Bridging Research and Practice to Implement Change in Teaching and Learning at Scale

Daniel C. Edelson (1), BSCS Science Learning, dedelson@bscs.org
Brian J. Reiser, Northwestern University, reiser@northwestern.edu
Katherine L. McNeill, Boston College, kmcneill@bc.edu
Renee Affolter, Boston College, raffolter.vtscience@gmail.com
Andrew E. Krumm, University of Michigan, aekrumm@umich.edu
William R. Penuel, University of Colorado - Boulder, william.penuel@colorado.edu

Abstract: This symposium explores the challenges of applying research at scale through the consideration of a case study of a collaboration between a consortium of learning sciences researchers from five institutions and science education leaders from 10 US states. The case is the OpenSciEd Project, which is an initiative to create and disseminate a 3-year, middle school science program. The OpenSciEd project has the explicit goal of meeting three competing goals in order to have impact at scale: implementing findings from research (research-based), meeting the practical constraints of American public school classrooms (practicality), and creating products that will be widely adopted in the market place (marketability). In this symposium, we will discuss how these three competing goals have influenced the organization of the project, the design of instructional materials, the design of professional learning experiences for teachers, and the collection and analysis of data in field trials.

Symposium background

Learning Sciences researchers often aspire to having a large-scale impact on educational practice. Achieving impact can depend on much more than having relevant and robust research findings. Studies of research utilization suggest that there are multiple paths to impact: practitioners can take up findings to make decisions regarding policies and programs; they can use research concepts to inform thinking about persistent problems of practice; and they can use tools that reflect key principles derived from a body of research (Weiss & Buculavas, 1980; Penuel, et al., 2017). In this symposium, we focus on this last path, through the exploration of a project that is developing and disseminating tools derived from research. In this case, the tools take the form of instructional materials and resources for teacher professional learning. Well-designed tools can be a particularly powerful means to communicate and help practitioners take up instructional shifts, particularly when they provide models for how to change practice in significant ways (Ikemoto & Honig, 2010). Instructional materials are tools that teachers “participate with” (Remillard, 2005) to make changes to practice. As teachers use them, their classrooms transform, and they can come to see new possibilities for student learning. In this symposium, we consider the OpenSciEd project, a project to develop an open content (2) science curriculum for middle school designed to:

- Enable schools to achieve the goals of the Next Generation Science Standards (NGSS; NGSS Lead States, 2013),
- Provide equitable learning opportunities for students from populations that have historically been underserved,
- Be adopted and successfully implemented by a large number of schools across the U.S.

The OpenSciEd project was launched in 2017 by the Carnegie Corporation of New York, a private foundation, and Achieve, the organization that worked with 26 states to develop the NGSS. Their objective was to meet the need they observed for high quality instructional materials for the NGSS. Their idea was to bring researchers with the expertise and experience to create high quality materials together with state-level leaders in science education who could collaborate with researchers in the creation of materials that would be marketable and practical to implement. The Carnegie Corporation, who has since been joined by three other funders, selected a consortium of five organizations (the Developers Consortium) to develop the OpenSciEd program and recruited state science education leaders from 10 states (the State Steering Committee) to oversee the work of the developers and to facilitate the field testing of the materials in their states.

In this symposium, members of the Developers Consortium will describe challenges that we, as learning scientists, are experiencing in this attempt to translate research into instructional materials and professional learning programs with the goal of large-scale—both broad and deep—impact. We will focus on the challenges of meeting the competing goals and constraints inherent in this work, including creating products that are:

- Research-based, remaining faithful to the richness and complexity of the research base supporting science learning in social contexts;
- Practical, ensuring that classroom and professional learning materials are feasible to implement across the diversity of classroom contexts found in American public schools; and
- Marketable, producing products and programs that are attractive to large numbers of schools and districts, and affordable for them to implement.

The four presentations in the symposium will describe how we have addressed these challenges in specific contexts: 1) the organization of the OpenSciEd project, 2) the design of instructional materials, 3) the design of professional learning programs for teachers, and 4) the collection and analysis of data in field.

Organization of the OpenSciEd project

Daniel C. Edelson, BSCS Science Learning

In many respects, the OpenSciEd project has been structured to expose and explicitly address the tensions between the three goals of being: research-based, practical, and marketable. The fact that the project founders selected researchers to lead the development reflects their commitment to creating a research-informed program. The fact that they sought out educational leaders from state educational agencies and educational service agencies to serve as a steering committee reflects their commitment to practicality and marketability. The commitments that they asked the members of both the Developers Consortium and the State Steering Committee to make and the decision-making structure that they put in place ensured that the goals and concerns of both researcher and practitioner perspectives would be carefully weighed in important decisions.

From its earliest stages, important decisions about the project have been negotiated by the members of the Developers Consortium and the State Steering Committee, with the objective of achieving compromise that would enable the project to achieve its end goal. This end goal was described in an early consensus document as:

> *A group of state education agencies, working with school districts, classroom educators, experienced science curriculum developers and the science education community, will create and field test a complete set of robust, research-based, open-source science instructional materials that are aligned to the Framework and NGSS and accessible to all students, while building demand for the materials and implementation supports in tandem. (OpenSciEd, 2018a)*

The mechanism for making these decisions became the development of guiding documents for the project. In the first few months of the project, the State Steering Committee and the Developers Consortium jointly created six guiding documents:

1. State commitments, laying out the expectation for all state partners, including their participation in a steering committee and implementation of a field test in their state;
2. Guidelines and guiding principles for the Developers Consortium, a high-level description of the components to be developed;
3. Overview of program design specifications, describing the contents to be included in the program's specifications
4. Professional learning design principles, describing key attributes that the professional learning programs and resources for OpenSciEd must possess
5. Data collection sampling strategy, laying out the basic strategy for collecting data in field trials
6. Operations and governance functions, describing a new organization to be established to oversee the development of the OpenSciEd program and its dissemination.

What brought these state leaders and researchers together was two shared beliefs: (1) high-quality, low cost instructional materials are necessary to enable schools and teachers to achieve the goals of the NGSS, and (2) there was no other mechanism in place that would meet this need. What enabled them to achieve consensus on these foundational documents was their recognition of the need to compromise in order to achieve this shared goal. State representatives were asked to represent the interests and concerns of their states but not to object to proposals based on the unique concerns of their states. Developers were asked to advocate for approaches that are supported by research evidence, but to recognize the need to compromise based on practical challenges of implementation across diverse public school contexts and the considerations of the marketplace.

Two early issues tested the willingness of participants to compromise and ultimately laid the foundation for the collaborative process that has played out over the subsequent years. The first issue required substantial concessions on the part of state representatives. This issue pitted practicality against marketability. In developing a scope and sequence for the program, the Developers Consortium analyzed the standards across the ten partner states and found that 4 of the 10 states specify grade levels at which specific standards are to be taught. (The NGSS does not specify a sequence for standards within the 3-year middle school grade band). We found that the likelihood that two states had assigned the same standard to the same grade was roughly 1 in 3. Recognizing that this marketability issue for their individual states could defeat the project before it even got started, all ten states agreed to collaborate with the developers to sequence instruction in a way that had the strongest justification in research on learning progressions.

The second issue forced a compromise by developers. It pitted the goal of being research-based against practical and market considerations. All state partners expressed a sense of urgency. They saw a limited window of opportunity to release these materials to the marketplace before schools and districts would be forced to adopt by policies in place. In fact, in November 2017 at the first convening of the partner states, they argued that the project would fail unless initial instructional materials could be in field testing within 12 months and released to the public within 24. For researchers with longstanding commitments to careful development of instructional materials and professional learning programs through co-design with teachers and multiple cycles of development, classroom testing, third-party review, and revision, this was a stunning request. However, recognizing the validity of the state partners' concerns, the developers agreed to a timeline that called for only one cycle of field trials and revision of each 6-week unit prior to public release. This timeline called for the release of the first three units within 18 months, followed by the release of three new units every 6 months over a three-year period.

These are only two of dozens of compromises that have been negotiated over the course of the project so far. They are the product of a structure designed to bring tensions to the surface, and they helped establish a culture that recognizes the need for compromise to achieve shared goals.

Design of OpenSciEd instructional materials

Brian J. Reiser, Northwestern University

Developing the program of instructional materials has involved two strands of work. First, we developed two foundational documents to guide design: a scope and sequence and a set of design specifications (OpenSciEd, 2019a, 2019b). The scope and sequence maps out how learning objectives are to be accomplished over the course of the three-year, 18-unit program. The design specifications provide instructional and usability guidelines for the materials. Second, using these foundational documents for guidance, we are developing the instructional units, conducting field trials, and revising them based on lessons learned from the trials. As of February 2020, we have developed and field tested twelve 3- to 6-week units, and we have revised and released five of them. Throughout, we have struggled with tensions being research-based, practical, and marketable. In this section, we focus on several challenges that emerge from tensions between the instructional shifts called for by the (research-based) reforms being targeted and established practices or policies (i.e., practicality and marketability).

Challenge 1: Develop a design approach backed on research synthesis rather than a single perspective

From its inception, the project has had the goal of creating a program that reflects the best available research. However, in contrast to many research-based educational programs that focus on implementing specific lines of research associated with its developers, state partners felt that the program should reflect research findings broadly and be synthetic, rather than parochial. As a result, the Developers Consortium is composed of multiple research and development groups and has been charged with engaging a broad research community in its work.

To achieve this goal, the development consortium proposed to recruit committees of experts representing multiple perspectives to develop program design specifications that would guide the work of the consortium. The consortium formed ten specification-writing teams, each focused on a key aspect of the materials. The specifications cover: the pedagogical approach to be used in the units (instructional model and classroom routines); strategies for embedding professional learning opportunities within the instructional materials; strategies for assessment; strategies for supporting equity and access; techniques for implementing particular aspects of the NGSS (e.g., individual science and engineering practices and crosscutting concepts); the integration of English language arts and mathematics; and how to meet practical needs and constraints (e.g., cost of materials and use of technology). Each team was led by one or two researchers with expertise in the team's assigned area. Every team included both researchers and practitioners, including at least one teacher nominated by a partner state. The design specifications were reviewed by the Developers Consortium and the State Steering Committee, and were revised, prior to formal approval by the Steering Committee. The specifications have been released to the public as a product of the OpenSciEd project (OpenSciEd, 2019b).

Challenge 2: Draw on research but include the practitioner voice in all aspects of design work

For the development of instructional materials, all our teams have been staffed with researchers, developers, and classroom teachers. Teams often worked with partner teachers to pilot candidate anchoring phenomena and solicit feedback on directions. Field trials enabled us to bring in perspectives of hundreds of teachers from the ten partner states, beginning with their feedback in professional learning, and most importantly through data collected from them and their students while they enacted the units. Additionally, we established a design process that included a series of meetings with State Steering Committee members at multiple points of the development timeline. Following field trials of that unit, the design team reviewed all feedback from student and teacher data from field trials and a review of the unit from Achieve on its alignment with NGSS and developed a revision plan. The design team then presented the revision plan to the State Steering Committee for feedback before implementing their revisions.

Challenge 3: Support the instructional shifts of NGSS but avoid creating materials that teachers perceive as too foreign or feel unprepared to enact

The partner states vary in how far along they were in their NGSS implementation. Several states had been early adopters of the standards and had offered several years of professional learning for large numbers of teachers. Others were more recent adopters or had not been able to implement much professional learning. Furthermore, the goal of the state partners was to fully test the usability of the project's instructional materials at scale, and therefore not to restrict participants in field trials to only experienced "early adopter" teachers. Indeed, following four days of professional learning on the pedagogical approach and the specifics of the units they were preparing to teach, field trials teachers reported the materials posed challenging shifts for their current pedagogical practice.

The core commitments of the pedagogical approach, as outlined in the design specifications, reflected the key shifts in the reforms (e.g., Windschitl & Stroupe, 2017). The Framework and NGSS advocate for a shift from "learning about" science that others have established to "figuring out" the science by building the knowledge as a learning community (Schwarz, Passmore, & Reiser, 2017). This necessitates engagement in science and engineering practices as a fundamental approach for building science knowledge, by making sense of phenomena and solving problems Second engaging learners in knowledge building practices means that students should see their science work as addressing questions and problems they have identified, rather than simply following instructions from teachers or textbooks. Thus, each lesson needs to be coherent from the students' perspective, rather than solely motivated because teachers and curriculum designers know the lesson would be useful (Reiser, Novak, & McGill, 2017). Addressing these shifts is ambitious. However, we needed to avoid developing materials that required teachers to be already familiar with and bought-in to the new pedagogical approaches, or that would require professional learning beyond what the state partners said was reasonable to expect. We aimed to develop materials that would support teachers in making the incremental shifts needed to make progress on these goals. To accomplish this, we pursued several strategies to attempt to balance innovations in pedagogy with broad usability.

First, we realized we needed to address the interacting issues of classroom climate, classroom discourse, and instructional design in both curriculum materials design and professional learning. The systemic nature of the pedagogical shifts meant teachers needed to work on discourse moves that would support the shifts in task design that gave students more agency in developing questions, engaging in argumentation, and developing and revising models. We identified a limited set of classroom routines that we have used consistently across all units. We have emphasized them in both instructional materials and professional learning. We also incorporated educative features into the teacher materials to call out these general routines and help teachers apply them in context.

Second, we placed an emphasis on guidance for assessment in the instructional materials and professional learning. Recognizing that teachers are often subject to school or district policies governing assessments and grading, we provide guidance on how to use students' work throughout the unit as formative assessment opportunities, and we organized working sessions during professional learning to work with peers on assessment and grading strategies using the instructional materials. These strategies focused on using assessment to support the NGSS instructional shifts reflected in the materials, rather than in ways that might be in tension.

A third approach to helping teachers take on these materials as incremental steps toward shifting their practice involved striking a balance between being responsive to student ideas and expecting teachers to improvise based on the particular student ideas in their classroom. Engaging students in knowledge building that meaningfully addresses their own questions requires skill in cultivating student questions and careful listening to students' ideas (Windschitl et al., 2017). However, it can be challenging for teachers to weave together what they hear from students with what they know is the sequence of phenomena and investigations planned in the instructional materials. While having students feel like their own questions are driving their investigations may be an ideal which some teachers succeed in supporting (Reiser, Novak, & McGill, 2017), the goal in the OpenSciEd pedagogy is somewhat more modest. The instructional materials ask teachers to elicit student questions, and then play a more directed role in helping students understand where they need to go next to investigate them. In general, while the instructional materials provide ample opportunity for students to engage in practices to develop knowledge, we attempted to provide sufficient supports for teachers to enact more scaffolded versions of these practices than explored in some smaller scale design-based research studies.

Design of OpenSciEd professional learning

Katherine L. McNeill and Renée Affolter, Boston College

Teachers are central to classroom learning, yet learning sciences research has paid less attention to teachers than to students and other stakeholders (Fishman, Davis, & Chan, 2014). In the context of science, the shifts in learning called for by the NGSS requires teachers to take on new roles (Windschitl & Stroupe, 2017). Teachers need professional learning experiences to support those changes (National Academies, 2015). Teachers, instructional leaders and other stakeholders in the K-12 education system may not recognize the teacher changes that need to occur, such as supporting sensemaking from the student perspective. Yet the adoption of new curriculum materials can create the need, space or resources for professional learning opportunities for teachers. Teachers may ask for support around the new curriculum materials. Furthermore, previous research suggests that the strongest student outcomes occur when teachers receive both new curriculum materials and professional learning (PL) (Lynch et al., 2019).

We use the ***practical*** needs of teachers for curriculum-based professional learning as an on ramp for PL that also addresses the professional learning needs of the teachers implementing ***research-based*** shifts in practice. Furthermore, we have attended to the challenges of ***marketability*** for a program that requires professional learning by attempting to meet school and district policies and norms for professional learning as much as possible. For example, one of the most challenging constraints for practicality and marketability is the amount of time that is available for teachers to participate in face-to-face PL. Although research suggests that PL is most effective when it is sustained over extensive time (Desimone, 2009), teachers and schools have limited time available for it. These design constraints have resulted in the Professional Learning Design Framework that guides our development of professional learning programs and resources for the OpenSciEd program (see Table 1).

<u>Table 1: OpenSciEd Professional Learning Design Framework</u>

Professional Learning Element	Description
1. Focus on equitable sensemaking	Threaded across all of the professional learning experiences is a focus on more equitable learning in science in which all students are known, heard and supported with access and opportunities for learning.
2. Frame experiences with a problem of practice	Each professional learning session begins with a "problem of practice" that frames the multi-day experience. These problems are grounded in challenges currently being experienced by the field test teachers in order to address an area of need.
3. Provide images of classroom instruction	Classroom videos and student artifacts are used when in the PL to illustrate classroom instruction with a range of students. These images highlight both what is possible and challenges that can arise that teachers can draw from for their own classrooms.
4. Offer the student perspective	Teachers are asked to experience 3D science instruction in "student hat", such as developing models and participating in sensemaking discussions. These experiences support teachers' understanding of the instructional model and increase their empathy for their students.

5. Engage teachers to collaborate and reflect with peers	All of the professional learning activities occur in a collaborative environment in which teachers work together to better understand the curricular innovation and to reflect on implications for their own classrooms.

In this section, we focus on Element 2 - *Frame Experiences with a problem of practice* – which directly targets the practical needs of teachers and schools. Table 2 presents a transcript from the four-day introductory professional learning workshop in which middle school teachers were learning about their first OpenSciEd unit, *How do things inside our bodies work together to make us feel the way we do?* The unit explores the case of M'Kenna who is having health issues and the students work to figure out what is causing those issues. The teachers had just completed an investigation from the student perspective in which they explored how dialysis tubing has properties that are similar to the surface of the small intestine. The teachers were then asked in "student hat" to share out what they had figured out so far about what is causing M'Kenna's health issues. In sharing out, one participant in "student hat" uses the word "absorbed" as part of their explanation. Another teacher (Participant 4) responds moving out of "student hat" to their own "teacher hat" stating "Yeah I can never assume that my kids have learned that," which is the beginning of the transcript below.

Table 2: Transcript from professional learning

Facilitator:	Yep. So let's assume not so how could we write it without the word "absorbed"?
Participant 1:	There's more glucose and complex carbohydrates in M'Kenna than the healthy person.
Participant 5:	Leftover or, or still there.
Participant 2:	Still in the intestines. Not in the middle.
Participant 1:	And should we all be writing this down in what we figure out?
Facilitator:	In a moment, like see how I crossed it out. 'Cause kids don't want to erase right? There's more glucose and complex carbs left in M'Kenna's small intestine than the healthy person. Is that okay?

In this section of the transcript, the participants work together to construct language that they think may align with what one of their students might say as they explain the phenomenon. The facilitator supports them in building this conceptual understanding without focusing on the scientific language. This move supports teachers in thinking about how to increase access to more students using the resources they bring rather than prioritizing academic language. This aligns with research suggesting the importance of students figuring out phenomena from their own perspective in order to support greater student coherence and epistemic agency (Zivic, et al., 2018). Yet about three minutes later in the same conversation we see another teacher push back on this suggesting they needed to use the word "absorb" because of practical needs. Participant 6 states

> *So my pushback in using that word is it, it is a 6th grade standard. And I do think that we should. If we're driving. If we're preparing them for 9th grade, 10th grade, preparing them for the [state test], I do think it's an appropriate time...I do think it's an appropriate time to start to scaffold up, put a little rigor in here, and use appropriate vocabulary. Like we can't just like shy away from like, we have to really start to do that. It's in preparation for bigger things to come.*

We see Participant 6 concerned about preparing middle school students for the state test and for high school so that it is important to "scaffold up" the academic or scientific language. This example illustrates a tension we see at times during the PL. The practical needs or perceived needs of teachers can be in conflict with research about student learning. In the professional learning design, we try to balance those tensions in order to continue to move teachers forward on their own learning trajectory while being aware of the practical influences on their classroom instruction.

Collection and analysis of data

Andrew E. Krumm, University of Michigan, and William R. Penuel, University of Colorado - Boulder

Quality implementation has long been cited in the learning sciences as critical to the success of any instructional intervention (e.g., Cobb & Jackson, 2012). To understand implementation and to inform the revision of instructional materials and professional learning opportunities more broadly, OpenSciEd organized an overarching field test based that draws on implementation science (IS) and quality improvement (QI) approaches (Bryk, Gomez, Grunow, & LeMahieu, 2015; Nilsen, 2015). While stemming from different intellectual traditions, IS and QI address a common concern when compared against more traditional research and evaluation: IS and QI are about making innovations work in specific contexts as opposed to determining whether an innovation can or does work (Grunow, Hough, Park, Willis, & Krausen, 2018).

The purpose of the OpenSciEd field test is not to support generalizable research claims or provide a high-stakes evaluation of materials of learning opportunities (Solberg, Mosser, & McDonald, 1997). Instead, the goal is to continuously improve materials and identify the types of capabilities needed by both teachers and schools to use the materials well (Peurach, 2016). One implication of this purpose is that the role of participants is different from a traditional research study. In a traditional research study, participants are kept at arm's length and are not active co-designers (Penuel & Gallagher, 2009). Another implication is for the instruments used to collect data (Yeager, Bryk, Muhich, Hausman, & Morales, 2013). In a traditional study or evaluation, instruments must adhere to disciplinary standards of validity and reliability for the claims that researchers want to make (e.g., broadly generalizable or high stakes). Given the stage of development for the instructional materials, the field test team privileged instruments that could be used easily *in* practice, *about* practice, and *by* educators that could generate data that was useful to developers and state leaders in serving multiple purposes. We used data collection instruments that were practical to administer to students and that were aligned with the instructional model used in the curriculum materials. Practicality was important because teachers were not offered a reward for providing data, and because we wanted teachers to focus their attention on implementation, not data collection. Aligning data collection to the instructional model could inform re-design in ways that general fidelity or stages of concern approaches to the study of implementation could not, because model-specific data could explicitly address elements of the instructional materials that developers conjectured were important to learning.

Therefore, the field test was shaped by multiple constituencies: developers (curriculum and professional learning) and state science leads. State leaders had "symbolic" uses for data generated through the field test that entailed helping stakeholders in a given state buy-in to the process; state leaders also had a "conceptual" use that involved helping to shift district and school leaders' thinking about quality science instructional materials. Developers, on the other hand, had more "instrumental" uses that entailed informing redesign of materials, teacher tools / guides, and professional learning opportunities (Weiss & Buculavas, 1980).

Early in 2018, members of the field test team proposed five principles by which the field test would operate, which were reviewed and modified by developers and state leads prior to their formal adoption for the project. These principles were informed by the field test team's prior work along with tools and processes from QI and IS. They were: (1) *Minimize burden*: Collect only data that is needed by someone, for a purpose agreed upon as important by different stakeholder groups. (2) *Foreground equity*: Address questions about students' access to prepared teachers, challenging activities, and experiences of the classroom. (3) *Support learning*: Learn from the pilot to improve equitable implementation and outcomes for all. (4) *Instrument practice*: Collect data directly from practice. (5) *Be meaningful and actionable to end user*: Who will be working with the data and what will they do with it? Given the improvement and implementation focus of the field test, the consortium paid particular attention to the burdensomeness of data collection approaches (see principle 1, above). The field test team worked with state leads to identify that the appropriate amount of time for interviews and surveys was 30 and 15 minutes, respectively. To further attend to burden, we only asked to teachers of multiple sections to collect data in one section, and we distributed data collection tasks across teachers so individual teachers did not have to conduct all data collection tasks.

There is evidence that developers have made use of data from the field test to inform both unit-level and specific lesson revisions. A recent survey (n = 7) of writers indicated that they used survey data to focus more effort on developing assessments teachers said they needed, provide more tips for sustaining student interest throughout units, and support engagement in science practices and crosscutting concepts. They relied heavily on student artifacts to make more specific changes to lessons. Among the changes they reported making were to clarify confusing directions, modify pacing guidance, and removing activities teachers found did not reliably work. These survey data provide some preliminary evidence of one key aim of data collection, namely to inform improvements to materials.

Endnotes
(1) Authors are listed in order they will present in the symposium.

(2) The term *open content* describes a work that is released to the public under a license that allows anyone to freely use, distribute, and modify the work.

References

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Education Press.

Cobb, P., & Jackson, K. (2012). Analyzing educational policies: A learning design perspective. *Journal of the Learning Sciences, 21*(4), 487–521.

Desimone, L.M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.

Fishman, B. J., Davis, E. A., & Chan, C. K. K. (2014). A learning sciences perspective on teacher learning research. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 707–725). New York, NY: Cambridge University Press.

Grunow, A., Hough, H., Park, S., Willis, J., & Krausen, K. (2018). *Towards a common vision of continuous improvement for California*. Getting Down to Facts II (Technical Report). Stanford, CA: Stanford University.

Ikemoto, G. S., & Honig, M. I. (2010). Tools to deepen practitioners' engagement with research: The case of the Institute for Learning. In C. E. Coburn & M. I. Honig (Eds.), *Research and practice in education: Building alliances, bridging the divide* (pp. 93-108). Lanham, MD: Rowman & Littlefield.

Lynch, K., Hill, H. C., Gonzalez, K. E & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis, 41*(3), 260-293.

National Academies of Sciences, Engineering, and Medicine. (2015). *Science teachers learning: Enhancing opportunities, creating supportive contexts.* Washington, DC: The National Academies Press.

NGSS Lead States (2013) *Next Generation Science Standards: For states, by states*. Washington, DC: The National Academies Press.

Nilsen, P. (2015). Making sense of implementation theories, models, and frameworks. *Implementation Science, 10*(53), 1-13.

OpenSciEd. (2018a) OpenSciEd project overview: Next generation open source science instructional materials. Unpublished document.

OpenSciEd. (2019a). OpenSciEd middle school scope and sequence. https://www.openscied.org/scope-sequence/

OpenSciEd. (2019b). OpenSciEd design specifications. https://www.openscied.org/design-specifications/

Penuel, W. R., Briggs, D. C., Davidson, K. L., Herlihy, C., Sherer, D., Hill, H. C., . . . Allen, A.-R. (2017). How school and district leaders access, perceive, and use research. AERA Open, 3(2), 1-17.

Penuel, W. R., & Gallagher, L. P. (2009). Preparing teachers to design instruction for deep understanding in middle school earth science. *Journal of the Learning Sciences, 18*(4), 461-508.

Peurach, D. J. (2016). Innovating at the nexus of impact and improvement: Leading educational improvement networks. *Educational Researcher, 45*, 421–429.

Reiser, B. J., Novak, M., & McGill, T. A. W. (2017). Coherence from the students' perspective: Why the vision of the Framework for K-12 science requires more than simply "combining" three dimensions of science learning. Paper commissioned for the BOSE workshop "Instructional materials for the NGSS."

Solberg, L., Mosser, G., & McDonald, S. (1997). The three faces of performance measurement: Improvement, accountability, and research. *The Joint Commission Journal on Quality Improvement*, *23*(3), 135–147.

Schwarz, C. V., Passmore, C. M., & Reiser, B. J. (2017). Moving beyond "knowing" science to making sense of the world. In C. V. Schwarz, C. M. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world through next generation science and engineering practices* (pp. 3-21). Arlington, VA: NSTA Press.

Weiss, C. H., & Bucuvalas, M. J. (1980). *Social science research and decision-making*. New York, NY: Columbia University Press.

Windschitl, M. A., & Stroupe, D. (2017). The three-story challenge: Implications of the Next Generation Science Standards for teacher preparation. *Journal of Teacher Education*, *68*(3), 251–261.

Windschitl, M., Thompson, J., Braaten, M., & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*, 96(5), 878-903.

Zivic, A., Reiser, B. J., Edwards, K. E., Novak, M., & McGill, T. A. W. (2018). Negotiating epistemic agency and target learning goals: Supporting coherence from the students' perspective. In J. Kay & R. Lukin (Eds.), 13th international conference of the learning sciences (ICLS) (Vol. 1, pp. 25-32). London, UK: International Society of the Learning Sciences.