Doing the Project and Learning the Content: Designing Project-Based Science Curricula for Meaningful Understanding

DAVID E. KANTER
Curriculum, Instruction, and Technology in Education Department, College of Education, and Biology Department, College of Science and Technology, Temple University, Philadelphia, PA 19122, USA

Received 30 July 2008; revised 23 September 2009; accepted 5 October 2009

DOI 10.1002/sce.20381
Published online 15 December 2009 in Wiley InterScience (www.interscience.wiley.com).

ABSTRACT: Project-based science curricula can improve students’ usable or meaningful understanding of the science content underlying a project. However, such curricula designed around “performances” wherein students design or make something do not always do this. We researched ways to design performance project-based science curricula (pPBSc) to better support the meaningful understanding of science content. Using existing curriculum design frameworks, we identified the learner’s need to “create the demand” for the science content, anticipating how to use it in the performance, and to “apply” the science content, both being necessary to ensure meaningful understanding. Designing the pPBSc I, Bio we discovered how these guiding principles manifested as curriculum design challenges. We generalized from the design of I, Bio and related literature design approaches for addressing each challenge. Finally, we measured the extent to which a pPBSc incorporating these design approaches developed meaningful understanding, 652 middle grades students using I, Bio completed pre- and posttests on the science content behind the I, Bio performance. Our findings provide preliminary evidence that a pPBSc that incorporates these design approaches is consistent with gains in meaningful understanding. We discuss how the results of this work can be used to improve systematic experiments on instructional supports.


INTRODUCTION

Educational researchers have suggested that project-based science (PBS) curricula may promote students’ understanding of standards-based science more effectively than other
curricula (Krajcik, McNeill, & Reiser, 2008). Because students are assessed on the basis of what they know and can do with the science concepts targeted by state and national standards, helping students learn standards-based science is an essential goal. Not only do educators want to change students’ alternative conceptions about standards-based science, which often resist change with traditional instruction (Driver, Squires, Rushworth, & Wood-Robinson, 1997), into more scientific understandings, but they also want students to be able to retrieve and use their improved understandings to solve novel problems (Ausubel, 1968; Bransford, Brown, & Cocking, 1999), coming to know standards-based science concepts as a network of ideas and conditions of their applicability (Bransford et al., 1999). Such meaningful understanding (Ausubel, 1968), equivalently the ability to use what one has learned to solve novel problems, thus showing evidence of effective remembering plus positive transfer (Bransford et al., 1999), contrasts with students effectively remembering but having poor transfer, or a kind of learning that is inert (Whitehead, 1929). Meaningful understanding of standards-based science does more than just address the goals of classroom accountability; it also addresses today’s need to educate scientifically literate individuals, who can use what they know and build on that knowledge to make decisions in the face of uncertainty (Talbert & McLaughlin, 1993). PBS curricula embed standards-based science learning in realistic projects, which, beyond improving what students remember, make such curricula good candidates for improving students’ meaningful understanding of standards-based science.

Two major types of projects can serve as PBS curriculum projects. An investigation project might, for example, require answering a puzzling question about why some finches survived a crisis on a Galápagos island while many others died off (Reiser et al., 2001) or ask students to investigate different causes of disease and how their bodies fight disease (Hug, Krajcik, & Marx, 2005). Studies of investigation PBS curricula suggest that they improve students’ meaningful understanding of science (Krajcik et al., 2008; Linn, Bell, & Davis, 2004; Linn & Clark, 1997; Marx et al., 2004; Rivet & Krajcik, 2004; Schneider, 2002). A performance project, on the other hand, might require designing paint, fins, and nose cones to make model rockets go as high as possible (Barron et al., 1998) or designing and building a miniature car and its propulsion system to go over several hills and beyond (Kolodner et al., 2003). In addition to students designing artifacts (e.g., Learning by Design (Kolodner et al., 2003), Design-Based Science (Fortus, Dershimer, Krajcik, Marx, & Mamlok-Naaman, 2004), or design-based learning (Apedoe, Reynolds, Ellefson, & Schunn, 2008)), performances can be other sorts of design projects that do not produce the same kind of material artifacts but whose intellectual activity is fundamentally the same: prescribing remedies for a sick patient or devising a new sales plan (Simon, 1999, p. 111). Performances are thus studies into the “sciences of the artificial,” wherein design has the broader definition of students devising courses of action aimed at changing existing situations into preferred ones (Simon, 1999). This is as contrasted with investigations, which are studies into the “natural sciences,” wherein students aim to learn about natural things, how they are and how they work. This same dichotomy is seen in the difference between Stokes’s Pasteur’s and Edison’s quadrants, wherein research is inspired by considerations of use, as contrasted to Bohr’s quadrant, wherein research is solely concerned with a quest for understanding (Stokes, 1997). Others have described a similar dichotomy (Clark & Linn, 1997; Dewey, 1913; Kanter et al., 2003; Schank & Korcuska, 1996).

Performance projects naturally engage students in the world (Roberts, 1995), the better to see the utility of the underlying science; the real-world relevance of such projects may excite students to extended participation in class. Thus, PBS curricula designed around performance projects should be good vehicles for building students’ meaningful understanding.
of standards-based science content and have been an active area of research. However, while some studies of specific performance PBS curricula (hereafter, pPBSc) suggest that they, too, improve meaningful understanding (Apedoe et al., 2008; Fortus et al., 2004; Holbrook, Gray, Fasse, Camp, & Kolodner, 2001; Kanter & Schreck, 2006; Kolodner et al., 2003; Puntambekar & Kolodner, 2005), other studies suggest that pPBSc may actually interfere with students’ ability to learn and learn to use science content. Sherin, Brown, and Edelson (2005) describe a pPBSc in which students analyze climate data to prepare briefings about the causes of the earth’s warming and what to do about it. The findings of Sherin et al. suggest that while students successfully completed the project, they maintained unscientific understandings of the nature of light and the structure of the earth–sun system and could not use this knowledge in their briefings. Petrosino (1998) found that students completed the making and launching of model rockets in a pPBSc, but without learning target ideas about experimentation and measurement sufficiently to evaluate what makes rockets better or worse. We contend that for all their potential, pPBSc can present an unresolved tension between the practical doing and the content learning, resulting in students failing to gain meaningful understanding.

The aim of this study was to explore approaches to designing pPBSc that might more reliably ensure meaningful understanding. Parts I and II of this paper identify approaches for designing pPBSc to better ensure students gaining meaningful understanding of science content while doing the performance. Existing curriculum design frameworks help identify guiding principles for developing the meaningful understanding of science content. While designing a specific pPBSc, I, Bio, design challenges (for curriculum designers) arising from applying these guiding principles were identified, thus revealing some specific ways that the tension between the doing and the learning may manifest. Then, the study considers how the design of I, Bio resolves these challenges—in view of other literature on curriculum design strategies—and generalizes approaches for designing pPBSc for addressing these challenges to meaningful understanding. Parts III and IV of this paper investigate whether these pPBSc design approaches do indeed support meaningful understanding as expected by evaluating the extent to which I, Bio supports gains in meaningful understanding of the science content that underlies its performance.

**PART I: GUIDING PRINCIPLES FOR MEANINGFUL UNDERSTANDING OF CONTENT**

Part I begins with guiding principles for promoting the meaningful understanding of science content. These principles are derived from existing science curriculum design frameworks. To improve the understanding of science, especially that which students have difficulty learning even after instruction, Stern and Roseman (2004) argue for instructional strategies that take account of learners’ ideas, engage learners with relevant phenomena, and promote thinking about these phenomena in order to learn the science a curriculum aims to teach. Most of the curricula Stern and Roseman reviewed were poor in providing such instructional strategies. This speaks to the guiding principle that learners need to be active in constructing their understanding of science content. Inquiry can support learners constructing scientifically correct conceptions from whatever alternative conceptions they possess by collaboratively designing and carrying out investigations, collecting and analyzing data, and reflecting and communicating results (Krajcik, Czerniak, & Berger, 1999; Sandoval & Reiser, 2004; Singer, Marx, Krajcik, & Chambers, 2000). This same guiding principle is part of the Learning-for-Use curriculum design framework, which explicitly requires that content ideas be “constructed” in memory via firsthand experience (Edelson, 2001).
However, if the science knowledge is only constructed, the learner may attain an improved general understanding but not necessarily be able to use this knowledge in new contexts. Knowledge construction is necessary but insufficient for meaningful understanding. Consequently, another guiding principle—motivating—precedes constructing. This principle, also a feature of the Learning-for-Use framework, posits that the learner must have a need to acquire the knowledge (Edelson, 2001). Motivating, here, is used in a narrow sense to refer to just the motivation to acquire the knowledge, and assumes that the learner is already engaged. In their framework, Stern and Roseman (2004) similarly suggest that for the science ideas being taught, instructional strategies must convey the purpose of the activity to the learner and relate this purpose to the larger curriculum unit and its other activities. For all the curricula Stern and Roseman reviewed, all but one were poor to fair in providing instructional strategies to convey a lesson purpose and all but one were poor to fair in providing instructional strategies to convey the unit purpose. Thus, before the science knowledge is constructed, the learner must have a reason to learn it, thus creating a context in memory for integrating the new knowledge, ultimately to build meaningful understanding. For pPBSc, motivating the learner takes the form of “creating demand” to use the content to make progress on the performance (Edelson, 2001). Again, “creating demand” does not concern a student’s emotional state but rather the learner anticipating a use for the content in advance of the learning. Learning content while anticipating its use in the performance is crucial to the learner being able to retrieve and apply the content in the future (Gagné, Briggs, & Wager, 1992).

The third guiding principle—after creating a demand for the science knowledge and then constructing it—is equally necessary for promoting meaningful understanding: organizing the content knowledge. According to the Learning-for-Use framework, the newly obtained knowledge must be organized, connecting new knowledge structures to old ones to support future retrieval and use (Edelson, 2001). This same guiding principle informs Stern and Roseman’s framework (2004), which requires curricula to provide instructional strategies, such as demonstrations or practice questions, that emphasize the utility of the science ideas, something that that all the science curricula they reviewed were poor in. For pPBSc in particular, organizing content knowledge comes in the form of “applying” the science ideas to complete the performance project (Edelson, 2001).

These curriculum design frameworks suggest that to help learners achieve a meaningful understanding of science content, a pPBSc must first create the learner’s demand to learn the science in anticipating its usefulness for the performance. Then, after the learner is supported in constructing an understanding of the content, the learner must be helped to organize the connections among content ideas to promote their future retrieval and use by applying the content to do the performance. The “circumstances in which [the] knowledge is constructed and subsequently used determine its accessibility for future use” (Edelson, 2001).

Exposing learners to features of a content domain as they arise naturally in project situations is touted as an effective way to achieve meaningful understanding (Bransford et al., 1999). However, pPBSc do not always ensure this understanding, as evidenced by the Sherin et al. and Petrosino studies discussed above. This failure to promote deep understanding may stem from challenges in addressing the “creating demand” and “applying” guiding principles. If a pPBSc is not designed to help the learner anticipate a utility to the project for the content to be learned, or to help the learner apply the content to make progress on the performance, the learner will not be any better prepared to retrieve and use the science in novel situations, that is build a meaningful understanding. In Part II, Design Challenge 1 describes a curriculum design challenge to “creating demand” for the
DESIGNING PROJECT-BASED SCIENCE TO TEACH CONTENT 529

science content. Design Challenge 2 and Design Challenge 3 describe curriculum design challenges to supporting the opportunity to “apply” the science content. These curriculum design challenges are specific ways we found the tension between the doing and the learning to be realized. We describe these curriculum design challenges as found while designing the pPBSc I, Bio in sufficient detail for another designer to recognize the unexpected ways in which any pPBSc might be failing to support learners “creating demand” for and “applying” the science content. For each design challenge, we identify general curriculum design approaches for its resolution.

PART II: CURRICULUM DESIGN CHALLENGES AND THEIR DESIGN APPROACHES

Part II identifies three curriculum design challenges encountered while designing the pPBSc I, Bio, a curriculum for middle grades biology and physical science content (Kanter, Kemp, & Reiser, 2001). The design work was funded by a National Science Foundation (NSF) Postdoctoral Fellowship in Science, Mathematics, Engineering, and Technology Education, a Research Experience for Teachers supplement to the NSF-funded Center for Bioengineering Educational Technologies, and the Quaker Oats Foundation, and conducted in cooperation with the NSF-funded Center for Learning Technologies in Urban Schools. The work involved a multidisciplinary team of middle grades science teachers, science education researchers, and content experts. The discovery of the challenges was purposeful. By examining the conversations among design team members over the multiple iterations of this pPBSc as it took shape, it was expected that the team would be able to notice curriculum design challenges to addressing the “creating demand” and “applying” guiding principles. The team also believed it would be able to identify general design approaches used for resolving these challenges. Of course, working in the context of designing one pPBSc circumscribes the findings, but preexisting curricula offered little from which to identify these challenges and approaches. This work is the first part of a program of design research. As discussed by Collins, Joseph, and Bielaczyc (2004), a design research methodology must begin with identifying critical elements of the design and how they interact, as well as how each critical element is addressed in the design. In the context of designing one pPBSc, we work to elucidate these critical elements (i.e., the general design approaches) for realizing the guiding principles. As per Collins et al., only in terms of these critical elements can implementations of alternative designs be evaluated and compared.

The science targeted by the I, Bio curriculum was based on national, state, and district science standards for the middle grades. The design team focused on the topic of energy interconversion, that is, how the human body’s organ systems, composed of specialized cells, work together to derive energy from food. Energy interconversion in living organisms is core to teaching life science, as can be seen by the prominence of this concept in Benchmarks for Science Literacy (American Association for the Advancement of Science [AAAS], 1993), the National Science Education Standards (National Research Council [NRC], 1996), and state standards (Kesidou & Roseman, 2003). This concept of energy interconversion in living organisms was also the focus of the AAAS Project 2061 middle school curriculum evaluation study discussed above (Stern & Roseman, 2004). We aimed to build a pPBSc that would improve middle school students’ meaningful understanding of the middle grades biology and physical science content on which this concept is based.
The Performance PBS Curriculum I, Bio

Before discussing the specific design challenges encountered, an overview of the final form of the pPBSc I, Bio and the science it was designed to teach is in order. Figure 1 presents the questions that frame the individual lessons, and the seven science concepts of which the curriculum aims to promote a meaningful understanding.

The first lesson asks the question, “What will it take to redesign our school lunch choices to meet our bodies’ needs?” Students figure out that to promote better health, they will need to learn how to measure the energy that food supplies to their bodies, measure the energy their activities use up, and then design their school lunch and activity choices to balance these two quantities. This is the performance project around which the curriculum was designed. Lessons 2–4 focus on students devising the means by which to measure the energy in food, and, in so doing, creating the demand for, constructing, and applying science content about energy forms and interconversions, chemical change (as related to oxidation), and the properties of matter (as relevant to measuring heat, i.e., specific heat). This first piece of the performance project culminates with students reinventing direct calorimetry. This technique requires burning the food inside an oxygen-filled chamber that is surrounded by water. The heat from the burning increases the temperature of the water, and this temperature change can be used to calculate the calories in food.

Lessons 5–7 focus on students devising the means by which to measure the energy consumed in activities, and, in so doing, creating the demand for, constructing, and applying science content about levels of biological organization, organ systems and their integrated function, and cellular respiration. This second piece of the performance project culminates with students reinventing indirect calorimetry. This technique uses a one-way valve to collect a subject’s expired air while doing a given activity. Using an oxygen and volume sensor, the amount of oxygen consumed is calculated (Consolazio, Johnson, & Pecora, 1963). This measurement indirectly determines in calories the energy used up by all the body’s working cells.

Figure 1. Overview of the performance PBS curriculum I, Bio and the standards-based science content of which it aims to promote a meaningful understanding.
The last lesson, Lesson 8, completes the performance project. Students redesign their school lunch and activity choices to balance the calories consumed in their school lunch choices with the calories used up doing their activities. The literature shows that the seven science concepts addressed in this pPBSc are challenging to teach (Arnaudin & Mintzes, 1985; Barak, Gorodetsky, Chipman, & Gurion, 1997; Dreyfus & Jungwirth, 1989; Driver et al., 1997; Gayford, 1986; Lijnse, 1990; Mintzes, 1984; Mintzes, Trowbridge, Arnaudin, & Wandersee, 1989; Novick, 1976; Nuñez & Banet, 1997; Songer & Mintzes, 1994). As such, I, Bio should be a good test of the effectiveness of the identified design approaches in building students' meaningful understanding.

**Design Challenge 1: Creating the Demand for Unfamiliar Content**

The first curriculum design challenge identified while designing the pPBSc I, Bio stemmed from the difficulty of addressing the guiding principal of creating the learner’s demand for the science content. Specifically, we encountered multiple instances where the requisite content for doing the performance was unfamiliar to the learner, in which case the learner could not anticipate how learning this unfamiliar content would be helpful in doing the performance. The designers knew that students would eventually need to apply this content to do the performance, but how could we get the learners to anticipate such utility when the content was unfamiliar to them? This was the first curriculum design challenge, unique to trying to “create demand” for content in pPBSc.

And yet, if we tried to resolve this challenge by simply telling the student to trust us that they would eventually need to use this content to do the performance, we would take away the student’s chance to anticipate a use for the content before learning it. Some learning of the content may still occur, but the likelihood of inert as opposed to meaningful understanding would be increased, since the student would not have constructed their knowledge of the content anticipating its utility, the circumstance that would support its future retrieval and use. Telling the student to learn the content because she or he would eventually need to apply it would not suitably resolve this design challenge. We needed to find ways to ensure that the student’s demand to learn the unfamiliar content was connected to their perceived utility of the content for making progress on the performance, even if what the student initially anticipated as the utility of the content would be different from the ultimate utility of the content to the performance.

We identified three different design approaches we used to address this design challenge. Each of the three design approaches used a different way to stimulate the student to learn the unfamiliar content by helping them generate expectations about the utility of learning the unfamiliar content for doing the performance. In the next six subsections, we describe the three design approaches—“Unpack the Task,” “Highlight an Incongruity,” and “Try to Apply”—as they appear in the flow of the curriculum. Presenting the design approaches as they appear in the flow of the curriculum will help the reader understand how each addressed the design challenge of creating demand for unfamiliar content.

**The “Unpack the Task” Design Approach.** In Lesson 2, students start to devise a means to measure the energy in food. We needed to create the student’s demand to learn unfamiliar content about energy forms and interconversions. We found that we used the design approach “Unpack the Task” to do this. We noticed that we asked students what they wanted to know to make progress on measuring the energy in food. When we used this design approach to ask students to unpack the task of devising a way to measure the energy in food, students would suggest the need to first figure out where to look inside the food for...
Figure 2. We used three design approaches (“Unpack the Task,” “Highlight an Incongruity,” and “Try to Apply”) to create a demand to learn unfamiliar content within the first piece of the performance project.

We found a different curriculum design strategy in the literature consistent with this design approach. In particular, students might generate their own subquestions to a driving question as a way to create a demand to learn new science content (Singer et al., 2000). Singer et al. mention employing subquestions of the driving question (including ones that students generate on their own) as a design approach to help students create the demand to learn specific science content to do the project. These might be subquestions to “what is the quality of air in my community?” like “what is air?” and further “are air and oxygen the same?”

The “Highlight an Incongruity” Design Approach. In Lesson 3, we identified “Highlight an Incongruity” as a different design approach for creating the learner’s demand for
unfamiliar content about chemical change, in particular oxidation. We noticed we used a very different approach to get the students to anticipate a utility for this unfamiliar content: we highlighted a seeming incongruity between students’ earlier progress on measuring the energy in food and what a student might reasonably believe. Having determined that the energy in food can be interconverted via burning into heat that can be measured, students were asked whether this heat energy from burning food could really be the same as the energy that the food gives them when they eat it. Students would say that there are no flames in their bodies, and this incongruity would cause students to call into question earlier progress about measuring the energy in food via burning. Students could then do a laboratory to resolve the incongruity, wherein they used a gaseous oxygen sensor to measure the oxygen concentration of the air in a jar before and after burning a tortilla chip and similarly measure the oxygen concentration in the ambient air compared to the air they exhaled. Students could use this unfamiliar content about oxidation to resolve the incongruity. Students could construct the new knowledge, that oxygen is needed and consumed during both combustion and respiration, and then apply this content to decide that the energy hidden in food and the energy food adds to the body’s energy stores are one and the same, and that it does make sense to measure the heat from burning food to measure how much energy the food adds to the body’s stores.

This design approach of “Highlight an Incongruity” should create the student’s demand to learn about the role of oxygen in combustion and respiration, anticipating that learning this unfamiliar content would help eliminate any doubt about how to proceed to measure the energy in food. This design approach should help students experience a demand to learn the unfamiliar content to resolve the incongruity about how to continue with the performance, this demand to learn the content being necessary to build its meaningful understanding. (See the second row of Figure 2.)

We were again able to find another curriculum design strategy in the literature consistent with this design approach, namely discrepant events designed to create conceptual conflict and the need for conceptual change (Nussbaum & Novick, 1982). The main difference is in how the incongruity calls into question earlier progress on the performance project, which would not necessarily be the case for all discrepant events.

**The “Try to Apply” Design Approach.** In Lesson 4, we identified the design approach “Try to Apply” as the last way we resolved the challenge of creating the learner’s demand for unfamiliar content, in this case properties of matter (in particular, specific heat). We saw how we asked students to try to apply the content they had already learned to measure the heat from burning food, but their content was insufficient to reliably measure heat, and students would fail at this task. Students could then do the heat collector laboratory, learning unfamiliar content about specific heat, anticipating that learning it should help them measure heat reliably, the final step in measuring the energy in food. Students had already made progress on measuring the energy they take in. They knew where to find the energy in food, how to interconvert the energy in food into heat via burning, and that they could measure this heat to determine how much energy food gives us. Students could collect the heat from a burning tortilla chip, but given how students defined heat (usually incorrectly, as the temperature of the water in the heat collector) their answers would vary with the initial temperature or the amount of water. Students could continue the heat collector laboratory to learn about specific heat, anticipating that this unfamiliar content should help them find a way to reliably measure heat independent of the initial temperature or the amount of water. They could construct a new understanding of specific heat and then apply this new content to measure, in calories, the heat from burning food as a measure of how much energy one gets from food, i.e., students could do direct calorimetry.
With the design approach “Try to Apply,” students would anticipate that learning unfamiliar content about specific heat would help them establish a reliable way to measure heat to determine the energy in food, even though they could not anticipate the ultimate utility of this content for measuring heat in calories. Trying to apply, but failing, should help the student demand to learn unfamiliar content to fix their failure, this demand being necessary to support the meaningful understanding of the content. (See the third row of Figure 2.)

The “Try to Apply” design approach encompasses Schank’s (1999) notion of learning being driven by failure. It is also similar to how Barron et al. (1998) helped students learn unfamiliar content through a video “toolbox” of resources indexed by specific obstacles that students might encounter. Students used the indexing in a just-in-time fashion to learn the content most useful in overcoming a given sticking point.

**Another Use of the “Unpack the Task” Design Approach.** The general nature of these design approaches is emphasized by how we found each being used twice in *J. Bio*. We found that we used the “Unpack the Task” design approach again in Lesson 5, wherein students are now trying to devise a way to measure the energy used up doing the body’s activities. We noticed that we again asked students what they wanted to know to make progress on making this measurement. We did this as a way to create the student’s demand to learn unfamiliar content about the levels of biological organization (organs, tissues, and cells). Students would suggest the need to figure out where to look for the work the body does as a reasonable first step toward measuring the energy it uses up. Students could then do a microscope investigation that asked whether there is something smaller in muscle that does the shortening of muscle contraction, i.e., muscle tissue and ultimately muscle cells. They would expect learning this unfamiliar content would help them answer this question about where the body’s work is done, a first step toward measuring the energy the body uses up. Doing this investigation, students could construct an understanding that it is the body’s cells that do the body’s work and then apply this content to figure out that it is the individual cells that are using up the body’s energy stores.

Just as before, “Unpack the Task” should help students anticipate a utility for learning the unfamiliar content about the levels of biological organization to make incremental progress on the performance, even though students could not anticipate using this content to add up the energy all the body’s working cells use up. Creating the student’s demand, this design approach should help build a meaningful understanding of this unfamiliar content. (See the first row of Figure 3.)

**Another Use of the “Highlight an Incongruity” Design Approach.** We used the “Highlight an Incongruity” design approach again in Lesson 6 to create the learner’s demand to learn unfamiliar content about how the body systems work and work together. We noticed how we highlighted a seeming incongruity between the body’s cells doing the work and the difficulty of getting oxygen and foodstuffs from the body’s energy stores to and into every one of the body’s working cells. Having determined that it is in the body’s cells where foodstuffs are combined with oxygen to power all the body’s work, and thus it is in the body’s cells where energy stores are used up, students were asked if foodstuffs and oxygen can really get to and into all the body’s working cells. Highlighting this incongruity would cause students to call into question their earlier idea about measuring the energy cells use up to measure the energy the body uses up. Students could then access anatomical charts and other resources to find a path for oxygen and foodstuffs from outside the body to and into a working muscle cell in the foot. Students would expect to use this unfamiliar content about
**PERFORMANCE PROJECT PIECE 2:** Devise a means by which to measure the energy used up doing activities and make this measurement for activity choices.

**Lesson 5:** Where in our bodies are energy stores used up?

<table>
<thead>
<tr>
<th>Design Approach: UNPACK THE TASK</th>
<th>Content Demanded and Constructed</th>
<th>Apply Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create the Demand to learn unfamiliar content linked to doing the performance by asking: Where in our bodies are the energy stores used up doing work, in order to measure energy used up there?</td>
<td>- levels of organization: organs, tissues, and cells</td>
<td>Decide that cells are where foodstuffs and oxygen are combined and cells are where the energy from our stores is used up doing the work of our daily activities. We need to measure the energy stores used up at all working cells.</td>
</tr>
</tbody>
</table>

**Lesson 6:** How do ingredients for energy use get to and into every working cell?

<table>
<thead>
<tr>
<th>Design Approach: HIGHLIGHT AN INCONGRUITY</th>
<th>Additional Content Demanded and Constructed</th>
<th>Apply Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create the Demand to learn unfamiliar content linked to doing the performance by asking: Can foodstuffs and oxygen really get from outside the body to and into all the body’s working cells?</td>
<td>- how organs work together in systems and how body systems work together</td>
<td>Decide that indeed makes sense to measure the energy stores used up by all working cells.</td>
</tr>
</tbody>
</table>

**Lesson 7:** What can we measure, and how, to measure the energy used up by all working cells?

<table>
<thead>
<tr>
<th>Design Approach: TRY TO APPLY</th>
<th>Additional Content Demanded and Constructed</th>
<th>Apply Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create the Demand to learn unfamiliar content linked to doing the performance by trying (but failing) to find a do-able way to measure the energy stores used up by all cells.</td>
<td>- cellular respiration</td>
<td>Re-invent indirect calorimetry. Measure in calories the energy used up doing activity choices.</td>
</tr>
</tbody>
</table>

**Figure 3.** We again used the same three design approaches (“Unpack the Task,” “Highlight an Incongruity,” and “Try to Apply”) to create a demand to learn additional unfamiliar content within the second piece of the performance project.

...the body systems to resolve the incongruity. Doing this activity, students could construct a knowledge of how the respiratory and digestive systems work and work together with the circulatory system to deliver foodstuffs and oxygen to and into every working cell in the body via capillary diffusion. Students could then apply this content to determine that it does indeed make sense that cells themselves use up the energy from the body’s stores during activities.

As before, the “Highlight an Incongruity” design approach should create students’ demand to learn unfamiliar content about the body’s systems to resolve the incongruity and eliminate doubt about totaling the energy used up by cells to measure the energy used up by the body. This design approach should help students anticipate a use for the content in resolving the incongruity about how to proceed on the performance, even though the students will ultimately apply this content to do indirect calorimetry. (See the second row of Figure 3.)

**Another Use of the “Try to Apply” Design Approach.** In Lesson 7, we found that we again used the “Try to Apply” design approach as students finished figuring out how to measure the energy used up from the body’s stores doing activities. Students tried to apply the content they had learned thus far by adding up the energy stores used up by all the body’s working cells to measure the body’s energy use, but this approach was not feasible. Students also suggested doing direct calorimetry with humans, but doing this in the classroom was impractical. Applying what content they already knew, students would fail at measuring the energy used up. Students could then do an activity to learn about balancing chemical reactions and the specific ratios in which foodstuffs and oxygen react, i.e., an equation for cellular respiration, anticipating that this unfamiliar content about cellular respiration would help them find a practical way to measure the energy the body uses up doing activities.

*Science Education*
Students could then construct this new content about cellular respiration and ultimately apply it to do indirect calorimetry, to measure in calories how much energy one uses up from one’s stores. The amount of oxygen a single working cell uses up is an indirect measure of the energy it uses up from the body’s stores and total oxygen consumption measured at the mouth is an indirect measure of all the energy used up by all the body’s working cells.

The “Try to Apply” design approach should again help the student anticipate a utility for learning unfamiliar cellular respiration content to fix their failure and devise a practical way to measure the energy the body uses up, even though they could not anticipate at the outset using this content to do indirect calorimetry. Regardless, the demand this design approach creates should help students build a meaningful understanding of this unfamiliar content. (See the third row of Figure 3.)

To summarize Design Challenge 1, we were able to identify the curriculum design challenge that resulted from following the guiding principle of creating demand for the science content in a pPBSc: the need to help the student generate a demand to learn the science content to do the performance when the student had no prior familiarity with the science content or what it might be useful for. Then, we were able to identify three different curriculum design approaches used in *I, Bio* to address this challenge: “Unpack the Task,” “Highlight an Incongruity,” and “Try to Apply.” These design approaches provided a different way to get the student to anticipate using unfamiliar content to make progress on the performance. These approaches provided ways to design pPBSc to create in the student an a priori demand to learn the unfamiliar content based on an expected usefulness, thus creating a context in memory for integrating the new science content before constructing it, necessary for building meaningful understanding.

These design approaches provide three specific ways to do what Beissner and Reigeluth (1994) described as connecting content learning with using the learning to complete a performance by “integrating” the learning of the content with the performance work. As such, we believe these three approaches to designing pPBSc to create demand should support a meaningful understanding of content. It is also worth noting that all three of these design approaches leverage the demands of the performance project to pique students’ curiosity to pursue what in essence become investigation subprojects, but in such a way that the students’ motivation to do these investigations lies beyond mere curiosity, but in the practical utility they anticipate for what is to be learned. Thus, these design approaches help draw the doing of the performance and the learning from an investigation closer together.

**Design Challenge 2: Applying All the Content**

In the work designing *I, Bio*, two additional challenges were identified that made it difficult to satisfy the guiding principle that the learner apply the science in a pPBSc. The first was ensuring that *all* the target science content is ultimately useful and used by the learner to complete the performance around which the curriculum is designed. It is not necessarily the case that doing a performance will require the application of all the science content given what the performance asks of the learner. This challenge is likely to be unresolved in the design of a pPBSc if the theme of the performance project makes it seem like completing it would necessarily require applying all the target content. As described above, it seemed like completing the global warming and model rocket performances would build students’ meaningful understanding of underlying science content, but this was not found to be the case.

Although the target science content may be relevant to doing a performance, students may find its application unnecessary. For example, while the best briefings about the earth’s
warming would require applying the science to explain findings or justify conclusions, learners may be able to complete a briefing that does not apply this science, given how the performance was defined. It might never have been the case that learners found themselves impeded in making progress on these performances for lack of understanding and applying the science content given how these performances were defined. Some learning of the content may still occur, but constructing the content without applying it to the performance increases the likelihood that the eventual understanding of this content will be inert as opposed to meaningful. The challenge for the pPBSc designer is greatest when certain science is very thematically relevant to a performance but inessential to apply to complete the performance. Ultimately, the choice of performance project will determine whether all the content needs to be applied. The design approach described next is a way to design pPBSc to ensure this.

**The “Analyze and Refocus or Identify an Alternative Performance” Design Approach.**

The first part of the design approach we used in I, *Bio* for addressing this design challenge of applying all the content was to conduct a rudimentary conceptual analysis of the performance (Reigeluth & Moore, 1999) to determine the scope of the content used to do the *I, Bio* performance. We initially expected that a biomedical engineering-based design performance would require students to apply all the target biology and physical science content. We expected building a medical device like an artificial heart, a performance that would be exciting for students, would require the application of all the target content. Engineering-based design performances had been shown to support learning in K-12 science classrooms (Baumgartner & Reiser, 1998; Holbrook et al., 2001; Kolodner, Crismond, Gray, Holbrook, & Puntambekar, 1998; Puntambekar & Kolodner, 2005). However, conducting a rudimentary conceptual analysis, we were able to recognize that this project would not require students to apply as much of the biology content in particular as we aimed to teach, even though the project was strongly thematically linked to this content. Breaking down the performance of building a heart, the design team realized that completing the parts of the performance would require students to apply the principles of mechanics, forces and motion, and electricity and circuits, respectively, but that they would apply little of the biology content or in particular the content about energy interconversion in living organisms. This demonstrates the subtle nature of this design challenge. What makes this design challenge hard to recognize, let alone resolve, is how performance projects can be compelling in their strong thematic relationship to the science content. These performances may also be ones in which students are enthusiastic to engage, but ultimately the content may be merely relevant but not necessary to apply to complete the performance.

The next aspect of the design approach was to see whether, without changing the essential performance, changing the focus of the performance in small ways could result in big changes in the content that must be applied to complete it. We found support in the literature for this aspect of the design approach: Petrosino was able to deepen students’ understanding of the target content when he changed the model rocket challenge in a small way, from “build the rocket and see how high it will go” to asking “would rockets go higher if they were painted or not, if they had 3 or 4 fins, or if they had a rounded or pointed nose cone” (Petrosino, 1998). When the performance project was refocused to require evaluating what makes rockets better or worse, students had to apply the target content about experimentation and measurement and their understanding improved. In our case, no small change to the focus of this performance such as building a different medical device than an artificial heart seemed to rectify the problem that not all of the target science content would be applied.
At this point in designing *I, Bio*, our rudimentary conceptual analysis having revealed that the content necessary to apply to complete a given performance was far less than that which we were targeting, and not being able to rectify this by any small changes to the focus of the performance, we found that we needed to be willing to reject the initial choice of performance and search broadly for alternative performances and reanalyze. This is the final aspect of the design approach for resolving the design challenge of applying all the content.

We searched broadly for other performances that might better require the application of all the target content. We analyzed a very different sort of performance project, one suggested by a teacher on the design team, that of students devising measurements to design their school lunch and activity choices to meet their bodies’ energy needs. As described above, this performance consists of making a measurement of the energy that food adds to students’ bodies and making a measurement of the energy activities use up from students’ bodies and designing school lunch and activity choices to balance these two measurements. We were surprised to find that this very different performance actually did a better job of requiring students to apply all the target content. Reinventing how to measure the energy in food via direct calorimetry requires applying content about the nature of energy in food, energy interconversions via combustion, and heat energy and how to measure it. Reinventing how to measure the energy used up doing activities via indirect calorimetry requires applying additional content about the levels of biological organization, how the body’s organs and organ systems work together to provide cells with food and oxygen for energy, and cellular respiration. All told, this entirely new choice of performance project had the unexpected result of requiring the student to apply all the target content.

We used the design approach we have described to arrive at the final *I, Bio* performance project (in which students devise a way to measure the energy in food and a way to measure the energy used up doing activities and redesign their school lunch and activity choices to make these two numbers equal). Using this design approach, we ensured that all the target science content would be applied in doing the performance, and thus the potential exists to build a meaningful understanding of all this science content. (See Figure 1 for a summary of all the content that is applied completing this performance project. See Figures 2 and 3 for the particulars of how each content idea is applied.)

This approach to designing pPBSc to address the challenge that all the science content is applied to support its meaningful understanding encompasses what Singer et al. (2000) mention as assessing their driving question to determine the extent of the content one must apply to answer it. Edelson (2001) also describes selecting the project to ensure that all the target content is applied. Our design approach details how one might go about doing this.

**Design Challenge 3: Applying All the Content in Time**

We identified a second challenge to “applying” the science content. Given the complex nature of the kinds of performances around which pPBSc are designed—performances that can take on the order of 10 weeks or more in the classroom to complete and that include many science content ideas—even if all the content is applied in the end, meaningful understanding might be reduced if too much time elapses between when the content was first demanded (and constructed) and when it is applied to do the performance. A pPBSc designed employing Gagné, Briggs, and Wager’s hierarchical approach to sequencing content (Gagné et al., 1992) might teach all the prerequisite content first from simple to complex and only then ask the learner to apply all this content to complete the performance as a “capstone” at the very end (Reigeluth & Moore, 1999). While a lengthy pPBSc designed this way would address the challenge of applying all the content, the content
initially learned would not be applied to the performance until the very end. If the content were forgotten by that time, promoting its meaningful understanding would be preempted. Time is a factor. With this “capstone” approach, cognitive load theory (Sweller, 1988) suggests that if the learner is not given the timely opportunity to apply the content and thus integrate the new information with prior knowledge in long-term memory, the learner will experience a heavy working memory load, detrimental to the meaningful understanding we are trying to promote (Kirschner, Sweller, & Clark, 2006). The working memory cognitive load can grow too great and some content can be forgotten before applying it to promote its meaningful understanding (Reigeluth & Stein, 1983). Thus, a design challenge unique to pPBSc, secondary to applying all of the science content, is that of applying all the content in time.

The “Piece Apart the Performance” Design Approach. Addressing this design challenge in I, Bio, we found we used a design approach of completing the performance project in pieces. After introducing the performance project, which requires balancing the energy in food with the energy used up doing activities to design school lunch and activity choices (Lesson 1), we noticed that we split the performance project into two pieces. Each of these two pieces focused on students devising the means by which to make one of these energy measurements and then doing so. The first piece of the performance project (Lessons 2–4) focused only on students devising direct calorimetry as a means by which to measure the calories in food and then calculating the calories they consume in their school lunch choices. Students would complete just this first piece of the performance project and in doing so apply just the content related to energy, energy interconversions, and specific heat to devise direct calorimetry and use it to measure the energy in food and in their school lunch choices. (See Figure 2.) This first piece of the performance project would take about 3–4 weeks to complete, but in piecing apart the project in this way, the student would apply a subset of all the target content just to devise direct calorimetry, well in advance of applying all the content to complete the entire 10-week performance.

The second piece of the performance project (Lessons 5–7) focused on students devising indirect calorimetry as a means by which to measure the calories it takes to do work and then calculating the calories used up doing daily activities. Students would apply the additional content related to how the body’s organ systems work together to meet all cells’ needs and cellular respiration to do this second piece of the performance project and devise and use indirect calorimetry to measure the energy used up doing their activity choices. (See Figure 3.) This second piece of the performance would take about 5 or 6 more weeks. The important thing to notice is that the earliest learned content was already applied previously in completing the first piece of the performance project, well in advance of when this earliest learned content is applied again, along with additional content to complete the second piece of the performance project.

Had the I, Bio performance project not been divided into pieces, completing the performance would still have required applying all the content. However, the learner would have had to wait until the end of the curriculum to do so (as one might do with a “capstone” project). This would have resulted in the content learned near the beginning of the curriculum not being retained to be organized by the time it was applied. “Piecing Apart the Performance” into even two content-discrete pieces would help apply all the content in time, supporting organizing it, and thus meaningful understanding, before students would forget the content due to their working memory cognitive load growing too great.

This design approach encompasses design strategies of solving a problem by breaking it into parts (Gagné et al., 1992), using ordered task decompositions to provide structure for
complex tasks (Quintana et al., 2004), or giving apprentices only parts of a task at a time on which to work (Lave & Wenger, 1991).

To summarize the last two subsections, in designing the pPBSc I, Bio to satisfy the guiding principle of applying the science content, as necessary to build meaningful understanding, we found the two design challenges of applying all the science content and doing so in time. Seeing these design challenges clearly, we were able to generalize ways of addressing them, drawing from I, Bio and design strategies in the literature: using the “Analyze and Refocus or Identify an Alternative Performance” and “Piece Apart the Performance” design approaches, respectively.

In this first part of the paper, we were able to identify specific curriculum design challenges that would stand in the way of the meaningful understanding of science content in pPBSc. We expect the three design challenges we identified to be an incomplete list, although we do expect them to apply to pPBSc other than I, Bio. We were subsequently able to generalize approaches for designing pPBSc to address each of these design challenges, and as such try to ensure meaningful understanding of content when learning with such curricula. Again, we expect our design approaches for resolving each challenge to be an incomplete list, but at the same time, these design approaches should be generally applicable to resolving these design challenges in pPBSc. One might expect to identify additional design challenges and other approaches examining other pPBSc. The identification of these design challenges and related design approaches for tackling them, as well as the detailed illustration of both the challenges and their approaches in the context of I, Bio, should help other curriculum designers.

In the next two parts of the paper, we make preliminary progress toward gathering empirical evidence that designing pPBSc with these approaches is consistent with students improving their meaningful understanding of the science content underlying the performance. In the Method and Results sections, we present a study of learning by students using the I, Bio curriculum, designed with our approaches for helping students to create the demand to learn the unfamiliar science content and apply all the science content and do so in a timely fashion. A gain in students’ meaningful understanding would suggest that employing these approaches to designing pPBSc can mitigate the tension between doing the performance and learning the science, improving our confidence that we can design pPBSc to more reliably promote students’ usable understanding of standards-based science content.

PART III: METHOD—MEANINGFUL UNDERSTANDING OWING TO THE DESIGN APPROACHES

The I, Bio performance project was that of designing school lunch and activity choices to balance a measurement of calories consumed in school lunch choices with a measurement of calories used up doing daily activities. As discussed above, the curriculum was designed with approaches to try to support students learning the underlying science content while making progress on the performance. Would we find that students using a curriculum designed with these approaches would gain in their meaningful understanding of the science content, or would our attentions to the design of the pPBSc do little to ensure this? During the 2002–2003 school year, the I, Bio curriculum was enacted by 12 teachers in 37 classrooms with 652 sixth-, seventh-, and eighth-grade students. Students’ grade levels varied depending on the grade teachers taught this science content in their school. The 12 participating teachers were selected due to their interest in using this pPBSc with their students. The student demographics were split almost evenly between Hispanic or
Latino (32.5%), White (32.2%), and African American (29.9%) students. Many (51.1%) of these students fell into the state’s low-income category, which includes students eligible to receive free or reduced-price lunches, or that come from families on public aid. (See Table 1 for demographic information on participating students.) We met with the 12 teachers for professional development 3 hours weekly during their enactment. It was in this context that we emphasized the curriculum activities designed to create the demand for and support the application of the science content. During these weekly professional development sessions, teachers discussed their students’ progress, confirming that their students completed the performance project and its activities.

We designed a pre- and posttest for use during the enactment. The test was designed with items at three levels of cognitive difficulty based on a revised Bloom’s taxonomy (Anderson et al., 2001). Level 1 items asked students to remember, providing evidence that they learned the content. Level 2 items asked students to understand and apply the content, but in circumstances similar to those in which they learned these ideas in the curriculum, providing evidence of a deeper but not yet meaningful understanding. Level 3 items asked students to apply the content in a new context and furthermore to analyze or evaluate, indicative of meaningful understanding as we have described it above. The test was designed to have a few items at each level for each of two clusters of the standards-based content we targeted, Cluster 1 around “cells and body systems” content and Cluster 2 around “biological energy” content. Cluster 1 focuses on students’ understanding of levels of organization: organs, tissues, and cells, how organs work together in body systems, and how three body systems work together in getting energy from food. Cluster 2 focuses on students’ understanding of energy forms, energy interconversion including oxidation and cellular respiration, and measuring energy including heat. By designing the test with items at these three levels of cognitive difficulty for each content cluster, we would be able to see the extent to which students achieved a meaningful understanding of the content, for which we would see gains in Level 3 items, beyond the extent to which they improved their general understanding of the content, for which we would see gains in Levels 1 and 2 items. The approaches employed in the design of I, Bio should support gains in meaningful
understanding in all the science content, which would appear as gains in Level 3 items in both content clusters.

Both the content clusters and the levels of cognitive difficulty were a priori categorizations based on written rubrics. These rubrics were used to develop items that met the criteria of being about particular content and at a particular level of cognitive difficulty. Three members of the research team used the rubrics to review all items’ assignments to a particular cluster and level. Item categorization or the items themselves were changed until consensus was reached on their categorization. In this way, the levels and the clusters are maintained to be conceptually valid.

Given the constraints on classroom time available for testing, the pre- and posttest had only 17 multiple-choice and short, constructed-response items, each worth a single point. Most of the Level 3 items were short, constructed responses. For example, one Level 3 item for content Cluster 1 on cells and body systems read,

Your heart and stomach are both made up of cells. The heart pumps blood while the stomach makes acid. Do you think the cells in your heart will be the same as or different from those found in the stomach? Explain your answer.

The stomach and its cells were not studied in the curriculum, but a student who had gained a meaningful understanding of levels of biological organization should be able to apply this understanding to describe distinct functions that cells in the stomach must be specialized to do (secrete) compared to cells in the heart (contract). Such constructed-response items were graded with a rubric. In this case, a response like “They would be different because different substances affect the cells differently” does not reflect any understanding that the sum of all cells’ work results in an organ’s distinct function, while responses like “The cells will be different because the stomach and the heart are two different organs with two different functions” or “Different because they function two different ways and each of them does different work in the body” show some understanding that the (muscle) cells in the heart are differently specialized to perform a different function than the (secretory) cells in the stomach.

In contrast, most of the Level 1 items, designed to see whether students remembered ideas learned in the curriculum, were multiple choice. A Level 1 item for the cells and body systems content cluster was, “Which is the most basic unit of living things?” Students had to choose the correct answer, cells, from among several distractors: tissues, organs, and bones.

In general, Level 2 items that asked students to apply a concept in a context similar to that in which they learned the concept were either multiple-choice or constructed-response items. One constructed-response Level 2 item for the cells and body systems content cluster, adapted from a publicly reported Michigan Educational Assessment Program item, was as follows:

Mindy has been training to run in a marathon. She often takes her dog Rusty with her when she goes running. Finally, race day arrives and Mindy registers for the race. After the race, Mindy eats a large meal. Mindy’s body needs and uses the nutrients from the meal. Name two body systems involved in getting nutrients from the food that Mindy eats and describe how the systems’ primary functions work together to do so.

For the constructed-response Level 2 items, students were again graded with a rubric. In this case, a response such as, “Blood and digestion. The blood carries the energy and the digestion takes away any other energy,” mentions the correct body systems but exhibits some wrong ideas about the digestive system and does not address how the systems interact.
while a response like “Digestive system and circulatory (blood). The small intestine takes out all the nutrients and energy so all that is left is waste and the blood takes energy to muscles and places where needed” correctly applies the specific functions and interactions of each system to the situation. The same items were used for both pre- and posttests. A single scorer with experience using the scoring rubric graded all the constructed-response items for all the tests. The scorer was not involved in item development and was blind to all items’ cluster and level assignments.

While we proceed to analyze our levels and clusters on the basis of their conceptual validity, even while the total number of items is small, we also report the Gutmann split-half measure of reliability of the levels and clusters. This measure is used since each level and cluster is a mix of multiple-choice and constructed-response items with unequal variances. For Levels 1, 2, and 3, reliabilities on the posttest were 0.4, 0.5, and 0.3, respectively. For Clusters 1 and 2, reliabilities were 0.6 and 0.4, respectively.

PART IV: RESULTS—STUDENTS’ MEANINGFUL UNDERSTANDING

Overall Gains in Students’ Content Understanding

We used a paired *t*-test to evaluate students’ learning gains for the content targeted by the *I, Bio* curriculum. The shift from 6.1 to 8.4 of 17 total points was modest but statistically significant (*t* = 25.2, (651), *p* < .001). Another way to examine gains such that they can be compared on a standard scale is to present effect sizes. These were calculated as the difference between the post- and pre-mean divided by the pooled standard deviation, i.e., the square root of the sum of the square of the standard deviation at pre and the square of the standard deviation at post. For the test as whole, the effect size (ES) shift was ES = 0.85. That said, students had some prior knowledge about human biology as evinced by their overall pretest score of 6.1, greater than the approximately 3 point reward expected with random guessing; however, this correct prior knowledge was primarily reflected as correct answers to Level 1 items from the cells and body systems content cluster on the pretest. Before experiencing the curriculum, students knew cells to be the most basic unit of living things and were able to state the correct function of blood.

We can compare effect size gains against the normal achievement gains children make from one year to the next to provide a frame of reference for interpreting the meaning of the effect size of the intervention. We draw on recent work on average annual effect size gains for science from nationally normed tests. The yearly gain in science expressed in terms of effect size in the transition between any two middle grades is on average 0.23 (Bloom, Hill, Black, & Lipsey, 2009). We can compare the effect size gains we find in our study for the test as a whole against this anticipated growth in science for any middle grades year. Student gains on the test as a whole were 3.7 times greater than this value.

Gains in Students’ Content Understanding by Content Cluster

Effect sizes also allow us to compare the relative magnitudes of gains, in this case the individual content cluster gains. For the cells and body systems content Cluster 1, the effect size of students’ gain was ES = 0.60 [*t* = 15.4, (651), *p* < .001]. For content Cluster 2 on biological energy, the effect size of students’ gain was ES = 0.85 [*t* = 17.4, (651), *p* < .001]. (See Table 2 for a summary of these results.) The tests used in the Bloom et al. study cover a wider range of science subject matter compared to our test item subsets, which narrowly focus on just cells and body systems and biological energy subject matter. Middle school students on average may improve more in some science content areas than others.
## TABLE 2
Pre- to Posttest Shift by Content Cluster

<table>
<thead>
<tr>
<th>Content Cluster</th>
<th>Maximum Score</th>
<th>N</th>
<th>Pretest Mean</th>
<th>Posttest Mean</th>
<th>Gain</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cells and body systems</td>
<td>8</td>
<td>652</td>
<td>2.9</td>
<td>3.8</td>
<td>+0.9* (+31%)</td>
<td>0.6</td>
</tr>
<tr>
<td>Biological energy</td>
<td>9</td>
<td>652</td>
<td>3.2</td>
<td>4.6</td>
<td>+1.4* (+44%)</td>
<td>0.85</td>
</tr>
</tbody>
</table>

*p < .001.

However, without more directly comparable cells and body systems- or biological energy-specific studies, the Bloom et al. study provides an average value for any science topic against which to compare our findings. For the cells and body systems content Cluster 1, the effect size of students’ gain was 2.6 times greater than the Bloom et al. value. For content Cluster 2 on biological energy, the effect size of students’ gain was 3.7 times greater.

All these results show that students’ understanding improved overall and for each content cluster. However, these results do not yet reveal the extent to which the carefully designed pPBSc was able to get students to learn content at high levels of cognitive difficulty, that is to say build their meaningful understanding of the science content. The analysis that follows addresses the extent to which the curriculum brought about students’ meaningful understanding of the content. To do this, we focus on the gains from the Level 3 cognitive difficulty items compared to the gains from cognitive difficulty Levels 1 and 2 across the test as a whole and also within each content cluster. Effect sizes will help us assess the relative magnitudes of the content cluster gains from pre- to posttest as a function of the cognitive difficulty of the items.

### Overall Gains in Students’ Meaningful Understanding of Content

Across both content clusters, the gains in meaningful understanding at Level 3 cognitive difficulty were of an $ES = 0.51$. This is an improvement in meaningful understanding across both content clusters, and this result speaks favorably to using our curriculum design approaches to resolve challenges to meaningful understanding with pPBSc. Gains in meaningful understanding of science content were supported. The tests used in the Bloom et al. study have a mix of items across cognitive difficulty levels such that the gain a middle school student makes on these tests would be expected to be greater than the gain for just the most cognitively difficult items. As such, the Bloom et al. study sets a high bar against which to compare the gain in meaningful understanding in this study. The gain in meaningful understanding at Level 3 cognitive difficulty in this study was 2.2 times the Bloom et al. value. Student gain at cognitive difficulty Level 2 was of an $ES = 0.98$, 1.9 times our Level 3 finding. This shows more improvement in students’ ability to apply content in situations similar to those in which the content was learned in the curriculum. Student gain at cognitive difficulty Level 1 was of an $ES = 0.35$, less than our Level 3 finding by a factor of 0.7. This gain still shows a positive impact on student learning at the lowest level of cognitive difficulty, but it is surprising to see a greater gain at cognitive difficulty Level 3 than at Level 1. This may be due to a ceiling effect on the Level 1 cells and body systems items as mentioned above. (See Table 3 for a summary of these results.) Level 2 and Level 1 gains speak less to our design approaches than to the care with which we provided opportunities in the curriculum to construct all the science content ideas.
### TABLE 3
Pre- to Posttest Shift by Level of Cognitive Difficulty

<table>
<thead>
<tr>
<th>Cognitive Difficulty Level</th>
<th>Maximum Score</th>
<th>N</th>
<th>Pretest Mean</th>
<th>Posttest Mean</th>
<th>Gain</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (remember)</td>
<td>4</td>
<td>652</td>
<td>2.6</td>
<td>3.0</td>
<td>+0.4∗ (+15%)</td>
<td>0.35</td>
</tr>
<tr>
<td>2 (apply in same context)</td>
<td>10</td>
<td>652</td>
<td>2.6</td>
<td>4.4</td>
<td>+1.8∗ (+69%)</td>
<td>0.98</td>
</tr>
<tr>
<td>3 (meaningful understanding)</td>
<td>3</td>
<td>652</td>
<td>0.8</td>
<td>1.3</td>
<td>+0.5∗ (+62%)</td>
<td>0.51</td>
</tr>
</tbody>
</table>

∗p < .001.

### Gains in Students’ Meaningful Understanding by Content Cluster

Within the cells and body systems content cluster, gains came from Level 3 items with $ES = 0.50$ [$t = 12.1, (651), p < .001$]. Gains came from Level 2 and 1 items with $ES = 0.54$ [$t = 10.6, (651), p < .001$] and $ES = 0.28$ [$t = 6.2, (651), p < .001$], respectively. (See Table 4 for a summary of these results).

This shift in meaningful understanding of cells and body systems can be illustrated by the difference between a student’s pre- and postresponses to a Level 3 item. To answer the following, “Immediately before and after running a 50-meter race, your pulse and breathing rates are taken. What changes do you expect to find and why?,” a student has to apply to the new context of exercise what they learned in the curriculum about the circulatory and respiratory systems and how they interact. A student responded this way on the pretest: “Because you are running and are short of breath. Since you are short of breath, your heart is pumping harder and faster to keep the flow of blood.” Such a response does not show an integrated understanding of how the respiratory system works harder during exercise to better oxygenate the blood, which the circulatory system works harder to deliver. On the posttest, this same student’s response—“This is because you need more oxygen to convert your stored energy. And your pulse quickens because the heart is pumping hard to direct blood with oxygen to your cells.”—shows an improved understanding.

We were very encouraged to find that students gained in their meaningful understanding of cells and body systems content at cognitive difficulty Level 3, $ES = 0.5$. This finding is consistent with how we tried to use our design approaches to resolve challenges to meaningful understanding of standards-based cells and body systems content while doing the performance project. We also found that students gained significantly in their understanding of this content at cognitive Levels 1 and 2, which again speaks less to our design approaches than to curriculum opportunities to construct these content ideas.

Within the biological energy content cluster, gains came from the Level 3 cognitive difficulty items with $ES = 0.24$ [$t = 5.3, (651), p < .001$]. Gains came from Level 2

### TABLE 4
Within Content Cluster 1, Cells and Body Systems, Pre- to Posttest Shift by Level of Cognitive Difficulty

<table>
<thead>
<tr>
<th>Cognitive Difficulty Level</th>
<th>Maximum Score</th>
<th>N</th>
<th>Pretest Mean</th>
<th>Posttest Mean</th>
<th>Gain</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (remember)</td>
<td>3</td>
<td>652</td>
<td>2.3</td>
<td>2.5</td>
<td>+0.2∗ (+9%)</td>
<td>0.28</td>
</tr>
<tr>
<td>2 (apply in same context)</td>
<td>3</td>
<td>652</td>
<td>0.2</td>
<td>0.6</td>
<td>+0.4∗ (+200%)</td>
<td>0.54</td>
</tr>
<tr>
<td>3 (meaningful understanding)</td>
<td>2</td>
<td>652</td>
<td>0.4</td>
<td>0.8</td>
<td>+0.4∗ (+100%)</td>
<td>0.50</td>
</tr>
</tbody>
</table>

∗p < .001.
and 1 items with $ES = 0.85$ [$t = 20.6, (651), p < .001$] and $ES = 0.28$ [$t = 5.4, (651), p < .001$], respectively. (See Table 5 for a summary of these results.) An example of how we assessed for meaningful understanding about biological energy was to ask a short constructed-response question about how a person’s energy use would change as a function of their body mass. To answer this question, students had to employ their understanding of biological energy in a context different from any they encountered in the curriculum.

Biological energy gains in Level 3 that assessed for meaningful understanding were significant at $ES = 0.24$. Achieving at this most cognitively difficult level provides further evidence that the $I, Bio$ pPBSc was designed in such a way that it supported gains in meaningful understanding of the target science. This is true even while Level 3 gains for cells and body systems content were about twice that of Level 3 gains for biological energy content.

The largest gains in content Cluster 2 on biological energy were from the Level 2 items. The gains in Level 1 biological energy items were also significant. Again, student learning gains in biological energy at cognitive difficulty Levels 1 and 2 may speak less to our design approaches than to opportunities we provided in the curriculum to construct these content ideas.

**PART V: DISCUSSION**

Students’ gains in meaningful understanding of content using the $I, Bio$ pPBSc—crafted using the design approaches that are the subject of the first part of this paper—is preliminary evidence that pPBSc can be designed in such a way that students can do the performance but at the same time develop a meaningful understanding of the related science content. We have shed light on some of these design approaches and have illustrated them in the context of their use in the $I, Bio$ curriculum, earmarked by the curriculum design challenges each addresses, the better for other curriculum designers to employ them. We mined our $I, Bio$ curriculum design and the literature to identify these design approaches.

The student learning outcomes are an indication of the relevance to science content learning of designing pPBSc with these approaches. The validity of our findings would be improved by comparing our student learning outcomes to those of students who studied a similar pPBSc designed without the specific design approaches, thus employing a control group. It would further improve the validity of our findings to vary the items between the pre- and posttests to prevent pretesting from contaminating posttest scores. Also, it is important to note that we are combining scores across students at different grade levels. One additional caveat about our findings is that we must recognize that teachers sometimes use curricula in a manner different from that which curriculum designers intend (Schneider, Krajcik,
DESIGNING PROJECT-BASED SCIENCE TO TEACH CONTENT

While in our present study all teachers received instruction about which curriculum activities addressed which design challenge, and all teachers reported doing these activities, we did not determine the extent to which teachers did these activities in a manner consistent with the intent of the curriculum design approach. Collecting this information would improve our ability to differentiate the impact of the curriculum design approaches themselves from teachers’ abilities to use them.

Another concern in designing pPBSc with our approaches is that students may experience the performances as fairly constrained. The design approaches constrain students from pursuing the project in whatever direction they see fit and as such may reduce students’ enthusiasm for doing the performance. In fact, a pPBSc designed with these approaches looks less like minimally guided instruction (Kirschner et al., 2006) and more like how Hmelo-Silver, Duncan, and Chinn (2007) describe problem-based and inquiry learning as having scaffolds that structure the complex task to reduce cognitive load. Beyond the extent to which PBS is already a guided curricular approach—different from Problem-based Learning, which can be open-ended or guided, from which it is derived—a pPBSc designed with these approaches further constrains students from pursuing the project in any direction they choose. As such, so structuring the performance might dampen the student enthusiasm that we were trying to leverage. However, it might also be the case that these design approaches strike a good balance between constraining the task enough to increase the likelihood of predictably covering the standards-based science content with reduced cognitive load (for which preliminary evidence has been presented) while at the same time maintaining a level of authentic pursuit of the performance from the perspective of the learner, and thus some of their natural enthusiasm for doing performance-oriented tasks. The extent to which this is so would have to be tested. Such a finding would support both sides of the recent debate on constructivist instruction (Tobias & Duffy, 2009) insofar as this particular constructivist-based instructional approach would succeed in promoting learning, but do so by incorporating carefully designed guidance or scaffolds.

A further limitation to using a performance PBS instructional approach in classrooms is that relatively few science content ideas are covered, albeit in depth, in the extended class time devoted to completing the performance. However, a recent empirical study (Schwartz, Sadler, Sonnert, & Tai, 2009) showed that depth of coverage, spending a month or more on at least one major science topic in high school, predicted better grades in that science in college. Providing greater depth is advocated by national science standards documents (AAAS, 1993; NRC, 1996). The recent National Research Council publication *Taking Science to School* (Committee on Science Learning Kindergarten Through Eighth Grade, 2007) also expressed the need to develop curricula that can support such opportunities for sustained engagement with the same science ideas over extended periods of time. Nevertheless, further research would be required to determine whether extended depth of coverage on fewer science ideas with pPBSc helps students learn more in later years.

It is also unclear whether there is an advantage of pPBSc designed with our approaches in building students’ meaningful understanding of science content as compared to direct or other modes of instruction, which also support students learning to transfer science content. The Klahr and Nigam study (2004) cited by Kirschner et al. (2006) shows that children who mastered the content via direct instruction were as skilled at applying their knowledge to a novel task, i.e., had meaningful understanding, as were students who mastered the content via discovery learning. In the Klahr and Nigam study, meaningful understanding was measured as students’ total number of valid critiques. But we are left to wonder if other modes of instruction could have helped students generate even more valid critiques. Is even more meaningful understanding possible? One could compare the extent
of science meaningful understanding attained with a pPBSc built with the design approaches discussed herein to that achieved with the direct instruction or discovery learning approaches explored in the Khlar and Nigam study. To conduct this experiment, these different curricular approaches would have to be designed to teach the same science content and a common measure of meaningful understanding would have to be employed. It may be found that pPBSc designed with our approaches, wherein standards-based science content learning is essential to students designing or making something in the world, educe even more meaningful understanding.

Overall, this work better positions us to address the call by Sweller, Kirschner, and Clark (2007) to design systematic experiments to test which instructional supports matter for learning and how much. We can only do this by being able to be explicit about the specific types of instruction guidance that different modes of instruction provide, regardless of what the different instructional approaches are called. Our curriculum design approaches provide an initial vocabulary of instructional supports for pPBSc that are important for content learning. We believe we have defined these at a general enough level of description that they are applicable to any pPBSc, independent of the science content being taught in that curriculum. These design approaches have been defined in sufficient detail, using examples from the curriculum, and also earmarked by the curriculum design challenges they address, to support a reliable use of this vocabulary across instances. One could then use this vocabulary to compare across pPBSc to notice when they are providing the same essential instructional supports even when the curriculum details make this similarity difficult to see. Similarly, one could begin to see when curricula that are representative of entirely different modes of instruction are providing the same essential instructional supports. We can continue to refine and grow this vocabulary of design approaches at a grain size that helps draw forth similarities in the deep structure of how different curricula are designed to support content learning, as opposed to every two such curricula being perceived as unique on this account. We will then be in a better position to do systematic experiments to determine the extent to which different curricular design approaches impact content learning, independent of the instructional mode to which a given curriculum is said to belong.

This work was supported in part by the following Programs of the National Science Foundation: the Postdoctoral Fellowships in Science, Math, Engineering, and Technology Education Program under award DGE-9906515; the Research in Education, Policy, and Practice Program under award REC-9720383; and the Engineering Research Centers Program under award EEC-9876363. The author gratefully acknowledges the assistance of Brian J. Reiser, Emily Kemp, Lindsey Own, Scott MacKenzie, H. David Smith, Kimberly Tester, and Jack Gallagher, without whose help this work would not have been possible.

REFERENCES


DESIGNING PROJECT-BASED SCIENCE TO TEACH CONTENT


