2002; Bybee, et al., 1989) and the importance of formative assessment. Each year the activities, questions and assessments get refined and improved based on student successes and challenges.

While *Energizing Physics* has a clear progression through physics ideas and building on previous skills and knowledge, the sequence is nontraditional. After a first chapter about speed and measurement, the course explores conservation of energy, building the rest of mechanics on a foundation of energy ideas. The sequence of chapters is listed in the table below. As discussed in Chapter 1 of this monograph, putting energy "front and center" allows it to be a central theme in the course. It also has the benefit of putting more challenging and abstract material, such as motion graphs, later in the course when students are more comfortable with physics. The *Energizing Physics* chapter sequence is shown below:

Chapter 1. Speed and Measurement
Chapter 2. Energy
Chapter 3. Work and Energy
Chapter 4. Forces and Motion (1 dimension)
Chapter 5. Electricity
Chapter 6. Thermal Energy
Chapter 7. Waves
Chapter 8. Multiple Objects, Multiple Dimensions

Energizing Physics has been built with the following goals:

Goal 1. Deep understanding of concepts and connections. We design questions that require application of multiple concepts in a rich way. During development, we sometimes had the experience that a quiz or exam question was unsuccessful, and so we went back to the activity and

included scaffolded questions to help students build a deeper understanding in the next iteration of the course.

Goal 2. Accessible to all. Introductory physics must be engaging and accessible to all students. It must meet them where they are mathematically and help them build skills and understanding of what physics is all about. While we still find that physics is very challenging for some students, all students need to have the chance for some memorable success. We were not willing to leave math out. Mathematical problem solving takes time and practice to build real skills, and must be scaffolded from less to more challenging problems as the year progresses. While the introductory physics course should provide a solid foundation for subsequent study of physics in high school or in college, the course should also be a good introduction for a student who never takes more physics. It should give each student skills and knowledge for citizenship and a take-away appreciation for physics.

Goal 3. Authentic science inquiry. As the *Framework* (NRC, 2012) and *NGSS* (Achieve, 2013) suggest, students must model phenomena, conduct experiments, analyze their own data, share and discuss evidence with classmates, and engage in sense-making. Activities throughout the curriculum should enable students to practice all of these skills. Furthermore, successful mastery of physics includes more than success in solving problems. Students must also be able to connect physics concepts with hands-on experiments and engineering design projects. To provide this sort of practice, *Energizing Physics* includes multi-day projects at the end of each chapter.

The overall sequence of chapters, as well as each chapter's internal progression, has been designed and refined with these goals in mind. The course starts by asking students to design a method for measuring speed, without using modern measuring devices or units. The need for measurement, unit conversions, and problem-solving strategies surfaces in the context of this accessible first task.

Reaching all three of the goals listed above requires a significant commitment of time. To help illustrate the choices we made in favor of depth over breadth, consider Chapter 3, Work and Energy.

This example gives a glimpse into the depth vs. breadth tradeoff. Early in our course development, we decided we weren't satisfied with traditional physics problems. While we might be able to get students to solve these problems as an exercise in algebra, it wouldn't necessarily mean they really understood the models of friction or work, and their connections. Also, students would not necessarily realize that these models can actually be applied to everyday objects, e.g., that coefficients of friction are real quantities describing their own experiences. Early on we added the slingshot project to foster a richer experience with this content. In the first years, we set the students loose with the project soon after introducing the necessary equations for work and friction. We observed that most students struggled to make the connections on their own, and that even if they knew the equations they did not understand the ideas deeply. Some groups were successful, often led by a few star students, but many students found that they were still confused when they encountered the project at the end of the chapter. Subsequent revisions of the chapter involved more activities that scaffolded students to the deeper understandings and connections needed for success on the project. The additional activities involved some of the experimental skills needed and more practice with concepts and connections. These additional activities take more time (at least another week of class) but have resulted in notable improvement in the number of students who undertake the final project with the knowledge and skills they need to participate fully with their teammates and succeed.

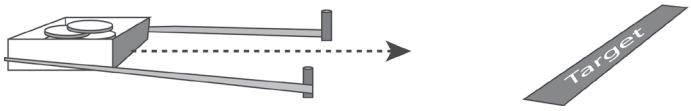
Traditional Physics Problem	Chapter 3 Introduction of *Energizing Physics*
Problem: You are launching a 30 gram box across the floor using a 50 N/m slingshot. The coefficient of friction between the box and the ground is 0.25. You stretch the launcher back 0.20 m. How far will the box slide? **Solution:** $EPE + W_f = 0$ $\frac{1}{2} k x^2 - (\mu mg)d = 0$ $\frac{1}{2}(50)(0.2)^2 - (0.25*0.03*9.8)d = 0$ $d = 13.6m$	At the conclusion of this chapter, you will become a target shooter. Your goal will be to launch a small box (jewelry, mints, etc) from a rubber band slingshot so that the box slides across the ground and stops a certain distance away as shown in the drawing. Your instructor will tell you **(1) how many pennies to put in the box** and **(2) how far it needs to slide before coming to a stop**. Over the next few weeks, you will prepare for this challenge by learning how to modify the $E_1 = E_2$ model to incorporate energy being added or subtracted from a system. Let's get started! 
Comment: Students in many physics classes can learn to solve this problem by memorizing the necessary equations and using algebra to "plug and chug." Success at doing so might indicate "understanding" and the teacher might move on to a new topic. Without actually doing the experiment the students do not have an opportunity to observe how well their model matches reality.	**Comment:** In the progression of Chapter 3, students learn the various equations (models) that relate the variables in this situation as they progress through the chapter. Midway through they experiment with how adding mass affects the friction of sliding objects, and discover the linear relationship they need for the project. When they reach the end of the chapter, they have to solve the traditional problem of relating stretch to sliding distance, but they have to do it in the context of their collected data. Success on the project requires a deep understanding of 1) the importance of context and 2) the relationship between a model and experimental reality.

Figure 1. The slingshot project.

Many of our chapters have had similar evolution. If we want students to engage in the scientific practices of Dimension 1, and then connect experiments to Dimension 2 concepts in deep ways, instruction will take longer. We have found that adding less complex experiments along the way helps students build skills towards the rich end-of-chapter projects. Cycling between experiments and problems and between group work and individual practice, allows reinforcement and chances for different learners to make connections at different speeds. We have also found that adding more practice problems, and spending group time working on them, helps more students to succeed on assessments. Doing science and building problem-solving skills takes more time, but yields results: more students can be successful with deep physics mastery.

Piloting the Curriculum

As part of the NSF-supported project (NSF grant # 1020385) "Bridging the Gap from High School to College Physics," we revised and prepared our *Energizing Physics* materials for use in five other schools. To support our goals as listed above, we incorporated the following new elements as part of the learning cycle of the activities:

1. A learning target for each activity that is transparent to students.
2. An entry-point question to engage and elicit prior understanding
3. A DYGIT (Did You Get It?) question which allows teacher and student to quickly and formatively assess class understanding of key ideas.
4. A reflection question that asks students to summarize the big idea(s) learned and apply the idea(s) in a new context.
5. Practice problems after each activity instead of the end of the chapter.
6. Quick quizzes that formatively assess each student before moving too far past the activity.

While we had been doing some of these things before, formalizing them throughout the curriculum was a successful part of our pilot project. These elements helped create the transparent and reflective learning environment we wanted, with a full learning cycle for each activity. We saw many more students succeed with the previously challenging problems, and students appreciated the wrapping up of each activity's learning cycle. We were helping students learn physics deeply, taking time to tie concepts together and connect them to data, to science. Our pilot teachers gave positive feedback about each of these elements.

Depth Does Cut into Breadth

In *Energizing Physics*, we have emphasized depth:

- Activities and projects involve all 8 of the *Framework*'s (NRC, 2012) scientific practices.
- Scaffolded practice problems lead all students to sophisticated multi-concept connections.
- Elements have been added to make the full learning cycle transparent and reflective.
- Formative assessment tasks allow the teacher and students to monitor progress and review or re-teach if necessary.

These depth elements match our goals and the goals of the *Framework* (NRC, 2012) and *NGSS* (Achieve, 2013), but they do take significant time. During this past pilot year, most classes were unable to get much past Chapter 5, Electricity. Let me be honest about what this means. Our course chooses deep understanding of a few topics instead of fast "coverage" of many topics. Our introductory students got limited or no exposure to waves, thermal energy, momentum and collisions, projectile motion, circular motion, magnetism, fluids, optics or any "modern" physics. In the year of the Higgs Boson, it didn't make our list. Many physics teachers would be uncomfortable leaving out so many classic topics of physics, and we respect these reservations. However, we have seen the fruits of our choices: more students get more of what we want for them, including experience with physics as an experimental science, a chance to see deep connections, and the chance to be successful with rich and challenging projects.

Concerns About the Future

We want to be honest that we too struggle with fully embracing a physics course that doesn't teach physics teacher favorites like projectile and circular motion, and doesn't have time for modern physics like relativity or particles. While continuing to revise *Energizing Physics* over the next few years, we hope to tweak a few things and increase our efficiency to be able to finish Chapter 6, Waves and Chapter 7, Thermal Energy. We feel that these are classic and important physics topics that are accessible and ubiquitous in

student lives. But we find it unlikely that we'd ever be able to add much else. This is where we have come to in our own "breadth vs. depth" trajectory.

We are excited about pushing our science students to engage in science practices and recognize cross-cutting connections. We feel like our choices in *Energizing Physics* are supported by these priorities. However, we are concerned about the future. The final draft of the *Next Generation Science Standards* (Achieve, 2013), includes several major physics concepts that are addressed in chapters 6, 7, and 8 of *Energizing Physics*. If we attempt to cover all of those concepts as well, we will have to give up some of the experiences that we know our students need to understand mechanics in depth.

Judging from prior education reforms, it is likely that there will soon be high stakes tests aligned with the *Next Generation Science Standards* (Achieve, 2013), which means that physics teachers will experience pressure to cover all of the core ideas in the NGSS. If this happens, teachers will not be able to take the time needed for depth, and a fundamental tension may develop between standards-based reform as proclaimed in the visionary *Framework* (NRC, 2012) and standards-based reform as driven by high stakes tests.

It is an exciting and critical time in science education. New ideas are spreading, new standards are coming out, and physics classrooms are changing (or will be soon). Our experience with *Energizing Physics* has led us to embrace inquiry and depth. Undoubtedly this results in less coverage, less breadth. We hope this chapter raises the issue again, and we'd like to hear from others about their opinions and solutions.

Acknowledgements

We are indebted to the teachers and students who participated in this study. This work was partially supported by a National Science Foundation grant to Portland State University (NSF grant # 1020385 Bridging the Gap Between High School and College Physics: An Exploratory Study). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

AAAS Project 2061. (1989). *Science for all Americans.* New York, NY: Oxford University Press.

AAAS Project 2061. (1993). *Benchmarks for Science Literacy.* New York, NY: Oxford University Press.

Achieve. (2013). *Next generation science standards.* Retrieved from http://www.nextgenscience.org

Bybee, R. (2002). *Learning science and the science of learning.* Arlington, VA: NSTA Press.

Bybee, R., Buchwalk, S., Crissman, D., Heil, D., Kuerbis, P., Matsumoto, P., & McInerney, J. (1989). *Science and technology education for the elementary years: Frameworks for curriculum and instruction.* Washington, D.C.: The National Center for Improving Instruction.

Hammer, D., & Elby, A. (2003). Tapping epistemological resources for learning physics. *The Journal of the Learning Sciences*, *12*(1), 53-90.

National Research Council (1996). *National Science Education Standards.* Washington, D.C.: National Academy Press.

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

Osowiecki, A., & Southwick, J. (in press). *Energizing Physics.* New York: Freeman. (The course is also described at energizingphysics.com)

Schwartz, M. S., Sadler, P. M., Sonnert, G, & Tai, R. H. (2008). Depth Versus Breadth: How Content Coverage in High School Science Courses Relates to Later Success in College Science Coursework. *Science Education*, *93*, 798-826.

Chapter 7

The road ahead

Cary Sneider, Portland State University, Portland, OR

Editors' Note: *Energizing Physics* guides teachers to use formative assessment in a number of ways that engage students in "doing physics," and that embed information about common student misconceptions within the teacher guide. Interesting questions for the reader are: How might such a course best support professional development for physics teachers? What sorts of experiences would teachers need to take full advantages of the structures built into the course?

The road ahead

This monograph can be viewed as a snapshot of an ongoing process to improve physics education by combining what we know about best practice in science teaching and curriculum with a growing appreciation for the potential of formative assessment. The overarching goal of the project has been to develop and test a formative assessment system to help physics teachers monitor their students' learning so that they can modify instruction to meet their students' needs *during instruction*. The purpose of that work is both to deepen students' conceptual understanding and to increase their self-confidence. In order to accomplish this goal we have created a system that integrates two very different approaches to formative assessment: 1) a process that occurs minute-by-minute in the classroom as students interact with each other and with their teacher and 2) a number of tools built into the curriculum, with guidelines for teachers to assess students' levels of accomplishment on each learning target. We have relied on our pilot teachers to invent and share ways that they are using the information that they obtain through formative assessment and to modify instruction accordingly. Now it is time to take stock and consider next steps. Most of our conversations have concerned the following issues:

- Alignment with NGSS
- Second Pilot Test
- Differentiated Instruction
- Professional Development
- Further Research

Alignment with NGSS

The previous chapter in this monograph described the conundrum that arises when a team of researchers and experienced physics teachers tackle the fundamental issues that have plagued high school physics for decades. With a spirit of "leave no physics student behind," the *Energizing Physics* course authors and pilot teachers have built in a sequence of activities, tools, and discussions to help all students succeed. The result is a very effective course that engages students in learning a limited number of physics

concepts in depth, so that the students can use what they have learned to solve interesting and very practical problems. However, depth has been achieved at the expense of breadth. Although the recent publication of Next Generation Standards (Achieve, 2013) includes the concepts that are featured in *Energizing Physics*, it also includes concepts that are in *Energizing Physics* chapters 6, 7, and 8, that few pilot teacher had time to teach in one year. We will need to examine the course carefully to see if some parts of chapters 1-5 can be left out to make room for teaching at least some of the topics from these final text chapters. If we find that is not possible, we may recommend that physics become a 1.5 or 2.0 year course so that all of the core ideas can be taught in-depth.

Second Pilot Test

We were very pleased to find that several commercial publishers of educational materials were interested in *Energizing Physics*. Near the end of our grant period the course authors signed a contract with Bedford, Freeman, & Worth Publishing Group, which also publishes *Living By Chemistry*, another high school science course with an embedded system of formative assessment. The new course publishers plan to conduct a larger pilot study in the 2013-2014 school year to test the improvements made as a result of the first pilot, and to identify any other elements that need to be further strengthened. The second pilot will provide an excellent opportunity for researchers who were not involved in the development of the course materials, and would therefore be unbiased, to conduct research on program effectiveness

Differentiated Instruction

The value of formative assessment is not realized until teachers are able to adjust their instruction based on the information they gain from the assessment. To that end we have relied on our pilot teachers to tell us how they are differentiating instruction. In most cases it has affected the pacing of the class, especially when it turns out that many of the students are confused and additional time needs to be spent teaching a concept. However, several of the teachers have been quite inventive, either structuring the class so that more advanced students help those who are less advanced, or dividing the class so that the instructor can spend more time with students who are struggling, while more advanced students use supplementary materials (including research projects called "interludes") in which they apply their knowledge and skills to additional real-world situations (Osowiecki, 2011). Expanding the repertoire of approaches to differentiated instruction and providing guidance to teachers based on the results of the pilot studies will be important areas for our future work.

Professional Development

The finding that many teachers may not be well-prepared to teach physics suggests that a major focus of our future work will be in developing and testing new approaches to professional development that results in increased pedagogical content knowledge. This will be especially challenging in light of school budget shortfalls, and reductions in recent years of grant funds from the federal level for professional development of teachers. Cost-effectiveness will therefore become an important criterion for such programs, along with other more traditional measures of efficacy.

Encourage Further Research

Although our work has emerged from the literature on physics education and formative assessment, our contributions to those fields so far have been limited. As we point out in Chapter 5, we have not been able to tease out the effects of different approaches or specific tools for formative assessment, nor have we

compared the effectiveness of *Energizing Physics* with courses taught using a non-mathematical approach or with a traditional mathematical physics course that covers more topics in the course of a year. However, we have prepared the ground for such studies by developing a cognitive-domain model with specific learning targets for student outcomes in physics, a reliable means for measuring students' levels of accomplishment of the learning targets, instruments that enable teachers and students to test their understanding during the study of a unit, and a pre- post-test instrument with excellent psychometric qualities for measuring what students have learned in an introductory physics course. As we pointed out in our introduction, these are elements that researchers need to examine the value of formative assessment in the context of a domain-specific program (Bennett, 2011). Other studies might examine the value of a professional development program on common misconceptions in physics for increasing both teacher knowledge and student outcomes. We invite other researchers to use these instruments to conduct their own studies, and we offer whatever assistance we can provide to help move those efforts forward.

References

Achieve. (2013). *Next generation science standards.* Retrieved from http://www.nextgenscience.org

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), p. 5-25.

Osowiecki, A. (2011). Idea bank: An environmentally responsible physics course. *The Science Teacher*, *78*(9), 68-9.