Association for Psychological Science

Psychological Science in the Public Interest
2021, Vol. 22(1) 8–43

DOI: 10.1177/1529100620973974
www.psychologicalscience.org/PSPI
SAGE

# The Curious Construct of Active Learning

**Doug Lombardi[1], Thomas F. Shipley[2], Astronomy Team, Biology Team, Chemistry Team, Engineering Team, Geography Team, Geoscience Team, and Physics Team***

[1]Department of Human Development and Quantitative Methodology, University of Maryland, College Park, and [2]Department of Psychology, Temple University
*All team members are listed in the transparency section at the end of the article.

**Abstract**

The construct of active learning permeates undergraduate education in science, technology, engineering, and mathematics (STEM), but despite its prevalence, the construct means different things to different people, groups, and STEM domains. To better understand active learning, we constructed this review through an innovative interdisciplinary collaboration involving research teams from psychology and discipline-based education research (DBER). Our collaboration examined active learning from two different perspectives (i.e., psychology and DBER) and surveyed the current landscape of undergraduate STEM instructional practices related to the modes of active learning and traditional lecture. On that basis, we concluded that active learning—which is commonly used to communicate an alternative to lecture and does serve a purpose in higher education classroom practice—is an umbrella term that is not particularly useful in advancing research on learning. To clarify, we synthesized a working definition of active learning that operates within an elaborative framework, which we call the *construction-of-understanding ecosystem*. A cornerstone of this framework is that undergraduate learners should be active agents during instruction and that the social construction of meaning plays an important role for many learners, above and beyond their individual cognitive construction of knowledge. Our proposed framework offers a coherent and actionable concept of active learning with the aim of advancing future research and practice in undergraduate STEM education.

**Keywords**

active learning, STEM education, student agency

College instructors in science, technology, engineering, and mathematics (STEM) have hailed *active learning* as a way to transform science education. A widely cited 2014 meta-analysis ($N = 225$ studies across eight STEM disciplines) provided a solid evidentiary basis for the benefits of active learning (Freeman et al., 2014). The authors concluded that "active learning leads to increases in examination performance" and decreases in failure rates compared with traditional lecture (Freeman et al., 2014, p. 8410). Further impetus for active learning has been added by evidence suggesting that active-learning strategies provide more equitable outcomes for under-represented students in undergraduate biology education (Ballen et al., 2017; Eddy & Hogan, 2014; Haak et al., 2011).

However, a variety of instructional activities fall under the active-learning moniker: participation in flipped classrooms, use of clickers or other student response systems, engagement with computerized conversational agents, and much more. For some, active learning describes any undergraduate STEM instruction other than traditional lecture in which students sit down, listen to an instructor, and take notes. Despite the uncertainty in exactly what is meant by active learning, such findings have led to changing the culture and design guidance for university classrooms to promote student interactions (see, e.g., Stains et al., 2018). The National Research Council (2012a) convened a committee on the status, contributions, and future directions of discipline-based education research (DBER) and found that "either implicitly or explicitly, the principle of active learning

**Corresponding Author:**
Doug Lombardi, Department of Human Development and Quantitative Methodology, University of Maryland
E-mail: lombard1@umd.edu

has had the greatest influence on DBER scholars and their studies" (p. 121).

The concept of active learning has been around for several decades. For example, *How People Learn* (Bransford et al., 2000) contained a two-page section on active learning suggesting that it involves students taking control of their learning through some level of metacognitive sense making, self-assessment, and reflection. *How People Learn II* (National Academies of Sciences, Engineering, and Medicine, 2018) did not add greater specificity, but it characterized active learning as nested within the two familiar pedagogical approaches of student-centered and inquiry-based learning (see, e.g., Kuh, 2008).

Given the uncertainty about the meaning of active learning—and about whether all activities under this moniker represent strong teaching strategies or only a subset are superior strategies that raise the entire portfolio—the following questions arise: What characterizes effective learning processes in undergraduate STEM fields, and what role does active learning play in these processes? What concrete guides have emerged from active-learning research that elucidate promising practices in undergraduate STEM instruction? Are there active-learning strategies that optimize learning in some situations but not others? Is lecture—the predominant teaching strategy in undergraduate STEM instruction (Stains et al., 2018)—inherently flawed, or are there some circumstances under which students can actively learn during lecture (see, e.g., Smith et al., 2011, which showed that student learning was deepened when lecture-based instruction was supplemented with active-learning methods)?

The purpose of this review is to examine what works, what does not, and what we still do not know with respect to active learning. It aims to bring clarity to the term and provide a framework for active learning, supported by learning theory and principles, in the undergraduate STEM context. The authors come from the fields of educational and cognitive psychology and from DBER in astronomy, biology, chemistry, engineering, geography, geoscience (an umbrella term for atmospheric, Earth, and ocean sciences), and physics. Thus, we leverage a cross-DBER alliance to build knowledge that does not exist in any single domain (Henderson et al., 2017; Shipley et al., 2017). Specifically, we approached the idea of active learning from historical perspectives of educational research emerging from psychology and from the STEM disciplines (i.e., DBER). We synthesized these two perspectives in a convergent design with the hope of more clearly characterizing the notion of active learning, and we present a framework to assist researchers and practitioners in differentiating which modes of instruction are likely to help particular learners and which are not.

In addition to summarizing and evaluating extant evidence and theoretical positions to guide practice, we identify five key takeaways and offer suggestions from each to advance undergraduate STEM instruction. We list these takeaways here and revisit them more explicitly toward the end of the article:

1. Multiple things from differing ontological categories (e.g., psychological and social constructs; instructional pedagogies, strategies, and design principles) have been tagged as active learning. Further, active learning is commonly used as an umbrella term to communicate an alternative to lecture, and as such, it serves a purpose in higher education classroom practice. However, active learning as an umbrella term is not particularly useful in advancing research on mechanisms of learning in undergraduate STEM.
2. This review unpacks many ideas that have been placed under the active-learning umbrella and presents them in a coherent and useful framework motivated by DBER findings. This framework, which we call the *construction-of-understanding ecosystem*, was developed through a collaboration of psychologists and interdisciplinary DBER teams.
3. Trying to capture meaningful regularities in learning, given the complex nature of learning systems, results in a tension between reporting inductive hypotheses that offer actionable guidance and evaluating studies against a "gold-standard" experimental metric of randomized controlled trials. Overall, the exact value of any given instructional strategy that falls under the active-learning moniker is still an open question. Therefore, we review the available DBER evidence, including both principle-based design analyses and outcome comparisons across courses.
4. The idea that undergraduate learners should be active agents during instruction is important and serves as a cornerstone of our framework for active learning.
5. An emergent pattern is that the social construction of meaning plays a key role for many learners, above and beyond the individual cognitive construction of knowledge. Traditional lecture relies mostly on the individual for sense making, which may privilege fewer learners (i.e., only those who can successfully make meaning on their own or establish social meaning making outside of the lecture).

Finally, we report on the process by which this review was constructed; that is, we reflect on our explicit attempt to transcend traditional academic silos to do cross-DBER science in constructing new knowledge about active learning, as well as novel and innovative knowledge about undergraduate learning processes. We acknowledge that there have recently been excellent reviews that have closely and critically examined active learning within the context of psychology instruction (see, e.g., Bernstein, 2018; Liu & Beaujean, 2017). But although we also acknowledge that individuals within a particular discipline (e.g., psychology) can be effective at thoroughly assimilating literature outside their communities, doing so is somewhat rare. More often, researchers stay insulated within a particular disciplinary literature because it is more accessible. Our novel contribution in the present review and analysis is that we have synthesized findings on STEM learning from a wide set of disciplines. In short, we aim to provide a coherent and actionable concept of active learning that incorporates these diverse literatures. We also aim to offer guidance on research and practice, as well as suggestions for how to construct knowledge across disciplines for an interdisciplinary research community.

## Past Views on Active Learning and Its Measurement

In this section, we present a convergent synthesis of historical psychological and DBER perspectives, detailing how many researchers have characterized the notion of active learning along two dimensions—as a generalized instructional process (a) for constructing knowledge and (b) for deepening engagement. We then describe how these two dimensions laid the foundation for our framework for active learning. Finally, we describe other characterizations of active learning as (a) antithetical to passive learning and (b) antithetical to lecture, and we discuss some of the issues they raise, particularly in relation to measurement.

### *Active learning as knowledge construction and deepening engagement*

Researchers began applying the metaphor of knowledge construction as a result of the cognitive revolution in the late 1950s and 1960s (see, e.g., Chomsky, 1959; Miller, 1956; Newell et al., 1958). In this metaphor, individuals build their knowledge and understanding by encoding information (e.g., sensory perceptions) into mental representations that, in turn, are stored in mind for later use (Mayer, 1992; Ormrod, 2017). Many debates about the nature of learning revolved around whether knowledge is constructed (a) cognitively and individually or (b) socially and culturally. In the former camp were proponents of cognitive constructivism, which emerged from theories about schemata (i.e., mental structures that organize knowledge; R. C. Anderson et al., 1977) and information processing (i.e., sensory information is processed before storage in long-term memory; J. R. Anderson, 1983). In the latter camp were proponents of social constructivism, which emerged from sociocultural (Bakhurst, 1995) and sociological (Blumer, 1986) perspectives.

Some researchers have adopted a cognitive-constructivist stance when examining active learning in undergraduate STEM education, whereas others have adopted a social-constructivist stance. For example, in a recent review of active learning in undergraduate science education, Arthurs and Kreager (2017) claimed that the opportunity to "actively engage with the subject matter as well as with other individuals" results in "deep learning and construction of knowledge," per Vygotsky's (1934/1968) social-constructivist perspective (Arthurs & Kreager, 2017, p. 2075). However, many active-learning researchers do not explicitly identify a singular constructivist stance, which may have contributed to confusion about the active-learning construct. Freeman et al. (2014), for example, said that "theories of learning that emphasize the need for students to construct their own understanding have challenged the theoretical underpinnings of the traditional, instructor-focused, 'teaching by telling' approach" (p. 8410), citing both cognitive-constructivist foundations (i.e., Piaget, 1926) and social-constructivist foundations (i.e., Vygotsky, 1978). Taken separately, these perspectives guide researchers to record and measure different aspects of learning situations, and individually, they offer incomplete accounts of the classroom environment (Marshall, 1996). Therefore, to capture the complexity of learning, researchers and instructors would do well to view it through an integrated, but not conflated, accounting of cognitive and social-constructivist stances.

Many researchers describe active learning as a way to increase student engagement in classroom, laboratory, or field experiences (see, e.g., Järvelä & Renninger, 2014). Perceived as the "the holy grail of learning," engagement is often related to positive learning outcomes, possibly because it increases students' attention, interest, optimism, and purposive self-regulation (Sinatra et al., 2015, p. 1). To unpack this multifaceted concept, some researchers have characterized engagement as domain specific and manifesting along four dimensions:

- social-behavioral, involving participation in learning activities with classmates (Linnenbrink-Garcia et al., 2011),
- cognitive, stemming from a willingness to engage in effortful thinking, purposiveness, strategy use,

and self-regulation in the learning process (Fredricks et al., 2004);

- emotional, involving positive feelings (e.g., enjoyment), a sense of belonging, and perceptions of learning tasks as valuable (Furrer & Skinner, 2003); and
- agentic, whereby learners recognize and implement their role as constructors of knowledge for both themselves and their peers (Reeve & Tseng, 2011).

High levels of social-behavioral, cognitive, and emotional engagement in science-related learning tasks may lead to higher academic achievement in science and greater sustained interest, persistence, and retention in STEM (see, e.g., Cleary & Zimmerman, 2012; Greene et al., 2004; Heddy & Sinatra, 2013; Reschly & Christenson, 2012).

Agentic engagement, which has been characterized as "students' proactive, constructive attempts to assert their agency and influence the flow of instruction" (Patall et al., 2019, p. 77), may be a particularly important component in undergraduate education. In a recently proposed framework, Patall et al. (2019) suggested that students' social-behavioral, emotional, and cognitive components of engagement deepen their agentic engagement by increasing competency in learning tasks, personal connectedness to the learning topics and subject, and autonomy in constructing meaning. Patall et al. (2019) also posited that increased levels of students' agentic engagement can help create classroom environments where the teachers increasingly support their students' autonomy for learning. Thus, increasing students' agentic engagement via autonomy, competence, and connectedness can lead to increased support for autonomy from teachers, which, in turn, can loop back to increase students' agentic engagement.

### *Active learning as the antithesis to passivity and lecture*

Chi (2009) acknowledged the potential vagaries of the active-learning construct. To establish greater precision and clarity—and, in particular, to differentiate active learning from passive learning—Chi and Wylie (Chi, 2009; Chi & Wylie, 2014) developed a framework of learners' overt activities in classroom settings. In this framework, situations that include activity are better than situations in which the student is not overtly doing anything (e.g., being passive; Chi, 2009). Although this framework focused on behaviors, it connected different types of learning to underlying cognitive processes. With the instructional goal of engaging the learner, activity in and of itself could result in a greater likelihood of engagement. Chi and Wylie (2014) made the case that active learning is better (i.e., it results in deeper cognitive engagement) than passive learning, but only incrementally so. In their framework, active learning (as defined by movement) is bettered by constructive learning (in which the learner is "producing outputs that contain ideas that go beyond the presented information"), which is in turn bettered by interactive learning (in which the learner is "dialoguing substantively on the same topic, and not ignoring a partner's contributions"; Chi, 2009, p. 77). Many types of instructional activities may fall under these categories, including learners' generation of self-explanation when prompted and peer instruction during group work. For example, students' participation in laboratory activities could be categorized as passive, active, constructive, or interactive depending on the nature of their instruction and their amount of cognitive engagement (Chi & Wylie, 2014). It is important to note, however, that the active-learning literature typically does not distinguish among the active, constructive, and interactive modes of engagement and incorporates strategies that encompass all three. Finally, Chi and Wylie's (2014) framework focuses primarily on the learner as an individual.

In line with Chi and Wylie (2014), we recognize the value of developing a better understanding of the overtly active construction of knowledge; however, in our posited framework, effective undergraduate learning contexts may involve situations in which (a) the learner is centrally responsible for sense making but might appear passive (e.g., during a lecture or study session in which the learner is carefully paying attention) and (b) there is collaborative responsibility for cooperative sense making (e.g., during a team-based session in which instructors and learners are coordinating learning; see, e.g., Liu & Beaujean, 2017).

Active learning is often contrasted with lecture. For example, in their meta-analysis, Freeman et al. (2014) compared lecture (i.e., "teaching by telling") to active learning (p. 8410), specifically focusing on studies conducted in undergraduate STEM fields (astronomy, biology, chemistry, computer science, engineering, geoscience, mathematics, natural resources or environmental science, nutrition or food science, physics, psychology, and statistics). For their analysis, Freeman et al. (2014) operationalized lecture as "continuous exposition by the teacher" during which students would be "taking notes and/or asking occasional and unprompted questions" (p. 8414).

Although Freeman et al.'s (2014) meta-analysis served as a comprehensive step in assessing the value of active learning, it left some important questions unanswered. For example, are all active-learning activities effective in promoting desired learning outcomes? (See, e.g., Bernstein, 2018, who suggested that research should go beyond asking whether active learning works and address what general factors might account for stronger or weaker learning activities.) Overall, Freeman et al.'s (2014) findings offer a strong impetus to refine our understanding of active learning to effectively guide future research and practice in undergraduate STEM instruction. However, because we do not have an accepted cognitive model to account for the observed benefits of active learning in undergraduate STEM education, a more detailed characterization of the construct is warranted.

A dichotomy that contrasts active learning with lecture would seem to overlook characteristics of lecture that could strongly promote learning. Many proponents of active learning see some value of the lecture mode in active-learning classrooms (see, e.g., Deslauriers et al., 2019; Smith et al., 2011; J. D. Walker et al., 2008). And although we stress that continuous lecture in a classroom setting, regardless of the lecturer's charisma and clarity, almost certainly results in less desirable learning outcomes compared with more active modes of instruction, we suggest that there may be situations and contexts in which lecture is useful and beneficial (i.e., a mode of active lecture; Bernstein, 2018). Such an approach is consistent with the guidance offered in *How People Learn II* (National Academies of Sciences, Engineering, and Medicine, 2018) about observational learning. The report specifically states that "people also learn by observing . . . others' behavior, attitudes, or emotional expressions, with or without actually imitating the behavior or skill" and that such observational learning exercises the "cognitive capacities for imitation, interpretation, and inference" (National Academies of Sciences, Engineering, and Medicine, 2018, p. 41). Likewise, Rogoff et al. (1996) suggested that students are learning all the time, and even if they are not fully engaged in a learning task, they might be learning about such things as norms of behavior, power structures, or epistemic authority.

It may be that traditional lecture-based learning limits the diversity of students who engage in STEM content and consider STEM careers. Steele suggested that underrepresented groups receive messaging—predominantly through lecture environments—that they do not belong in STEM and other academic settings (Steele, 1995; Steele & Aronson, 1997). However, Lundberg et al. (2018) found that Lantinx students were more successful learners when they had high-quality interactions with faculty (including rapid and supportive feedback indicating that they could improve and that the instructor knew they were capable) and interactions outside of the classroom. Ovink and Veazey (2011) also found that students from groups that are underrepresented in the sciences were more likely to persist in their studies if they were part of a community—in this case, a university-sponsored program in which they collectively developed academic skills both in and out of the classroom. These studies imply that deep learning may be best facilitated in situations where students have opportunities to construct knowledge both socially (e.g., through participation in a collaborative group) and individually (e.g., through listening to a lecture)—for example, a class in which groups of students first differentiate knowledge about a phenomenon through contrasting cases and then listen to a follow-up lecture that provides the scientific explanation for the phenomenon (Schwartz & Bransford, 1998).

## *Active learning and measurement*

One important consideration for any formal theory of active learning is the measurement of learning—for instance, through concept inventories (CIs), unit tests, and examinations. Measurements of learners' responses to tasks (of various types) are typically used to support claims about the efficacy of instructional strategies or curricular transformations. For example, Freeman et al. (2014) noted that studies of active learning that used a CI score as the outcome variable tended to report greater effect sizes than studies in which learning was measured with instructor-authored exams. In explaining this difference, Freeman et al. (2014) suggested that CIs assess "higher-level cognitive skills" better than instructor-authored exams, which tend to focus on "content mastery or the ability to solve quantitative problems" (p. 8411; see also Momsen et al., 2010). This difference may be a function of the rigorous design process undertaken for the development of many CIs (Bailey, 2020). The selection of an appropriate assessment is a critical component of both instructional and research design. CIs tend to be effective for providing a broad overview of students' current understanding of the topic, relative to their understanding before instruction, which allows instructors to target particular areas of confusion and identify changes over time; however, they are not generally intended to measure individual understanding in the same way as, say, a final exam. CIs can also be useful for comparisons across instructors, programs, or schools (Bailey, 2020; Smith et al., 2019), but CIs, unit tests, and examinations are not (and should not be treated as) interchangeable. Different types of measures, such as open-ended prompts, may provide insights into learning complexities involving concepts that are in the process of construction but are not fully aligned with scientific understanding (see, e.g., Prevost et al., 2016).

Other assessments that emphasize factual recall, whether instructor or community authored, may not reflect whether students can use their knowledge in practices characteristic of work in science (M. M. Cooper et al., 2017; M. M. Cooper & Stowe, 2018; Laverty et al., 2016; Stowe & Cooper, 2019; Underwood et al., 2017). Many students perform very well on such assessments without a robust and useful understanding of foundational processes—for example, students may do well on chemistry assessments without understanding forces between molecules (M. M. Cooper et al., 2015; Williams et al., 2015). To elicit evidence of what students know and can do requires explicitly asking them for the reasoning underpinning constructed or selected explanations, predictions, and models. Multitiered CIs begin to elicit such reasoning (Caleon & Subramaniam, 2010; Chandrasegaran et al., 2007), but other assessment strategies may be better suited to address this issue.

Alignment within the design and interpretation of active-learning research is of critical importance. Researchers and instructors should be sure that learning goals are guiding the selection of learning strategies and that assessments are appropriate for those learning goals (M. M. Cooper et al., 2017; M. M. Cooper & Stowe, 2018; Laverty et al., 2016; McDaniel et al., 2016; Stowe & Cooper, 2019; Underwood et al., 2017). In addition, within a curriculum, courses have different but interrelated pedagogical goals to support a well-rounded education. Thus, different measures may be required across the larger curriculum to determine the extent to which this wider range of goals have been met. Finally, an important consideration is how data from CIs and other measures are analyzed. Focusing only on correctness or measures of gain may limit researchers' and instructors' understanding of learning (e.g., the type and nature of incorrect answers students give reveals little about the various types of thinking that could have led to the answers; Brewe et al., 2016; Wallace & Bailey, 2010; Wells et al., 2019; Yang et al., 2020).

## Active Learning in Current Instructional Practice

The history of the term *active learning* and associated assessment in psychology and DBER offers context for how active learning is currently practiced in undergraduate STEM education. An overview of that history emerged from the DBER teams' white papers and helped us to formulate the various components of the active-learning framework that we present in a later section ("Active Learning: A Synthetic View"). It should be noted that most of the studies in the active-learning literature focus on a single course or a set of common courses (e.g., multiple introductory astronomy courses). This pattern suggests that the state of active learning captured in any one study is a snapshot and may not reflect what happens over a complete program (e.g., moving from large introductory courses that are primarily lecture-based to upper-division courses in which students conduct independent research). Likewise, studies that have compared lecture-based learning to active learning may not have considered corequisite labs, recitations, or discussion sections, or they may have considered those ancillary components of a course independently from the large lecture.

### *Characterizing current lecture-based and active-learning practices*

As researchers have recognized the need to better understand the nature of student-instructor interactions in the classroom, observation protocols have been developed to increase the precision of their descriptions. Examples include the Reformed Teaching Observation Protocol (RTOP; Sawada et al., 2002), the Teaching Dimensions Observation Protocol (Hora et al., 2013), the Classroom Observation Protocol for Undergraduate STEM (COPUS; Smith et al., 2013), the Practical Observation Rubric to Assess Active Learning (Eddy et al., 2015), and the Measurement Instrument for Scientific Teaching (Durham et al., 2017). These protocols document the nature and frequency of multiple instructional practices, include categories of active-learning strategies, and can be helpful both for research purposes and to provide feedback to instructors on their practices.

Relative to middle and high school, college involves more lecturing, even in the first year (Akiha et al., 2018). A large-scale survey of class time across STEM courses found that lecturing accounted for about 75% of the time sampled; courses allotting 80% or more of class time to lecture represented approximately 60% of the surveyed courses in chemistry, physics, and engineering and approximately 50% of those in biology, geoscience, and math (Stains et al., 2018). Within individual fields, the best data on classroom practices comes from Henderson and Dancy (2009), who surveyed 722 physics faculty in the United States and reported on the percentage who were aware of and who were currently using research-based instructional strategies (RBIS; e.g., use of clickers). Almost all physics faculty (87%) were familiar with one or more RBIS, and about half of all faculty were familiar with six or more. Among the most commonly used strategies, several were lecture based (peer instruction, ranking tasks, and interactive lecture demonstrations), although one, cooperative group problem solving, could be implemented in a variety of settings. Likewise, this review's biology team collected survey data and found that biology instructors ($N$ = 105) reported using specific strategies such as metacognition (45% of survey respondents), discussion (34%), and group work (29%; Driessen et al., 2020). Finally, Stains et al. (2018) found that STEM faculty employed varied teaching practices within a given course, but lecture remained the prominent mode of instruction for undergraduate classes.

These findings suggest that the most common approach taken by faculty is to make small changes to lectures to promote interactivity. Note that the biology team also found that it can be hard to evaluate the efficacy of specific strategies because in manuscripts about active learning in classrooms, specific strategies are lumped into general categories, such as group work and group discussion (Driessen et al., 2020). Having a clear understanding of what strategies are being employed and how they are being tested—as well as the alignment between the two—will be important going forward.

Although Henderson and Dancy (2009) documented the widespread knowledge and use of RBIS in physics, they also found that faculty often substantially modify the implementation of those methods (e.g., skipping a clicker vote after posing a question), and about a third of the time, faculty discontinue using an RBIS because they claim it did not work for them or their particular situation (Henderson et al., 2012). They further found that, contrary to popular beliefs about RBIS in physics, RBIS knowledge and use were unrelated to faculty age, institutional type, and percentage of job responsibilities related to teaching (Henderson et al., 2012). Prather et al. (2009), in looking at the relationship between strategy use and scores on the Light and Spectroscopy Concept Inventory, found that participating instructors reported spending between 0% and 49% of class time using active-learning strategies. The researchers made the following two inferences from their results: First, students must spend at least 25% of class time on active-learning strategies to make the highest gains (defined as normalized gain $< g > \geq 0.30$), and second, variation in the results suggests that the quality of implementation of active-learning strategies may be an important factor in determining if a class will experience a gain in scores. Another important factor may be variations in buy-in from the students, or fidelity of implementing in a way that promotes knowledge construction by the student. A review by Henderson and Dancy (2007) of how and under what circumstances instructors adopt more active-learning practices found that such changes were most likely to occur when programs were guided to make changes by administration and were less likely to be spontaneously adopted through diffusion from peer adopters. Henderson et al. (2011) recommended strategies that revolved around developing reflective teachers and a shared vision among stakeholders. These strategies are more likely to generate a cultural shift toward more active learning (Grunspan et al., 2018).

The geoscience education research community has conducted two long-term assessments of teaching practices: a nationwide survey deployed in 2004, 2009, and 2012 (Manduca et al., 2017) and direct observations of teaching practices using the RTOP (Teasdale et al., 2017; Viskupic et al., 2019). The national survey has shown a growing percentage of instructors reporting increases in their use of active-learning strategies; as of 2012, 51% of responding geoscience faculty reported spending more than 20% of class time on student activities, questions, and discussions (Manduca et al., 2017). Direct observations of 204 undergraduate classes supported those self-report data (Teasdale et al., 2017). Teasdale et al. (2017) also found no relationship between instructional practices as measured by the RTOP and demographic factors, including instructor gender, instructor rank, class size, institution type, and course type. Furthermore, 23 of 74 direct observations were identified as teacher-centered according to RTOP criteria (Viskupic et al., 2019), suggesting some use of more traditional, rather than active, learning strategies.

## *Current supports for active-learning practices*

In an effort to provide students with active-learning opportunities in the classroom, institutions and professional societies have been developing new infrastructure. This includes community-reviewed resources for classroom activities, such as those provided by the American Association of Physics Teachers (ComPADRE, www.compadre.org/; PhysPort, www.physport.org/), the National Association of Geoscience Teachers (Teach the Earth, https://serc.carleton.edu/teachearth/index.html), and a consortium of professional biological societies (*CourseSource*, https://www.coursesource.org/). By surveying the content and assessment tools in such repositories and peer-reviewed journals, one may infer instructor goals. Likewise, education and cognitive science can provide theoretical foundations for instructional design. Thus, developments in teaching are likely to benefit from both DBER research and design insights from the scholarship of teaching and learning—indeed, progress is likely to occur when the two fields mutually draw upon each other (Shipley et al., 2017).

Faculty view improving their students' chances of learning as an important goal. How do faculty learn about active learning, particularly if they did not experience it themselves as students? Many disciplines' professional societies provide resources and professional development to improve pedagogy (e.g., the American

Association of Physics Teachers, American Astronomical Society, and American Physical Society cohost a New Faculty Workshop to promote interactive instructional strategies; Henderson & Dancy, 2008). Manduca et al.'s (2017) national surveys found that faculty were more likely to use active-learning practices if they had participated in On the Cutting Edge, a national professional development program. Viskupic et al. (2019) found that geoscience instructors who had engaged in professional development focused on integrating content with active-learning strategies were more frequently observed teaching student-centered classes than those who had not engaged in professional development (33% vs. 13%; $p < .001$). Further, instructors who had completed topically aligned professional development were about 5 times more likely to teach a student-centered class than instructors who had not. It is therefore important for the resources to be well aligned to the content of courses to allow for the translation of suggestions to practice.

Another way instructional modes have changed has been though changing classroom architecture. Classrooms have changed as institutions have moved from lecture-based science education to group-project-based science education (Stains et al., 2018). These new classrooms no longer have rows of seats facing a central screen or lectern; instead, tables and chairs are arranged in clusters with technological support equipment for group work (e.g., SCALE-UP, which originated at North Carolina State University; Beichner et al., 2006). The changes are designed to promote small group work such as occurs in the flipped classroom model, described below, in which students complete readings or watch videos outside of class meetings and engage in more interactive tasks (labs, small group work, etc.) during class.

## Active Learning: A Synthetic View

Our examination of current instructional practice and past learning theories revealed an underdeveloped view of active learning. In short, active learning as an umbrella term lacks usefulness in advancing research; however, we recognize that it is commonly used to communicate an alternative to lecture and therefore serves a purpose in undergraduate STEM education classroom practice. Only by specifying what aspects of learning are being studied, and doing so in a way that reflects whether strategies are discipline specific or can transcend disciplines, will the research community more thoroughly and effectively build upon the foundation created by past work to advance our understanding of undergraduate STEM learning. This review therefore offers a guide for research that can build across disciplines to evaluate different approaches to support learning.

To begin this endeavor, each of the contributing DBER teams initially developed a synthesis of active learning within their particular domains and, in doing so, provided a notion of how their domains defined active learning (see Table 1 for a summary of each domain's ideas). All of the teams acknowledged the difficulty in defining the notion of active learning within their domain and used various approaches to capture a definition of the construct. For example, when reviewing the literature on active learning in biology education research from 2016 through 2018, the biology team found that a majority of studies (53%) did not provide a definition for the term but instead listed examples of specific active-learning strategies ($N = 148$ publications from three popular biology education journals: *CBE—Life Sciences Education, CourseSource*, and the *Journal of Microbiology & Biology Education*). It is noteworthy that the second most common approach (30%) in the literature was to leave the term undefined. Likewise, the astronomy team surveyed a broad review of research published in the *Proceedings of the International Astronomical Union* (Bretones & Megid Neto, 2011) and found that more than 87% of the studies did not describe a theoretical framework for their research. The geoscience team, alternatively, used Sinatra et al.'s (2015) four-component characterization of engagement in science learning (comprising cognitive, emotional, behavior, and agentic components) to sort active-learning studies within their domain (LaDue et al., 2020).

Overall, the definitions used by the disciplines fit well within the historical views of active learning as the antithesis to lecture, as the antithesis to passivity, as a cognitive or social process of knowledge construction, or as engagement; see Table 1 for each discipline's definition). We drew from each discipline to construct a synthetic working definition: *Active learning is a classroom situation in which the instructor and instructional activities explicitly afford students agency for their learning* (also included at the bottom of Table 1 and discussed in more detail in subsequent sections).

Our shared view is that knowledge construction in undergraduate science education should parallel knowledge construction and engagement in the scientific enterprise. Therefore, active-learning environments are best seen as contexts for knowledge construction and meaning making in which there is a prominent role for both social (i.e., involving the science learning community, which operates analogously to the community of scientific practitioners) and individual (i.e., involving the learner, who operates in parallel with the scientist) components. The foundation of meaning is aligned with engagement around ideas (e.g., pondering different

**Table 1.** Initial DBER Team Notions and Resulting Synthetic Definition of Active Learning

| DBER domain | Notion of active learning | Comments and elaborations |
|---|---|---|
| Astronomy | Active learning involves instructional strategies that allow students to work independently or in small collaborative groups to learn or reinforce content. Having students engage in actively constructing their understanding (often in a social setting with their peers) is ultimately more effective than having them passively listen to the instructor tell them what they should understand. | The astronomy team noted similarities between astronomy and physics education research, but they also noted that the different contexts (i.e., predominantly general education courses for nonscience majors in astronomy vs. courses for STEM majors in physics) allow for some differences across the two fields. |
| Biology | Active learning is an interactive and engaging process that may be implemented through the employment of strategies that involve metacognition, discussion, group work, formative assessment, practicing core competencies, live-action visuals, conceptual class design, worksheets, or games. | The biology team reviewed the literature in biology education research and surveyed the community using two prompts: "In your own words, define the term 'active learning' in the context of undergraduate biology classrooms" and "List the active-learning techniques that you use in biology classrooms." The literature review and survey results were complementary, but they also showed wide variation around the characterization of active learning in biology education research (Driessen et al., 2020). |
| Chemistry | Active learning emerges from instructional practices that promote students' engagement in the learning process to facilitate their construction and use of knowledge. | The chemistry team showed how notions of active learning in their field were grounded in social constructivism, or the idea that knowledge is co-constructed via interactions with other people (Vygotsky, 1978; Windschitl, 2002). The team also noted that as the idea of active learning spread in the late 1990s and early 2000s, it was quickly presumed to be *the* approach by which improvements in student learning should be achieved. |
| Engineering | Active learning includes course-related activities that engage students in the learning process, with students doing meaningful learning activities within the context of a class setting and thinking about what they are doing during these activities. | The engineering team examined various definitions of active learning (i.e., Felder & Brent, 2009, 2016; Freeman et al., 2014; Prince, 2004) and concluded that there was appreciable overlap. The team thought that (a) engagement, (b) thinking (i.e., reflection), and (c) classroom and topic context were the essential elements of active learning. In other words, as long as an activity incorporates all these elements, then it constitutes active learning. |
| Geoscience | Active learning involves situations in which students are engaged in the knowledge-building process. Engagement is manifest in many forms, including cognitive, emotional, behavioral, and agentic, with cognitive engagement being the primary focus in effective active learning. | The geoscience team examined both inductive notions of active learning (an idea of learning that emerged from empirical DBER studies) and theoretical notions of active learning (Chi & Wylie, 2014) and engagement (Sinatra et al., 2015). They systematically categorized geoscience education research studies using components of Sinatra et al.'s (2015) engagement framework to test their definition and were able to successfully categorize these studies within four engagement components (LaDue et al., 2020). |
| Physics | Active learning encompasses any mode of instruction that does not involve passive student lectures, recipe labs, and algorithmic problem solving (i.e., traditional forms of instruction in physics). It often involves students in working in small groups during class to interact with peers and/or the instructor. | The physics team left the concept of active learning fairly broad (i.e., as the antithesis to lecture-based learning) but noted a relatively recent review of active learning within the physics community that defined 13 characteristics of active-learning instructional methods (Meltzer & Thornton, 2012) and a review of course attributes to promote student learning (Mestre, 2001). Conceptions of active learning in the two reviews overlapped on the following key characteristics: (a) opportunity for qualitative reasoning based on physics concepts; (b) opportunity for construction and sense making of physics knowledge; (c) opportunity to engage in the process of doing science; (d) opportunity to apply knowledge flexibly across multiple contexts; (e) formative assessment techniques; (f) help with organizing content knowledge according to some hierarchy; and (g) instruction on metacognitive strategies. |
| Synthetic Definition | | |
| Active learning is a classroom situation in which the instructor and instructional activities explicitly afford students agency for their learning. In undergraduate STEM instruction, it involves increased levels of engagement with (a) direct experiences of phenomena, (b) scientific data providing evidence about phenomena, (c) scientific models that serve as representations of phenomena, and (d) domain-specific practices that guide the scientific interpretation of direction observations, analysis of data, and construction and application of models. | | |

Note: DBER = domain-based education research; STEM = science, technology, engineering, and mathematics.

aspects of the world and designing solution to problems)—in this case, scientific explanations (e.g., theories, hypotheses)—that are constructed through the process of abductive inferences (i.e., coming to the best of all plausible alternative explanations on the basis of evaluations of their coherence with lines of scientific evidence; Lombardi et al., 2016; Peirce, 1931; Walton, 2001). Further, our ideas reflect current science education reform as presented in *A Framework for K-12 Science Education* (National Research Council, 2012b), which views science learning through a multidimensional perspective that integrates practices and overarching conceptual ideas about scientific content.

Our working definition offers a framework, which we call the *construction-of-understanding ecosystem*, visualized in Figure 1. For clarity of explanation, we contrast the strict flow of meaning among the actors in lecture-based learning (Fig. 1a) with the variety of potential flows of meaning in active-learning settings (Fig. 1b). Situating active learning within an ecosystem framework affords a way to think about the construct of active learning that captures the variety of uses of that term, without being limited by them, and the complexity of interactions in the system that is undergraduate STEM learning. Each of the individual relationships identified in Figure 1b is supported by theoretical or empirical work in at least one—often more—of the disciplines. Our aim in developing Figure 1 is not so much to direct research; we do not see it as a theoretical proposal, given that the extant research that informs it does not allow for clear predictions about interactions among the relationships. We provide the organizational visualization to make salient the systems nature of learning and the various ways active learning might be pursued. The potential for endogenous interactions, such as feedback loops and delays, within the ecosystem means that it is important to take a systems approach (Meadows, 2008; Richardson, 1991) within active-learning research. The ecosystem structure highlights the importance of careful articulation in planning and presenting research, so that investigations are positioned within a constructive territory, and it helps ensure coverage of the entire territory to support both research and practice. The figure's aim is thus to support communication about how to deepen student learning in undergraduate STEM instruction.

## *A working definition of active learning*

Active learning is a classroom situation in which the instructor and instructional activities explicitly afford students agency in learning. In undergraduate STEM instruction, active learning involves higher levels of engagement with (a) direct experiences of phenomena, (b) scientific data providing evidence about phenomena, (c) scientific models that serve as representations of phenomena, and (d) domain-specific practices that guide scientific interpretation of observations, analysis of data, and construction and application of models.

Our synthetic working definition of active learning focuses on formal educational contexts within a curriculum—that is, a range of courses including classes at various levels as well as other organized didactic situations, such as laboratories and field experiences. In all cases, instructors and instructional activities explicitly afford opportunities for students to increase their agency during learning. Such granting of opportunities for agency may run counter to traditional undergraduate STEM instruction, in which many instructors either implicitly or explicitly control the formal learning environment.

In our conception of agency, we rely on Bandura's (2001) definition: "To be an agent is to intentionally make things happen by one's actions. . . . The core features of agency [intentionality, forethought, self-reactiveness, and self-reflectiveness] enable people to play a part in their self-development, adaptation, and self-renewal with changing times" (p. 2). In the context of undergraduate instruction, active learning represents classroom contexts in which students may—to a greater extent than in lecture-based learning—(a) exercise agency to be interactive learners, (b) have forethought to set achievement goals, (c) react accordingly in the classroom to achieve these goals, and (d) evaluate how well they are progressing (or have progressed) to achieve their desired learning outcomes. We extend Bandura's definition, which is focused more on individual self-regulation, by incorporating the affordance of agency as a central construct: Active learning positions students to experience greater levels of autonomy via an environment in which the instructor effectively supports that increased autonomy (Patall et al., 2019).

A potential consequence of increased autonomy is that active learning could lead to deeper cognitive, social-behavioral, and emotional engagement when students are effective agents in meaning construction by interacting with their peers and teachers. Meaning construction, in turn, may lead to increased knowledge, the primary outcome of learning (National Academies of Sciences, Engineering, and Medicine, 2018). By supporting agency and autonomy, active-learning environments may enable what many call student-centered instructional situations (Lou & Jaeggi, 2019). However, well-designed instructor guidance and scaffolding are critical for making student-centered learning more effective than traditional (e.g., lecture-based) learning (Alfieri et al., 2011; Hmelo-Silver et al., 2007; Kirschner et al., 2006; Mayer, 2004).

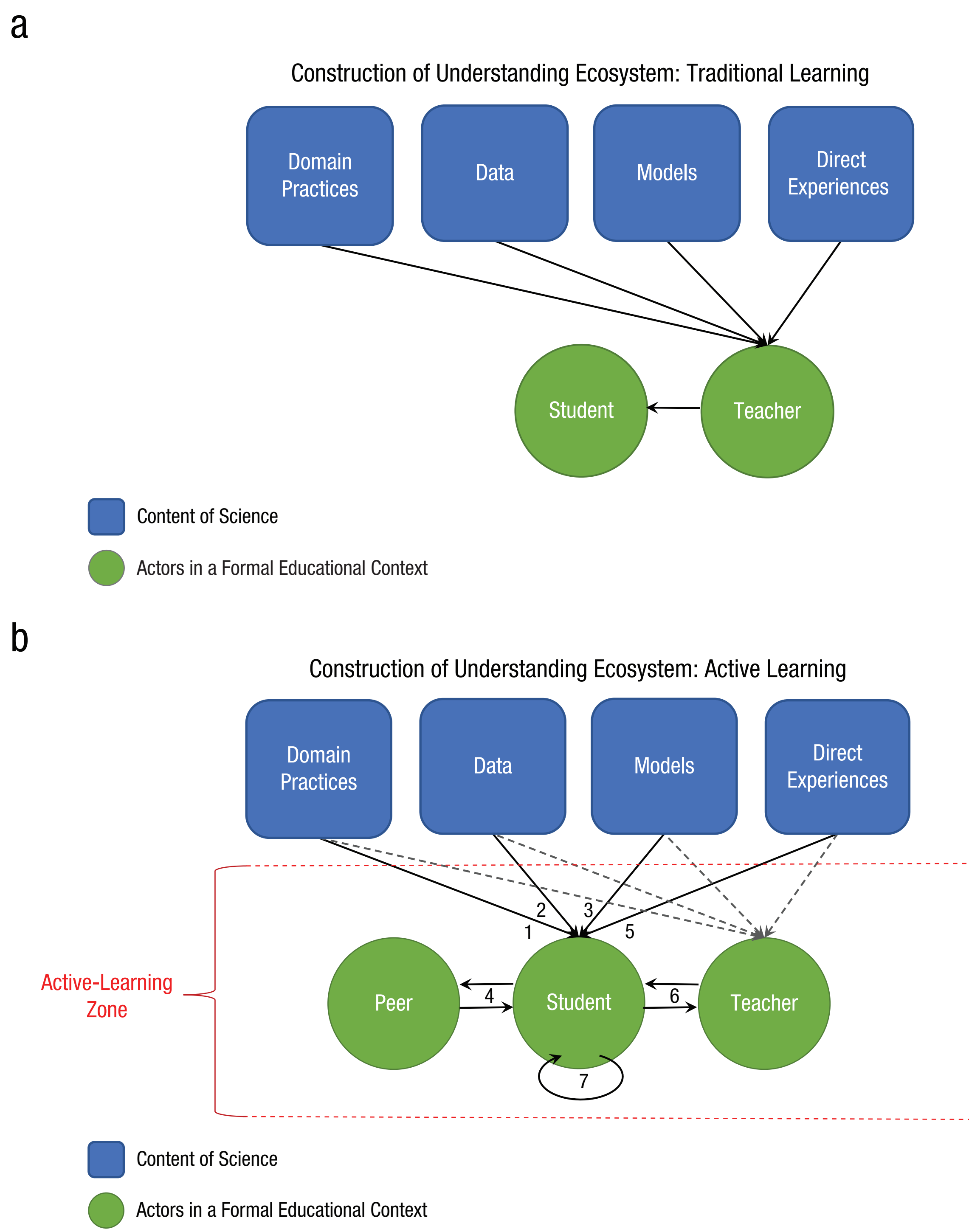


**Fig. 1.** The construction-of-understanding ecosystem. Panel (a) illustrates traditional learning situations, in which the teacher makes sense of science content and transmits this information to the students. Panel (b) illustrates an active-learning environment, in which students are agents in making sense of domain practices (Pathway 1), data about phenomena (Pathway 2), and scientific models (Pathway 3) together with their peers (Pathway 4). Students also explicitly consider the science they have personally experienced in an active-learning environment (Pathway 5), and the teacher acts to facilitate students' sense making (Pathway 6) given their expert knowledge of science content (dashed arrows). Finally, the student self-reflects (Pathway 7) on their understanding.

We contrast active learning with situations in which meaning is explicitly constructed for the student by a teacher. In this traditional teacher-centered environment, students could still attempt to exercise agency to construct meaning. But such agency would not be explicit in the instruction, or if it is, only superficially so. Therefore, in traditional learning situations, teachers may be the only ones explicitly exercising agency for meaning construction. Although we acknowledge that a good lecturer can support agency (see, e.g., Rogoff et al., 1996), here we are imagining an "end-member" lecturer who may not exist in such an extreme form in any lecture hall yet captures some lecturers' belief that their primary job is to deliver content. Students may share this belief, and those students may have less motivation, engagement, and self-regulation when experiencing active learning in the classroom than when experiencing instruction via a dynamic lecturer (Deslauriers et al., 2019). In this extreme form, traditional learning has low levels of teacher support for autonomy and therefore may dampen the autonomy that undergraduate STEM students need for deep learning. Learning may still occur via perception and observation, habit and conditioning, and implicit pattern recognition, but such modes of learning are "knowledge lean," in contrast to the "knowledge-rich" learning of factual information and scientific explanations (e.g., models, hypotheses, and theories) required for success in STEM education (National Academies of Sciences, Engineering, and Medicine, 2018).

## *Components of the learning ecosystem*

The figure illustrates two approaches to knowledge construction within the learning ecosystem, which comprises various nodes connected via pathways as arrows (Fig. 1). The first (Fig. 1a) is the traditional approach, which represents undergraduate instruction in which class meetings consist predominantly of lecture and note-taking. In these class meetings, the teacher makes sense of direct observations of and data about phenomena, integrates data and scientific models providing explanations and predictions of phenomena, and reviews the domain-specific practices that scientists use to make observations, collect data, and construct and interpret models.[1] The teacher thus transmits their sense making to students.

By contrast, the second approach to knowledge construction represents an active-learning system of undergraduate instruction (Fig. 1b), in which students are granted agency for their learning by the teacher. In this way, students make sense of direct observations of and data about phenomena, construct and work with scientific models explaining and predicting phenomena, and participate in domain-specific practices similar to those used by scientists.

It is important to note that the four sense-making nodes—direct observations, data about phenomena, scientific models, and domain-specific practices—represent knowledge, some of which the learner may already have stored in long-term memory. For example, students may have learned about the natural world through past experiences involving direct observations (National Academies of Sciences, Engineering, and Medicine, 2018). However, some of this observational learning may have led to conceptions that are inconsistent with scientific knowledge—naive theories, alternative conceptions, or misconceptions. There is some debate about the stability of conceptions, with some researchers suggesting that conceptions reside in memory structures that are relatively coherent (Vosniadou & Skopeliti, 2014) and others suggesting that conceptions are fragmented pieces of elemental knowledge generally lacking a larger-scale cognitive structure (diSessa, 2013). We propose a relatively agnostic view: that active-learning classrooms should support students in recognizing when such prior experiences are relevant as they weave these experiences together with data and models to make sense of scientific understandings.

Framing the activation of personal experiences and accompanying conceptual resources as important to meaning making shifts from a knowledge-deficit perspective (i.e., one focused on fixing all the "misconceptions" students have) to a view in which all students enter learning environments with a range of potentially productive ideas that can be useful for learning (Hammer et al., 2005). In an active-learning classroom, these prior conceptions and knowledge fragments are activated in a way that allows students to build upon their prior knowledge, even if that knowledge is inconsistent with scientific understanding (Merrill, 2002). (However, to be effective, the activation of such prior knowledge must account for students' cognitive load and motivation; Garcia & Pintrich, 1994; Kalyuga, 2007.) Likewise, students should integrate understanding of scientific ideas with social-behavioral and cognitive engagement in the practices of science and engineering, and with an emphasis on the discipline-specific nature of these practices (National Research Council, 2012a). Examples of discipline-specific practices include open-ended problem solving in engineering, relating atomic and molecular structure to observable properties in chemistry, analyzing interactions and changes over a wide expanse of scale in physics, conducting investigations using experimental and observational approaches in biology, and constructing explanations about past events using incomplete observational field data in geoscience.

In active-learning environments, interactions between the student and teacher are dynamic and dependent on the context of the class meeting (e.g., student questions may influence the teacher's input to the student). Further, interactions with peers inform expectations about how to learn as well as opportunities to construct meaning in the classroom. On an individual level, active learning also involves a feedback loop via self-regulated learning and metacognition. Finally, active-learning environments retain the instructor's expertise in sense making, but they do so in an indirect manner that transitions the role of the instructor to a facilitator of students' knowledge construction.

To illustrate the contrast between lecturer-as-constructor and student-as-constructor, consider how a hypothesis and supporting data might be introduced to a student. In lecture-telling mode, an instructor might present a hypothesis (e.g., that humans are causing current climate change), present data from an experiment or observation that are consistent with that hypothesis (e.g., a graph plotting long-term global temperatures and carbon emissions), and then explain how the hypothesis fits those data (e.g., showing how human emissions of carbon have been modeled and related to global temperature). In the active-learning mode, the instructor might present the hypothesis and then ask students to predict the data, observation, or outcome of an experiment before showing them the data. Or, to highlight the need to make sense of unfamiliar data, the instructor might show the data and ask students to look for patterns and consider how they might be explained. Because it can be difficult for students to simultaneously assimilate new data and construct explanations, an instructor could scaffold such an activity by presenting multiple hypotheses and ask students to identify the one that best aligns with the data. Using an example from climate change education, having students evaluate how well various lines of scientific evidence or forms of data support a model and discuss alternative explanations for rising global temperatures (i.e., as induced by human activity vs. natural solar activity) before explaining the scientifically accurate model has resulted in appreciable shifts in students' understanding of climate and weather compared with more traditional instruction (Lombardi et al., 2013, 2018). In this way, instruction sets the stage for construction of meaning by the students as they integrate the data and theory.

In the contrast between active and passive learning within formal education settings, we recognize a continuum of more or fewer, richer or poorer, opportunities to construct meaning. Indeed, when instructors report bringing active learning into their classrooms, they commonly add activities to lectures (e.g., Henderson & Dancy, 2007). These are opportunities to pause in the telling and offer students an opportunity to be agents of knowledge construction. Such opportunities may activate more of the pathways indicated by the solid arrows in Figure 1b. For example, after making a scientific claim, an instructor could initiate a think-pair-share activity in which students search recent notes for evidence supporting this claim, then compare the lines of evidence they found with another student, and finally engage with the whole class in consolidating these major lines of evidence through discussion. Such an activity could be followed by a lecture-based tutorial focusing on argumentation (i.e., the role of evidence, claims, and reasoning) in scientific knowledge construction within a particular domain (e.g., geoscience) or about a particular topic (e.g., causes of current climate change).

## Relationships and strategies within the ecosystem

In research on active learning across the various DBER literatures, studies have often focused on one, or a subset, of the solid arrows in Figure 1b. Here, we briefly review some of this research, highlighting each of the relationships illustrated in Figure 1b. We recognize that science practice involves all of the component categories featured in Figure 1b and that education typically conveys some of this complexity. Our categorization of studies is, therefore, somewhat artificial; nevertheless, we seek to present research in sections where it illuminates meaning construction within that category.

The following sections also highlight various strategies for increasing students' agentic engagement in the learning process, although those strategies vary in the degree to which they affect agency. Decisions about instructional design should therefore account for both the level of agentic engagement that a particular strategy promotes and the extent of change that instructors must make to their curriculum and teaching. For example, Process Oriented Guided Inquiry Learning (POGIL; described in the next section) is a fruitful strategy that originally emerged from chemistry instruction and is a relatively well-developed instructional package that may be implemented with minimal effort (Moog & Spencer, 2008); thus, it may represent a relatively low-"cost," high-impact strategy.

***Engaging with data.*** A range of techniques have been developed to facilitate students' engagement in domain-specific practices (Pathway 1 in Fig. 1b) with disciplinary data (Pathway 2 in Fig. 1b). Introduced above, POGIL is an instructional strategy in which groups of students construct conceptual understanding via analysis and explanation of trends that emerge from data (see, e.g., Brown, 2010; Moog et al., 2009; Moog & Spencer, 2008). Presenting data before interpretation offers students an opportunity to consider explanations for trends they notice. This opportunity might also be achieved by reversing the typical sequence of lecture and laboratory activities (e.g., Maria et al., 2011). In this way, the teacher is showing, then telling, instead of the more familiar telling, then showing. An additional learning benefit to presenting data before interpretation is that this matches the typical discovery path in the practice of science (Shipley & Tikoff, 2017).

POGIL activities are specifically designed to support students as they construct and use concepts. POGIL is now embraced by a national network of educators (Moog & Spencer, 2008). However, despite the large number of adopters, there are relatively few studies on the effectiveness of POGIL. Conclusions from a recent meta-analysis of the 21 extant studies assessing the impact of POGIL (L. Walker & Warfa, 2017) were similar to those from Freeman et al.'s (2014) meta-analysis: Students enrolled in classes featuring POGIL had marginally higher grades (a relatively small effect size of 0.29) and substantially higher odds of passing (odds ratio = 2.2) than students taught through more traditional modes. Overall, POGIL reduced the risk of failing a course by 38% (L. Walker & Warfa, 2017). One reason for the relatively small number of studies may be that it is challenging to design and carry out cross-sectional efficacy studies to evaluate active-learning activities in a way that distinctly isolates their effects on learning. Clearly this is an important area for the investigation of learning outcomes. We note that POGIL activities aim to engage students in practices characteristic of work in science and that it is important to understand the success of this objective. Future studies should carefully consider assessing the impact of POGIL interventions on students' learning of important aspects of STEM practice (M. M. Cooper & Stowe, 2018).

A related approach to active learning is to provide students an opportunity to wrestle with new data and learn how to collect it. Efforts to reform the physics laboratory experience to include considerations of what data to collect and how to respond to potentially unexpected data include the *RealTime Active Learning Laboratories* (Sokoloff et al., 2004) and *Workshop Physics* (Laws, 2004), as well as the Investigative Science Learning Environment and a textbook based on it called *The Physics Active Learning Guide* (Van Heuvelen & Etkina, 2005). These curricula encourage students to engage in the practice of science, not just follow step-by-step instructions for conducting investigations.

Other disciplines have also developed pedagogies to bring students into the workflow of practicing scientists, where students are confronted with data that have not yet been interpreted by anyone. These activities may be collectively referred to as *course-based undergraduate research experiences* (CUREs), as they are called in biology, where students engage in opportunities to perform authentic research in laboratory environments. Examples of CUREs include the SEA-PHAGES program (http://seaphages.org), the Small World Initiative (www.smallworldinitiative.org), the CUREnet network (http://curenet.cns.utexas.edu), the Genomics Education Partnership (http://gep.wustl.edu/index.html), the UT Austin Freshman Research Initiative (http://cns.utexas.edu/fri/research-streams), and the Community College Undergraduate Research Initiative (http://ccuri.org). Likewise, in astronomy, students can collect and analyze data from robotic or remote telescopes, such as the Global Telescope Network (http://gtn.sonoma.edu), the Faulkes Telescope Project (www.faulkes-telescope.com), or the MicroObservatory Network (https://mo-www.cfa.harvard.edu/MicroObservatory). (For an overview of use and a partial summary of research on learning from telescope data, see Fitzgerald et al., 2018.) Geoscience educators often promote active learning by infusing learning experiences with data (e.g., from satellites, radiosondes, ocean sensors, and seismic monitors) that engage students in flexible problem solving and building knowledge through authentic tasks. These data are often complex and offer natural opportunities to develop multiple hypotheses, which in turn requires systematically evaluating potential explanations using new data (Chamberlin, 1890). Multiple hypotheses are effectively multiple runnable models (i.e., models that can be used to make predictions or explain observations), similar to those used by scientists to investigate natural phenomena and thus important for engaging students in observation-prediction cycles that mimic scientific practice (Shipley & Tikoff, 2017).

Like POGIL, CUREs are difficult to study. It is hard to broadly characterize what is learned in CUREs, and research reports on learning suggest a variety of outcomes. This is an area of burgeoning research in several domains, particularly in biology (Auchincloss et al., 2014; Bangera & Brownell, 2014; Brownell & Kloser, 2015; Corwin et al., 2015), community college settings (Hewlett, 2009), computer-based research projects (Kirkpatrick et al., 2019; Lukes et al., 2019), and non-majors' courses (Ballen et al., 2018). Among upper-division undergraduates studying immunology, K. M. Cooper et al. (2019) observed higher cognitive and emotional ownership in the context of a CURE relative to a traditional lab course. Current work in biology seeks to understand how to adapt CUREs effectively for non-majors. Ballen et al. (2018) isolated two core components of CUREs: the experience of discovery and the production of data broadly relevant to the scientific or local community. They did not find significant impacts of these components on nonbiology majors' academic performance, science self-efficacy, sense of project ownership, or perceived value of the laboratory experience.

Outside of biology, Wooten et al. (2018) found improvements in students' perceived confidence in participating in science but no clear increase in understanding of science practices after participating in an astronomy CURE. In the same program, Dobaria (2018) found, through student essays and interviews, that students expressed enjoyment, increased self-efficacy, and other positive affective outcomes. In geoscience, Kortz and van der Hoeven Kraft (2016) found that students who participated in CUREs for introductory courses outside of their majors increased their appreciation for scientists and science in general, even if they had no intentions to major in science themselves. These researchers also reported increases in students' interest, content learning, and skill development.

CUREs are designed to offer authentic laboratory and field-based activities to more students by embedding them in an instructional setting. One rationale for considering how to work to design large-scale authentic science experience is the documented value of apprenticeship-based research experiences for undergraduates (REUs). Students who participated in a National Weather Center REU were significantly more committed to attending graduate school than nonparticipants (Gonzalez-Espada & LaDue, 2006), and students in a biology REU were more likely than unselected applicants to pursue PhDs and more likely to publish if they went on to graduate school (A. E. Wilson et al., 2018). Such experiences may offer important cues for future CURE design.

***Engaging with models.*** An important aspect of a research experience is the mental challenge of aligning data with theories. Here we refer to applications of theory to science by using the umbrella term *models* (Pathway 3, Fig. 1b, between models and student). Models may be physical, computational, theoretical, or mental; in any case, they explicitly represent accounts of structures or processes or implicitly describe how a theory aligns with some data set (Ruppert et al., 2019; B. Y. White, 1993). Classic models in science include the meiotic model in biology, which explains the process of allele segregation and gamete formation, and the Copernican heliocentric model in astronomy, which predicts the orbital positions of Earth and other planets in our solar system. In our usage, models must be runnable, allowing students to make predictions or explain observations and thus participate in making sense of data (Brewe, 2008; Nersessian, 2002). Some representations, such as photos, are not models because they cannot be used to make valid explanations and predictions (Clement, 2000; Lehrer & Schauble, 2000). However, other representations, such as three-dimensional ball-and-stick representations of atoms and molecules, may be models because their physical constraints allow predictions about allowable configurations, which may or may not accurately capture the atom's actual behavior (Fortus et al., 2016; Giere, 2004; Schwarz & White, 2005).

Understanding the relationship between a model and data is central to understanding science, and is challenging. How data fit with a model may not be self-evident, and thus aligning them requires support (Chinn & Brewer, 2001). One reason the fit is not always clear is that models may not account for all observations, but the reasons for this can vary by discipline. For example, physics experiments can offer reliable reproduction, but data may still vary because of instrument noise or variation in the execution of procedures, and thus not come out exactly as anticipated by a hypothesis—a prediction made by a model. Geoscience and ecology models, by contrast, may be limited because multiple processes can be superimposed in unique ways in each location on Earth and therefore obscure the unique contribution of a single geoscience or ecological process.

A key to understanding the relationship between data and models is being able to use a model to make predictions. This requires reasoning by analogy: using a model of the world to make a prediction that can then be applied to the world (Dumas et al., 2013). A common way to reason by analogy requires aligning an unfamiliar domain with a familiar domain to allow inferences in the new domain based on the knowledge structure of the familiar one (Gentner, 1983). For spatial inferences, scientists may employ physical models (e.g., stream tables in geoscience).

Forming an analogy to use a model can be difficult, and instructors can support students' analogical alignment in a number of ways. Many of the effective active-learning teaching techniques described in this section are good ways to support analogical learning. For instructors considering how to implement analogical reasoning, guiding students in understanding the strengths and limitations of a particular analogical mapping, so that the analogy does not instill conceptual errors, is of paramount importance (Jee et al., 2010).

Each STEM domain has considered how models should be employed and what their value is. One way to support aligning data and models is with simulations, which allow students to see representations of objects and processes that are far removed from everyday experience. The bulk of research around models in astronomy has focused on the efficacy of individual simulations or simulation types. For example, Ruzhitskaya and Speck (2011) created a simulation of stellar parallax and found that students who engaged with it had higher posttest scores than those who completed a similar, paper-based activity. In physics, a method that has been widely adopted in high school classrooms and is starting to make its way into universities is *modeling instruction*, in which students collaborate to construct, test, and apply conceptual and mathematical models (Brewe, 2008; Hestenes, 1987). In geoscience, physical models are widely recognized as a tangible way for students to observe and manipulate simulated processes (McConnell et al., 2017) and build their knowledge of geoscience concepts, especially through analogical reasoning that relates new information to something that is already understood (Jee et al., 2010; Nottis, 1999; Sibley, 2009). Students may make inferences about targeted phenomena using their knowledge of the model. Such analogies are particularly important for understanding phenomena that occur over large temporal or spatial scales (Cheek et al., 2017; Czajka & McConnell, 2018; Karlstrom et al., 2008; Resnick et al., 2017a, 2017b).

In addition to supporting analogical alignment, engaging students in cycles of observation and prediction is a way to structure their interactions with model systems (Liew & Treagust, 1995; Monaghan & Clement, 2000; R. White & Gunstone, 1992), including physical models (Warfa et al., 2014), augmented reality models, computational models, and models represented through simulations, demonstrations (Arthurs, 2019), and classroom experiments (Coştu et al., 2010). Reasoning with models also builds students' conceptions of the authentic process of science (Dolphin et al., 2018).

Computational modeling rests on existing conceptual models that are represented through computer coding and allow for variation of conditions. Students are able to test hypotheses by manipulating conditions and examining model output, thus replicating a predict-observe-explain process that is well described in the science education literature (Gunstone & White, 1981; Liew & Treagust, 1995; Monaghan & Clement, 1999) and practices and methods of inquiry used by experts (Shipley & Tikoff, 2017). By examining a model's output, students build upon prior conceptions and construct new knowledge. For example, in geoscience classrooms, software called STELLA (short for Structural Thinking Experiential Learning Laboratory with Animation) is often used to model dynamic Earth systems. The software enables students to ask questions and then experiment with the model to find answers (Bice, 2001).

Simulations do not require students to develop a model but do allow them to discover rules and data patterns through exploration. Animated, interactive, research-based computer simulations developed by the Physics Education Technology Project (PhET; https://phet.colorado.edu) connect real-life phenomena with underlying science and seek to make the conceptual models of experts accessible to students (Perkins et al., 2006). One method for using these simulations is to provide students with a sequence of challenging questions to elicit predict-observe-explain processes (Wieman et al., 2010), which enables students to test their working hypotheses about phenomena. Several of the PhET simulations also target concepts from chemistry (e.g., atomic structure, gas laws), geoscience (e.g., glaciers, radioactive dating, wave interference), and other domains and can be used to elicit simple to complex pattern recognition and prediction. Luo et al. (2016) used the Web-based Interactive Landform Simulation Model—Grand Canyon (WILSIM-GC) in introductory physical geography laboratory sections to model time-intensive processes that formed the Grand Canyon. Although they found no significant differences in test performance between students who read written material explaining the geological processes and students who worked with the model, exploratory analyses suggested that the simulation group had a deeper understanding of causal processes underlying the evolution of landscapes.

Some evidence suggests that one added benefit of working with spatial models is that it can improve general spatial reasoning skills (e.g., mental rotation) in chemistry (Stieff et al., 2012; Uttal & Cohen, 2012) and physics (Kozhevnikov et al., 2007). Hannula (2019) found that from the beginning to the end of a sophomore-level course on field mapping, which involves working with and creating models of what is under the surface of the Earth, most students' penetrative thinking skills improved (Cohen's $d = 0.54$, medium effect). Such skills are important for visualizing relationships between three-dimensional structures and spatiotemporal processes (Shipley et al., 2013), switching between temporal and spatial scales (Cheek, 2013; Czajka & McConnell, 2018; Ormand et al., 2014; Resnick et al., 2017b), and understanding the relationships among the many parts of an interconnected system (Kastens & Manduca, 2012, 2017).

***Collaborative construction.*** Active knowledge construction may occur within a single student or among multiple students. A variety of approaches have been developed to support knowledge construction by students working with their peers (Pathway 4, Figure 1b). A survey of biology instructors indicated that of the most common active-learning strategies, 63% focused on peer interactions in the form of discussion (34%) or group work (29%; note that instructors could select more than one option, resulting in a total greater than 100%). Indeed, peer interactions are often used in conjunction with classroom polling technologies such as clickers (Dufresne et al., 1996).

At least since Piaget (1926) and Vygotsky (1978), researchers and practitioners in education have recognized the value of cooperation in learning situations. Cooperation requires establishing a common understanding. Mismatches can be revealed through communication when students try to align their understandings. For example, from Piaget's (1926) perspective, the need for a workable common understanding motivates work to accommodate one's own understanding and assimilate another's. Thus, communication with peers supports (a) individual construction of a concept (because conversation requires specificity, one must construct missing pieces or recognize the need for more information); (b) joint construction, whereby different people may grasp different elements that together offer a more complete understanding; and (c) evaluation of how well oneself and others understand what is being communicated, with ongoing opportunities to reevaluate.

To illustrate how communication might support learning, consider research on the value of nonverbal spatial communication in STEM. Sketching and gesturing (in which spatial relations may be transiently sketched in the air) have been found to support understanding and communication of complex spatial relations in chemistry (Stieff et al., 2012), physics (Sattizahn et al., 2015), and geoscience (e.g., Gagnier et al., 2017; Gobert & Clement, 1999). Both sketching and gesturing may be employed in small group settings for communication or by an individual when working on a problem. Sketching constrains spatial representations in ways that may support the slow construction of understanding, thus expanding students' existing notions into more robust mental representations through assimilation and accommodation of new concepts. Gesturing provides learners with a way to offload information and reduce cognitive load, thereby facilitating cognitive engagement with increasingly complex concepts (Roth, 2001). In contrast to sketching, in which meaning can be constructed slowly and ideas have a permanence that allows for repeated querying, gesturing is transient and thus allows repeated construction of the same spatial model, which may be good in circumstances where an erroneous mental model committed to paper would interfere with learning.

Various peer activities have been developed to structure and facilitate cooperative communication. The activities can be applied whether students are learning the basic concepts of a science or more complex, systems-level qualities of the science. The following are noteworthy examples (to avoid redundancy, they are reviewed in this section only):

- *Peer-led team learning* (PLTL) activities, akin to traditional problem-solving activities, engage students in helping their peers work through problem sets (Mazur, 1997). For example, in chemistry, multiple-choice questions or worksheets are done in groups (Cardellini, 2006; M. M. Cooper et al., 2008; Mahalingam et al., 2008).
- *Flipped classrooms* task students with listening to lectures at home and solving problems during class (Eichler & Peeples, 2016; Hartman et al., 2015). Such courses have shown notable improvements in students' learning over traditional lecture-based courses in physics (Hake, 1998) and astronomy (Lazendic-Galloway et al., 2016).
- *Think-pair-share*, developed by Lyman (1981), poses questions to students, who respond in three stages. They first think about the question on their own, perhaps writing down some of their ideas, before pairing with a neighbor to discuss what each came up with. After a short discussion period, select groups share their ideas with another pair or with the larger group. For example, an instructor might ask students to think-pair-share about how seasons occur on Earth before providing details through the lecture.
- *Peer instruction* (Mazur, 1997) involves presenting multiple-choice conceptual questions (ConceptTests, i.e., questions that use scientific principles in a concrete example—e.g., "The force that holds molecules together is. . ."—with distractor choices that include known misconceptions) and having students vote for the best response using clickers, colored cards, or a show of hands. Depending on the outcome of the vote, the questions may then be discussed by small groups of students for a few minutes, after which a second vote is held. Peer instruction was originally developed in physics but has been adopted by other disciplines as well (see, e.g., Green, 2003, in astronomy and McConnell et al., 2006, in geoscience).
- *Interactive lecture demonstrations* provide students with opportunities to make a prediction about an in-class demonstration and discuss it with their peers (Sokoloff & Thornton, 1997). Learning Physical Science, a guided-inquiry, conceptual physical science course developed for large lecture environments (Goldberg et al., 2012), is one example of a course that uses interactive lecture demonstrations (e.g., via clicker use) throughout its scope and sequence.
- *Concept mapping*, or "cartooning," engages groups of students in the construction of a schematic diagram that captures important relationships or processes in a discipline to focus attention on the key concepts and integrate those

concepts with others (Johnson & Reynolds, 2005). In this way, concept mapping also allows spatial mental models to be externalized for inspection, reflection, and correction (Gagnier et al., 2017).

- *Lecture tutorials* (Arthurs & Kreager, 2017; Kortz & Smay, 2010; Kortz et al., 2008; McConnell et al., 2017; McDermott & Shaffer, 1998; Prather et al., 2013) adopt a Socratic dialogue approach to help students confront their naive ideas, engage in critical reasoning, and construct robust explanatory models for common topics (Prather et al., 2004). For example, in astronomy education, a short lecture on a topic may be followed by 10- to 20-minute lecture tutorials among students working in small collaborative groups (Brogt, 2007; Prather & Wallace, 2020). A sequence of questions scaffolds the students' discourse on the topic, and built-in mock debates give students a chance to check their conceptual understanding.
- *Ranking tasks* (Hudgins et al., 2006) present four to eight diagrams showing variations of a basic physical situation and require students to collaboratively order or rank the diagrams on the basis of specific features. For example, in astronomy education, students might be presented with diagrams displaying six phases of the Moon at various locations in the sky. Then they rank the diagrams according to the time of day at which each is phase is visible, from earliest to latest.

Research on these various collaborative approaches has generally shown improvements in learning. Crouch and Mazur (2001), in a review of 10 years of research on peer instruction, reported that "after discussion, the number of students who give the correct answer to a ConcepTest [conceptual question] increases substantially, as long as the initial percentage of correct answers . . . is between 35% and 70%" (p. 972). S. B. Wilson and Varma-Nelson (2016), in a study of PLTL in chemistry, noted that course grades improved in 13 of the 39 studies examined. Attitudes toward science were also reported to improve in some, though not all, of the cases surveyed. Notably, peer leaders guiding PLTL group work often report increases in content knowledge, confidence, and self-efficacy (Gafney & Varma-Nelson, 2007). However, it seems that time on task may be a significant contributor to the success of PLTL students.

Some studies have examined collaborative approaches incorporated into a more traditional instructional design. In astronomy, students who engaged in structured peer activities performed 20% above those who received lectures alone (LoPresto & Murrell, 2009; Prather et al., 2004). Bailey and Nagamine (2012) found significant improvements from pretest to posttest in courses using both peer instruction and lecture tutorials ($\eta_p^2 = .11$). Finally, Smith et al. (2011) showed that a combination of peer instruction and instructor explanation was more effective than either strategy alone, particularly for subject area majors, who saw nearly a 30% advantage in unit test scores with the combined strategies.

Two concerns arise in the context of integrating and evaluating nonlecture pedagogies into large-enrollment courses: First, most reports represent tweaks to fairly traditional learning environments rather than a rethink of what students should know and be able to do, and second, the effectiveness of the various strategies is hard to disentangle from confounded variables. These concerns apply broadly to work across the ecosystem, but they are particularly acute in the context of evaluating the role of communication among students. Consider evaluating the effectiveness of flipped classrooms, which have been enthusiastically adopted in chemistry (Christiansen et al., 2017; Eichler & Peeples, 2016; Fautch, 2015; Flynn, 2015; Hartman et al., 2015). It is exceedingly difficult to design a study that can persuasively demonstrate the efficacy of flipped classrooms relative to other collaborative learning options.

In a recent study, Casselman et al. (2020) attempted to address this issue by focusing on a short, 2-day course. Students were randomly assigned to a flipped learning condition, a traditional lecture condition, or a control condition with no instruction. Students in the flipped learning condition and in the traditional lecture condition completed separate 2-day organic stereochemistry units led by the same instructor. Although both groups of students performed better than the control group, the results indicated that there was no difference between them in learning gains from pretest to posttest (Casselman et al., 2020). A full-scale study of an entire course would have to randomly assign students to either a flipped or control condition and ensure that the instructors in both environments were equally invested in the approach they were using, which may be impractical in most university settings. Extant literature on flipped classrooms tends to suffer from one or more fundamental design issues, including comparison of nonsimilar student groups, failure to account for instructor bias, small sample sizes, and comparison of flipped classes to traditionally taught classes rather than classes taught using evidence-based practices (M. M. Cooper & Stowe, 2018). There is not sufficient evidence to conclude that flipping a classroom is any more beneficial than integrating collaborative learning opportunities into another course format.

Finally, we note that an indirect benefit of actively involving peers in learning is that when peers are from marginalized groups, other disempowered students can see themselves as part of the discipline. Many STEM disciplines suffer from an image problem, perceived as stale science done by White, heteronormative males (K. M. Cooper & Brownell, 2016; Sexton et al., 2014). Considering the lack of diversity in science (Bernard & Cooperdock, 2018; Hartten & LeMone, 2010, 2014) and its impending workforce needs as the baby boomer generation retires (Summa et al., 2017), active learning that fosters more inclusion and agency may be essential to the health of the STEM disciplines. Empirical studies from K–12 science education have shown that project-based curricula, which have many aspects reflective of our active-learning framework, can serve to reduce achievement gaps among underrepresented minority groups (see, e.g., Geier et al., 2008). This is certainly an area where more research is warranted within undergraduate STEM education.

***Seeing the science in the world.*** The theories and practices of each of the STEM disciplines can be applied to the everyday lives of students, in which they may directly experience the science in the world (Pathway 5, Fig. 1b)—for example, by using trigonometry to calculate distances and angles when building an object or understanding why food changes color and texture when vinegar or baking soda is added. In K–12 STEM education, there is significant interest in designing learning environments that support students' construction and critique of explanations or models of observable phenomena (Schwarz et al., 2017). Such environments often feature units structured around *anchoring phenomena* that students experience, wonder about, and collaboratively move toward understanding via engagement in a series of scaffolded experiences. In undergraduate geoscience, Pugh et al. (2019) found that students who made observations of the world and recognized how they might be explained by concepts from their classes were more likely to stay in their major than those who do not report this experience. In biology, female undergraduate students learned more about the biological impacts of climate change when the studied examples were local than when they were global (Theobald et al., 2015).

***Student-instructor interactions.*** Students may contribute to the flow of a class by asking questions that guide the teacher's input (Pathway 6, Fig. 1b). The opportunity to co-construct instruction may be structured, as occurs in *robust inquiry learning*, in which students develop a question, design an experiment, and interpret the results (Grissom et al., 2015; Ryker & McConnell, 2017), or when students are charged with leading a component of field learning based on a particular site (Todd & Goeke, 2012). Another way to support student-guided instruction comes from the Just-in-Time Teaching approach, in which students answer questions online before class and the instructor uses their responses to adapt the instruction (Novak et al., 1999). Thus, it is through a combination of autonomy support and structure that effective student and teacher interactions occur and may maximize active learning (Jang et al., 2010).

A well-structured collaborative instructional context not only promotes interactions among students but may support productive interactions between students and their instructor. Henderson and Dancy (2011) suggested that "once an instructor understands the importance of social interactions for learning, they may be more likely to incorporate this aspect into their own reinventions or inventions" (i.e., by redesigning an undergraduate STEM course to focus more on active learning; p. 7). Consistent with this idea, Smith et al. (2013) observed that instructors in active-learning settings tended to move around the classroom, participate in focused discussions with students, and pose questions. Including near peers in an instructional team is also a way to promote increased interactions. For example, in PLTL, small groups of students are led by a fellow student who maintains direct contact with the class instructor (Hockings et al., 2008). Such near-peer strategies may change the nature of who is considered a teacher, from the perspective of both students and instructors.

***Student self-reflection.*** Perhaps counterintuitively, active-learning activities may support learning by slowing the rate at which instructors present information. As textbooks get longer to accommodate the accumulation of knowledge and instructors feel increased pressure to lecture more or faster, slowing down may not be perceived to be desirable. However, students who are given more time to process information may be more likely to construct and reconstruct meaning for richer understanding.

A familiar rhetorical move for many teachers is to slow down by pausing to ask a question—for example, "Why do hurricanes rarely happen near the equator?" In lecture mode, the instructor would immediately go on to provide an answer, discussing the Coriolis effect. This quick transition provides little time and opportunity for students to process the information, and many students may not realize that they do need to think deeply during the pause because the instructor will soon provide an answer. If students do not deeply engage and reflect on the question, they may not be conscious of their understanding or lack of understanding. Student response systems, or clickers, are often employed to ensure that students engage with such questions (e.g., the biology team's survey found that 49% of faculty employed clickers in their lectures). Use of student response systems is one of the more frequently employed active-learning strategies in large-enrollment classrooms (Henderson & Dancy, 2007; Lewin et al., 2016; Prather et al., 2009). When instructors explicitly connect questions from lectures to students' progress in the course, particularly in advance of larger assessments such as tests, it can lead to improved self-efficacy and greater knowledge gains by helping students identify which topics they may need to study more (Bailey et al., 2017). Thus, self-reflection (Pathway 7, Fig. 1b) can occur whenever there is an opportunity to actively engage in knowledge construction.

Under what circumstances is self-reflection most likely to occur? There is a hint to the answer in a review by MacArthur and Jones (2008) considering when clicker questions do and do not appear to be effective in chemistry coursework. Notably, studies that reported positive impacts from clicker use required some sort of collaborative engagement around the questions to which students were responding. It may therefore be the case that it was co-construction of knowledge that resulted in improved learning, not the employment of a response system per se.

## Opportunities and Challenges of the Specific STEM Disciplines

In this section, we discuss how DBER research from the various disciplines informs these larger models of learning and resulted in our conceptualization of the construction-of-understanding ecosystem. Research across DBER varies in part because of differences in the function of STEM courses in the academy (e.g., introductory astronomy is typically a terminal science course for nonscience majors, whereas introductory physics often serves as a prerequisite not only for physics majors but for students in other disciplines, such as premedicine or engineering). The disciplines also vary in how well students can connect the science to their lived experience. One of the hallmarks of the geosciences are societally relevant topics, such as natural hazards, natural economic resources, and climate change. These examples offer an opportunity for increased social-behavioral and emotional engagement, in which groups of individuals may have similar responses and feelings about topics and events of shared relevance (Linnenbrink-Garcia et al., 2011). Topics of societal importance may have shared meaning, above and beyond individual meaning, which may promote agency in the learning context (Hancock et al., 2019). For example, in the context of recent environmental activism, Bandura and Cherry (2019) showed how increased social-behavioral and emotional engagement led to increased self-efficacy among youth, which motivated them to understand and act on climate change mitigation and adaptation (i.e., through agency). Individual socioemotional factors may well interact with active-learning strategies that rely on engagement, thus complicating characterizations of what works well for whom.

Aside from the specific knowledge each student brings to a formal education setting, they also bring their prior beliefs about how to learn. Students' views of how to learn are an important moderating variable. Students who believe they learn best from lecture will be dissatisfied with and unlikely to participate in a class that extensively employs active-learning pedagogy (Deslauriers et al., 2019)—even if the active-learning activities would have improved their performance in the class. For example, Deslauriers et al. (2019) found that undergraduate students learned more about physics (i.e., had higher post-instructional scores on two multiple-choice tests about statics and fluids) in an active-learning phase of instruction but *felt* they learned more (i.e., had higher self-reported feelings of learning) in a lecture phase. The best way to address this issue awaits further research. For now, instructors who use active-learning strategies might try to mitigate such effects by telling their students at the beginning of the course about research on the benefits of active learning and importance of students' beliefs about how to learn.

One factor that influences instructors' decisions to implement active-learning strategies is students' level of satisfaction in their course; for example, students may resist or openly oppose reformed instructional strategies (Gaffney & Gaffney, 2016). Henderson and Dancy (2007) found that the dominant barriers to the use of RBIS were situational factors that strongly favored traditional instruction. Unless instructors can effectively communicate to students why the course is being taught in a reformed way and students buy into the active-learning approach, they will not "become active participants rather than simply tolerant participants" (Mestre, 2001, p. 48).

### *Distinguishing between historical and experimental sciences*

From the perspective of active learning, there is an important distinction between historical sciences (e.g., geology, astronomy, archeology), which use observations about the present to make inferences about the past, and experimental sciences (e.g., chemistry, physics), which use predictions of experimental outcomes to test theories. Although practicing science in most fields requires a mix of observation and experimentation, the balance between the two differs across fields; for example, geology and astronomy involve predominantly observational methods, whereas chemistry involves predominantly experimental methods.

For the experimental sciences, a focal educational challenge is learning to design experiments, whereas for the historical sciences, a significant challenge is aligning classroom concepts with observations that allow inferences about the past. Small-scale studies in geoscience education have developed promising active-learning interventions that focus on skills that are critical for inferring past processes from current structures, such as understanding three-dimensional structures (Atit et al., 2015) and large spatial and temporal magnitudes (Cheek, 2013; Czajka & McConnell, 2018; Resnick et al., 2017a). However, large-scale design-based research is needed to understand how to support the active learning of complex spatiotemporal reasoning skills necessary to infer past events from current event traces.

## *Laboratory learning*

Laboratory experiences are widely believed by both the chemistry and physics communities to be important for student learning. This commitment is evident in the dedication of teaching resources, material, equipment, and specialized space to provide students with lab experiences. However, a recent survey of the literature by Bretz (2019) found no credible evidence to support that investment of resources. Furthermore, a direct experimental test by Holmes et al. (2017) found no difference in outcomes on exam questions between students who had and had not experienced the lab as a conjugate to lecture. The authors of both studies stressed the need for more testing and questioned the presumption that there is something critical to learning an experimental science that is particularly likely to occur by participating in the lab, or is most effectively accomplished with lab experiences. As noted in our review of the research on CUREs above, there is little evidence that laboratory experiences have an impact on assessments that reflect exam content, but they may indirectly affect learning by supporting agency and the development of procedural skills. Clearly, there is a need for additional research on the role and value of lab experiences within the ecosystem of STEM learning.

## *Field and place-based learning*

Whereas much research has focused on learning in classrooms, relatively little has considered the challenges and opportunities of learning in the field, a common component of geoscience education in particular, as well as biology education. Learning in the field is fraught with practical challenges: weather issues, land-rights issues, inclusion and accessibility, and expense. The investment warrants careful analysis of the potential benefits. Research in this area has tended to focus on understanding the challenges of students' initial field experiences and supporting students. As noted by O'Connell et al. (2020), there is both an empirical and a pedagogical need for well-designed tools to evaluate the impact of field-based learning on students. The challenge of working and learning in a place one has never been in before has been captured by the concept of *novelty space*, which is the notion that when students are in a new space, there are cognitive, geographic, and psychological barriers to overcome before learning can occur. Orion and Hofstein (1994) found that when students' novelty spaces were minimized (e.g., through awareness of what to expect in a new area and recognition of the social-emotional effects of being in new places), learning was optimized. Furthermore, agency is important in developing a sense of place (Jolley et al., 2018).

Place-based learning can engage students emotionally through their existing or emerging sense of place (Semken & Freeman, 2008). A geoscience-specific theoretical framing of emotional engagement identifies students' attachment to place and connections to the aesthetic aspect of Earth as important dimensions worthy of study (van der Hoeven Kraft et al., 2011). Researchers have studied the value of place-relevant context in geoscience education. For example, Hispanic students enrolled in a course focusing on the geology of Mexico were more likely to take additional geology courses and declare geology majors than Hispanic students enrolled in a structurally similar traditional physical geology course (Hammersley et al., 2012). In the biological sciences, ethnographic observations of students experiencing a 10-day field course in the Galapagos Islands showed increases in scientific literacy and social cohesion among students (Mason et al., 2018). Other work (Barrow et al., 2016) showed that a majority of early-career scientists in biology and ecology believed that understanding natural history—the observational study of organisms in their environment—was essential to their current work. However, although the scientists placed a high value on natural-history-based skill sets, many felt that their institutions' support for this type of research and training was waning.

Sense of place may be most acute in indigenous communities, where a blend of traditional knowledge with modern science tools can facilitate learning (Cohn et al., 2014; Riggs, 2005) and implementation should align with the traditional values of teamwork and community-building (Ward et al., 2018). These findings highlight the complexity of developing individualized, and thus more effective, active-learning strategies that can be deployed at scale.

Given the importance of spatial reasoning for fieldwork, it is not surprising that geologists are highly skilled at both reasoning about physical objects (Resnick & Shipley, 2013) and navigation (Weisberg et al., 2014). Hambrick et al. (2012) found that spatial ability predicted successful field mapping for low-knowledge geologists, but not high-knowledge geologists. This finding indicates that spatial ability may be more important for novice geologists who have less experience (e.g., have not seen many different outcrops in the field). Studies of field-mapping behavior have shown that novices often unnecessarily revisit outcrops (Riggs et al., 2009) and that students who more fully explored a field area were more likely to recognize important aspects of the geologic structure being mapped (Baker & Petcovic, 2016). In addition, students who made spatial-navigation decisions about which specific geologic sites to visit to illustrate different features or processes experienced both emotional benefits (bonding with their classmates) and cognitive gains (greater geologic knowledge; Todd & Goeke, 2012).

Finally, we note that the practice of emphasizing field-based experiences presents obstacles for some potential learners (e.g., people with young children or mobility or vision impairments). A burgeoning area of research is focused on developing inclusive practices for field-based learning (Carabajal et al., 2017; Gilley et al., 2015; Hendricks et al., 2017). The same concerns apply more generally when active learning is interpreted literally as requiring movement or overtly active behaviors (Chi, 2009) that could act as a barrier to including students with disabilities.

### *Systems thinking*

For instructors in each STEM discipline, emphasizing the value of the discipline's research for society offers an opportunity to link natural sciences with social sciences. Such activities can be used to engage students in thinking about complex systems. For example, understanding how Earth systems and human systems jointly produce climate change requires a systems-thinking approach that captures myriad influences, controls, feedbacks, reservoirs, and rates of change. Systems thinking has always been a key component of geoscientific thinking, but is challenging to teach and learn because systems are ill-structured—that is, many problems in geoscience have multiple potential solutions, no agreed-upon correct answer, a lack of consensus on how to assess solutions, no or weak analogs to draw from, or uncertain components (Holder et al., 2017; Scherer et al., 2017). Recent work has begun to develop and study active-learning strategies that promote systems thinking, specifically around topics of sustainability in the Earth system (Gosselin et al., 2019). Likewise, there has recently been a movement to foreground systems thinking in chemistry education (York et al., 2019). However, virtually all systems of societal import are tremendously complex at the atomic/molecular level, and there has been scant work toward developing systems-thinking learning environments and assessments in chemistry.

## Guidance on the Employment of Active Learning in STEM Undergraduate Educational Research and Practice

It is clear that there is no single instructional strategy that will work across all situations and that the idea of active learning, as an umbrella term for effective undergraduate STEM instruction, provides little specificity and direction for either research or practice. In examining the perspectives that have emerged from psychological and DBER-based research on learning, as well as how active learning is currently manifest in DBER instruction, we developed the construction-of-understanding ecosystem. Our aim was to develop a framework that provides coherence and utility for directing future research and practice. What also emerges from our review is that vagaries in understanding and uncertainty among faculty may also stem from the wide range of empirical evidence in active-learning research. Certainly, there is a need for experimental and quasi-experimental investigation, along with a greater emphasis on replication, to increase the rigor of research and the quality of findings (National Research Council, 2012a). We hope that the construction-of-understanding ecosystem will provide a basis for some theoretical development and impetus for such rigor by highlighting the variety of points in the system where learning could be influenced (Smaldino, 2019) and the potential for interactions among those influences over time. Better measures of what students know and can do are also needed to provide greater understanding of the impact active-learning environments have on students. As these are developed, it is critical that STEM learning objectives are kept in mind to fully align assessments with instruction. For example, constructed response assessments, in which students predict, explain, or model phenomena, may provide more robust measurements of learning. We acknowledge that such assessments are time-consuming to develop, validate, and code and that the need for more accurate and precise measurement, as well as theory-driven replication, is not unique to undergraduate STEM education or DBER but permeates all educational and psychological research.

Active-learning environments represent a change in the learning ecosystem and as a result are unlikely to have simple linear effects. To conclude that active learning simply improves learning, as if learning is a single dimension, is to oversimplify and lose much of what is important. Broadly, lecture-based instruction that focuses on a domain's concepts results in learning those concepts; in contrast, active-learning environments that focus on synthesizing concepts result in learning how to synthesize. "Higher-order learning" may be difficult to achieve in a single course and may require a well-designed, integrated curriculum, such as a larger multi-course program. It is important to reflect on pedagogical goals when considering whether or not to change the status quo and, concomitantly, to reflect on the multiple loci of change, because which changes will be practical and influential may differ over time within a system. Most importantly, it is clear that effective undergraduate instruction must explicitly encourage students to be active agents in their learning, both for themselves and for the formal learning environment in which they are situated. Increasing agency allows students to be more self-reflective, such that they learn by taking stock of what they know and how that knowledge was constructed, and similarly reflect on how scientific knowledge is constructed via interactions of discipline-based practices, data collection and analysis, and model development and interpretation. Such activities are likely to result in learning that can feed back to influence how students approach new learning challenges. Although we advocate for wide-ranging consideration of what might be changed to promote STEM learning, it is not immediately evident that changes should be made at all potential loci within the ecosystem, or that changes will additively advance learning.

Interpreting active-learning situations, in which meaning is constructed both cognitively and socially, offers an organizational framework for the science of learning, design-based research, and scholarship of teaching and learning to integrate theory and observations toward understanding learning across the life span. An implication of a life span perspective is that knowledge construction in an undergraduate STEM education system should parallel knowledge construction in the system of scientific practice. Thus, it is important to view active-learning environments as those in which knowledge construction and meaning-making processes play a prominent role for both the social (i.e., the science learning community, which operates in parallel to the scientific community) and the individual (i.e., the learner, who operates in parallel with the scientist) components. Even if students sitting in a lecture appear passive, some may be learning via cognitive processes activated by this observational mode. However, there is little evidence that learning occurs best when isolated in an individual's mind, and there is fairly compelling evidence that learning is well supported in a social context. The educator's challenge is to structure the learning context so that critical concepts are foregrounded and students engage with one another on the concepts of the discipline.

The principles of promoting knowledge construction might be expected to apply across the age span of learners and across diverse contexts. But although introductory STEM courses are an important focus, given their size and their potential downstream impacts, having a better sense of whether and how strategies need to be adapted as students progress is critical. The vast majority of DBER research has been aimed at introductory courses, which in some situations may be predominantly made up of nonscience majors. What little research there is on more advanced STEM students and coursework has not generally addressed active learning (e.g., Baleisis, 2009; Gross & Lopez, 2009). Future work could investigate active learning in upper-division courses, small classes, community colleges, minority-serving institutions, and more. Further, active-learning research has traditionally focused almost exclusively on performance outcomes rather than longitudinal or individual-differences outcomes. We therefore know little about the long-term consequences of participating in student-centered teaching, for example, nor do we know if some active-learning strategies support learning for certain students but interfere with learning for others. Finally, more work needs to be done on closing demographic gaps in outcomes to increase diversity in STEM (National Academies of Sciences, Engineering, and Medicine, 2016). Explicitly promoting agentic engagement may support adaptive functioning skills essential for the transition into academic culture. Gaining greater autonomy, agency, and independence in the learning process can be a means for members of historically disempowered and underrepresented groups to enhance their academic success and opportunity (Eccles et al., 1993; Hofstra et al., 2020; Rose & Rudolph, 2006; Wang & Degol, 2013).

## Conclusion

We conclude by reflecting on the process by which we constructed this article: using cross-DBER science to construct new knowledge about active learning, as well as novel and innovative knowledge about undergraduate learning processes, through cross-domain connections and synthesis. To initiate that process, the first two authors identified leaders within each of the DBER fields included in this review: astronomy, biology, chemistry, engineering, geography, geoscience, and physics. We asked each of these leaders to write a white paper presenting a synthesis of active learning within their domain, soliciting help from others in their discipline as necessary. These leaders and their team members wrote the white papers, and submitted them to the first and second authors.

In reviewing these white papers, it became apparent that "active learning" meant a wide variety of things (e.g., psychological and social constructs, instructional pedagogies and strategies, and design principles) and that we needed focus and clarity. We specifically asked the teams to help us to come to consensus on a useful definition and conceptualization of active learning. We also asked them about the evidence supporting the notions that (a) active learning results in improved learning compared with traditional lecture and (b) active learning results in learning different things than lecture. The goal was to come to a better understanding of the nature and characteristics of outcome variables examined in each DBER area. For example, we wondered: Were the measurements of outcomes ecologically valid in each discipline? Is learning how to learn an outcome? Is learning how to engage in the practices of science an outcome in each discipline? And does the way each discipline's DBER literature defines and describes active learning match how that discipline's instructors define and practice active learning? We then synthesized the original white papers and the feedback we received from these queries into an initial draft manuscript for internal review.

The synthesis encapsulated by the constellation of interconnections in Figure 1b was an emergent product that arose from our attempt to align the white papers; indeed, the key takeaways from this review arose from combining DBER work across fields. Each discipline brought to the table a history of research that reflected the challenges instructors experienced in student learning and the core intellectual practices of the discipline. The disciplines that rely on advancing science by observing the world in all its complexity (e.g., geoscience and ecology) ensured that we presented what was knowable in the context of the complex system of students trying to learn science, in line with a tradition of design-based education. Simultaneously, disciplines that rely on experimental control of all relevant variables (e.g., chemistry) ensured that we evaluated research to reflect specific factors that could be isolated and conclusively identified as critical, or on which future work was needed, in line with a tradition of reserving confident conclusions to those developed from randomized-controlled-trial outcomes. We aimed to balance what was known about active learning collectively with the need, going forward, to differentiate affordances and outcomes among the various activities. The individual DBER fields brought numerous insights; most notably, from geology and ecology, the importance of models and field experience and the challenge when models do not fit data; from chemistry, the importance of identifying which aspects of disciplinary practice are still missing from science education; and from physics, a well-researched set of resources (practices, strategies, and tools) and data about what is happening in classrooms that other fields could draw from to guide research. Finally, variation across disciplines also required us to think through how to integrate social and cognitive aspects of learning, foregrounding the importance of agency.

Those familiar with working across disciplines will not be surprised that the insights we gleaned are a tribute to the commitment of the project's members to respectfully seek meaning and clarity in what can be said about active learning though engaged work that was not always easy. This interdisciplinary collaboration was largely possible because of extramural funding to support interdisciplinary work, and the emergent product reflects the value of supporting "trading zones" where teams can work across disciplines (Galison, 1997).

In sum, active learning is firmly embedded within the context of undergraduate STEM education. Active learning is a commonly used term to communicate an alternative to lecture, and as such, it serves a shorthand purpose in higher education. However, active learning can mean many things, often different things, to many people. Therefore, as an umbrella term, active learning is not a useful concept for advancing research on effective undergraduate STEM learning. To facilitate more effective research, we have offered a framework for the collective that the term *active learning* encompasses, which we hope will be useful for deepening students' undergraduate STEM learning. An active-learning approach does not preclude lecture from being an effective instructional form, but it is unlikely that lecture, in and of itself, will deepen STEM understanding for the diversity of students in the world. Therefore, whenever lecture is incorporated into other instructional strategies that fall within the broader context of active learning, it must be done in such a way as to increase students' agency in meaning making and knowledge construction. This meaning-making process is both socially derived via peer and instructor collaboration and feedback and cognitively constructed through processes that allow students to reflect on what and how they have learned. Active-learning environments in this form hold the promise to open STEM to more students, particularly those who have been disempowered and underrepresented via traditional modes of instruction.

## Transparency

*Editor:* Nora Newcombe
*Astronomy Team Members*
Janelle M. Bailey, Temple University
Paulo S. Bretones, Universidade Federal de São Carlos
Edward E. Prather, University of Arizona
*Biology Team Members*
Cissy J. Ballen, Auburn University
Jennifer K. Knight, University of Colorado, Boulder
Michelle K. Smith, Cornell University
*Chemistry Team Members*
Ryan L. Stowe, University of Wisconsin-Madison
Melanie M. Cooper, Michigan State University
*Engineering Team Member*
Michael Prince, Bucknell University
*Geography Team Members*
Kinnari Atit, University of California, Riverside
David H. Uttal, Northwestern University
*Geoscience Team Members*
Nicole D. LaDue, Northern Illinois University
Peggy M. McNeal, Towson University
Katherine Ryker, University of South Carolina
Kristen St. John, James Madison University
Kaatje J. van der Hoeven Kraft, Whatcom Community College
*Physics Team Member*
Jennifer L. Docktor, University of Wisconsin-LaCrosse
*Author Contributions*
Teams wrote white papers summarizing work in their discipline. D. Lombardi and T. F. Shipley synthesized an initial draft from the white papers. Each team reviewed, corrected, and supplemented the manuscript. All authors assisted in editing the entire paper and approved the final version for submission.
*Declaration of Conflicting Interests*
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.
*Funding*
Preparation of this manuscript was supported in part by the National Science Foundation (Science of Learning Collaborative Network, Grant 1640800; National Robotics Initiative, Grant 1734365; Future of Work at the Human Technology Frontier, Grant 1839705; Discovery Research PreK-12, Grants 1721041 and 2027376; Research Coordination Network in Undergraduate Biology Education Grant RCN-UBE-1919462). Any opinions, findings, conclusions, or recommendations expressed are those of the authors and do not necessarily reflect the National Science Foundation's views.

## Note

1. We view the nodes of domain-specific practices, data about phenomena, scientific models, and direct observations as ontologically distinct (i.e., we view them as different "things"). Domain-specific practices are cultural and normative expressions within a particular discipline that guide the interpretation of data collection and analysis, construction and application of models, and direct observations (Pickering, 1995). Data relate to material evidence about phenomena recognized as relevant to a scientific discipline; typically, these are quantified or categorized (Chinn & Brewer, 2001). Models are representations of phenomena used for explanation and prediction that may take mathematical, physical, visual, analogical, and/or conceptual forms (Ruppert, Duncan, & Chinn, 2019). Direct experiences are students' and teachers' perceptions of and experiences with natural and engineered environments (Mintzes et al., 2000).

## References

Akiha, K., Brigham, E., Couch, B. A., Lewin, J., Stains, M., Stetzer, M. R., Vinson, E. L., & Smith, M. K. (2018). What types of instructional shifts do students experience? Investigating active learning in science, technology, engineering, and math classes across key transition points from middle school to the university level. *Frontiers in Education*, *2*, Article 68. https://doi.org/10.3389/feduc.2017.00068

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. https://doi.org/10.1037/a0021017

Anderson, J. R. (1983). *The architecture of cognition*. Harvard University Press.

Anderson, R. C., Spiro, R. J., & Montague, W. E. (Eds.). (1977). *Schooling and the acquisition of knowledge*. Erlbaum.

Arthurs, L. A. (2019). Using student conceptions about groundwater as resources for teaching about aquifers. *Journal of Geoscience Education*, *67*(2), 161–173. https://doi.org/10.1080/10899995.2018.1561111

Arthurs, L. A., & Kreager, B. Z. (2017). An integrative review of in-class activities that enable active learning in college science classroom settings. *International Journal of Science Education*, *39*(15), 2073–2091. https://doi.org/10.1080/09500693.2017.1363925

Atit, K., Gagnier, K., & Shipley, T. F. (2015). Student gestures aid penetrative thinking. *Journal of Geoscience Education*, *63*(1), 66–72. https://doi.org/10.5408/14-008.1

Auchincloss, L. C., Laursen, S. L., Branchaw, J. L., Eagan, K., Graham, M., Hanauer, D. I., Lawrie, G., McLinn, C. M., Pelaez, N., Rowland, S., Towns, M., Trautmann, N. M., Varma-Nelson, P., Weston, T. J., & Dolan, E. L. (2014). Assessment of course-based undergraduate research experiences: A meeting report. *CBE—Life Sciences Education*, *13*(1), 29–40. https://doi.org/10.1187/cbe.14-01-0004

Bailey, J. M. (2020). Measuring students' understanding in astronomy with research-based assessment tools. In C. Impey & S. R. Buxner (Eds.), *Astronomy education, Vol. 1: Evidence based instruction for introductory courses* (pp. 10-1–10-19). IOP Publishing. https://doi.org/10.1088/2514-3433/ab2b42

Bailey, J. M., Lombardi, D., Cordova, J. R., & Sinatra, G. M. (2017). Meeting students halfway: Increasing self-efficacy and promoting knowledge change in astronomy. *Physical Review Physics Education Research*, *13*(2) Article 020140. https://doi.org/10.1103/PhysRevPhysEducRes.13.020140

Bailey, J. M., & Nagamine, K. (2012). Experiencing conceptual change about teaching: A case study from astronomy. *American Journal of Physics*, *80*(6), 542–551. https://doi.org/10.1119/1.3699064

Baker, K. M., & Petcovic, L. H. (2016). Sensitivity to landscape features: A spatial analysis of field geoscientists on the move. *Journal of Geoscience Education*, *64*(3), 242–253. https://doi.org/10.5408/15-110.1

Bakhurst, D. (1995). On the social constitution of mind: Bruner, Ilyenkov, and the defence of cultural psychology. *Mind, Culture, and Activity*, *2*(3), 158–171. https://doi.org/10.1080/10749039509524697

Baleisis, A. (2009). *Joining a discourse community: How graduate students learn to speak like astronomers* [Doctoral dissertation, The University of Arizona]. SAO/NASA Astrophysics Data System. https://ui.adsabs.harvard.edu/abs/2009PhDT.........9B

Ballen, C. J., Thompson, S. K., Blum, J. E., Newstrom, N. P., & Cotner, S. (2018). Discovery and broad relevance may be insignificant components of course-based undergraduate research experiences (CUREs) for non-biology majors. *Journal of Microbiology & Biology Education*, *19*(2), 1–9. https://doi.org/10.1128/jmbe.v19i2.1515

Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K. R. (2017). Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning. *CBE—Life Sciences Education*, *16*(4), Article 56. https://doi.org/10.1187/cbe.16-12-0344

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bandura, A., & Cherry, L. (2019). Enlisting the power of youth for climate change. *American Psychologist*, *75*(7), 945–951. https://doi.org/10.1037/amp0000512

Bangera, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE—Life Sciences Education*, *13*(4), 602–606. https://doi.org/10.1187/cbe.14-06-0099

Barrows, C. W., Murphy-Mariscal, M. L., & Hernandez, R. R. (2016). At a crossroads: The nature of natural history in the twenty-first century. *Bioscience*, *66*(7), 592–599.

Beichner, R. J., Saul, J. M., Abbott, D. S., Morse, J. J., Deardorff, D. L., Allain, R. J., Bonham, S. W., Dancy, M. H., & Risley, J. S. (2006). The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) project. In E. F. Redish & P. J. Cooney (Eds.), *Research-based reform of university physics*. American Association of Physics Teachers. https://www.compadre.org/Repository/document/ServeFile.cfm?ID=4517&DocID=183

Bernard, R. E., & Cooperdock, E. H. (2018). No progress on diversity in 40 years. *Nature Geoscience*, *11*(5), 292–295. https://doi.org/10.1038/s41561-018-0116-6

Bernstein, D. A. (2018). Does active learning work? A good question, but not the right one. *Scholarship of Teaching and Learning in Psychology*, *4*(4), 290–307. https://doi.org/10.1037/stl0000124

Bice, D. M. (2001). Using STELLA models to explore the dynamics of earth systems: Experimenting with Earth's climate system using a simple computer model. *Journal of Geoscience Education*, *49*(2), 170–181. https://doi.org/10.5408/1089-9995-49.2.170

Blumer, H. (1986). *Symbolic interactionism: Perspective and method*. University of California Press.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school* (Expanded ed.). National Academy of Sciences.

Bretones, P. S., & Megid Neto, J. (2011). An analysis of papers on astronomy education in proceedings of IAU meetings from 1988 to 2006. *Astronomy Education Review*, *10*(1), Article 010102. https://doi.org/10.3847/AER2010010

Bretz, S. (2019). Evidence for the importance of laboratory courses. *Journal of Chemical Education*, *96*(2), 193–195. https://doi.org/10.1021/acs.jchemed.8b00874

Brewe, E. (2008). Modeling theory applied: Modeling instruction in introductory physics. *American Journal of Physics*, *76*(12), 1155–1160. https://doi.org/10.1119/1.2983148

Brewe, E., Bruun, J., & Bearden, I. G. (2016). Using module analysis for multiple choice responses: A new method applied to Force Concept Inventory data. *Physical Review Physics Education Research*, *12*(2), Article 020131. http://doi.org/10.1103/PhysRevPhysEducRes.12.020124

Brogt, E. (2007). A theoretical background on a successful implementation of lecture-tutorials. *Astronomy Education Review*, *6*(1), 50–58. https://doi.org/10.3847/AER2007005

Brown, P. J. (2010). Process-oriented guided-inquiry learning in an introductory anatomy and physiology course with a diverse student population. *Advances in Physiology Education*, *34*(3), 150–155. https://doi.org/10.1152/advan.00055.2010

Brownell, S. E., & Kloser, M. J. (2015). Toward a conceptual framework for measuring the effectiveness of course-based undergraduate research experiences in undergraduate biology. *Studies in Higher Education*, *40*(3), 525–544. https://doi.org/10.1080/03075079.2015.1004234

Caleon, I., & Subramaniam, R. (2010). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education*, *32*(7), 939–961. https://doi.org/10.1080/09500690902890130

Carabajal, I. G., Marshall, A. M., & Atchison, C. L. (2017). A synthesis of instructional strategies in geoscience education literature that address barriers to inclusion for students with disabilities. *Journal of Geoscience Education*, *65*(4), 531–541. https://doi.org/10.5408/16-211.1

Cardellini, L. (2006). Fostering creative problem solving in chemistry through group work. *Chemistry Education Research and Practice*, *7*(2), 131–140. https://doi.org/10.1039/B5RP90019K

Casselman, M. D., Atit, K., Henbest, G., Guregyan, C., Mortezaei, K., & Eichler, J. (2020). Dissecting the flipped classroom: Using a randomized controlled trial experiment to determine when student learning occurs. *Journal of Chemical Education*, *97*(1), 27–35. https://doi.org/10.1021/acs.jchemed.9b00767

Chamberlin, T. C. (1890). The method of multiple working hypotheses. *Science*, *15*(366), 92–96. https://doi.org/10.1126/science.ns-15.366.92

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry*

*Education Research and Practice*, *8*(3), 293–307. https://doi.org/10.1039/B7RP90006F

Cheek, K. A. (2013). How geoscience novices reason about temporal duration: The role of spatial thinking and large numbers. *Journal of Geoscience Education*, *61*(3), 334–348. https://doi.org/10.1016/0005-2728(75)90129-2

Cheek, K. A., LaDue, N. D., & Shipley, T. F. (2017). Learning about spatial and temporal scale: Current research, psychological processes, and classroom implications. *Journal of Geoscience Education*, *65*(4), 455–472. https://doi.org/10.5408/16-213.1

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Chinn, C. A., & Brewer, W. F. (2001). Models of data: A theory of how people evaluate data. *Cognition and Instruction*, *19*(3), 323–393. https://doi.org/10.1207/S1532690XCI1903_3

Chomsky, N. (1959). A review of B. F. Skinner's *Verbal Behavior*. *Language*, *35*(1), 26–58.

Christiansen, M. A., Lambert, A. M., Nadelson, L. S., Dupree, K. M., & Kingsford, T. A. (2017). In-class versus at-home quizzes: Which is better? A flipped learning study in a two-site synchronously broadcast organic chemistry course. *Journal of Chemical Education*, *94*(2), 157–163. https://doi.org/10.1021/acs.jchemed.6b00370

Cleary, T. J., & Zimmerman, B. J. (2012). A cyclical self-regulatory account of student engagement: Theoretical foundations and applications. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 237–257). Springer.

Clement, J. (2000). Model based learning as a key research area for science education. *International Journal of Science Education*, *22*(9), 1041–1053. https://doi.org/10.1080/095006900416901

Cohn, T. C., Swanson, E., Whiteman Runs Him, G., Hugs, D., Stevens, L., & Flamm, D. (2014). Placing ourselves on a digital Earth: Sense of place geoscience education in crow country. *Journal of Geoscience Education*, *62*(2), 203–216. https://doi.org/10.5408/12-404.1

Cooper, K. M., Blattman, J. N., Hendrix, T., & Brownell, S. E. (2019). The impact of broadly relevant novel discoveries on student project ownership in a traditional lab course turned CURE. *CBE—Life Sciences Education*, *18*(4), Article ar57. https://doi.org/10.1187/cbe.19-06-0113

Cooper, K. M., & Brownell, S. E. (2016). Coming out in class: Challenges and benefits of active learning in a biology classroom for LGBTQIA students. *CBE—Life Sciences Education*, *15*(3), Article ar37. https://doi.org/10.1187/cbe.16-01-0074

Cooper, M. M., Cox, C. T., Nammouz, M., Case, E., & Stevens, R. (2008). An assessment of the effect of collaborative groups on students' problem-solving strategies and abilities. *Journal of Chemical Education*, *85*(6), 866–872. https://doi.org/10.1021/ed085p866

Cooper, M. M., Posey, L. A., & Underwood, S. M. (2017). Core ideas and topics: Building up or drilling down? *Journal of Chemical Education*, *94*(5), 541–548. https://doi.org/10.1021/acs.jchemed.6b00900

Cooper, M. M., & Stowe, R. L. (2018). Chemistry education research—from personal empiricism to evidence, theory, and informed practice. *Chemical Reviews*, *118*(12), 6053–6087. https://doi.org/10.1021/acs.chemrev.8b00020

Cooper, M. M., Williams, L. C., & Underwood, S. M. (2015). Student understanding of intermolecular forces: A multimodal study. *Journal of Chemical Education*, *92*(8), 1288–1298. https://doi.org/10.1021/acs.jchemed.5b00169

Corwin, L. A., Graham, M. J., & Dolan, E. L. (2015). Modeling course-based undergraduate research experiences: An agenda for future research and evaluation. *CBE—Life Sciences Education*, *14*(1), Article es1. https://doi.org/10.1187/cbe.14-10-0167

Coştu, B., Ayas, A., & Niaz, M. (2010). Promoting conceptual change in first year students' understanding of evaporation. *Chemistry Education Research and Practice*, *11*(1), 5–16. https://doi.org/10.1039/C001041N

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*(9), 970–977. https://doi.org/10.1119/1.1374249

Czajka, C. D., & McConnell, D. (2018). An exploratory study examining undergraduate geology students' conceptions related to geologic time and rates. *Journal of Geoscience Education*, *66*(3), 231–245. https://doi.org/10.1080/10899995.2018.1480826

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences, USA*, *116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

diSessa, A. A. (2013). A bird's-eye view of the "pieces" vs "coherence" controversy (from the "pieces" side of the fence). In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 31–48). Routledge.

Dobaria, A. S. (2018). *Impacts of a course-based undergraduate research experience in introductory astronomy using robotic telescopes* [Master's thesis, San Francisco State University]. CSU Institutional Repository. http://dspace.calstate.edu/bitstream/handle/10211.3/213656/AS362018PHYSD63.pdf

Dolphin, G., Benoit, W., Burylo, J., Hurst, E., Petryshen, W., & Wiebe, S. (2018). Braiding history, inquiry, and model-based learning: A collection of open-source historical case studies for teaching both geology content and the nature of science. *Journal of Geoscience Education*, *66*(3), 205–220. https://doi.org/10.1080/10899995.2018.1475821

Driessen, E. P., Knight, J. K., Smith, M. K., & Ballen, C. J. (2020). Demystifying the meaning of active learning in postsecondary biology education. *CBE—Life Sciences Education*, *19*(4), Article ar52. https://doi.org/10.1187/cbe.20-04-0068

Dufresne, R. J., Gerace, W. J., Leonard, W. J., Mestre, J. P., & Wenk, L. (1996). Classtalk: A classroom communication system for active learning. *Journal of Computing in*

*Higher Education*, *7*(2), 3–47. https://doi.org/10.1007/bf02948592

Dumas, D., Alexander, P. A., & Grossnickle, E. M. (2013). Relational reasoning and its manifestations in the educational context: A systematic review of the literature. *Educational Psychology Review*, *25*(3), 391–427. https://doi.org/10.1007/s10648-013-9224-4

Durham, M. F., Knight, J. K., & Couch, B. A. (2017). Measurement Instrument for Scientific Teaching (MIST): A tool to measure the frequencies of research-based teaching practices in undergraduate science courses. *CBE—Life Sciences Education*, *16*(4), Article ar67. https://doi.org/10.1187/cbe.17-02-0033

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, *48*(2), 90–101. https://doi.org/10.1037//0003-066x.48.2.90

Eddy, S. L., Converse, M., & Wenderoth, M. P. (2015). Portaal: A classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE—Life Sciences Education*, *14*(2), Article ar23. https://doi.org/10.1187/cbe.14-06-0095

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE—Life Sciences Education*, *13*(3), 453–468. https://doi.org/10.1187/cbe.14-03-0050

Eichler, J. F., & Peeples, J. (2016). Flipped classroom modules for large enrollment general chemistry courses: A low barrier approach to increase active learning and improve student grades. *Chemistry Education Research and Practice*, *17*(1), 197–208. https://doi.org/10.1039/C5RP00159E

Fautch, J. M. (2015). The flipped classroom for teaching organic chemistry in small classes: Is it effective? *Chemistry Education Research and Practice*, *16*(1), 179–186. https://doi.org/10.1039/C4RP00230J

Felder, R. M., & Brent, R. (2009). Active learning: An introduction. *ASQ Higher Education Brief*, *2*(4), 1–5.

Felder, R. M., & Brent, R. (2016). *Teaching and learning STEM: A practical guide*. John Wiley & Sons.

Fitzgerald, M., Cutts, R., Salimpour, S., & Slater, S. (2018). A brief overview of robotic telescopes, student research and education research in the English literature. *RTSRE Proceedings*, *1*(1), 1–19. https://doi.org/10.32374/rtsre.2017.001

Flynn, A. B. (2015). Structure and evaluation of flipped chemistry courses: Organic & spectroscopy, large and small, first to third year, English and French. *Chemistry Education Research and Practice*, *16*(2), 198–211. https://doi.org/10.1039/C4RP00224E

Fortus, D., Shwartz, Y., & Rosenfeld, S. (2016). High school students' meta-modeling knowledge. *Research in Science Education*, *46*(6), 787–810. https://doi.org/10.1007/s11165-015-9480-z

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. https://doi.org/10.2307/3516061

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, USA*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, *95*(1), 148–162. https://doi.org/10.1037/0022-0663.95.1.148

Gaffney, J. D. H., & Gaffney, A. L. H. (2016). Student satisfaction in interactive engagement-based physics classes. *Physical Review Physics Education Research*, *12*(2), Article 020125. https://doi.org/10.1103/PhysRevPhysEducRes.12.020125

Gafney, L., & Varma-Nelson, P. (2007). Evaluating peer-led team learning: A study of long-term effects on former workshop peer leaders. *Journal of Chemical Education*, *84*(3), 535–539. https://doi.org/10.1021/ed084p535

Gagnier, K. M., Atit, K., Ormand, C. J., & Shipley, T. F. (2017). Comprehending 3D diagrams: Sketching to support spatial reasoning. *Topics in Cognitive Science*, *9*(4), 883–901. https://doi.org/10.1111/tops.12233

Galison, P. (1997). *Image & logic: A material culture of microphysics*. The University of Chicago Press.

Garcia, T., & Pintrich, P. R. (1994). Regulating motivation and cognition in the classroom: The role of self schemas and self-regulatory strategies. In D. Schunk & B. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 127–154). Erlbaum.

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., & Clay-Chambers, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *Journal of Research in Science Teaching*, *45*(8), 922–939. https://doi.org/10.1002/tea.20248

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*(2), 155–170. https://doi.org/10.1016/S0364-0213(83)80009-3

Giere, R. N. (2004). How models are used to represent reality. *Philosophy of Science*, *71*(5), 742–752. https://doi.org/10.1086/425063

Gilley, B., Atchison, C., Feig, A., & Stokes, A. (2015). Impact of inclusive field trips. *Nature Geoscience*, *8*(8), 579–580. https://doi.org/10.1038/ngeo2500

Gobert, J. D., & Clement, J. J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching*, *36*(1), 39–53. https://doi.org/10.1002/(sici)1098-2736(199901)36:1<39::aid-tea4>3.0.co;2-i

Goldberg, F., Price, E., Robinson, S., Boyd-Harlow, D., & McKean, M. (2012). Developing the learning physical science curriculum: Adapting a small enrollment, laboratory and discussion based physical science course for large enrollments. *Physical Review Special Topics-Physics Education Research*, *8*(1), Article 010121. https://doi.org/10.1103/PhysRevSTPER.8.010121

Gonzalez-Espada, W. J., & LaDue, D. S. (2006). Evaluation of the impact of the NWC REU program compared with other undergraduate research experiences. *Journal* *of Geoscience Education*, *54*(5), 541–549. https://doi.org/10.5408/1089-9995-54.5.541

Gosselin, D. C., Egger, A. E., & Taber, J. J. (2019). *Interdisciplinary teaching about Earth and the environment for a sustainable future*. Springer.

Green, P. J. (2003). *Peer instruction for astronomy*. Prentice Hall.

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2004). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. *Contemporary Educational Psychology*, *29*(4), 462–482. https://doi.org/10.1016/j.cedpsych.2004.01.006

Grissom, A. N., Czajka, C. D., & McConnell, D. A. (2015). Revisions of physical geology laboratory courses to increase the level of inquiry: Implications for teaching and learning. *Journal of Geoscience Education*, *63*(4), 285–296. https://doi.org/10.5408/14-050.1

Gross, N. A., & Lopez, R. E. (2009). Advanced undergraduate and early graduate physics students' misconception about solar wind flow: Evidence of students' difficulties in distinguishing paradigms. *Astronomy Education Review*, *8*(1), Article 010102. https://doi.org/10.3847/AER2009013

Grunspan, D. Z., Kline, M. A., & Brownell, S. E. (2018). The lecture machine: A cultural evolutionary model of pedagogy in higher education. *CBE—Life Sciences Education*, *17*(3), Article es6. https://doi.org/10.1187/cbe.17-12-0287

Gunstone, R. F., & White, R. T. (1981). Understanding of gravity. *Science Education*, *65*(3), 291–299. https://doi.org/10.1002/sce.3730650308

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, *332*(6034), 1213–1216. https://doi.org/10.1126/science.1204820

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74. https://doi.org/10.1119/1.18809

Hambrick, D. Z., Libarkin, J. C., Petcovic, H. L., Baker, K. M., Elkins, J., Callahan, C. N., Turner, S. P., Rench, T. A., & LaDue, N. D. (2012). A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. *Journal of Experimental Psychology: General*, *141*(3), 397–403. https://doi.org/10.1037/a0025927

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. Transfer of learning from a modern multidisciplinary perspective. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 89–119). Information Age Publishing.

Hammersley, L. C., Levine, R., Cornwell, K., Kusnick, J. E., & Hausback, B. P. (2012). The geology of Mexico: A quantitative evaluation of a course designed to increase the number of Hispanic students participating in the geosciences at California State University, Sacramento. *Journal of Geoscience Education*, *60*(2), 189–198. https://doi.org/10.5408/11-243.1

Hancock, T. S., Friedrichsen, P. J., Kinslow, A. T., & Sadler, T. D. (2019). Selecting socio-scientific issues for teaching. *Science & Education*, *28*(6–7), 639–667. https://doi.org/10.1007/s11191-019-00065-x

Hannula, K. A. (2019). Do geology field courses improve penetrative thinking? *Journal of Geoscience Education*, *67*(2), 143–160. https://doi.org/10.1080/10899995.2018.1548004

Hartman, J. R., Dahm, D. J., & Nelson, E. A. (2015). ConfChem conference on flipped classroom: Time-saving resources aligned with cognitive science to help instructors. *Journal of Chemical Education*, *92*(9), 1568–1569. https://doi.org/10.1021/ed5009156

Hartten, L. M., & LeMone, M. A. (2010). The evolution and current state of the atmospheric sciences "pipeline." *Bulletin of the American Meteorological Society*, *91*(7), 942–956. https://doi.org/10.1175/2010BAMS2537.1

Hartten, L. M., & LeMone, M. A. (2014). Corrigendum: The evolution and current state of the atmospheric sciences "pipeline." *Bulletin of the American Meteorological Society*, *95*(5), Article 1.

Heddy, B. C., & Sinatra, G. M. (2013). Transforming misconceptions: Using transformative experience to promote positive affect and conceptual change in students learning about biological evolution. *Science Education*, *97*(5), 723–744. https://doi.org/10.1002/sce.21072

Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching*, *48*(8), 952–984. https://doi.org/10.1002/tea.20439

Henderson, C., Connolly, M., Dolan, E. L., Finkelstein, N., Franklin, S., Malcom, S., Rasmussen, C., Redd, K., & John, K. S. (2017). Towards the STEM DBER alliance: Why we need a discipline-based STEM education research community. *Journal of Engineering Education*, *106*, 349–355. https://doi.org/10.1002/jee.20168

Henderson, C., & Dancy, M. H. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics-Physics Education Research*, *3*(2), Article 020102. https://doi.org/10.1103/PhysRevSTPER.3.020102

Henderson, C., & Dancy, M. H. (2008). Physics faculty and educational researchers: Divergent expectations as barriers to the diffusion of innovations. *American Journal of Physics*, *76*(1), 79–91. https://doi.org/10.1119/1.2800352

Henderson, C., & Dancy, M. H. (2009). Impact of physics education research on the teaching of introductory quantitative physics in the United States. *Physical Review Special Topics-Physics Education Research*, *5*(2), Article 020107. https://doi.org/10.1103/PhysRevSTPER.5.020107

Henderson, C., & Dancy, M. H. (2011, February). *Increasing the impact and diffusion of STEM education innovations*. Invited paper for the National Academy of Engineering, Center for the Advancement of Engineering Education Forum, Impact and Diffusion of Transformative Engineering Education Innovations. https://www.nae.edu/File.aspx?id=36304

Henderson, C., Dancy, M. H., & Niewiadomska-Bugaj, M. (2012). Use of research-based instructional strategies in introductory physics: Where do faculty leave the innovation-decision process? *Physical Review Special Topics-Physics Education Research*, *8*(2), Article 020104. https://doi.org/10.1103/PhysRevSTPER.8.020104

Hendricks, J. E., Atchison, C. L., & Feig, A. D. (2017). Effective use of personal assistants for students with disabilities: Lessons learned from the 2014 accessible geoscience field trip. *Journal of Geoscience Education*, *65*(1), 72–80. https://doi.org/10.5408/16-185.1

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics*, *55*(5), 440–454. https://doi.org/10.1119/1.15129

Hewlett, J. (2009). The search for synergy: Undergraduate research at the community college. In B. D. Cejda & N. Hensel (Eds.), *Undergraduate research at community colleges* (pp. 9–18). Council on Undergraduate Research.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark. *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368

Hockings, S. C., DeAngelis, K. J., & Frey, R. F. (2008). Peer-led team learning in general chemistry: Implementation and evaluation. *Journal of Chemical Education*, *85*(7), 990–996. https://doi.org/10.1021/ed085p990

Hofstra, B., Kulkarni, V. V., Galvez, S. M. N., He, B., Jurafsky, D., & McFarland, D. A. (2020). The diversity-innovation paradox in science. *Proceedings of the National Academy of Sciences, USA*, *117*(17), 9284–9291. https://doi.org/10.1073/pnas.1915378117

Holder, L. N., Scherer, H. H., & Herbert, B. E. (2017). Student learning of complex Earth systems: A model to guide development of student expertise in problem-solving. *Journal of Geoscience Education*, *65*(4), 490–505. https://doi.org/10.5408/17-261.1

Holmes, N. G., Olsen, J., Thomas, J. L., & Wieman, C. E. (2017). Value added or misattributed? A multi-institution study on the educational benefit of labs for reinforcing physics content. *Physical Review Physics Education Research*, *13*(1), Article 010129. https://doi.org/10.1103/PhysRevPhysEducRes.13.010129

Hora, M. T., Oleson, A., & Ferrare, J. J. (2013). *Teaching Dimensions Observation Protocol (TDOP) user's manual*. Wisconsin Center for Education Research. http://tdop.wceruw.org/Document/TDOP-Users-Guide.pdf

Hudgins, D. W., Prather, E. E., Grayson, D. J., & Smits, D. P. (2006). Effectiveness of collaborative ranking tasks on student understanding of key astronomy concepts. *Astronomy Education Review*, *5*(1), 1–22. https://doi.org/10.3847/AER2006001

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology*, *102*(3), 588–600. https://doi.org/10.1037/a0019682

Järvelä, S., & Renninger, K. A. (2014). Designing for learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 668–685). Cambridge University Press.

Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C., Shipley, T. F., Tikoff, B., Ormand, C. J., & Sageman, B. (2010). Commentary: Analogical thinking in geoscience education. *Journal of Geoscience Education*, *58*(1), 2–13. https://doi.org/10.5408/1.3544291

Johnson, J. K., & Reynolds, S. J. (2005). Concept sketches—using student- and instructor-generated, annotated sketches for learning, teaching, and assessment in geology courses. *Journal of Geoscience Education*, *53*(1), 85–95. https://doi.org/10.5408/1089-9995-53.1.85

Jolley, A., Kennedy, B. M., Brogt, E., Hampton, S. J., & Fraser, L. (2018). Are we there yet? Sense of place and the student experience on roadside and situated geology field trips. *Geosphere*, *14*(2), 651–667.

Kalyuga, S. (2007). Enhancing instructional efficiency of interactive e-learning environments: A cognitive load perspective. *Educational Psychology Review*, *19*(3), 387–399. https://doi.org/10.1007/s10648-007-9051-6

Karlstrom, K., Semken, S., Crossey, L., Perry, D., Gyllenhaal, E. D., Dodick, J., Williams, M., Hellmich-Bryan, J., Crow, R., Bueno, N., & Ault, C. (2008). Informal geoscience education on a grand scale: The trail of time exhibition at Grand Canyon. *Journal of Geoscience Education*, *56*(4), 354–361. https://doi.org/10.5408/informal_geoscience_education_

Kastens, K. A., & Manduca, C. A. (2012). *Earth and mind II: A synthesis of research on thinking and learning in the geosciences* (Vol. 486). Geological Society of America.

Kastens, K. A., & Manduca, C. A. (2017). Using systems thinking in the design, implementation, and evaluation of complex educational innovations, with examples from the integrate project. *Journal of Geoscience Education*, *65*(3), 219–230. https://doi.org/10.5408/16-225.1

Kirkpatrick, C., Schuchardt, A., Baltz, D., & Cotner, S. (2019). Computer-based and bench-based undergraduate research experiences produce similar attitudinal outcomes. *CBE—Life Sciences Education*, *18*(1), Article ar10. https://doi.org/10.1187/cbe.18-07-0112

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kortz, K. M., & Smay, J. J. (2010). *Lecture tutorials for introductory geoscience*. W. H. Freeman and Company.

Kortz, K. M., Smay, J. J., & Murray, D. P. (2008). Increasing learning in introductory geoscience courses using lecture tutorials. *Journal of Geoscience Education*, *56*(3), 280–290. https://doi.org/10.5408/1089-9995-56.3.280

Kortz, K. M., & van der Hoeven Kraft, K. J. (2016). Geoscience education research project: Student benefits and effective design of a course-based undergraduate research experience. *Journal of Geoscience Education*, *64*(1), 24–36. https://doi.org/10.5408/15-11.1

Kozhevnikov, M., Motes, M. A., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive*

*Science*, *31*(4), 549–579. https://doi.org/10.1080/15326900701399897

Kuh, G. (2008). *High-impact educational practices: What are they, who has access to them, and why they matter*. Association of American Colleges and Universities.

LaDue, N. D., McNeal, P. M., Ryker, K., St John, K., & van der Hoeven, K. J. (2020). *Using an engagement lens to model active learning in the geosciences*. Manuscript submitted for publication.

Laverty, J. T., Underwood, S. M., Matz, R. L., Posey, L. A., Carmel, J. H., Caballero, M. D., Fata-Hartley, C. L., Ebert-May, D., Jardeleza, S. E., & Cooper, M. M. (2016). Characterizing college science assessments: The three-dimensional learning assessment protocol. *PLOS ONE*, *11*(9), Article e0162333. https://doi.org/10.1371/journal.pone.0162333

Laws, P. (2004). *Workshop physics activity guide* (Vol. 1–4). John Wiley & Sons.

Lazendic-Galloway, J., Fitzgerald, M., & McKinnon, D. H. (2016). Implementing a studio-based flipped classroom in a first year astronomy course. *International Journal of Innovation in Science and Mathematics Education*, *24*(5), 35–47.

Lehrer, R., & Schauble, L. (2000). Developing model-based reasoning in mathematics and science. *Journal of Applied Developmental Psychology*, *21*(1), 39–48. https://doi.org/10.1016/s0193-3973(99)00049-0

Lewin, J. D., Vinson, E. L., Stetzer, M. R., & Smith, M. K. (2016). A campus-wide investigation of clicker implementation: The status of peer discussion in STEM classes. *CBE—Life Sciences Education*, *15*(1), Article ar6. https://doi.org/10.1187/cbe.15-10-0224

Liew, C. W., & Treagust, D. F. (1995). A predict-observe-explain teaching sequence for learning about students' understanding of heat. *Australian Science Teachers Journal*, *41*(1), 68–71.

Linnenbrink-Garcia, L., Rogat, T. K., & Koskey, K. L. (2011). Affect and engagement during small group instruction. *Contemporary Educational Psychology*, *36*(1), 13–24. https://doi.org/10.1016/j.cedpsych.2010.09.001

Liu, S. N. C., & Beaujean, A. A. (2017). The effectiveness of team-based learning on academic outcomes: A meta-analysis. *Scholarship of Teaching and Learning in Psychology*, *3*(1), 1–14. https://doi.org/10.1037/stl0000075

Lombardi, D., Bailey, J. M., Bickel, E. S., & Burrell, S. (2018). Scaffolding scientific thinking: Students' evaluations and judgments during Earth science knowledge construction. *Contemporary Educational Psychology*, *54*, 184–198. https://doi.org/10.1016/j.cedpsych.2018.06.008

Lombardi, D., Nussbaum, E. M., & Sinatra, G. M. (2016). Plausibility judgments in conceptual change and epistemic cognition. *Educational Psychologist*, *51*(1), 35–56. https://doi.org/10.1080/00461520.2015.1113134

Lombardi, D., Sinatra, G. M., & Nussbaum, E. M. (2013). Plausibility reappraisals and shifts in middle school students' climate change conceptions. *Learning and Instruction*, *27*, 50–62. https://doi.org/10.1016/j.learninstruc.2013.03.001

LoPresto, M. C., & Murrell, S. R. (2009). Using the Star Properties Concept Inventory to compare instruction with lecture tutorials to traditional lectures. *Astronomy Education Review*, *8*(1), Article 010105. https://doi.org/10.3847/AER2009014

Lou, A. J., & Jaeggi, S. M. (2019). Reducing the prior knowledge achievement gap by using technology-assisted guided learning in an undergraduate chemistry course. *Journal of Research in Science Teaching*, *57*(3), 368–392. https://doi.org/10.1002/tea.21596

Lukes, L. A., Ryker, K., Millsaps, C., Lockwood, R., Uhen, M. D., George, C., Bentley, C., & Berquist, P. (2019). Leveraging a large database to increase access to undergraduate research experiences. *Scholarship and Practice of Undergraduate Research*, *2*(4), 4–13. https://doi.org/10.18833/spur/2/4/6

Lundberg, C. A., Kim, Y. K., Andrade, L. M., & Bahner, D. T. (2018). High expectations, strong support: Faculty behaviors predicting Latina/o community college student learning. *Journal of College Student Development*, *59*(1), 55–70. https://doi.org/10.1353/csd.2018.0004

Luo, W., Pelletier, J., Duffin, K., Ormand, C., Hung, W. C., Shernoff, D. J., Zhai, X., Iverson, E., Whalley, K., Gallaher, C., & Furness, W. (2016). Advantages of computer simulation in enhancing students' learning about landform evolution: A case study using the Grand Canyon. *Journal of Geoscience Education*, *64*(1), 60–73. https://doi.org/10.5408/15-080.1

Lyman, F. (1981). The responsive classroom discussion: The inclusion of all students. In A. S. Anderson (Ed.), *Mainstreaming digest* (pp. 109–113). University of Maryland College of Education.

MacArthur, J. R., & Jones, L. L. (2008). A review of literature reports of clickers applicable to college chemistry classrooms. *Chemistry Education Research and Practice*, *9*(3), 187–195. https://doi.org/10.1039/B812407H

Mahalingam, M., Schaefer, F., & Morlino, E. (2008). Promoting student learning through group problem solving in general chemistry recitations. *Journal of Chemical Education*, *85*(11), 1577–1581. https://doi.org/10.1021/ed085p1577

Manduca, C. A., Iverson, E. R., Luxenberg, M., Macdonald, R. H., McConnell, D. A., Mogk, D. W., & Tewksbury, B. J. (2017). Improving undergraduate STEM education: The efficacy of discipline-based professional development. *Science Advances*, *3*(2), Article e1600193. https://doi.org/10.1126/sciadv.1600193

Maria, A. H., Millam, E. L., & Wright, C. L. (2011). Using a differential scanning calorimeter to teach phase equilibria to students of igneous and metamorphic petrology. *Journal of Geoscience Education*, *59*(2), 63–70. https://doi.org/10.5408/1.3580766

Marshall, H. H. (1996). Implications of differentiating and understanding constructivist approaches. *Educational Psychologist*, *31*(3–4), 235–240. https://doi.org/10.1080/00461520.1996.9653270

Mason, N. A., Brunner, R. M., Ballen, C. J., & Lovette, I. J. (2018). Cognitive and social benefits among underrepresented first-year biology students in a field course: A case study of experiential learning in the Galápagos. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *30*(3), 1–19.

Mayer, R. E. (1992). Cognition and instruction: Their historic meeting within educational psychology. *Journal of Educational Psychology*, *84*(4), 405–412. https://doi.org/10.1037/0022-0663.84.4.405

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, *59*(1), 14–19. https://doi.org/10.1037/0003-066X.59.1.14

Mazur, E. (1997). *Peer instruction: A user's manual*. Prentice Hall.

McConnell, D. A., Chapman, L., Czajka, C. D., Jones, J. P., Ryker, K. D., & Wiggen, J. (2017). Instructional utility and learning efficacy of common active learning strategies. *Journal of Geoscience Education*, *65*(4), 604–625. https://doi.org/10.5408/17-249.1

McConnell, D. A., Steer, D. N., Owens, K. D., Knott, J. R., Van Horn, S., Borowski, W., Dick, J., Foos, A., Malone, M., McGrew, H., & Greer, L. (2006). Using Conceptests to assess and improve student conceptual understanding in introductory geoscience courses. *Journal of Geoscience Education*, *54*(1), 61–68. https://doi.org/10.5408/1089-9995-54.1.61

McDaniel, M. A., Stoen, S. M., Frey, R. F., Markow, Z. E., Hynes, K. M., Zhao, J., & Cahill, M. J. (2016). Dissociative conceptual and quantitative problem solving outcomes across interactive engagement and traditional format introductory physics. *Physical Review Physics Education Research*, *12*(2), Article 020141. https://doi.org/10.1103/PhysRevPhysEducRes.12.020141

McDermott, L. C., & Shaffer, P. S. (1998). *Tutorials in introductory physics* (Preliminary ed.). Prentice Hall.

Meadows, D. H. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing.

Meltzer, D. E., & Thornton, R. K. (2012). Resource letter ALIP-1: Active-learning instruction in physics. *American Journal of Physics*, *80*(6), 478–496. https://doi.org/10.1119/1.3678299

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development*, *50*(3), 43–59. https://doi.org/10.1007/BF02505024

Mestre, J. P. (2001). Implications of research on learning for the education of prospective science and physics teachers. *Physics Education*, *36*(1), 44–51. https://doi.org/10.1088/0031-9120/36/1/308

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (Eds.). (2000). *Assessing science understanding: A human constructivist view*. Academic Press.

Momsen, J. L., Long, T. M., Wyse, S. A., & Ebert-May, D. (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE—Life Sciences Education*, *9*(4), 435–440. https://doi.org/10.1187/cbe.10-01-0001

Monaghan, J. M., & Clement, J. (1999). Use of a computer simulation to develop mental simulations for understanding relative motion concepts. *International Journal of Science Education*, *21*(9), 921–944. https://doi.org/10.1080/095006999290237

Monaghan, J. M., & Clement, J. (2000). Algorithms, visualization, and mental models: High school students' interactions with a relative motion simulation. *Journal of Science Education and Technology*, *9*(4), 311–325. https://doi.org/10.1023/a:1009480425377

Moog, R. S., Creegan, F. J., Hanson, D. M., Spencer, J. N., Straumanis, A., Bunce, D. M., & Wolfskill, T. (2009). POGIL: Process-Oriented Guided-Inquiry Learning. In N. J. Pienta, M. M. Cooper, & T. J. Greenbowe (Eds.), *Chemists' guide to effective teaching: Vol. II* (pp. 90–107). Prentice Hall.

Moog, R. S., & Spencer, J. N. (2008). POGIL: An overview. In R. S. Moog & J. N. Spencer (Eds.), *Process oriented guided inquiry learning (POGIL)* (Vol. 994, pp. 1–13). American Chemical Society.

National Academies of Sciences, Engineering, and Medicine. (2016). *Barriers and opportunities for 2-year and 4-year STEM degrees: Systemic change to support diverse student pathways*. The National Academies Press. https://doi.org/10.17226/21739

National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: Learners, contexts, and cultures*. The National Academies Press. https://doi.org/10.17226/24783

National Research Council. (2012a). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. The National Academies Press. https://doi.org/10.17226/13362

National Research Council. (2012b). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press. https://doi.org/10.17226/13165

Nersessian, N. (2002). The cognitive basis of model-based reasoning in science. In P. Carruthers (Ed.), *The cognitive basis of science* (pp. 133–153). Cambridge University Press.

Newell, A., Shaw, J. C., & Simon, H. A. (1958). Elements of a theory of human problem solving. *Psychological Review*, *65*(3), 151–166. https://doi.org/10.1037/h0048495

Nottis, K. E. K. (1999). Using analogies to teach plate-tectonics concepts. *Journal of Geoscience Education*, *47*(5), 449–454. https://doi.org/10.5408/1089-9995-47.5.449

Novak, G. M., Patterson, E. T., Gavrin, A. D., & Christian, W. (1999). Just in Time Teaching. *American Journal of Physics*, *67*(10), 937–938. https://doi.org/10.1119/1.19159

O'Connell, K., Hoke, K., Berkowitz, A., Branchaw, J., & Storksdieck, M. (2020). Undergraduate learning in the field: Designing experiences, assessing outcomes, and exploring future opportunities. *Journal of Geoscience Education*. Advance online publication. https://doi.org/10.1080/10899995.2020.1779567

Orion, N., & Hofstein, A. (1994). Factors that influence learning during a scientific field trip in a natural environment.

*Journal of Research in Science Teaching*, *31*(10), 1097–1119. https://doi.org/10.1002/tea.3660311005

Ormand, C. J., Manduca, C., Shipley, T. F., Tikoff, B., Harwood, C. L., Atit, K., & Boone, A. P. (2014). Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study. *Journal of Geoscience Education*, *62*(1), 146–154. https://doi.org/10.5408/13-027.1

Ormrod, J. E. (2017). *How we think and learn: Theoretical perspectives and practical implications*. Cambridge University Press.

Ovink, S. M., & Veazey, B. D. (2011). More than "getting us through:" A case study in cultural capital enrichment of underrepresented minority undergraduates. *Research in Higher Education*, *52*(4), 370–394. https://doi.org/10.1007/s11162-010-9198-8

Patall, E. A., Pituch, K. A., Steingut, R. R., Vasquez, A. C., Yates, N., & Kennedy, A. A. (2019). Agency and high school science students' motivation, engagement, and classroom support experiences. *Journal of Applied Developmental Psychology*, *62*, 77–92. https://doi.org/10.1016/j.appdev.2019.01.004

Peirce, C. S. (1931). *Collected papers of Charles Sanders Peirce*. Harvard University Press.

Perkins, K., Adams, W., Dubson, M., Finkelstein, N., Reid, S., Wieman, C., & LeMaster, R. (2006). PhET: Interactive simulations for teaching and learning physics. *The Physics Teacher*, *44*(1), 18–23. https://doi.org/10.1119/1.2150754

Piaget, J. (1926). *The language and thought of the child* (M. Gabain & R. Gabain, Trans.). Routledge and Kegan Paul.

Pickering, A. (1995). *The mangle of practice: Time, agency, and science*. University of Chicago Press.

Prather, E. E., Rudolph, A. L., Brissenden, G., & Schlingman, W. M. (2009). A national study assessing the teaching and learning of astronomy. Part I. The effect of interactive instruction. *American Journal of Physics*, *77*(4), 320–330. https://doi.org/10.1119/1.3065023

Prather, E. E., Slater, T. F., Adams, J. P., Bailey, J. M., Jones, L. V., & Dostal, J. A. (2004). Research on a lecture-tutorial approach to teaching introductory astronomy for non-science majors. *Astronomy Education Review*, *3*(2), 122–136. https://doi.org/10.3847/AER2004019

Prather, E. E., Slater, T. F., Adams, J. P., & Brissenden, G. (2013). *Lecture-tutorials for introductory astronomy* (3rd ed.). Pearson Addison-Wesley.

Prather, E. E., & Wallace, C. S. (2020). Lecture-tutorials in introductory astronomy. In C. Impey & S. Buxner (Eds.), *Astronomy education, Vol. 1: Evidence-based instruction for introductory courses* (3-1–3-21). IOP Publishing.

Prevost, L. B., Smith, M. K., & Knight, J. K. (2016). Using student writing and lexical analysis to reveal student thinking about the role of stop codons in the central dogma. *CBE—Life Sciences Education*, *15*(4), Article ar65. https://doi.org/10.1187/cbe.15-12-0267

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, *93*(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Pugh, K. J., Phillips, M. M., Sexton, J. M., Bergstrom, C. M., & Riggs, E. M. (2019). A quantitative investigation of geoscience departmental factors associated with the recruitment and retention of female students. *Journal of Geoscience Education*, *67*(3), 266–284. https://doi.org/10.1080/10899995.2019.1582924

Reeve, J., & Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology*, *36*(4), 257–267. https://doi.org/10.1016/j.cedpsych.2011.05.002

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3–19). Springer.

Resnick, I., Davatzes, A., Newcombe, N. S., & Shipley, T. F. (2017a). Using relational reasoning to learn about scientific phenomena at unfamiliar scales. *Educational Psychology Review*, *29*(1), 11–25. https://doi.org/10.1007/s10648-016-9371-5

Resnick, I., Newcombe, N. S., & Shipley, T. F. (2017b). Dealing with big numbers: Representation and understanding of magnitudes outside of human experience. *Cognitive Science*, *41*(4), 1020–1041. https://doi.org/10.1111/cogs.12388

Resnick, I., & Shipley, T. F. (2013). Breaking new ground in the mind: An initial study of mental brittle transformation and mental rigid rotation in science experts. *Cognitive Processing*, *14*(2), 143–152. https://doi.org/10.1007/s10339-013-0548-2

Richardson, G. P. (1991). *Feedback thought in social sciences and systems theory*. University of Pennsylvania Press.

Riggs, E. M. (2005). Field-based education and indigenous knowledge: Essential components of geoscience education for Native American communities. *Science Education*, *89*(2), 296–313. https://doi.org/10.1002/sce.20032

Riggs, E. M., Lieder, C. C., & Balliet, R. (2009). Geologic problem solving in the field: Analysis of field navigation and mapping by advanced undergraduates. *Journal of Geoscience Education*, *57*(1), 48–63. https://doi.org/10.5408/1.3559525

Rogoff, B., Matusov, E., & White, C. (1996). Models of teaching and learning: Participation in a community of learners. In D. Olson & N. Torrance (Eds.), *Handbook of education and human development: New models of learning, teaching, and schooling* (pp. 388–414). Basil Blackwell.

Rose, A. J., & Rudolph, K. D. (2006). A review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin*, *132*(1), 98–131. https://doi.org/10.1037/0033-2909.132.1.98

Roth, W.-M. (2001). Gestures: Their role in teaching and learning. *Review of Educational Research*, *71*(3), 365–392. https://doi.org/10.3102/00346543071003365

Ruppert, J., Duncan, R. G., & Chinn, C. A. (2019). Disentangling the role of domain-specific knowledge in student modeling. *Research in Science Education*, *49*(3), 921–948. https://doi.org/10.1007/s11165-017-9656-9

Ruzhitskaya, L., & Speck, A. (2011). Peer interaction: Help or distraction? In *American Astronomical Society Meeting Abstracts*, *217*, 405–407.

Ryker, K. D., & McConnell, D. A. (2017). Assessing inquiry in physical geology laboratory manuals. *Journal of Geoscience Education*, *65*(1), 35–47. https://doi.org/10.5408/14-036.1

Sattizahn, J. R., Lyons, D. J., Kontra, C., Fischer, S. M., & Beilock, S. L. (2015). In physics education, perception matters. *Mind, Brain, and Education*, *9*(3), 164–169. https://doi.org/10.1111/mbe.12085

Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The Reformed Teaching Observation Protocol. *School Science and Mathematics*, *102*(6), 245–253. https://doi.org/10.1111/j.1949-8594.2002.tb17883.x

Scherer, H. H., Holder, L., & Herbert, B. (2017). Student learning of complex Earth systems: Conceptual frameworks of Earth systems and instructional design. *Journal of Geoscience Education*, *65*(4), 473–489. https://doi.org/10.5408/16-208.1

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475–522. https://doi.org/10.1207/s1532690xci1604_4

Schwarz, C. V., Passmore, C., & Reiser, B. J. (2017). *Helping students make sense of the world using next generation science and engineering practices*. NSTA Press.

Schwarz, C. V., & White, B. (2005). Meta-modeling knowledge: Developing students understanding of scientific modeling. *Cognition and Instruction*, *23*(2), 165–205. https://doi.org/10.1207/s1532690xci2302_1

Semken, S., & Freeman, C. B. (2008). Sense of place in the practice and assessment of place-based science teaching. *Science Education*, *92*(6), 1042–1057. https://doi.org/10.1002/sce.20279

Sexton, J. M., O'Connell, S., Banning, J. H., & Most, D. E. (2014). Characteristics and culture of geoscience departments as interpreted from their website photographs. *Journal of Women and Minorities in Science and Engineering*, *20*(3), 168–178. https://doi.org/10.5408/11-249.1

Shipley, T. F., McConnell, D., McNeal, K. S., Petcovic, H. L., & St John, K. E. (2017). Transdisciplinary science education research and practice: Opportunities for GER in a developing STEM discipline based education research alliance (DBER-A). *Journal of Geoscience Education*, *65*(4), 354–362. https://doi.org/10.5408/1089-9995-65.4.354

Shipley, T. F., & Tikoff, B. (2017). The role of geoscience education research in the consilience between science of the mind and science of the natural world. *Journal of Geoscience Education*, *65*(4), 393–398. https://doi.org/10.5408/16-252.1

Shipley, T. F., Tikoff, B., Ormand, C. J., & Manduca, C. (2013). Structural geology practice and learning, from the perspective of cognitive science. *Journal of Structural Geology*, *54*, 72–84. https://doi.org/10.1016/j.jsg.2013.07.005

Sibley, D. F. (2009). A cognitive framework for reasoning with scientific models. *Journal of Geoscience Education*, *57*(4), 255–263. https://doi.org/10.5408/1.3559672

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist*, *50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924

Smaldino, P. (2019). Better methods can't make up for mediocre theory. *Nature*, *575*(7781), Article 9. https://doi.org/10.1038/d41586-019-03350-5

Smith, M. K., Brownell, S. E., Crowe, A. J., Holmes, N. G., Knight, J. K., Semsar, K., Summers, M. M., Walsh, C., Wright, C. D., Couch, B. A., & Couch, B. A. (2019). Tools for change: Measuring student conceptual understanding across undergraduate biology programs using Bio-MAPS assessments. *Journal of Microbiology & Biology Education*, *20*(2), 1–5. https://doi.org/10.1128%2Fjmbe.v20i2.1787

Smith, M. K., Jones, F. H., Gilbert, S. L., & Wieman, C. E. (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE—Life Sciences Education*, *12*(4), 618–627. https://doi.org/10.1187/cbe.13-08-0154

Smith, M. K., Wood, W. B., Krauter, K., & Knight, J. K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE—Life Sciences Education*, *10*(1), 55–63. https://doi.org/10.1187/cbe.10-08-0101

Sokoloff, D. R., & Thornton, R. K. (1997). Using interactive lecture demonstrations to create an active learning environment. *The Physics Teacher*, *35*(6), 340–347. https://doi.org/10.1119/1.2344715

Sokoloff, D. R., Thornton, R. K., & Laws, P. W. (2004). *RealTime physics: Active learning laboratories*. Wiley.

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan, M. K., Jr., Esson, M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., Mcdonnell, L. M., McKay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, T. M., . . . Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, *359*(6383), 1468–1470. https://doi.org/10.1126/science.aap8892

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613–629. https://doi.org/10.1037//0003-066x.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Stieff, M., Ryu, M., Dixon, B., & Hegarty, M. (2012). The role of spatial ability and strategy preference for spatial problem solving in organic chemistry. *Journal of Chemical Education*, *89*(7), 854–859. https://doi.org/10.1021/ed200071d

Stowe, R. L., & Cooper, M. M. (2019). Assessment in chemistry education. *Israel Journal of Chemistry*, *59*, 1–11. https://doi.org/10.1002/ijch.201900024

Summa, L., Keane, C., & Mosher, S. (2017). Meeting changing workforce needs in geoscience with new thinking about undergraduate education. *GSA Today*, *27*(9), 60–61. https://doi.org/10.1130/GSATG342GW.1

Teasdale, R., Viskupic, K., Bartley, J. K., McConnell, D., Manduca, C., Bruckner, M. Z., Farthing, D., & Iverson, E. (2017). A multidimensional assessment of reformed teaching practice in geoscience classrooms. *Geosphere*, *13*(2), 608–627. https://doi.org/10.1130/GES01479.1

Theobald, E. J., Crowe, A., HilleRisLambers, J., Wenderoth, M. P., & Freeman, S. (2015). Women learn more from local than global examples of the biological impacts of climate change. *Frontiers in Ecology and the Environment*, *13*(3), 132–137. https://doi.org/10.1890/140261

Todd, C. E. D., & Goeke, E. R. (2012). Incorporating student-led field trips and learner-centered teaching in a capstone geology course. *Journal of Geoscience Education*, *60*(3), 268–276. https://doi.org/10.5408/11-255.1

Underwood, S. M., Posey, L. A., Herrington, D. G., Carmel, J. H., & Cooper, M. M. (2017). Adapting assessment tasks to support three-dimensional learning. *Journal of Chemical Education*, *95*(2), 207–217. https://doi.org/10.1021/acs.jchemed.7b00645

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why and how. *Psychology of Learning and Motivation*, *57*, 147–181. https://10.1016/b978-0-12-394293-7.00004-2

van der Hoeven Kraft, K. J., Srogi, L., Husman, J., Semken, S., & Fuhrman, M. (2011). Engaging students to learn through the affective domain: A new framework for teaching in the geosciences. *Journal of Geoscience Education*, *59*(2), 71–84. https://doi.org/10.5408/1.3543934a

Van Heuvelen, A., & Etkina, E. (2005). *The physics active learning guide*. Pearson/Addison-Wesley.

Viskupic, K., Ryker, K., Teasdale, R., Manduca, C., Iverson, E., Farthing, D., Bruckner, M. Z., & McFadden, R. (2019). Classroom observations indicate the positive impacts of discipline-based professional development. *Journal for STEM Education Research*, *2*, 201–228. https://doi.org/10.1007/s41979-019-00015-w

Vosniadou, S., & Skopeliti, I. (2014). Conceptual change from the framework theory side of the fence. *Science & Education*, *23*(7), 1427–1445. https://doi.org/10.1007/s11191-013-9640-3

Vygotsky, L. S. (1968). *Thought and language* (A. Kozulin, Trans.). MIT Press. (Original work published 1934).

Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.

Walker, J. D., Cotner, S. H., Baepler, P. M., & Decker, M. D. (2008). A delicate balance: Integrating active learning into a large lecture course. *CBE—Life Sciences Education*, *7*(4), 361–367. https://doi.org/10.1187/cbe.08-02-0004

Walker, L., & Warfa, A. R. M. (2017). Process oriented guided inquiry learning (POGIL®) marginally effects student achievement measures but substantially increases the odds of passing a course. *PLOS ONE*, *12*(10), Article e0186203. https://doi.org/10.1371/journal.pone.0186203

Wallace, C. S., & Bailey, J. M. (2010). Do concept inventories really measure anything? *Astronomy Education Review*, *9*(1), Article 010116. http://doi.org/10.3847/AER2010024

Walton, D. (2001). Abductive, presumptive and plausible arguments. *Informal Logic*, *21*(2), 141–169. https://doi.org/10.22329/il.v21i2.2241

Wang, M. T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review*, *33*(4), 304–340. https://doi.org/10.1016/j.dr.2013.08.001

Ward, E. G., Dalbotten, D., Watts, N. B., & Berthelote, A. (2018). Using place-based, community-inspired research to broaden participation in the geosciences. *GSA Today*, *28*(8), 1943–2690. https://doi.org/10.1130/GSATG366GW.1

Warfa, A. R. M., Roehrig, G. H., Schneider, J. L., & Nyachwaya, J. (2014). Collaborative discourse and the modeling of solution chemistry with magnetic 3D physical models—impact and characterization. *Chemistry Education Research and Practice*, *15*(4), 835–848. https://doi.org/10.1039/C4RP00119B

Weisberg, S. M., Schinazi, V. R., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2014). Variations in cognitive maps: Understanding individual differences in navigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 669–702. https://doi.org/10.1037/a0035261

Wells, J., Henderson, R., Stewart, J., Stewart, G., Yang, J., & Traxler, A. (2019). Exploring the structure of misconceptions in the Force Concept Inventory with modified module analysis. *Physical Review Physics Education Research*, *15*(2), Article 020122. https://doi.org/10.1103/PhysRevPhysEducRes.15.020122

White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, *10*, 1–100. https://doi.org/10.1207/s1532690xci1001_1

White, R., & Gunstone, R. (1992). Prediction-observation-explanation. In R. White & R. Gunstone (Eds.), *Probing understanding* (pp. 44–64). The Falmer Press.

Wieman, C. E., Adams, W. K., Loeblein, P., & Perkins, K. K. (2010). Teaching physics using PhET simulations. *The Physics Teacher*, *48*(4), 225–227. https://doi.org/10.1119/1.3361987

Williams, L. C., Underwood, S. M., Klymkowsky, M. W., & Cooper, M. M. (2015). Are noncovalent interactions an Achilles heel in chemistry education? A comparison of instructional approaches. *Journal of Chemical Education*, *92*(12), 1979–1987. https://doi.org/10.1021/acs.jchemed.5b00619

Wilson, A. E., Pollock, J. L., Billick, I., Domingo, C., Fernandez-Figueroa, E. G., Nagy, E. S., Steury, T. D., & Summers, A. (2018). Assessing science training programs: Structured undergraduate research programs make a difference. *BioScience*, *68*(7), 529–534. https://doi.org/10.1093/biosci/biy052

Wilson, S. B., & Varma-Nelson, P. (2016). Small groups, significant impact: A review of peer-led team learning research with implications for STEM education researchers and faculty. *Journal of Chemical Education*, *93*(10), 1686–1702. https://doi.org/10.1021/acs.jchemed.5b00862

Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research*, *72*(2), 131–175. https://doi.org/10.3102/00346543072002131

Wooten, M. M., Coble, K., Puckett, A. W., & Rector, T. (2018). Investigating introductory astronomy students' perceived impacts from participation in course-based undergraduate research experiences. *Physical Review Physics Education Research*, *14*(1), Article 010151. https://doi.org/10.1103/PhysRevPhysEducRes.14.010151

Yang, J., Wells, J., Henderson, R., Christman, E., Stewart, G., & Stewart, J. (2020). Extending modified module analysis to include correct responses: Analysis of the Force Concept Inventory. *Physical Review Physics Education Research*, *16*(1), Article 010124. https://doi.org/10.1103/PhysRevPhysEducRes.16.010124

York, S., Lavi, R., Dori, Y. J., & Orgill, M. (2019). Applications of systems thinking in STEM education. *Journal of Chemical Education*, *96*(12), 2742–2751.