Introduction to Paper Set: Teaching for Science Literacy at Scale

Charles W. Anderson

Prepared for the annual meeting of the National Association for Research on Science Teaching, Portland, OR,
March, 2020

Contents

This research is supported in part by a grant from the National Science Foundation: Sustaining Responsive and Rigorous Teaching Based on *Carbon TIME* (NSF 1440988). 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 visit our website: http://carbontime.bscs.org/

Introduction to Paper Set: Teaching for Science Literacy at Scale

This paper introduces a paper set focused on the challenges of achieving three-dimensional learning as defined by the *Next Generation Science Standards (NGSS* Lead States, 2013) in diverse classrooms. Data for the five papers in this set come from 145 teachers and 25,000 students in three states, participants in a design-based implementation research project (DBIR; Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; Penuel, Fishman, Cheng, & Sabelli, 2011).

The five papers in this set are organized around three main themes, each noted in the title:

1. **Teaching…** The papers in this set focus primarily on classroom teaching and learning.
2. **…for Science Literacy…** We focus in particular on the goal of environmental science literacy, which we define as *the capacity to participate in evidence-based discussions about socio-ecological issues and to make decisions that are informed by science.* We elaborate on this goal below: here we note that our core goal focuses on the three-dimensional practices of informed citizens (described by Feinstein, 2011, as "competent outsiders").
3. **…at Scale.** This project enables us to study how a large, diverse sample of teachers and students respond to a curriculum innovation. We seek to study the full range—the less successful as well as the more successful - teachers and students.

In this introduction we briefly describe the project, then elaborate on the three themes of the paper set by discussing design challenges that we have encountered and our responses to those challenges—the foci for the papers in the set.

The Carbon TIME Project

The papers in this session report data from *Carbon: Transformations in Matter and Energy* (*Carbon TIME*: http://carbontime.bscs.org/), a DBIR project for middle and high school classrooms focusing on carbon cycling at multiple scales. This project builds on science educators' broad consensus around the *Framework for K-12 Science Education* (NRC, 2012) and the *NGSS*, including the framework of three-dimensional science learning in those documents.

In this paper set we report analyses of a large data set documenting how a diverse population of students and teachers responded to a systemic reform project that includes (a) assessments and teaching materials that support responsive and rigorous teaching; (b) sustained professional development; and (c) professional networks that encourage commitment and build expertise. The core of this system includes six teaching units focusing on carbon cycling and energy flow at multiple scales, from cellular processes such as photosynthesis and cellular respiration to global carbon cycling and energy flow. Four of these units (*Systems and Scale, Plants, Animals, Decomposers*) focus on investigations of macroscopic-scale processes and explanations using atomic-molecular models. The other two units (*Ecosystems, Human Energy Systems*) focus on ecosystem and global-scale processes.

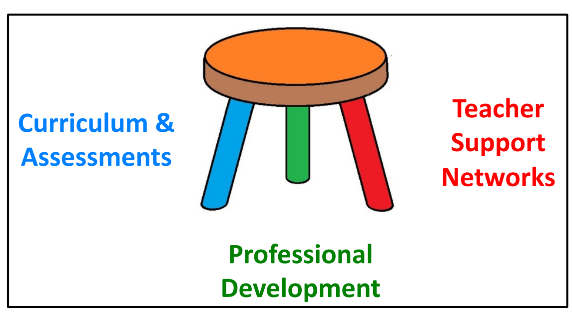

Figure 1: Three legs of the stool

These units are associated with an online assessment system, professional development, and teacher support networks, a combination that we call “three legs of the stool, arguing that each leg is necessary but not sufficient to achieve the goal described in our title: teaching for science literacy at scale. Papers in this set address two of those legs: curriculum and assessments and teacher support networks.

Data collection. We present analyses of data from the final four years of a five-year study involving 145 teachers and over 25,000 students. Data include a total of 197,000 student assessments, 881 teacher surveys, plus interviews with teachers, classroom video, examples of student work, and interviews with students in case-study classrooms.

The 94 participating schools include urban, suburban, and rural schools. There are 26 middle schools and 68 high schools. The percentage of students in a school receiving free and reduced lunches ranges from 3% to 99%, with a mean of 41%. The percentage of underrepresented minority students in the participating schools ranges from 0% to 100%, with a mean of 43%. Data sources for the papers in this session are summarized in Table 1.

Table 1: Data Sources for Papers in this Paper Set

Data Source	***First Year (2015-6)***	***Second Year (2016-7)***	***Third Year (2017-8)***	***Fourth Year (2018-9***	***Addi-tional Data****	***Papers Using These Data***
Full Data Set (105 participating middle and high school teachers)						
Student tests (~8/student per year)	21,058	60,878	76,783	35,903	244	3
Teacher surveys (~3/teacher per year)	169	294	322	152		
Teacher interviews about planning and school communities			66			5
Case Study Data Set (17 cases involving 14 teachers: 5 middle school, 9 high school)						
Student interviews (~4 focus students/class)	40	107		5*	52	2, 4
Teacher interviews (~4/teacher)	22	47		7*		2, 5
Classroom videos (~10 lessons/teacher, 2 videos/lesson)	195			8*	162	2, 4
Student work (~10/focus student)	472	450				4
* Paper 2 is based on data collected in agriscience classes during the 2018-9 school year.						

Like other DBIR projects, our research involves both design goals and research or knowledge-building goals. This paper set reports on progress in designing educational systems and in building understanding of how those systems work. In the next sections we first describe three key design challenges and our responses, then three key research goals and our approaches.

Overview of papers in this set. Figure 2 below shows how the papers in this set fit into a logic model for the project as a whole:

- Papers 1 and 2 address issues of curriculum design, including the iterative development process that we have used throughout this DBIR project.
- Papers 3, 4, and 5 all address issues associated with implementation at scale.
    - Paper 3 presents quantitative analyses of student assessment data
    - Paper 4 analyzes video data from five case study classrooms where teachers and students achieved different levels of success
    - Paper 5 analyzes interviews with the same five teachers, seeking to understand how they came to have different levels of classroom resources and the consequences for their classroom teaching.

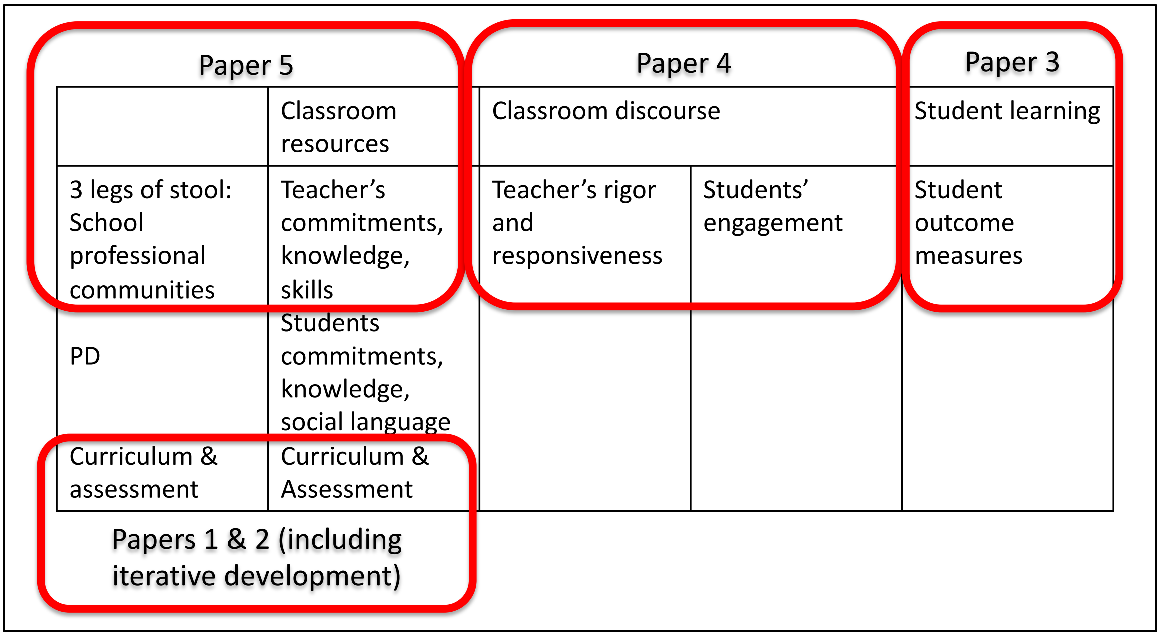

Figure 2: Logic model for papers in this set

In the remainder of this introduction we provide a brief overview of the five papers, focusing on how they connect to one another and the goals of the project as a whole, as well as on the contributions these papers make to the ongoing discussions in the field of science education about how to teach for science literacy at scale.

Papers 1 and 2: Design Principles for Science Literacy

Papers 1 and 2 (available at https://carbontime.bscs.org/conference-presentations) are as follows:

- Paper 1: *Designing Curriculum to Support the Literacy Aspects of Science Literacy,* by Kirsten D. Edwards and Charles W. Anderson
- Paper 2: *Utilizing Three-Dimensional Science Learning and Situated Instruction to Increase the Adoption of Sustainable Knowledge and Practice Among Rural Agriscience Students*, by Craig Kohn

These papers each report on a part of the overall curriculum design process associated with *Carbon TIME.* Paper 1 focuses in particular on the challenges of assessing and scaffolding students' written explanations; Paper 2 focuses on the development of an agriscience curriculum that shares some goals and methods with *Carbon TIME.* We start by positioning these papers, and the *Carbon TIME* project as a whole, within the larger NGSS-based science education reform movement.

The NGSS articulate goals for three-dimensional learning that pose difficult design challenges for classroom teaching. In this section we discuss some key design challenges that this reform movement faces and our responses to those design challenges. We start our discussion with a pair of quotes that represent key design principles for NGSS-based classroom teaching.

> "By centering science education on phenomena that students are motivated to explain, the focus of learning shifts from learning about a topic to figuring out why or how something happens." (NGSS, 2016).

Recommendation 1: "Science investigation and engineering design should be the central approach for teaching and learning science and engineering." (National Academies of Sciences, Engineering, and Medicine, 2019, page S-4)

The quotes above come from a much larger body of literature on design principles for NGSS-based teaching (e.g., Furtak & Penuel, 2019; Miller, et al., 2018; OpenSciEd, 2020; Penuel & Reiser, 2018). We agree with the design principles articulated in these articles, but we also argue that they are incomplete: They do not emphasize some design features that are essential for the purposes that we focus on in this paper set: Teaching for Science Literacy at Scale.

In the remainder of this section we first describe a design goal for classroom teaching that we share with the reform literature, then discuss some key features of *Carbon TIME* development and research that we emphasize more strongly than the reform literature cited above.

Shared design goal: Assessing and scaffolding students' three-dimensional engagement with phenomena

Here is a definition of teachers' work that is consistent with the reform literature cited above: Teachers are responsible for *assessing and scaffolding students' three-dimensional engagement with phenomena.* We elaborate on three key themes in the reform literature that we share and build on in this paper set.

1. Productive disciplinary engagement: Engle & Conant (2002; Engle, 2012) define "engagement" is a useful way. They make it clear that student engagement does not happen in every classroom, implicitly contrasting engagement with other patterns of classroom discourse that are common in science classrooms, including:

- *Disengagement:* Classrooms where teachers struggle to maintain order and capture students' attention
- *Procedural display:* Students can learn to produce correct responses to one-dimensional questions in ways that are not personally meaningful to them. This is what Bloome, Puro, & Theodorou (1989) call "procedural display."
- *Doing school:* Windschitl describes doing school as follows: "Doing school is not just a set of enactments; it is a frame that emphasizes teacher control, curricular coverage, consumption of knowledge, and individualism in learning." (Windschitl, 2019, p. 8). Doing school is accomplished in "traditional science classrooms" where students successfully engage in one-dimensional learning (Covitt et al., 2018; Johnson et al., 2017).

Engle and Conant (pages 402-3) distinguish three levels or forms of engagement, all desirable but hard to achieve:

- *Engagement* emphasizes students' authentic expression of their own ideas, interest in classroom discourse, and communication with others.
- *Disciplinary engagement* emphasizes "that there is some contact between what students are doing and the issues and practices of a discipline's discourse." Students are working toward conventionally correct performances that are meaningful to students.
- *Productive disciplinary engagement* combines students' personal engagement with disciplinary standards: "students' engagement is productive to the extent that they make intellectual progress, or, in more colloquial language, 'get somewhere.'"

The science education reform literature emphasizes, and we agree, that neither procedural display nor "doing school" that focuses on one-dimensional performances is sufficient to achieve three-dimensional learning goals. If we want to make progress at scale, that means making progress toward productive disciplinary engagement in thousands of science classrooms. As Engle and Conant point out, productive disciplinary engagement includes both (a) students' engagement in authentic expression of their own ideas and (b) conventionally correct scientific performances that are meaningful to students. This is what we mean by *students' three-dimensional engagement with phenomena.*

2. Shared storylines in classroom learning communities: We share the conviction that three-dimensional learning is both an individual and a communal process, and that productive disciplinary engagement requires shared storylines about students' learning (Cobb, 1994; Driver, et al., 1994; OpenSciEd, 2020; Penuel & Reiser, 2018). Niels Bohr's description of "the task of science" in scientific communities also describes an essential task of science classroom communities:

> The task of science is both to extend our experience and reduce it to order, and this task represents various aspects, inseparably connected with each other. Only by experience itself do we come to recognize those laws which grant us a comprehensive view of the diversity of phenomena. As our knowledge becomes wider we must always be prepared, therefore, to expect alterations in the points of view best suited for the ordering of our experience. (Niels Bohr, quoted in Hawkins, 1990, p. 100)

So classroom teachers have the responsibility to engage their students in investigations that will "extend their experience" with a "diversity of phenomena" and help them "reduce it to order." This is a form of productive disciplinary engagement that begins with students' authentic expression of their own ideas, then helps them to encounter new experiences that lead to "alterations in the[ir] points of view." This core teaching strategy is advocated for and enacted throughout the reform literature.

3. Importance of rigor and responsiveness in teachers' assessing and scaffolding practices: We also see, and agree with, consensus in the reform literature that teaching for students' three-dimensional engagement requires what Thompson, et al. (2016) describe as "rigor and responsiveness in classroom activity." This challenge for teachers—enabling students' three-dimensional engagement with phenomena through rigorous and responsive teaching—is the primary focus of this paper set.

Relevance to the *Carbon TIME* project. The *Carbon TIME* project is designed to provide teachers with a "tool kit" that they can use to accomplish these design goals. Teaching materials and supporting educator resources are available on the website (https://carbontime.bscs.org/). See, for example, the educator resources describing Three-dimensional learning in *Carbon TIME*, Strategies for sustaining rigor and responsiveness in Carbon TIME, and Purposes of assessment in *Carbon TIME*. As described below, all of the papers in this set discuss issues associated with enacting these design goals at scale.

What's missing? Additional important design features

Our work on the *Carbon TIME* project is consistent with work on NGSS-based classroom teaching in its emphasis on productive disciplinary engagement as a means to three-dimensional science learning. There are ways, though, in which our reading of the literature and our own experiences lead us to conclude that the reform documents cited above are incomplete: They advocate for design principles and research foci that will be necessary but not sufficient to achieve science literacy at scale.

We summarize below three ways in which the papers in this set add elements that we see as missing or rarely emphasized in the reform documents cited above.

1. Focus on the interpretation and application practices of informed citizens. Let's look again at the first recommendation from the NRC report, *Science and engineering for grades 6-12: Investigation and design at the center*.

Recommendation 1: "Science investigation and engineering design should be *the central approach* for teaching and learning science and engineering." (National Academies of Sciences, Engineering, and Medicine, 2019, page S-4, emphasis added)

As noted above, we fully agree with what this recommendation explicitly includes—a focus on investigations playing a central role in science teaching and learning, but what about what it implicitly *excludes*? Recommending investigation and design as *the* central approach positions professional scientists and engineers as *the* primary models for students' classroom practices. We see a similar pattern in much of the related literature, such as the NGSS storyline units (https://www.nextgenstorylines.org/ , Penuel & Reiser, 2018) or the OpenSciEd project (OpenSciEd, 2020).

What's missing, we feel, is more attention to the practices of people described by Feinstein (2011) as "competent outsiders"—people who are not scientists but who use science effectively in their daily lives and participate in civil society as informed citizens. Competent outsiders do not normally conduct investigations, though they do use scientific knowledge and engage in science practices as they figure out how to live their lives and act as citizens. Paper 2 describes the development of a curriculum for agriscience students who are preparing to be competent outsiders in this sense.

This leads us to look carefully at how classroom teaching can include investigations—essential to the preparation of competent outsiders as well as science professionals—but also include assessing and scaffolding other ways that students can acquire and use science knowledge and practice. Paper 1 describes how our work has led us to focus on students' *literacy* practices—reading and writing as key elements of productive disciplinary engagement. Most students (including professional scientists outside of their specializations) will be consumers and users, rather than creators, of scientific knowledge; we want to help students master reading and writing performances that will prepare them for their roles as informed citizens. While investigation is a key instructional element in *Carbon TIME* units, reading; writing, and revising are also integral to assessing and scaffolding productive disciplinary engagement in science classrooms.

2. Disciplinary engagement and conventionally correct performances as important goals. Let's look again at the quote above from the NGSS website and consider how it can be interpreted to exclude teaching goals and strategies that we consider important:

"By centering science education on phenomena that students are motivated to explain, the focus of learning shifts from *learning about* a topic to *figuring out* why or how something happens." (NGSS, 2016, emphasis added).

"Figuring out" can be interpreted broadly or narrowly:

- A broad interpretation of "figuring out" includes all three of Engle and Conant's forms of engagement as ways that students actively make sense of the world, in contrast with procedural display or doing school
- A narrow interpretation of "figuring out" focuses primarily or exclusively on students' personal engagement as the pathway to productive disciplinary engagement, emphasizing students' authentic expression of their own ideas and interests, while, marginalizing direct instruction in disciplinary conventions, knowledge, and practices.

It appears to us that the science education reform literature often favors the narrower interpretation, positioning direct instruction in conventionally correct science as backsliding toward "learning about." For example, E. Miller, et al. (2018, p. 1053) warn that "while we want students to do science, we seem to mean that students should mimic practices others have selected as important to learn, and content others have selected as foundational." Similarly, Furtak and Penuel (2019, p. 174) warn that "If students are merely 'going through the motions' to engage in science activities to come to an inevitable conclusion, we are shortcutting opportunities for them to authentically engage in the knowledge production that the Framework originally intended."

We recognize that the quotes above are taken out of the context of nuanced arguments about epistemic agency and scientific practice, but we are not apologetic that *Carbon TIME* units move toward "inevitable conclusions" where students master "practices others have selected as important to learn, and content others have selected as foundational." The NRC *Framework* (2012) identifies scientific knowledge and practices that are an important and valuable cultural legacy; as educators we are entrusted with passing that legacy on to our children.[1]

Thus we are intentional in identifying "conventionally correct performance" as an important goal. The quotes above associate "an inevitable conclusion" with "mimic[ing] practices" or "going through the motions." Both prior research and our own experience suggest that students can master conventionally correct performances in ways that are personally meaningful and even passionate. Heath & McLaughlin (1993), for example, describe how dance and music performance troupes can play a key role in helping inner city youth establish agency and strong personal identities. Lave & Wenger (1991) describe how newcomers are socialized into communities of practice through legitimate peripheral participation. Vygotsky (1980) describes how students learn from mentors through shared practices in the Zone of Proximal Development.

Following Gee (2005; see also Bowers, 2008; Covitt & Anderson, 2018), we see science as an "academic social language" or discourse. "To acquire an academic social language, students must be willing to accept certain losses and see the acquisition of the academic social language as a gain." (Gee, 2004, p. 21). Gee goes on to elaborate on the tradeoff that teachers face in mediating between students' lifeworld or vernacular languages and scientific language.

> If we "situate science talk" in children's lives as a way to bridge to and enhance the acquisition of one or more academic forms of language, that is a good thing. If we do so in such a way that these children cannot later read or discuss texts in content areas and can't get out of high school, while more privileged children can, that's a bad thing. (Gee, 2004, p. 41)

We have argued elsewhere (Anderson, 2018; H. Miller, et al., 2018) that crosscutting concepts identify conventional "rules" that play an essential role in scientific practices, and that mastering those conventions is necessary for students to participate in scientific discourse. These conventions enable students to describe, analyze, and communicate about phenomena with a depth and precision that is impossible without them.

These ideas affect both our assessing and scaffolding principles. Our assessments are based on learning progression research (Covitt & Anderson, 2018; Jin & Anderson, 2012; Mohan, Chen, & Anderson, 2009) in which the goals are defined as conventionally correct inquiry and explanation performances. Papers 3, 4, and 5 all use data from these assessments to make claims about the *Carbon TIME* curriculum and teacher effectiveness. Papers 1 and 2 both describe conventionally correct performances and our work with teachers to design, field

---

[1] We recognize that the scientific canon is a product of social and economic systems that have exploited the knowledge and practice of marginalized communities without crediting their accomplishments. These legacies continue to do social and environmental damage today. Our definition of scientific literacy includes both informing all students about the causes and consequences of these legacies and enabling all students to use and judge their products for the benefit of themselves and their fellow citizens. Scientifically literate citizens need canonically correct, sophisticated understandings of science to make decisions around personal well-being and environmental and social issues. See, for example, the Storyline Readings in each unit.

test, and revise teaching tools that help students master conventionally correct performances in personally meaningful ways. We argue that these tools can support teaching practices are both rigorous and responsive.

3. Apprenticeship as an important teaching strategy. The reform literature advocates that investigations should play a central role not only as a goal for student practices, but also as an instructional strategy. Again, we agree with what these recommendations explicitly include, but are concerned about what they implicitly exclude. Particularly when we are working toward goals associated with environmental science literacy—defined as practices of competent outsiders—we believe that strategies associated with cognitive apprenticeship (Collins, Brown, and Newman, 1989) are appropriate.

Carbon TIME units are built around an instructional model that scaffolds students' practices in three roles: questioner, investigator, and explainer. The questioner and investigator roles engage students in investigations and constructing arguments from evidence; the explainer role emphasizes helping students to figure out explanations of phenomena that make effective use of scientific models and principles.

Scaffolding students for these roles involves creating classroom discourse communities that include elements of cognitive apprenticeship, as described by Lave & Wenger (1991) and Collins, Brown, and Newman (1989). These authors describe learning as a culturally embedded activity that requires what Yackel & Cobb (1996) describe as classroom norms or what Windschitl (2019) analyzes as a conversational infrastructure. We describe key elements of the cognitive apprenticeship model and of the conversational infrastructure in educator resources available on the *Carbon TIME* website (Carbon TIME Project, 2019a; 2019b).

Summary of key design goals and strategies. Table 2 below summarizes key ways in which we add design goals and strategies to the current consensus literature on design principles to achieve three-dimensional learning as defined by NGSS.

Table 2: Key *Carbon TIME Design Goals and Features*

<table>
<tr><th>Issue</th><th>The reform literature emphasizes…</th><th>Carbon TIME adds…</th><th>Relevant Carbon TIME features</th></tr>
<tr><td>Quality of student engagement</td><td colspan="2">Students' engagement with phenomena and models (including disciplinary/productive disciplinary engagement)</td><td rowspan="3">Unit driving questions
Engagement with multiple phenomena
Content storyline: connecting observations, patterns, explanations
Formative & summative assessments
Process tools</td></tr>
<tr><td>Shared storylines in classroom learning communities</td><td colspan="2">Students extend their experience and reduce it to order, making sense of phenomena using increasingly complex models</td></tr>
<tr><td>Teachers' assessing & scaffolding practices</td><td colspan="2">Rigor: Assessing & scaffolding 3D performances
Responsiveness to students' ideas and interests</td></tr>
<tr><td>Adult communities of practice that exemplify goals</td><td>Practices of scientific communities as a goal</td><td>Practices of informed citizens (competent outsiders) as a goal</td><td>Practice storyline: Students as questioners, investigators, explainers</td></tr>
<tr><td>Indicators of successful classroom discourse</td><td>Authentic expression of students' ideas</td><td>Conventionally correct performance</td><td>Discourse routine</td></tr>
<tr><td>Organizing framework for classroom practices</td><td>Investigation: Students extend their experience and reduce it to order</td><td>Apprenticeship: Modeling, coaching, fading by knowledgeable mentors</td><td>Instructional model</td></tr>
</table>

Connections to papers in this set. The right-hand column of Table 2 lists features of *Carbon TIME* units and assessments that are relevant to the issues discussed above. All can be found on the website: https://carbontime.bscs.org/ . These features have been developed and field tested over many years.

Papers 1 and 2 both describe aspects of that development process. Paper 1 describes the development and field testing of scaffolds for students' written explanations. Paper 2 focuses on the development an field testing of a related agriscience curriculum, following the same design principles. Paper 3 describes results of assessments focusing on environmental science literacy. Paper 4 describes challenges that teachers encounter in developing norms and discourse routines that engage students in meaningful three-dimensional practices and the effectiveness of apprenticeship-oriented teaching strategies. Paper 5 looks at commitments of teachers to instructional practices associated with investigation and apprenticeship,

Paper 3: Assessing Students' Learning from Carbon TIME

This is a design-based implementation research (DBIR) project, meaning that we are particularly interested in issues that arise when we enact reform programs in our large, diverse educational system. We have written before about the design challenges associated with accomplishing three-dimensional learning at scale (Anderson, et al., 2018). Table 1 above shows that we have been successful at collecting and analyzing pre-post data on students' three-dimensional learning at scale in diverse classrooms. (The development and validation of this system is described in other presentations at this conference: Draney & Bathia, 2020; Thomas, 2020.)

Paper 3 (Factors affecting students' learning about Carbon TIME, by Qinyun Lin, Ken Frank, Charles W. Anderson, Karen Draney, Shruti Bathia, & Jay Thomas, available at https://carbontime.bscs.org/conference-presentations) reports on patterns in this large data set, including 197,000 student pretests and posttests. Key results from our analyses of the full student pre-post data set include the following:

1. *Carbon TIME* curriculum and professional development make a big difference, in two ways:
    a. Overall student achievement: It's not quite right to say that "all the children in *Carbon TIME* classrooms are above average," but we came pretty close. Comparing *Carbon TIME* posttest data with baseline data (posttests from the classrooms of *Carbon TIME* teachers the year before they started participating in the project), 91.4% of *Carbon TIME* students scored above the baseline median.
    b. Lowering the achievement gap: Students with lower pretest scores showed significantly higher learning gains than students with higher pretest scores in *Carbon TIME* classrooms.
2. Teachers make a significant difference: Teachers accounted for more of the variance in students' posttest achievement than pretest scores (i.e., students with low pretest scores in more successful classrooms did better on the posttest than students with high pretest scores in less successful classrooms).
3. School resources make a small difference: Students in schools with more organizational resources (indicated by percentage of free and reduced lunch or percentage of white and Asian students) did slightly better than students in schools with fewer organizational resource.

Papers 4 and 5: Learning from More and Less Successful Classrooms

Papers 4 and 5 expand on the second of these large-scale findings: If teachers make a significant difference in student achievement, then what are the characteristics of discourse in the more successful classrooms? How do those classrooms help students learn? What's happening in the less successful classrooms, and why?

We need to understand how large-scale reform efforts affect the sense-making practices of teachers and students in individual classrooms, so Paper 4 (Carbon TIME Classroom Discourse and Its Connections to Student Learning, by Beth A. Covitt, Christie Morrison Thomas, Qinyun Lin, Elizabeth de los Santos, & Charles W. Anderson) reports on analyses of classroom discourse in five more and less successful case-study classrooms, helping us to understand both how the actions of teachers and students made sense in their circumstances and the consequences for students' three-dimensional learning. Paper 5 (*Carbon TIME* Teacher Orientations and Contexts: Making Connections to Classroom Discourse and Student Learning, by Christie Morrison Thomas, Beth A. Covitt, Brian Hancock, Qinyun Lin, Stefanie Marshall & Charles W. Anderson) reports on interviews with the same five teachers, analyzing how their personal commitments and access to resources affected their classroom teaching strategies.

We point in particular to two kinds of findings enabled by these analyses. We point in particular to two kinds of findings, concerning (a) the challenges of rigor and responsiveness, and (b) classroom and organizational resources from the comparative qualitative analyses in these papers.

The challenges of achieving both rigor and responsiveness

It's hard to get funding for design-based research by reporting on our failures, but successful large-scale implementation depends in part on careful analysis of what's happening in classrooms where things aren't going as we had hoped.

One theme that runs through our analyses concerns how hard it is to achieve both rigor and responsiveness in science classrooms. In resource-rich classrooms, we see teaching and learning that are both rigorous (students are achieving high-level scientific performances) and responsive (students are expressing and discussing their authentic ideas about phenomena)—in Engle and Conant's terms, the students show productive disciplinary engagement. Papers 4 and 5 document the strength of conviction and extensive work needed by teachers to develop *discourse routines* that support teaching in depth (what Windschitl, 2019, analyzes as a "conversational infrastructure").

But we also study classrooms that work in different ways. For example, *Carbon TIME* and other reform curricula are designed to support deep student engagement with a small number of big ideas, reflecting a decades-long consensus among science educators that the American science curriculum is "a mile wide and an inch deep"—that most science classes cover too much content with too little attention to teaching for mastery (e., National Research Council, 2007, Chapter 8; Valverde & Schmidt, 1997).

And yet…. teaching for breadth over depth persists. Windschitl describes this pattern of teaching as "doing school:" "Doing school is not just a set of enactments; it is a frame that emphasizes teacher control, curricular coverage, consumption of knowledge, and individualism in learning." (Winschitl, 2019, page 8). Lemke (1990) and Hess and Azuma (1991) describe similar patterns. Hess and Azuma contrast "sticky-probing" Japanese classrooms with "quick and snappy" American classrooms, emphasizing that both styles of classroom discourse have deep cultural roots, helping to explain why teaching for breadth over depth has persisted in the U.S. for decades in spite of the academic consensus that depth is preferable.

Papers 4 and 5 investigate the causes and consequences of these different approaches to teaching. Our work with a diverse group of secondary science teachers provides empirical support for the academic consensus that teaching in depth leads to better learning outcomes. This work also helps us to understand how difficult it is to support students' productive disciplinary engagement and how less effective patterns of classroom discourse enable teachers with fewer classroom resources to do their jobs with some success.

Learning from the more successful teachers: Classroom enactment of productive disciplinary engagement. We see clear evidence that the case study teachers with higher learning gains are assessing and scaffolding their students' three-dimensional engagement with phenomena. Some key points:

- As discussed below, these are resource-rich classrooms, where the teachers showed exceptional skill in orchestrating classroom discourse and managing student work, being insightful about students, and devoting time and energy to planning and student feedback. These teachers also tended to have close and supportive relationships with their *Carbon TIME* case study coaches.
- The more successful teachers were still "doing school:" In Windschitl's terms, "teacher control, curricular coverage, consumption of knowledge, and individualism in learning" were essential obligations that came with these teachers' jobs, so the more successful teachers figured out how to assess and scaffold three-dimensional engagement while also meeting these obligations.
- Apprenticeship strategies are important: Key scaffolding strategies used by the more successful teachers included modeling, coaching, and fading conventionally correct performances (Collins, Brown, & Newman, 1989).

Learning from the less successful teachers. The case study teachers with lower learning gains were more likely to incorporate *Carbon TIME* materials into their routines for "doing school," so the students in their classrooms were more likely to perform procedural display and less likely to show three-dimensional engagement with phenomena. Some key points:

- They were missing essential resources: As discussed below, these teachers did not have all the resources that they needed to assess and scaffold their students' three-dimensional engagement with phenomena.
- "Doing school" is not nothing: Although these classrooms showed lower learning gains, the analyses in Paper 3 show that learning gains were higher than baseline even in less successful classrooms. So "doing school" with *Carbon TIME* materials was better than "doing school" with only the teachers' previous teaching materials.
- "Doing school" worked for teachers and students. When teachers had faster content coverage, interesting facts, and one-dimensional expectations for student performance, it was possible to negotiate what Doyle (1983) refers to as "the performance for grade exchange" in ways that appeared to be satisfactory for both teachers and students.

Classroom and organizational resources to achieve rigor and responsiveness

In a previous publication (Anderson, et al., 2018) we discussed material, human, and social organizational resources needed to support rigorous and responsive teaching. In this paper set we analyze how more successful teachers use those resources in their classrooms, as well as how teachers with fewer resources develop discourse routines that work in their classrooms. Using the budget analogy, high levels of rigor and responsiveness are *expensive.* They require:

- High levels of classroom resources: teacher's knowledge and skills, students' prior knowledge and motivation to learn, curriculum and teaching tools, support from school professional communities, etc.
- High levels of continuing effort from teachers: careful planning, grading complex performances, professional learning, etc.
- High levels of continuing effort from students: writing paragraphs, reading for understanding, effort to answer Three Questions about phenomena, etc.

Papers 4 and 5 analyze some classrooms where teacher and students can "afford" three-dimensional rigor and responsiveness: they both have resources and show continuing effort. We also analyze other classrooms with fewer resources, in which teachers and students figure out ways to make procedural display and doing school "work" for them, enabling everyone to get through the day with civility and some apparent satisfaction.

Table 3, below, describes indicators of different levels of rigor and responsiveness in the case study classrooms that we studied in Papers 4 and 5. Science educators often assign pejorative labels to the beliefs and practices in the "Low" column; we think it is important to recognize these beliefs and practices as sincere attempts by teachers to do their work with limited resources.

Table 3: General Indicators of Levels of Rigor and Responsiveness

Level of Classroom Resources	Low	Intermediate	High
Quality of Student Engagement	Disengaged Procedural display	Doing school around 1D performance	Engagement with phenomena (Productive) disciplinary engagement
Rigor	Low expectations Making tests/tasks easier Students don't do work	Covering content Grading for completion	Standards & coaching for 3D performance
Responsiveness	Accepting students' limited abilities Accepting students' lack of motivation to strive to meet high standards	Eliciting but not engaging students' ideas Quick and snappy Motivation with novelty	3D assessment Sticky probing Motivation to learn

Causes and consequences of different levels of rigor and responsiveness

Causes: Paper 5 analyzes interviews with teachers who show higher and lower levels of rigor and responsiveness in their classroom teaching. The interviews with the more successful teachers document their personal learning about assessing and scaffolding strategies, their skill in leveraging resources from their students, their schools, and the *Carbon TIME* project, and their sustained effort to improve their teaching and serve their students.

Interviews with the less successful teachers show their teaching strategies were reasonable in terms of their personal knowledge and skills, their school contexts, and resources available to them. Some of these teachers acknowledged the limitations of their classrooms and explain their struggles to achieve higher levels of rigor and responsiveness. Other teachers doubted the possibility of all their students achieving three-dimensional engagement with phenomena, advocating "constructivist" teaching strategies that are responsive but not rigorous or criticizing *Carbon TIME* materials as inappropriate for their students.

So we saw a set of "basic needs and obligations" that seem more or less inevitable for science teachers in American public schools (Kennedy, 2016; Smith, 1996). These needs include:

- Maintaining a professional identity that includes a sense of efficacy and positive relationships with students and colleagues
- Fulfilling their basic responsibilities as science teachers (Windschitl's doing school)
- Enabling most students to be successful in the performance for grade exchange
- Keeping students interested

We saw that teachers with more resources were able to meet these basic needs while also assessing and scaffolding students' three-dimensional learning. Table 4 summarizes some of the ways that high-resource teachers went "beyond the basics."

Table 4: Different Approaches to Basic Needs and Obligations of Science Teaching

Basic need	Low resource approach	High resource approach
Professional identity (including efficacy, relationships with students and colleagues)	Separate public discourse with colleagues from private discourse with students	Discourse with colleagues about assessing and scaffolding students' 3D engagement
"**Doing school** is not just a set of enactments; it is a frame that emphasizes teacher control, curricular coverage, consumption of knowledge, and individualism in learning." (Windschitl, 2019, p. 8)	• Control with clear rules, effective management • Coverage of 1D facts and skills • Consuming knowledge to get a grade • Individual work and accountability	• Control through student agency & engagement • Coverage of 3D performances • Consuming knowledge to act as informed citizens • Accountability for participation in classroom discourse
Performance for grade exchange	Student success through clear, low expectations	Student success through productive disciplinary engagement
Student interest	Quick and snappy: Keep up pace and novelty	Sticky probing: Intrinsic rewards of 3D learning

Consequences for student learning. Paper 3 explains how we constructed value-added measures of student learning in *Carbon TIME* classrooms. Although using these measures to evaluate the success of individual teachers is problematic (for reasons discussed in Paper 3), the data from these analyses support (a) the general claim that teachers make a difference—there are large differences in learning in different classrooms—and (b) that students in classrooms showing high levels of rigor and responsiveness are more successful in three-dimensional learning.

Conclusion

To return to the beginning of this introduction, this paper set focuses on the problem of *teaching for science literacy at scale.* The introduction presents a critique of what we see as the current consensus about teaching strategies to achieve this goal. The papers in the set focus on what the *Carbon TIME* project has to offer as we seek to advance the field.

Critique of consensus teaching strategies: If the goal is teaching for science literacy at scale, then the consensus strategies for classroom teaching are:

- **Incomplete** (necessary but not sufficient). In particular, we need to consider
  - Literacy goals, preparing students for informed citizenship
  - Apprenticeship strategies, achieving conventionally correct performances through modeling and coaching
- **Impractical** for diverse classrooms with limited resources. In particular, we need to consider:
  - Resource limitations for many classroom teachers
  - Basic needs and obligations that all teachers must meet

What *Carbon TIME* has to offer. While we critique the the current consensus as incomplete and impractical, our data tell us that *Carbon TIME* is **NOT** a program that is either complete or practical (i.e., necessary and sufficient for teaching for science literacy at scale). We do have:

- Examples of tools and strategies that are useful for teaching for scientific literacy, including:
  - Assessments of 3D performances related to science literacy
  - Scaffolds for students acting as questioners, investigators, and explainers
- Lessons learned from both more and less successful classrooms, including:
  - Basic needs and obligations that all teachers must meet
  - Resources that enable some teachers to go beyond meeting basic needs so that they successfully assess and scaffold students' three-dimensional engagement with phenomena

The papers in this set have implications for both the design and the implementation of classroom resources to support the development of science literacy in students. We discuss how the goal of science literacy has implications for the design of classroom teaching materials and strategies, including (a) the importance of preparing all students to be "competent outsiders," (b) the importance of conventionally correct performances as a goal, and (c) the importance of apprenticeship as a teaching strategy.

Our implementation research shows that both the *Carbon TIME* curriculum and teachers have significant effects on student learning, at scale, in diverse classrooms. We also show how rigorous and responsive teaching depends on other organizational resources, and how difficult it will be for all classrooms to move beyond doing school and procedural display. The papers provide both an analysis of the challenges ahead and ideas about how to meet those challenges. These papers(available at https://carbontime.bscs.org/conference-presentations) include:

1. *Designing Curriculum to Support the Literacy Aspects of Science Literacy,* by Kirsten D. Edwards and Charles W. Anderson
2. *Utilizing Three-Dimensional Science Learning and Situated Instruction to Increase the Adoption of Sustainable Knowledge and Practice Among Rural Agriscience Students*, by Craig Kohn
3. *Factors affecting students' learning about Carbon TIME,* by Qinyun Lin, Ken Frank, Charles W. Anderson
4. *Carbon TIME Classroom Discourse and Its Connections to Student Learning,* by Beth A. Covitt, Christie Morrison Thomas, Qinyun Lin, Elizabeth de los Santos, & Charles W. Anderson
5. *Carbon TIME Teacher Orientations and Contexts: Making Connections to Classroom Discourse and Student Learning*, by Christie Morrison Thomas, Beth A. Covitt, Brian Hancock, Qinyun Lin, Stefanie Marshall & Charles W. Anderson

Please visit the *Carbon TIME* website (https://carbontime.bscs.org/) for the full paper set and to learn more about resources available.

References

Anderson, C. W. (2018, December). Crosscutting concepts as epistemic heuristics in learning progression research and curriculum design. Written for the Summit for Examining the Potential of Crosscutting Concepts to Support Three-dimensional Learning, Arlington, VA.

Anderson, C. W., de los Santos, E. X., Bodbyl Roels, S., Covitt, B., Edwards, K. D., Hancock, J. B., Lin, Q., Morrison Thomas, C., Penuel, W., Welch, M. (2018). Designing educational systems to support enactment of the *Next Generation Science Standards. Journal of Research in Science Teaching, 55(7),* 1026-1052.

Bloome, D., Puro, P., & Theodorou, E. (1989). Procedural display and classroom lessons. *Curriculum Inquiry, 19(3)*, 265–291.

Bowers, C. A. (2008). Why a critical pedagogy of place is an oxymoron. *Environmental Education Research*, *14*(3), 325-335.

Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: Rethinking transfer: A simple proposal with multiple implications. *Review of research in education*, *24*(1), 61-100.

Carbon TIME Project (2019a). Three-dimensional learning in *Carbon TIME*.

Carbon TIME Project (2019b). Strategies for sustaining rigor and responsiveness in *Carbon TIME.*

Carbon TIME Project (2019c). Purposes of assessment in *Carbon TIME*.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, *23*(7), 13-20.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003, Jan/Feb). Design experiments in education research. *Educational Researcher. 32(1)*: 9-13.

Collins, A., Brown, J. S., Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.) *Knowing, learning, and instruction: Essays in honor of Robert Glaser.* Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 453-494.

Covitt, B. A., Dauer, J. M., and Anderson, C. W. (2017). The role of practices in scientific literacy. In C. Schwarz, B. Reiser, & C. Passmore (Eds.). *Bringing next generation science & engineering practices into our K-12 Classrooms: Moving beyond "knowing" science to making sense of the world.* (pp. 59-84) Washington, DC: NSTA Press.

Covitt, B. & Anderson, C. (2018). Assessing scientific genres of explanation, argument, and prediction. In A. L. Bailey, C. Maher, & L. Wilkinson (Eds.) *Language, literacy, and learning in the STEM disciplines: How language counts for English learners,* pp. 206-230. New York, NY: Routledge.

Doyle, W. (1983). Academic work. Review of Educational Research, 53(2), 159-200.

Draney, K., & Bathia, S. (2020, March). Assessment frameworks: Improving the evidence we collect about student cognition. Presented at the annual meeting of the National Association for Research in Science Teaching, Portland, OR.

Driver, R., Asoko, H., Leach, J., Scott, P., & Mortimer, E. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, *23*(7), 5-12.

Engle, R. A. (2012). The productive disciplinary engagement framework: Origins, key concepts, and developments. In *Design research on learning and thinking in educational settings* (pp. 170-209). Routledge.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and instruction*, *20*(4), 399-483.

Feinstein, N. (2011). Salvaging science literacy. *Science Education*, *95*(1), 168-185.

Furtak, E. M., & Penuel, W. R. (2019). Coming to terms: Addressing the persistence of “hands-on” and other reform terminology in the era of science as practice. *Science Education*, *103*(1), 167-186.

Gee, J. P. (2004). Language in the science classroom: Academic social languages as the heart of school-based literacy. In Yerrick, R. K., & Roth, W. M. (Eds.) *Establishing scientific classroom discourse communities: Multiple voices of teaching and learning research* (pp. 28-52). Routledge.

Hawkins, D. (1990). Defining and bridging the gap. In E. Duckworth; J. Easley, D. Hawkins, & A. Henriques (Eds.), *Science education: A minds-on approach for the elementary years* (pp. 97-139). Hillsdale, NJ: Lawrence Erlbaum Associates.

Heath, S. B., & McLaughlin, M. (1993). *Identity and inner-city youth*. Teachers College Press.

Hess, R. D., & Azuma, H. (1991). Cultural support for schooling: Contrasts between Japan and the United States. *Educational Researcher*, *20*(9), 2-9.

Jin, H., & Anderson, C. W. (2012). A learning progression for energy in socio-ecological systems. *Journal of Research in Science Teaching, 49(9),* 1149–1180.

Kennedy, M. (2016). Parsing the practice of teaching. *Journal of Teacher Education*, *67*(1), 6-17.

Latour, B., & Woolgar, S. (1986). An anthropologist visits the laboratory. In *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. (Cited by 74,046)

Miller, E., Manz, E., Russ, R., Stroupe, D., & Berland, L. (2018). Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. *Journal of Research in Science Teaching*, *55*(7), 1053-1075.

Miller, H., Freed, A., Dauer, J., Doherty, J., and Anderson, C. W. (2018. The role of crosscutting concepts in developing a three-dimensional learning progression framework. *Carbon TIME* Project (need to put on website)

Mohan, L., Chen, J., and Anderson, C. W. (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching, 46(6),* 675-698.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. National Academies Press.

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

National Research Council. (2014). *Developing assessments for the next generation science standards.* National Academies Press.

National Research Council (2019). *Science and engineering for grades 6-12: Investigation and design at the center*. National Academies Press.

NGSS Lead States. (2013). Next Generation Science Standards: For states, by states. Washington, DC: National Academies Press. https://www.nextgenscience.org/

NGSS Website (September, 2016). Using phenomena in NGSS-designed lessons and units. Retreived from http://nextgenscience.org/sites/default/files/Using%20Phenomena%20in%20NGSS.pdf

OpenSciEd (2020). OpenScied Design Specifications. https://www.openscied.org/design-specifications/

Penuel, W. R., Fishman, B. J., Haugan Cheng, B., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher, 40*(7), 331-337.

Penuel, W. R. and Reiser, B. (2018). Designing NGSS-aligned curriculum materials. Paper commissioned for the report *Engaging Middle and High School Students in Science and Engineering: New Approaches to Investigation and Design* from the Board on Science Education, National Academies of Science, Engineering, and Medicine.

Smith III, J. P. (1996). Efficacy and teaching mathematics by telling: A challenge for reform. *Journal for Research in Mathematics Education*, 387-402.

Thomas, J. (2020, March). Using machine learning to interpret NGSS tasks at scale. Presented at the annual meeting of the National Association for Research in Science Teaching, Portland, OR.

Thompson, J., Hagenah, S., Kang, H., Stroupe, D., Braaten, M., Colley, C., & Windschitl, M. (2016). Rigor and responsiveness in classroom activity. *Teachers College Record*.

Valverde, G.A., and Schmidt, W.H. (1997). Refocusing U.S. math and science education. *Issues in Science and Technology, 14(2),* 60-66.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard University Press. (Cited by 106,333)

Windschitl, M. (2019). Disciplinary Literacy Versus Doing School. *Journal of Adolescent & Adult Literacy*, *63*(1), 7-13.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for research in mathematics education*, 458-477.