Received: 17 May 2018 | Revised: 5 April 2019 | Accepted: 27 April 2019

DOI: 10.1002/sce.21525

LEARNING

Science Education WILEY

Toward coherence in curriculum, instruction, and assessment: A review of learning progression literature

Hui Jin[1] | Jamie N. Mikeska[1] | Hayat Hokayem[2] | Elia Mavronikolas[1]

[1]Student and Teacher Research Center, Educational Testing Service, Princeton, New Jersey

[2]Andrews Institute of Mathematics and Science Education, Texas Christian University, Fort Worth, Texas

Correspondence
Hui Jin, Student and Teacher Research Center, Educational Testing Service, Princeton, NJ 08541.
Email: hjin@ets.org

Abstract

To evaluate the recent advances in learning progression (LP) research and identify future research directions, we reviewed LP literature from 2006 to 2018. Through a systematic search of Web of Science databases and key journals, we located 130 LP articles published between 2006 and 2018. Among these articles, we reviewed 86 studies. The review was framed around three types of coherence that LPs can provide in a system of curriculum, instruction, and assessment: developmental, vertical, and horizontal coherence. Developmental coherence refers to the idea that LPs, as cognitive models, should describe development in students from intuitive thinking to scientific thinking. It provides a foundation for building the horizontal coherence (the alignment among curriculum, instruction, and assessment) and the vertical coherence (the linkage between classroom and large-scale assessments). The results of our review suggest significant advances in enhancing the developmental coherence. More specifically, existing LPs have captured the mechanisms of knowledge development and integrated knowledge and practice in different ways. Regarding the horizontal coherence, while the methodology for the development and validation of LPs has been established, limited attention has been given to LP-based interventions and teachers' understanding and use of LPs. Only one study explored LP's role in building vertical coherence. The review reveals a great need for future research (a) to develop LPs for scientific reasoning that cut across multiple science topics and disciplines, (b) to use LPs in instructional interventions, teacher education, and professional development, and (c) to use LPs to link classroom assessments with large-scale assessments.

KEYWORDS

assessment, curriculum, instruction, learning progression

1 | INTRODUCTION

During the past decade, learning progression (LP) research has emerged as a critical area for science education. LPs, "descriptions of the successively more sophisticated ways of thinking about a topic that can follow one another as children learn about and investigate a topic over a broad span of time" (National Research Council [NRC], 2007, p. 219), can play a pivotal role in promoting teaching and learning, especially given their potential to enhance coherence in a system of curriculum, instruction, and assessment.

Over the last 12 years, significant development and research work has focused on science LPs. In 2006, Smith, Wiser, Anderson, and Krajcik published a hypothetical LP for matter. Later, in 2009, researchers in science education, developmental psychology, and educational measurement gathered at the University of Iowa for the Learning Progressions in Science Conference, where they discussed defining, assessing, modeling, and using LPs. The product of the conference was an edited book *Learning Progressions in Science: Current Challenges and Future Directions* (Alonzo & Gotwals, 2012). Also in 2009, a special issue of the *Journal of Research in Science Teaching* was devoted to science LPs; it consisted of articles on general LP approaches (Duncan & Hmelo-Silver, 2009; Lehrer & Schauble, 2009), measurement issues in LP research (Steedle & Shavelson, 2009; Wilson, 2009), hypothetical LPs (Duncan, Rogat, & Yarden, 2009; Schwarz et al., 2009), and empirically developed and validated LPs (Mohan, Chen, & Anderson, 2009; Songer, Kelcey, & Gotwals, 2009). Around the same time, assessment frameworks such as the Berkeley Evaluation and Assessment Research (BEAR) Assessment System (see examples in Wilson, 2009) and Bayesian networks (see examples in West et al., 2012) were applied to provide methodological guidance for the development and validation of LPs.

In 2011, Duschl et al. published a seminal review of this early work (Duschl, Maeng, & Sezen, 2011), articulating how different views of learning (i.e., the "fix-it" view vs. the "work with it" view) were used to capture student development in LPs. Since Duschl et al. (2011), many new LPs have been developed; many empirical studies have been conducted to apply LP to classroom teaching and teacher learning. This trend has created a need for reviewing and synthesizing the research evidence and associated literature. In particular, an examination of different mechanisms of development captured in these LPs would deepen our understanding of student learning; a summary and evaluation of various LP-based approaches to curriculum, instruction, and assessment would provide significant implications for teaching and learning in science classrooms. To address this need, we reviewed LP articles published between 2006 and 2018.

Although others consider the LP to include the learning pathway and associated assessment (e.g., Mohan et al., 2009) or the combination of the learning pathway, assessment, and curriculum/instruction (e.g., Berland & McNeill, 2010), for the purposes of this review, we use LP to refer to the learning pathway alone. We do so in an effort to distinguish the cognitive model (i.e., LP) from the components associated with the model (e.g., curriculum, instruction, and assessment). Additionally, we use the plural, LP approaches, to denote the variety of perspectives and to acknowledge there is no single LP approach. There are, however, terms common across LP studies. An LP usually contains an *upper anchor* specifying the ultimate learning goals and expectations for students when they leave the progression, a *lower anchor* describing the most naïve ideas and thinking of students when they enter the progression, and *intermediate levels of achievement* linking these two anchors (see Duncan & Hmelo-Silver, 2009). To obtain a fine-grained description of student development, many researchers develop LPs that have multiple *progress variables*. Progress variables are the abilities, skills, and competencies that are important for students' success and are targeted by the curriculum (Wilson, 2009).

2 | FORMULATION OF RESEARCH QUESTIONS

As noted by Lehrer and Schauble (2015), LP approaches were initially presented in two NRC reports, *Taking Science to School* (NRC, 2007) and *Systems for State Science Assessment* (NRC, 2005), with an emphasis on a developmental perspective for building coherence across curriculum, instruction, and assessment. More specifically, for an LP-based educational system to be effective in promoting learning, it is important to consider three coherences—developmental coherence of productive learning across time; horizontal coherence across curriculum, instruction, and assessment; and vertical coherence between classroom assessments and large-scale assessments.

Coherence has been the central tenet of science education reforms, including the post-Sputnik curriculum reform (DeBoer, 1991), the science literacy movement (American Association for the Advancement of Science, 1993), and the development of the Next Generation Science Standards (NGSS; NGSS Lead States, 2013; NRC, 2012). Although a prominent feature of science education reform, coherence has shifted its meaning over time (Sikorski & Hammer, 2017) and has been used in many ways within the science education literature (Schmidt, Wang, & McKnight, 2005). We suggest viewing coherence from two aspects: the nature of learning and the approaches used to promote learning. Developmental coherence is about the nature of learning. It entails an LP addressing the logical structure of science disciplines, students' prior knowledge and experiences, and the integration of knowledge and practices. In contrast, horizontal coherence and vertical coherence are about the approaches used to promote productive learning for all students. These approaches may target the horizontal alignment of curriculum, instruction, and assessment or the vertical linkage between classroom and large-scale assessments. It is important to note that developmental coherence is the foundation for horizontal and vertical coherences (Fulmer, Tanas, & Weiss, 2018). Based on the above discussion, our research questions are:

1. Developmental coherence
   - **a.** What theoretical perspectives do researchers use to interpret and capture student development?
2. Horizontal coherence
   - **a.** How do researchers use literature and assessment data to develop and validate LPs?
   - **b.** How do researchers use LPs to guide curriculum, instruction, and assessment and to support teacher learning?
3. Vertical coherence
   - **a.** How do researchers use LPs to align classroom assessments with large-scale assessments?

3 | CONCEPTUAL FRAMEWORK

The review was framed around three types of coherence that LPs can provide in a system of curriculum, instruction, and assessment: developmental, horizontal, and vertical coherence. We drew upon theories and findings from developmental psychology, science education, and educational assessment to unpack the three coherences. Duschl et al. (2011) propose two fundamental questions for evaluating LPs: "How well developed is the identification of the foundational knowledge that facilitates and advances pathways of reasoning and understanding? How thorough is the description of the teacher mediated learning pathways?" (Duschl et al., 2011, p. 173). We used these two questions to guide the unpacking of developmental coherence and horizontal coherence.

3.1 | Developmental coherence

The unpacking of the developmental coherence is guided by Duschl et al.'s first question for evaluating LPs: How well developed is the identification of the foundational knowledge that facilitates and advances pathways of reasoning and understanding? To answer this question, we refer to literature in curricular coherence and developmental psychology.

Scientists rely on principle-oriented and logically structured knowledge to solve problems and explain phenomena across a wide variety of contexts. As such, the logical structure of scientific knowledge, which can be illustrated in science concept maps, is crucial for student learning. Scientists also conduct scientific inquiry. During post-sputnik era, science process skills (e.g., making observations, controlling variables, and making inferences) were identified as critical skills for students to learn how to do science like scientists (Sanderson & Kratochvil, 1971). In the science curriculum reform in 1960s, significant efforts were taken to develop curriculum materials that describe the logical structure of scientific knowledge and include the process skills (DeBoer, 1991). However, presenting scientific knowledge as logically connected concepts, principles, and ideas does not guarantee that students will understand that knowledge; teaching process skills separately from content knowledge does not guarantee that students will develop the ability to conduct scientific inquiry. Two important lessons on curricular coherence were learned from the prior reforms. The first lesson is that curricular coherence should not be treated strictly as a matter of disciplinary structure; instead, it must take into account students' prior knowledge and experiences (Sikorski & Hammer, 2017). The second lesson is that a coherent curriculum should include scientific inquiry; more important, teaching inquiry skills separately from content knowledge does not lead to students' mastery of scientific inquiry (Gabel, 2006). Together, these two lessons provide important guidelines for answering Duschl et al.'s first question. To address foundational knowledge that facilitates and advances pathways of reasoning and understanding, LPs must consider students' prior knowledge and integrate knowledge and practices. In the paragraphs that follow, we discuss soultions to these two lessons.

During the 1980s and early 1990s, a perspective shift from domain-general to domain-specific theories occurred in developmental psychology. This perspective shift addresses the first lesson learned from prior science education reforms. Evidence from multiple sources, such as universal grammar (Chomsky, 1965), constraints in making induction (Keil, 1981), and expert knowledge (Chi, Feltovich, & Glaser, 1981) suggested "the mind is less an all-purpose problem solver than a collection of enduring and independent subsystems designed to perform circumscribed tasks" (Hirschfeld & Gelman, 1994, p. 4). Naïve theories are domain-specific theories that explain how children develop different domain-specific mechanisms to understand biological (Carey, 1986; Keil, 1992), physical (Smith, Carey, & Wiser, 1985; Vosniadou, 1994), and psychological (Premack & Premack, 1995) phenomena. These mechanisms are theory-like, containing coherent and abstract ideas to explain observations and phenomena; they are in marked contrast to the scientifically accepted mechanisms for explaining the same phenomena (Gopnik & Wellman, 1994). As such, learning can be viewed as a process of transitioning from intuitive thinking (e.g., naïve theories, alternative conceptions) toward scientific thinking. This view provides a rationale for attending to students' prior knowledge and experience.

To explore how students' prior knowledge can be addressed in LPs, we draw upon literature in developmental psychology. Learning of scientific knowledge occurs under different conditions of prior knowledge (Carey, 1991; Chi, 2008). When prior knowledge is missing, learning consists of adding new knowledge. When prior knowledge is correct but incomplete, learning can be conceived as filling the gap. Learning under these two conditions is a process of *knowledge enrichment*, which does not involve restructuring existing knowledge. However, when prior knowledge conflicts with knowledge that has not yet been learned, conceptual change must happen to situate new scientific knowledge into what already exists in the learner's mind. Developmental psychologists maintain that students' pre-existing ideas about a topic or within a domain are embedded in a coherent theory-like structure, and, consequently, instruction targets a shift from naïve theories toward scientific theories (Carey, 1985; Vosniadou, Vamvakoussi, & Skopeliti, 2008). As such, conceptual change is a process of *theory transformation*. A third perspective views development as a process of *knowledge integration*. That is, learners often hold conflicting ideas about scientific phenomena (Linn & Hsi, 2000; Minstrell & Stimpson, 1996) and instruction should involve helping students examine, change, link, and organize to promote students' conceptual understanding of scientific concepts and principles (Linn, Clark, & Slotta, 2003). These three theoretical perspectives provide a base for researchers to identify foundational knowledge that facilitate and advance paths of reasoning and understanding.

As discussed above, the second lesson learned from the prior reforms suggests that the integration of knowledge and practices is a key feature of inquiry in both science disciplines and school science; it should be addressed in curriculum, instruction, and assessment. Therefore, it is important for LPs to include the integration of knowledge and practices. More important, students' foundational understanding regarding the integration of knowledge and practices should be identified; such foundational understanding facilitates and advances students' pathways toward scientific inquiry.

The integration of knowledge and practices is widely discussed in policy documents (e.g., NRC, 1996, 2000, 2012, 2014) and research articles (e.g., Osborne, 2010). In *Inquiry and the National Science Education Standards: A guide for Teaching and Learning* (NRC, 2000), a story of a practicing scientist's investigation and a story of an inquiry-based science lesson are provided to characterize scientific inquiry in science and classroom settings. When conducting scientific investigations, the scientist and the students connect multiple scientific skills in a path "from formulating scientific questions, to proposing, evaluating, and then communicating explanations" (NRC, 2000, pp. 27–28). More important, knowledge and inquiry are integrated because inquiry (practices) is used as a means to construct and apply knowledge. In NRC Framework (NRC, 2012) and NGSS (NGSS Lead States, 2013), scientific inquiry is specified into eight scientific practices. The relationship between knowledge and practices is conceptualized in the NRC's vision for science education—"students, over multiple years of school, actively engage in scientific and engineering practices and apply crosscutting concepts to deepen their understanding of the core ideas in these fields" (NRC, 2012, pp. 8–9). Therefore, one way to address the integration of practices and knowledge is to describe the process–product relationship between them. Researchers may also create other approaches to the integration of knowledge and practices.

In summary, developmental coherence emphasizes that LPs should describe the progression from foundational understanding and thinking toward scientific understanding and thinking. First, LPs must chart the development of students' thinking from the intuitive to scientific. Three theoretical perspectives (knowledge enrichment, knowledge integration, and theory transformation) can be used to investigate the mechanisms of this knowledge development. Second, the development of scientific practices should no longer be separated from the development of scientific knowledge. As such, it is important to examine the approaches that researchers used to achieve this integration. We use the above ideas to examine how current science LPs are used to describe knowledge development and practice development in ways that enhance developmental coherence.

3.2 | Horizontal coherence

Duschl et al. (2011) state that evaluating LPs should consider the following question: How thorough is the description of the teacher mediated learning pathways? We use this question to guide our unpacking of the horizontal coherence in a system of curriculum, instruction, and assessment. To ensure a thorough description of learning pathways, curriculum, instruction, and assessment ought to align to assist students' progress along pathways toward the upper anchor of an LP. To ensure teacher-mediated learning pathways, teachers should be at the center of the system of curriculum, instruction, and assessment. Therefore, it is meaningful to investigate how researchers (a) use literature and assessment data to develop and validate LPs, (b) use LP-based curriculum and instruction to promote student learning, and (c) use LP approaches to promote teaching and teacher learning. In the paragraphs that follow, we identify criteria for evaluating these research activities and critical questions to be answered in conducting these activities.

An LP study at its simplest level may only develop an LP. An LP must be validated to ensure that it captures the key characteristics of developmental coherence before it is widely used for curriculum, instruction, and teacher professional development. Anderson (2008) indicated that an LP's student development stories must be compatible with current research and be validated by empirical data from real students. Therefore, for studies that use literature and/or empirical assessment data to develop and validate LPs, we explored the following question: How do researchers achieve compatibility with current research? How do researchers achieve empirical validation?

In complex LP studies, curriculum, and instruction are used to transition students toward the upper anchor of an LP, with the LP's assessments applied to measuring student learning outcomes. Two questions are important for such studies: To what extent do the curriculum and instruction address the developmental trend described in the LP? To what extent do the LP-based curriculum and instruction produce desired learning outcomes?

With teachers enacting rather than delivering curriculum, the effectiveness of an LP-based system of curriculum, instruction, and assessment relies on teachers' understanding and use of the LP. Studies on teachers' use of LPs in daily instruction usually assess teachers' understanding of LPs, investigate teachers' use of LPs in designing and teaching lessons, and provide professional development opportunities for teachers to learn about LPs. Therefore, when analyzing publications in this group, we investigate the following question: How do researchers use LPs to assess teacher knowledge? How do researchers use LPs to help teachers improve their teaching practice?

In summary, horizontal coherence is built upon developmental coherence. To enhance horizontal coherence, researchers use assessments to develop and validate LPs, use LP-based curriculum and instruction to promote science learning, and investigate the roles of LPs in promoting teaching and teacher learning. As indicated in the critical questions, this part of our review targets the research strategies, gaps, and achievements in these research activities.

3.3 | Vertical coherence

An assessment system includes well-aligned classroom assessments and large-scale assessments (Bennett, 2010; NRC, 2014). In *Developing Assessments for the Next Generation Science Standards*, NRC (2014) suggested that assessments at different scales should work together to assess student learning of complex constructs with breadth and depth and over time. As such, there are three challenges with vertical coherence: assessing complex constructs, assessing learning over time, and assessing both breadth and depth of knowledge and practices. We elaborate on each challenge in the context of LP research. First, students learn knowledge and practices in an integrated manner. Therefore, a crucial question is: How do LP-based assessment tasks allow students to demonstrate their understanding of both knowledge and practices? Second, while classroom assessments provide information about intermediate understanding, large-scale assessments evaluate the extent to which students reach the expected sophisticated understanding. This begs the following question: What should be considered in determining intermediate and final learning goals for LP-based assessments? Finally, large-scale assessments evaluate learning over long periods such as several years. How do such large-scale assessments cover a wide range of knowledge and practices in depth? In this review, we focus on how researchers deal with these three challenges in building vertical coherence.

4 | METHOD

The Evidence for Policy and Practice Initiative created a systematic review approach (Gough, Oliver, & Thomas, 2012). In this approach, systematic searching is carried out to identify relevant articles. These articles are categorized to map the research, allowing researchers to determine which studies to include in the follow-up synthesis. Finally, the results of the included studies are synthesized to generate important findings. We used this systematic review approach in this review.

4.1 | Searching

In dealing with the tension between comprehensive and selective searching, we considered the following point from the Web of Science:

> *It would appear that to be comprehensive, an index of the scholarly journal literature might be expected to cover all journals published. It has been demonstrated, however, that a relatively small number of journals publish the majority of significant scholarly results. This principle is often referred to as Bradford's Law. (Testa, 2018, "Why Be Selective?" section).*

Following Bradford's Law, we used Web of Science and expert recommended journals as the databases for the search. This searching method, although not exhaustive, allowed us to include the majority of significant results in science LP research during a 12-year period.

First, using Web of Science, we searched for the topic of learning progression(s) in the category of "Education & Educational Research" for the period between 2006 and 2018. Second, we generated a list of major science education journals and several other journals that are well known for publishing high-quality LP research. We obtained feedback from three experts in LP research, one of whom is an initiator of science LP research, about which journals to include in our search. Based on this expert feedback, we added more journals to our list. As a result, the final list contained 14 journals: *Journal of Research in Science Teaching*, *Science Education*, *International Journal of Science Education*, *Studies in Science Education*, *Research in Science Education*, *Science & Education*, *Educational Researcher*, *Measurement: Interdisciplinary Research and Perspectives*, *Applied Measurement in Education*, *Review of Educational Research*, *Journal of the Learning Sciences*, *Science*, *American Educational Research Journal*, and *Cognition & Instruction*. (Web of Science did not have issues for *Measurement: Interdisciplinary Research and Perspectives*, a relatively new journal, for the entire search period.) We used the key word "learning progression(s)" to search the journals in the list. The two searches resulted in 518 unique articles.

Next, we screened the 518 articles, using both inclusion and exclusion criteria to set the boundaries for the review (Gough et al., 2012). In line with our research questions, the inclusion criteria included

- Articles addressing LPs in science education.
- Articles written in English.
- Articles reporting on studies conducted from any geographical location.
- Theoretical pieces or empirical studies on science LPs.
- Empirical studies, for which the participants were pre-K children, K–12 students, undergraduate students, or preservice/inservice teachers who teach K–12 students.

The following exclusion criteria were used in this review:

- Articles about using the findings from an empirical study for future LP research, but the study per se was not about science LPs.
- Articles citing LP literature, but the focus was not on science LPs.
- Articles reporting on LPs in subject areas other than science (e.g., English language arts, mathematics).
- Articles addressing LP approaches in a generic manner rather than focusing on science LPs.

The screening resulted in 130 articles on science LPs.

4.2 | Mapping the terrain

With 130 articles, we mapped the terrain of science LP research. We categorized the articles in terms of different areas of focus: (a) development of theoretical LPs, (b) empirical research of student learning, (c) research of teaching and teacher learning, (d) review articles, and (e) position articles.

Researchers often differentiate two types of LPs: hypothetical LPs and validated LPs (Corcoran, Mosher, & Rogat, 2009; Wilson, 2009). A hypothetical LP is grounded in literature and sometimes in empirical data, but it is not validated. Researchers may label an LP developed based on empirical data as a hypothetical LP when validity evidence is not sufficient to justify the meaning of, the distinctions among, and the order of the LP levels (e.g., Plummer & Maynard, 2014). To facilitate the synthesis of the LP articles, we used a different categorization that distinguished between theoretical and empirical LPs. LPs developed based on ideas from the literature were theoretical LPs, although in some of these LPs, students' responses were used as examples to illustrate the LP levels. LPs developed from student assessment data were empirical LPs, although the LP levels may not be fully validated. Accordingly, we categorized articles as either development of theoretical LPs or empirical research of student learning. Studies using literature to develop hypothetical LPs were categorized as development of theoretical LPs. Studies using assessment data to develop LPs, even though the LPs may not be fully validated, were categorized as empirical research of student learning. This latter category also included studies that implemented LP-based curriculum and instruction to promote student learning.

In addition to the above categories, we sought studies that examined teachers' understanding and use of LPs and studies that developed LPs for teacher knowledge. We categorized these articles as research of teaching and teacher learning. We identified one LP article as a review article and many others as position articles about various issues in LP research. After these five categories were established, two researchers used these categories to code each LP article independently. The reliability check resulted in a Cohen's κ of 0.91, indicating almost perfect agreement (Landis & Koch, 1977). We discussed the articles we coded differently and reached agreement on the final categorization codes.

4.3 | Research synthesis

At the stage of research synthesis, we used the 86 articles in the first three categories (development of theoretical LPs, empirical student learning research, and research of teaching and teacher learning) for in-depth analysis. These articles either provide empirical findings or proposed concrete ideas about developing, validating, or using an LP for a specific topic in science. In line with the conceptual framework, the in-depth analyses targeted the three coherences. We found only one article about vertical coherence. Therefore, we discuss the analyses for the developmental and horizontal coherences, but not vertical coherence.

The analyses of developmental coherence addressed the knowledge development and the practice development described in the LPs for student learning. Except for three articles that used LPs to describe teachers' development, all other 83 articles used LPs to describe students' development. These articles describe 46 LPs because sometimes two or more articles described the same LP. We analyzed how these LPs describe knowledge development and practice development. Regarding knowledge development, two researchers coded the LPs in terms of the mechanisms of development—knowledge enrichment, theory transformation, or knowledge integration. The reliability check for this coding resulted in Cohen's κ of 0.80, indicating substantial agreement (Landis & Koch, 1977). Regarding practice development, we identified and coded the strategies used to integrate practice and knowledge. The reliability check for this coding resulted in Cohen's κ of 0.77, indicating substantial agreement (Landis & Koch, 1977). We discussed the articles that received different codes and reached agreement on the final codes.

In the conceptual framework section, we identify research activities that use LPs to increase horizontal coherence (using literature and data to develop and validate LPs, using LP-based curriculum and instruction to promote student learning, and use LP approaches to promote teaching and teacher learning); for each research activity, we identified critical questions. In analyzing articles on horizontal coherence, the first three authors each read the articles and produced a written analysis to answer the critical questions. First, we found out whether the articles contained a student LP. For articles that include a student LP, we examined how researchers achieved compatibility with current research and empirical validity. Second, we studied articles categorized as empirical research of student learning as to whether and how researchers used curriculum and instruction to help students move toward the upper anchor of the LP and to what extent the LP-based interventions produced desired learning outcomes. Finally, for articles in the category of research of teaching and teachers, we considered how researchers investigated teachers' understandings and use of LPs. Based on the analyses from the three authors, we generated findings of horizontal coherence.

5 | RESULTS

In this section, we first provide an overview of LP research based on the results from mapping the terrain. Then, we discuss the findings of the three coherences derived from the research synthesis.

5.1 | Overview of LP research

We sorted 130 science LP articles into five categories. Forty-three position articles covered a wide range of topics, including methodological issues, practicality concerns, and theoretical foundations. One review article (Duschl et al., 2011) analyzed early LP research. Fourteen articles described theoretical LPs. In these articles, the LP levels were supported by ideas from the literature on student learning (e.g., Smith, Wiser, Anderson, & Krajcik, 2006) and illustrated by exemplar responses from students (e.g., Duncan et al., 2009; Lehrer & Schauble, 2012).[1] Sixteen articles report on research of teaching and teacher learning. Among them, three articles report on the development of LPs for teacher learning (Adadan & Oner, 2014; Lyon, 2013; Schneider & Plasman, 2011), whereas the rest of the articles report on teachers' understanding and use of student LPs.

Fifty-six articles reported on empirical studies about student learning. Table 1 provides descriptive information of these articles. As shown in Table 1, most articles in this category were sorted within or across three school levels; only one LP covered all four school levels (elementary school, middle school, high school, and college). Most studies included K–12 students, one study included preschool children, and eight studies included college students. There was a spread from smaller scale or moderate studies with 500 or fewer participants, to large-scale studies that collected data from more than 1,000 students. Studies that collected large-scale data all used quantitative methods such as item response theory (IRT) to analyze written assessment data.

5.2 | Developmental coherence

Researchers developed 46 LPs, among which, 34 LPs focus on knowledge development and 15 LPs focus on practice development. Note that three LPs describe both knowledge development and practice development. A summary of these LPs is provided as the Supporting Information material (Table S1).

5.2.1 | Knowledge development

Results showed that across the knowledge LPs, each one drew upon one of three theoretical perspectives—knowledge enrichment, knowledge integration, or theory transformation. Knowledge LPs using these theoretical perspectives are presented in Tables 2–4, respectively. Overall, the majority of these LPs used the perspectives of

[1]Catley, Lehrer, and Reiser (2005) reported on earlier thoughts about the same LP in 2005.

TABLE 1 Empirical studies that develop or validate learning progressions

Types of descriptive information	Descriptive information and number of studies				
Range of school levels	One school level 32	Two school levels 15	Three school levels 8	Four school levels 1	
Participants included in the study	Preschool children 1	Elementary students 20	Middle school students 32	High school students 29	College students 8
Sample sizes	Fewer than 100 participants 19	100–499 participants 16	500–999 participants 4	1,000–4,999 participants 15	5,000 and more than 5,000 2

Note: The total number of articles in the category of empirical student learning studies is 56. However, there are overlaps among the columns in the row "Participants included in the study" because some studies included students from two or more school levels.

knowledge enrichment and theory transformation. The knowledge integration perspective was used in only three LPs. Below we elaborate upon the key features within each perspective.

Knowledge enrichment refers to constructing advanced knowledge based on less sophisticated ideas. Note that students' alternative ideas or misconceptions are not explicitly addressed in this type of development. Instead, the focus is on adding new knowledge to students' existing knowing without major restructuring occurring. To capture such development, researchers focus on the logical structure of knowledge and the difficulty levels of different concepts, principles, and ideas. The difficulty levels are usually identified based on quantitative analysis of assessment data. For example, Neumann, Viering, Boone, and Fischer (2013) developed an LP for energy in physics. In the LP, four energy concepts are arranged in terms of difficulty levels—forms of energy, energy transfer, energy degradation, and energy conservation. Similarly, Johnson and Tymms (2011) developed an LP for matter, which is a map that illustrates the difficulty levels of various matter concepts, principles, and ideas in chemistry. For both LPs, the difficulty levels were determined by IRT analysis of student assessment data.

Knowledge integration LPs describe one or more of the following developmental mechanisms. First, moving up the levels, misconceptions gradually disappear, while more scientific ideas are developed. Second, moving up the levels, more and more linkages are established among relevant scientific ideas. Third, a specification of the ideas activated in different contexts is presented. Unlike knowledge enrichment, knowledge integration includes alternative and nonnormative ideas at lower levels. We identified three LPs that describe development as a process of knowledge integration. Lee and Liu (2010) developed an LP based on students' responses to items selected and modified from TIMSS and NAEP items. The LP focuses on student understanding of three energy concepts—energy sources, energy transformation, and energy conservation—across physical sciences, life sciences, and earth sciences. The levels of the LP include providing nonnormative ideas, providing one normative idea, linking two normative ideas, and producing at least two links among three or more normative ideas. As such, the LP describes the development as a process of creating normative ideas and making more and more links among those ideas. Alonzo, Steedle, and Shavelson (Alonzo & Steedle, 2009; Steedle & Shavelson, 2009) developed an LP for force and motion, which was later validated by Fulmer, Liang, and Liu (2014). The LP contains four levels of achievement. Each level has a set of ideas that may be either scientific or nonnormative. An elaboration of the ideas activated in different contexts (e.g., force and no motion; force and motion) are also provided. Moving up the levels, more and more scientific ideas appear and nonnormative ideas gradually disappear. Stevens, Delgado, and Krajcik (2010) developed an LP for the nature of matter. The LP is presented in a matrix that contains four levels of achievement cutting across two variables (atomic structure and electrical forces). Each level, visually presented as a panel cutting across the two variables, contains a set of ideas. Moving up the levels, more and more scientific ideas are generated

TABLE 2 Learning progressions using knowledge enrichment perspective

Topics	Learning progression
Energy and matter	**1.** Energy (Neumann et al., 2013; Opitz, Harms, Neumann, Kowalzik, & Frank, 2014; Yao, Guo, & Neumann, 2017) **2.** Energy (Herrmann-Abell & DeBoer, 2018) **3.** Concept of a substance (Johnson & Tymms, 2011) **4.** Structure of matter (Morell et al., 2017)
Water	**5.** Explaining hydrologic phenomena (Forbes et al., 2015)[a]
Ecosystems	**6.** Understanding simple food chains (Allen, 2003) **7.** Developing nonlinear reasoning in ecology (Hovardas, 2016) **8.** Complex reasoning about biodiversity (Gotwals & Songer, 2010, 2013; Songer et al., 2009)[a]
Genetics	**9.** Genetics (Elmesky, 2013) **10.** Genetics (Duncan, Castro-Faix, & Choi, 2016; Duncan, Choi, Castro-Faix, & Cavera, 2017; Duncan et al., 2009; Shea & Duncan, 2013; Todd & Kenyon, 2016; Todd, Romine, & Whitt, 2017) **11.** A modern genetics progress web (Todd, & Romine, 2016, Todd and Romine, 2017)
Sea level rise	**12.** Sea level rise (Breslyn et al., 2016)
Acid-base	**13.** Acid-base chemistry (Romine, Todd, & Clark, 2016)
Thermochemistry	**14.** Thermochemistry (Chen, Zhang, Guo, & Xin, 2017)
Stellar structure and evolution	**15.** Stellar structure and evolution (Colantonio, Galano, Leccia, Puddu, & Testa, 2018)
Buoyancy	**16.** Concept of buoyancy (Paik, Song, Kim, & Ha, 2017).

[a]Learning progressions that use separate progress variables to describe knowledge development and epistemological development.

and more and more linkages are established. The LP, therefore, describes students' development as a process of connecting and integrating ideas into a complex model of matter.

Theory transformation describes development as a radical change, in which intuitive theories restructure into scientific theories. Unlike knowledge enrichment and knowledge integration, a wide variety of ideas has been proposed to explain the mechanisms of theory transformation. Rooted in fundamental cognitive and learning theories, theory transformation LPs describe three developmental mechanisms. First, students use intuitive reasoning patterns, such as vitalistic reasoning (Inagaki & Hatano, 2004), teleological reasoning (Lombrozo & Carey, 2006), and impetus theory (McCloskey, Washburn, & Felch, 1983), to understand phenomena in the material world. Such reasoning patterns are very different from and sometimes incommensurable with the specialized ways of reasoning in science. Therefore, theory transformation may take the form of reasoning change, where students develop from intuitive reasoning to counter-intuitive scientific reasoning. As an example, the LP for water in socioecological systems (Gunckel, Covitt, Salinas, & Anderson, 2012) describes the development from force-dynamic reasoning toward a specialized way of scientific reasoning that focuses on driving forces and constraining factors. The LP for natural selection[2] contrasts the transformationist reasoning at lower levels with the scientific

[2]Mayr (1982) decomposed natural selection into five facts and three inferences. Based on this study, Furtak et al. developed a set of LPs to describe student progression in understanding the facts and inferences. To make Furtak et al.'s LPs compatible with the LPs developed by other researchers, we refer to their LPs as progress variables, which are combined into one LP for natural selection. decomposed natural selection into five facts and three inferences. Based on this study, Furtak et al. developed a set of LPs to describe student progression in understanding the facts and inferences. To make Furtak et al.'s LPs compatible with the LPs developed by other researchers, we refer to their LPs as progress variables, which are combined into one LP for natural selection.

TABLE 3 Learning progressions using theory transformation perspective

Topics	Learning progression
Energy and matter	**1.** Energy in socio-ecological systems (Jin & Anderson, 2012; Jin, Johnson, Shin, & Anderson, 2017; Jin, Shin, Johnson, Kim, & Anderson, 2015; Jin & Wei, 2014; Jin, Zhan, & Anderson, 2013) **2.** Matter and atomic-molecular theory (Smith et al., 2006)[a] **3.** Matter concept (Hadenfeldt et al., 2016; Hadenfeldt et al., 2014) **4.** Matter in socio-ecological systems (Jin, Johnson, Shin, & Anderson, 2017; Jin, Shin, Johnson, Kim, & Anderson, 2015; Jin, Zhang, & Anderson, 2013; Mohan et al., 2009) **5.** Structure of matter (Talanquer, 2009)
Water	**6.** Water in socioecological systems (Covitt et al., 2018; Gunckel, Covitt, & Salinas, 2018; Gunckel, Covitt, Salinas, & Anderson, 2012)
Ecosystems	**7.** Complex ecosystems (Hokayem & Gotwals, 2016)
Evolution	**8.** Natural selection (Furtak, 2012; Furtak & Heredia, 2014; Furtak et al., 2014, Furtak et al., 2018) **9.** Three-dimensional learning of the patterns of evolution (Wyner, & Doherty, 2017)
Human nutrition	**10.** Human nutrition (Cabello-Garrido, Espana-Ramos, & Blanco-Lopez, 2018)
Celestial motion	**11.** Celestial motion—explanation of seasons (Plummer & Krajcik, 2010; Plummer & Maynard, 2014) **12.** Celestial motion—the apparent motions of the sun, the moon, and the stars moon phases (Plummer, 2014) **13.** Change of seasons, solar and lunar eclipses, and moon phases (Testa, Galano, Leccia, & Puddu, 2015)
Solar system	**14.** Formation of solar system (Plummer et al., 2015)
Chemical thinking	**15.** Chemical thinking (Sevian & Talanquer, 2014)

[a]Learning progressions that use separate progress variables to describe knowledge development and epistemological development.

TABLE 4 Learning progressions using knowledge integration perspective

Topics	Learning progression
Energy and matter	**1.** Energy (Lee & Liu, 2010) **2.** Nature of matter (Stevens et al., 2010)
Forces and motion	**3.** Force and motion (Alonzo & Steedle, 2009; Fulmer, 2015; Fulmer et al., 2014; Steedle & Shavelson, 2009; von Aufschnaiter & Alonzo, 2018)

reasoning used in natural selection at the upper anchor (Furtak, 2012; Furtak, Morrison, & Kroog, 2014). Second, research of concept development in the history of science suggests two important mechanisms of theory transformation—conceptual differentiation and conceptual coalescence (Smith et al., 1985; Wiser & Carey, 1983). We found that one LP describes development using the idea of conceptual differentiation. The LP for matter has a variable that describes how two differentiated new concepts, weight, and density, are generated from an initial concept of *felt weight* that conflates weight and density. Third, conceptual change researchers differentiate between accommodation and assimilation (Posner, Strike, Hewson, & Gertzog, 1982). While accommodation represents the radical change in theories, assimilation is incorporating new information into an existing framework or knowledge structure. It is very common that students assimilate new ideas into their existing intuitive theories, resulting in *synthetic models* (Vosniadou, 1994). Some LPs describe a developmental process from intuitive reasoning, to synthetic models, and to scientific reasoning. For example, the LP for natural selection (Furtak &

Heredia, 2014) has a Level 3 that blends the scientific theory of natural selection and the naïve theory about need-based changes.

5.2.2 | Practice development

Overall, 15 practice LPs cover six out of the eight NGSS practices. Many LPs have been developed to describe student development in modeling and argumentation (see Table S1). Two NGSS practices that are not covered are asking questions and planning and carrying out investigations. We analyzed practice LPs from two aspects (a) whether two separate sets of variables are used to describe practice and knowledge and (b) how practice and knowledge are integrated. The results are presented in Table 5.

We found two LPs were developed without a connection to scientific content knowledge: an LP on error in science (Allchin, 2012) and an LP for communication (Hsin, Chien, Hsu, Lin, & Yore, 2016). This may be due to the nature of the practices. For example, communication and errors in science, when being analyzed at a meta-level, are not closely connected to content. We also note that both LPs are theoretical LPs. The LP for errors in science focuses on students' beliefs about errors in scientific experiments (Allchin, 2012), which is a component of analyzing and interpreting data (an NGSS practice). Developed on literature only, it describes a transition from recognizing that scientists make mistakes to identifying sources of errors in scientific experiments and remedying those errors. The LP for communication was first developed based on literature and then revised based on a survey that collected experts' evaluation and feedback of the initial LP. Therefore, both LPs are theoretical in nature; they have not been used with students. If these LPs are later used to guide student assessment or classroom learning activities, they will need to be applied to specific content areas. In such situations, integration of practice and knowledge will become necessary.

The remaining 13 LPs suggest three approaches to the integration of practice and knowledge (see Table 5). Eight LPs describe the development of practice in a context of scientific content knowledge. We label these LP as generic LPs. The achievement levels of these LPs only describe the scientific practices, but the assessment tasks and classroom activities associated with these LPs are about both practices and knowledge. For example, drawing upon the theory of cognitive load and Toulmin's argument model, Osborne et al. (2016) developed an LP for argumentation in understanding matter concepts. Moving up the levels of the LP, students include more and more argumentation elements (e.g., claim, reasoning, evidence, and rebuttal) and make more connections among the elements in their arguments, indicating an increasing cognitive load and a better understanding of the logic structure of scientific arguments. As such, the levels of the LP are not content-specific. In the study, this generic LP is used in the context of matter concepts—the assessment tasks require students to construct and evaluate arguments about phenomena related to matter. This LP can also be applied to other content contexts.

Four LPs are content-specific. These LPs describe the development of practice in ways that incorporate components of scientific content knowledge. Here, we use four contrasting cases to explain how the content-specific LPs differ from the generic LPs. The four cases include two generic and two content-specific LPs for modeling. One generic LP describes students' mastery of modeling practice across topics that are not closely connected; these topics are evaporation and condensation, light, and friction (Bamberger & Davis, 2013; Schwarz et al., 2009). This LP focuses on general aspects of modeling (e.g., models are tools for constructing knowledge), so that a comparison across content topics is possible (see Fortus, Shwartz, & Rosenfeld, 2016, for an example). The other generic LP (Forbes, Zangori, & Schwartz, 2015) has two sets of progress variables, one about understanding the water cycle and the other about modeling practice. The description of the levels for modeling practice is generic and can be easily applied to other content topics.

The two content-specific LPs are an LP for modeling in the Earth system (Rivet & Kastens, 2012) and an LP for modeling in evolution (Lehrer & Schauble, 2012). These LPs were developed based on the idea that modeling is the mapping of relationships and functions between the model and the target system (Gentner & Gentner, 1983). The LP for modeling in the Earth system describes the development in mapping the physical models of the Earth system

TABLE 5 Integration of practices and knowledge in practice learning progressions

Integration of practice and knowledge	No. of LPs	Set of variables: One set of variables focusing on practice	Set of variables: Two separate sets of variables on practice and knowledge separately
Practice LP not connected to knowledge	2	**1.** Error in science (Allchin, 2012) **2.** Communication (Hsin et al., 2016)	N/A
Generic practice LP (describes practice development in a content area)	8	**1.** Scientific modeling (Bamberger & Davis, 2013; Fortus et al., 2016; Pierson, Clark, & Sherard, 2017; Schwarz et al., 2009) **2.** Argumentation in science (Osborne et al., 2016) **3.** Scientific argumentation (Berland & McNeill, 2010) **4.** Use of evidence in decision-making contexts (Bravo-Torija & Jiménez-Aleixandre, 2018) **5.** Science explanation (Yao, & Guo, 2017) **6.** Quantitative reasoning in environmental science (Mayes et al., 2014)	**7.** Complex reasoning about biodiversity (Gotwals & Songer, 2010, 2013; Songer et al., 2009) **8.** Explaining hydrologic phenomena (Forbes et al., 2015)
Content-specific practice LP (incorporates components of content knowledge)	5	**1.** Modeling in the earth systems (Rivet & Kastens, 2012) **2.** Modeling practices in evolution (Lehrer & Schauble, 2012) **3.** Chemical reasoning in constructing arguments in thermodynamics (Moon, Courtney, Cole, & Towns, 2017) **4.** Model-based inquiry (Hernandez, Couso, & Pintó, 2015)	**5.** Matter and the atomic-molecular theory (Smith et al., 2006)

with the observation of moon phases (Rivet & Kastens, 2012). The LP for modeling in evolution describes student development in four progress variables in biological diversity: individual change, variability, population change, and ecology (Lehrer & Schauble, 2012). For each progress variable, knowledge components are incorporated in the practice levels. For example, the variable of population change contains four achievement levels: identifying attributes of a population based on measurement, using tables to produce qualitative descriptions of changes in attribute over time, describing change over time (linear change and distribute on) for a single variable, and using functions to characterize population growth. In these levels, the modeling practice (measurement, tables and graphs, and functions are used to map the model and real-world phenomena) and the knowledge components (population growth and change over time) are intertwined.

One LP, an LP for matter (Smith et al., 2006), uses a process-product approach to integrate knowledge and practices. The LP uses two separate sets of variables to describe practice development and knowledge development. The knowledge variables (structures and properties of matter and interactions of matter) describe development as a transformation from a theory of material kinds toward a scientific theory of matter. The practice variables describe development in three scientific practices—measurement, modeling, and argumentation. The knowledge and practice variables are integrated in a particular way: The development of the conceptual understanding of matter is achieved through the development of and engagement in the scientific practices. For example, the development in measurement skills (practice) facilitates the transition from reasoning in terms of felt weight to reasoning in terms of measured weight on the variable of structure and properties of matter (knowledge).

5.2.3 | Summary of research on developmental coherence

Developmental coherence emphasizes using LPs as cognitive models to describe student development of scientific knowledge and practices. For knowledge LPs, we found that researchers used three different perspectives to describe knowledge development. All three perspectives address both science and cognition, but their emphases differ. The knowledge enrichment perspective emphasizes *the logic of science*. That is, some scientific ideas are conceptually more challenging than others; the difficulty levels of the scientific ideas can be determined by quantitative analyses of student assessment data. The knowledge integration perspective highlights *the syntax of science*. Scientific concepts, principles, and ideas are connected in specialized ways, whereas students may struggle to make the connections. Therefore, the development should address how the connections among those concepts, principles, and ideas can be built over time. The theory transformation perspective focuses on *cognition*. Students hold domain-specific naïve theories or alternative conceptions, which are very different from scientific theories. Therefore, development should describe the transformation from naïve theories to scientific theories.

For practice LPs, researchers use three different ways to integrate practices and knowledge. Generic LPs describe general developmental trends that may appear in different content topics. These LPs integrate practices and knowledge by identifying *trends across content*. Content-specific LPs contain upper anchors that combine the components of scientific practices and scientific ideas. The approach to integration is *intertwined practice and knowledge*. The matter LP developed by Smith et al. (2006) describes the specific skills in practices (measurement, modeling, and argumentation) that are needed to develop understanding of matter concepts at a certain LP level. It uses a *process-product* approach for the integration.

5.3 | Horizontal coherence

Regarding horizontal coherence, we investigated how researchers (a) use assessment and literature to develop and validate LPs, (b) use curriculum and instruction to promote student development toward the upper anchor of the LP, and (c) study teachers' knowledge and use of LPs.

5.3.1 | Developing and validating LPs

We explored the strategies used to achieve the two criteria for LPs: compatibility with current research and empirical validation. We found that researchers have used an iterative process to achieve these two criteria (e.g., Mohan et al., 2009; Songer et al., 2009). Researchers often begin with developing a theoretical LP based on relevant literature. Using the theoretical LP as a guiding framework, they develop assessment tasks (e.g., clinical interview tasks and written assessment tasks) to elicit student thinking (e.g., Hokayem & Gotwals, 2016). They analyze the assessment data and use the analysis results to redevelop, revise, and refine the LP and associated assessment tasks. This cycle of development, implementation, and revision is repeated until sufficient validity evidence is obtained to show that the LP captures student thinking and the developmental trend and is compatible with existing literature. For an illustration of this iterative process, see Breslyn, McGinnis, McDonald, and Hestness (2016).

To achieve the compatibility with the field, researchers have used theories in science disciplines, philosophy of science, and learning sciences to develop the LP. Literature in science disciplines and philosophy of science is often used for the development of the upper anchor. Literature in cognitive and learning sciences (e.g., conceptual change) has been used often for developing the lower anchor and the intermediate levels. For the development of the 46 student LPs, all researchers discussed both literature in science disciplines and literature in cognitive and learning sciences. Moreover, several researchers reviewed the conceptual change literature about a science topic in a systematic manner; they used the review results to develop a theoretical LP for that topic (Alonzo & Steedle, 2009; Hadenfeldt, Liu, & Neumann, 2014; Plummer, 2014). As an example, Hadenfeldt et al. (2014) reviewed the literature about student understanding of matter. Based on the review, they identified four categories of understanding matter and developed a theoretical LP to describe students' development in those categories. This LP was then used to guide an empirical study that used student assessment data to develop an LP for matter (Hadenfeldt, Neumann, Bernholt, Liu, & Parchmann, 2016). Such work, although time-consuming, allows in-depth interpretation of both the scientific aspect and the cognitive aspect of the construct of interest and therefore provides a solid foundation for the development of the LP.

Three issues should be considered for the achievement of empirical validation. First, an LP is a cognitive model that describes learning over a broad time span; the LP may cross several grades (e.g., from grade four to grade six) or even school levels (i.e., elementary school, middle school, high school, and college). To develop such models, a wide variety of responses, including the most naïve responses and the most advanced responses, must be collected. Researchers used the following strategies to achieve this goal. When developing an LP for K–12 students, some researchers included undergraduates in the assessment to obtain responses at the upper anchor (e.g., Stevens et al., 2010). Some researchers included different instructional treatments in one study with the expectation that students receiving instruction would be more likely to provide responses at higher levels, while students receiving partial or no instruction would be more like to provide responses at lower levels (e.g., Plummer, 2014; Todd & Kenyon, 2016).

Second, both qualitative and quantitative techniques are needed to validate the meaning of, the order of, and the distinctions among the LP levels. Results suggest that researchers examined all three aspects of LP validation. Gotwals and Songer (2013) used both think-aloud interviews and clinical interviews to revise and validate the LP levels. The think-aloud interview data provided validity evidence for the response process—the thinking process that generates the answers, while the clinical interview data provided validity evidence for the products—the responses generated from a thinking process. As such, the interview data allowed the researchers to validate the meaning of the LP levels—to what extent the LP levels capture the salient patterns of student thinking. Wright maps, which are developed using Rasch models, are often used to show that the LP levels are differentiated from each other and the order of the levels is appropriate (e.g., Neumann et al., 2013; Rivet & Kastens, 2012). Undifferentiated levels in a Wright map often suggest that the levels do not capture major ideas from students or the assessment items do not effectively elicit and differentiate among student thinking. Wright maps may also suggest that a new variable or a subdimension should be added to the LP. For example, Morell, Collier, Black, and Wilson (2017) identified the existence of subcategories within a variable of their initial matter LP.

Third, the revision and validation of an LP are sometimes conducted by a group of researchers who were not involved in the development of the original LP. For example, Todd and Kenyon (2016) used the LP of modern genetics developed by Duncan et al. (2009) to conduct an intervention study. Based on the assessment data, they revised the original LP by collapsing or teasing apart some levels of the LP. Fulmer et al. (2014) conducted two studies to validate the LP for force and motion developed by Alonzo and Steedle (2009). Although they found validity evidence for higher levels, the data suggest that the lower levels are not well differentiated. In the above studies, a research team not involved in the original studies conducted a *conceptual replication* (Schmidt, 2009). The team verified the original conceptual model (i.e., the LP) and provided constructive suggestions for revising that model. Unlike replication studies in medical and other scientific experiments (see Makel et al., 2016), these replications studies did not replicate all major research conditions (e.g., the cultural contexts and the ages of the participants). However, such replications are still valuable because researchers not involved in the original research often bring fresh perspectives that improve the LPs.

5.3.2 | Using curriculum and instruction to support student development

We analyzed the 56 articles in the category of empirical research of student learning to understand (a) the alignment between the LP and curriculum/instruction and (b) the effectiveness of the LP-based intervention. Regarding the alignment between the LP and curriculum, we found three patterns. First, among the 56 studies, 31 studies do not contain evidence of the alignment between curriculum/instruction and the LP. In these studies, either no information about curriculum and instruction is provided, or curriculum and instruction are not intentionally designed or adapted to scaffold students' progression toward the upper anchor. More often than not, teachers used their own materials and taught the topics in their own manner.

Second, in 13 articles, the curriculum was not developed based on an LP, but evidence of alignment was provided. Some researchers examined an existing curriculum and briefly described its alignment with the LP. For example, in a study conducted in Germany, the four levels of an LP for energy were the four learning goals in the existing curriculum in the same order (Neumann et al., 2013). Some researchers adapted the existing curriculum to ensure that the curriculum was aligned with the LP to a certain degree. For example, in a study on an LP for explaining hydrologic phenomena (Forbes et al., 2015), two lessons on modeling were incorporated into the teachers' regular teaching because part of the LP is about students' modeling skills and modeling was not included in the teachers' original lessons. In one study (Plummer & Krajcik, 2010), the researchers first developed a curriculum based on preliminary findings of student understanding, and then they implemented the curriculum in an intervention and used the assessment data from the intervention to develop the LP.

Third, 12 articles describe LP-based interventions, where a curriculum developed based on an LP was implemented in science classes; and pre- and post-assessments are used to measure student learning gain in the intervention. Songer and her colleagues developed an LP that has two progress variables, including a content variable describing the development in biodiversity and an inquiry variable depicting development in constructing evidence-based explanation (Gotwals & Songer, 2010, 2013; Songer et al., 2009). Based on this LP, the researchers developed a sequence of curricular activities that targeted the focal points described in the levels of the content variable. The activities contained a scaffolding tool that helped students develop evidence-based explanations as defined in the upper anchor of the inquiry variable. Todd and Kenyon (2016) developed inquiry-based units to provide targeted instruction to help students achieve the higher levels of a genetics LP. The units were taught in coordination with the participating teachers' typical classroom instruction. The 12 articles reported mixed results about LP-based interventions. Some studies found significant learning gains (e.g., Anderson et al., 2018; Plummer & Krajcik, 2010; Plummer & Maynard, 2014; Songer et al., 2009), while other studies found significant learning gains on some but not all variables of the LP (Bamberger & Davis, 2013; Forbes et al., 2015). One early study found significant learning gains on the LP but the total percentage of responses at the upper anchor was less than 10% (Jin, Zhan, & Anderson, 2013). Given that limited information on curriculum and instruction is provided in these articles, it is difficult to decide why some LP-based interventions are more successful than others are.

5.3.3 | LP approaches to teaching and teacher learning

While 56 studies investigated student learning, only 16 studies investigated teachers' learning and teaching practice. More interestingly, all of the 16 studies were conducted with secondary teachers or preservice secondary teachers, with no study addressing teachers' use of LPs at the elementary level. This suggests that, while researchers have developed many LPs, the advance in research on learning has just begun to stimulate an advance in research on teaching in the same area. Among the 16 studies, three studies developed LPs for teachers' pedagogical content knowledge (Adadan & Oner, 2014; Lyon, 2013; Schneider & Plasman, 2011), while the rest 13 studies investigated teachers' understanding and use of a student LP. We discuss the results from the analysis of the latter because they provided information on teachers' role in building horizontal coherence among curriculum, instruction, and assessment. This group includes two studies assessed teacher knowledge as it relates to a student

LP (Gunckel, Covitt, & Salinas, 2018; Jin, Shin, Johnson, Kim, & Anderson, 2015); 10 studies investigated how teachers or preservice teachers used LPs for formative assessment (Alonzo, 2018; Covitt, Gunckel, Bess, & Syswerda, 2018; Furtak, 2012; Furtak, Bakeman, & Buell, 2018; Furtak, Circi, & Heredia, 2018; Furtak & Heredia, 2014; Furtak et al., 2014; Furtak et al., 2016; von Aufschnaiter & Alonzo, 2018; Zhai, Li, & Guo, 2018); and one study (Jin, Johnson, Shin, & Anderson, 2017) explored how teachers' classroom discourse helped students move up the LP levels.

Although LPs are purported to be useful tools for teaching science, they present significant challenges for teachers in both content knowledge and pedagogical content knowledge components, especially in understanding student thinking and designing follow-up instruction. Studies that assessed teachers' content knowledge found that achieving the upper anchor of the student LP was challenging for many teachers (Gunckel et al., 2018; Jin et al., 2015). Eliciting and interpreting student thinking at different LP levels is even more challenging (Furtak et al., 2014). For example, one study found that teachers struggled to connect intuitive ways of thinking within and across students; the complexity of the data and the different dimensions of the LP sometimes served as roadblocks toward the goal of having teachers use the LP productively (Furtak et al., 2014). Identifying student ideas at different levels is only one part of the challenge. Teachers must also consider what feedback to provide to help students move toward the upper anchor. Researchers found that teachers struggle with providing follow-up instructional moves that target students' ideas at a certain LP level (Furtak, 2012). Moreover, although LPs are intended to guide teachers to view students' ideas as funds of knowledge for the development of sophisticated understanding (Duschl et al., 2011), teachers tend to rely on traditional instructional approaches that often correct students' intuitive ideas described on the lower LP levels (Covitt et al., 2018; Gunckel et al., 2018; Jin et al., 2017).

Within the category of teaching and teacher learning, a large portion of the studies explored how LPs could be used to promote teachers' use of formative assessment. One key argument across this set of papers is the importance of ensuring that classroom assessments are linked directly to LPs to achieve the most productive use (Covitt et al., 2018; Furtak, 2012; Furtak & Heredia, 2014; Furtak et al., 2014; von Aufschnaiter & Alonzo, 2018). These papers argued persuasively for LPs to support classroom assessment practices whereby teachers use the LP to make sense of student thinking and determine subsequent instructional moves based on their interpretations.

Several recent studies used innovative approaches to promote teacher learning and produced empirical evidence of the effectiveness of those approaches. Von Aufschnaiter and Alonzo (2018) engaged preservice physics teachers in using an LP for forces and motion to analyze videos of student learning (cognitive interviews with students and classroom instruction involving students). They found that, by using the LP, preservice teachers were able to attend to specific aspects of student learning (e.g., ideas described on different LP levels) rather than providing a holistic evaluation. The preservice teachers also began to think about student-centered instructional approaches that target the identified intuitive ideas. Furtak et al. (2018) used a formative assessment design cycle, a five-step process, to help teachers use an LP for natural selection to design and enact formative assessment tasks. In the process, the researchers' support gradually fades and teachers take more and more responsibilities in using the LP to design and enact formative assessment. In another study by Furtak and Heredia (2014), the researchers engaged teachers in the learning community that helped to develop the LP; and that experience had a significant contribution to the teachers' understanding and use of the LP in their teaching. Gunckel et al. (2018) used an LP for water to develop educative teaching materials. The materials contained two important educative components. First, the materials contained a set of formative assessment tasks. For those tasks, the guidance for teachers explained how to analyze students' responses at different LP levels and suggested instructional foci for students performing at different LP levels. Second, the materials also contained a set of tools for reasoning. For those tools, the guidance to teachers explained how the tools could be incorporated into an instructional sequence and how the tools could be used to support students at each LP level.

5.3.4 | Summary of research on horizontal coherence

Research efforts have been distributed unequally across the three research activities for horizontal coherence. Considerable research has been devoted to the development of LPs. As a result, the methodology for developing an LP is well-established. To achieve the compatibility with the current research, researchers may conduct systematic review and include literature from a wide range of fields such as science, the philosophy and history of science, and cognitive and learning sciences. To enhance the empirical validity, researchers used iterations, independent conceptual replications, and sampling strategies. More important, both qualitative and quantitative techniques have been used to validate the meaning of, the distinctions among, and the order of LP levels.

Relatively fewer studies have been conducted to explore the use of LPs for curriculum, instruction, and teacher learning. Although some LP-based interventions have generated positive learning outcomes, we know very little about how researchers align curriculum and instructional approaches with LPs. LP research on teaching and teacher learning suggest that teachers need support both in developing understanding of the upper anchor and in using LPs for classroom assessment and instruction. A few researchers have used innovative ideas such as video analysis, educative materials, and professional learning communities to help teachers learn and use LPs. However, more efforts are needed, especially at the elementary school level, to help teachers understand and use LPs in their day-to-day teaching.

5.4 | Vertical coherence

In the conceptual framework section, we discuss three challenges in building vertical coherence between classroom assessments and large-scale assessments. These three challenges are assessing complex constructs, assessing learning over time, and assessing both the breadth and depth of knowledge and practices. We found one article discussing vertical coherence. Anderson et al. (2018) reported on a design-based implementation research project that supported teaching and learning of carbon cycling in science classrooms. The project was meant to help students achieve the three-dimensional learning goals of the NGSS. Because the article aimed to cover several major achievements of the project, only one section contributed to vertical coherence. This section provided some information about how the researchers dealt with the first two challenges. First, the project assessed learning over time by using three different types of assessments. Teachers used formative assessments to scaffold students in class and used periodical assessments to monitor student learning within several weeks or months. A large-scale assessment was used to evaluate the learning outcomes. The assessment was used with more than 850,000 students across several states. Second, a variety of item types, including constructed-response items, forced-choice items, and multiple-choice items, were used to provide students with diverse opportunities to demonstrate their understanding of integrated knowledge and practices. New technology such as online assessments and automated scoring made it possible to administer the assessments at different scales and to score a very large sample of responses.

6 | DISCUSSION

In this section, we first respond to a major critique to LP approaches. Then we discuss the implications of the major results of the review. Finally, we suggest future research directions.

6.1 | Response to the critique of LP approaches

Within the domain-specific learning perspective, two competing views exist. One group of researchers maintain that students' ideas about a topic or within a domain are embedded in a coherent theory-like structure (Carey, 1985; Reiner, Slotta, Chi, & Resnick, 2000; Vosniadou et al., 2008). The aforementioned naïve theories are a key part of this view. The other group of researchers has argued that students hold many p-prims or facets, which are idiosyncratic and fragmented ideas; these ideas are loosely organized and are activated in specific contexts (diSessa, 1993; Minstrell & Stimpson, 1996). While the initiators of LPs heavily drew upon naïve theories to conceptualize LPs (e.g., Smith et al., 2006), researchers with the knowledge-in-pieces view are skeptical. One major critique comes from Hammer and Sikorski (2015), who argued that most LPs tend to present development as linear and sequential, and such presentations do not capture the dynamics and contextual factors of learning.

The debate regarding whether students' ideas are theory-like or fragmented has been ongoing for several decades, yet empirical data support both views. Even in a quasi-replication condition, researchers found evidence supporting both of the opposing views (diSessa, Gillespie, & Esterly, 2004; Ioannides & Vosniadou, 2002). We think the controversy is largely due to the different grain sizes that the researchers use to investigate student learning. For example, when analyzing a student's interview or written responses, the researcher may focus on the salient patterns of the student's reasoning and disregard insignificant evidence of reasoning. The instructional approaches will focus on helping students make the transition from those salient reasoning patterns toward scientific reasoning. Conversely, the researcher may aim to compile a comprehensive list of the student's ideas and design instructional approaches that deal with those loosely connected ideas in the list. There are conceptual reasons and empirical evidence for both views to be correct. Therefore, taking a pragmatic standpoint, we suggest that the evaluation of LP approaches focus on whether enough empirical evidence is available to support the effectiveness of LP approaches in promoting teaching and learning. Currently, many LPs have been validated based on longitudinal and large-scale data. Several studies have generated significant learning gains from LP-based instructions (e.g., Anderson et al., 2018). These studies provide ample evidence for the effectiveness of LP approaches. LP research does shed useful insights for promoting teaching and learning of science in schools. However, we should caution not to treat LP as the only or the best approach to science education.

6.2 | Instructional implications of LPs

Different types of LPs have different implications for instruction. For the knowledge LPs, we use three LPs for energy as examples to explain the different instructional implications of the knowledge enrichment LPs, the knowledge integration LPs, and the theory transformation LPs. Neumann et al. (2013) developed an LP for energy in physics. The LP describes student development as a process of knowledge enrichment and resolved two important problems in traditional physics curriculum. In traditional physics curriculum, energy is first introduced as "the ability to do work." Subsequently, key energy concepts are introduced in a particular sequence: energy transformation/transfer, energy conservation, and energy degradation (heat dissipation). This instructional sequence has at least two problems. First, the abstract concept of energy is defined by using another abstract concept: work. As such, the energy definition does not provide any useful information for students to grasp the scientific meaning of energy. Second, ample evidence suggests that curriculum materials using the above instructional sequence do not help students develop a coherent understanding of energy. In particular, students seldom recognize the connections between energy conservation and energy degradation (i.e., the total quantity of energy is conserved, but the amount of useful energy decreases), and they tend to treat these two principles as contradictory (Duit, 1984, 2014). Neumann et al.'s LP begins with forms of energy, a concept that students can use to connect real world experience (e.g., light, motion, and warmth) with the abstract energy concepts. This helps students develop a preliminary understanding of what energy is. In the LP, energy degradation appears before energy conservation because students must first know that heat dissipation always happens before they understand that heat must be considered when conserving energy. Therefore, this LP, which has been validated based on empirical assessment data, suggests a lesson sequence that resolved the two problems in the traditional approach. The LP presents the logic of science—the relationships among the energy concepts. It also attends to cognition—there are both conceptual reasons and empirical evidence regarding why some concepts/principles appear earlier.

Jin & Anderson (2012) developed an LP for energy in socioecological systems. The LP contains four levels, with Levels 1, 2, and 3 each describing an intuitive reasoning pattern and Level 4 describing the scientific reasoning. More specifically, Level 1 is force-dynamic reasoning, which explains natural phenomena in terms of enablers (e.g., water, soil, and air), agents (e.g., a living tree), and results (e.g., a small plant growing into a big tree). Level 1 reasoning is very different from Level 4 reasoning that traces matter and energy across biochemical processes. This LP focuses on cognition because the three lower levels describe intuitive reasoning patterns. The lower levels cannot be used as learning goals. When using this LP, teachers think about how each lower level differs from the upper anchor and create learning activities that help students transform their intuitive reasoning into scientific reasoning.

Lee and Liu (2010) developed an LP for energy across three subject areas, including physical sciences, life sciences, and earth sciences. The LP has the following levels: providing nonnormative ideas, providing one normative idea, linking two normative ideas, and producing at least two links among three or more normative ideas. These levels are not specific about the content topics. Thus, this LP provides a preliminary description of development in understanding energy concepts. In order for the LP to provide a detailed illustration of the syntax of science, we suggest continued work on specifying what links are built among what ideas. By doing so, the LP could provide practical guidance for teachers to help students develop the conceptual understanding of energy as the students learn about the energy across disciplines.

Similarly, different integration approaches used in practice LPs have varying implications for teaching scientific practices. The generic LPs describe trends across content. They are best suited for monitoring student learning of practices across content topics. Using these LPs, teachers can determine whether their students develop a better understanding of a scientific practice as they move from one content topic to the next. Content-specific LPs illustrate the intertwined practices and knowledge. They help teachers understand what exactly a scientific practice looks like in the context of a science topic and how students develop competence in the scientific practice over time. The matter LP by Smith et al. (2006) uses a process-product approach to integration of practices and knowledge. Such LPs provide useful information for designing learning activities, where students learn content knowledge through their engagement in scientific investigations. Process-product LPs layout which inquiry skills are needed for students to achieve a certain level of conceptual understanding of science content.

6.3 | Future research directions

We recommend future directions in LP research to address gaps in developmental, horizontal, and vertical coherences.

6.3.1 | Developmental coherence: Scientific reasoning that unifies disciplinary knowledge and practices

In science, knowledge and practices are unified through scientific reasoning. The mastery of scientific reasoning enables students to think like scientists when learning knowledge, analyzing phenomena, and solving problems. However, learning scientific reasoning can be challenging because it is often unfamiliar or even counterintuitive to students. Although many LPs have been developed to target scientific reasoning within one content topic and one scientific practice, we know little about the pathways that describe how scientific reasoning is developed across science topics or even science disciplines.

Inspired by Osborne, Rafanelli, and Kind (2018), we differentiate between two types of scientific reasoning. First, scientific reasoning can be about an ontic entity (e.g., pattern, matter, and energy) and therefore is useful within a certain range, such as several content topics or a scientific practice. For example, a specialized way of reasoning is used to understand matter and energy across all science disciplines: matter cycling and energy flowing within and across biological, physical, and chemical processes. Teaching this reasoning consistently across multiple science topics help students develop a coherent understanding of energy (e.g., Fortus, Adams, Krajcik, & Reiser, 2015). LP approaches can be useful for helping students develop such coherent understanding. Teachers can use LP-based periodical assessments to monitor student progress as them move from one energy topic to another. Second, scientific reasoning can also have broad coverage and application. Osborne and colleagues (Osborne et al., 2018) proposed using six "styles of scientific reasoning" (e.g., mathematical deduction and experimental exploration) to unify science curriculum. These styles of reasoning were used in the history of science for conducting scientific investigations and for learning scientific knowledge across disciplines. Compared to reasoning that is only associated with ontic entities, the styles of reasoning have a much broader range of application. They are essential for the learning of not only content knowledge, but also procedural knowledge and epistemic knowledge. More important, they are specialized ways of reasoning that are essential for all major science disciplines. As such, it is important for students to master these styles of scientific reasoning through their learning of knowledge and practices across science disciplines. We encourage researchers to investigate how to use LPs to describe the development achieved through learning across disciplines and how to use LP-based curriculum, instruction, and assessment to promote productive learning of the styles of reasoning.

6.3.2 | Horizontal coherence: Teachers' use of LP approaches in action research

Horizontal coherence targets the connection between curriculum, instruction, and assessment, with teachers at the center of the system. Although several studies have investigated how teachers use LPs as the products generated by researchers (e.g., Furtak et al., 2018; Gunckel et al., 2018; von Aufschnaiter & Alonzo, 2018), no research has investigated teachers' use of LP approaches.

To contribute to this line of research, we propose using LP approaches in teachers' action research. An example comes from the learning trajectory research in mathematics education. When teaching a class of preservice elementary mathematics teachers, Simon (1995) developed a learning trajectory and a lesson sequence on the multiplicative relationship involved in the area of a rectangle. In the teaching experiment, Simon repeated a cycle that contained the following steps: identifying/hypothesizing students' conceptual difficulties, designing mathematics problems to address the conceptual difficulties, observing and analyzing students' discussions on solving the problem, and developing a new problem to be solved in the following lesson. Sometimes, Simon was surprised by the intuitive ideas that surfaced in the class discussion. In such situations, Simon changed his instruction following the class discussion accordingly. Toward the end of the five-week teaching segment, Simon developed a learning trajectory of preservice teachers' understanding and a lesson sequence associated with this learning trajectory. Simon's approach is a good example of how K–12 teachers can develop synchronized LP and lesson sequences in their own classrooms. Teacher educators may provide professional development opportunities for teachers to learn LP approaches. Teachers may learn how to use data from students' work to develop an LP and how to use the LP to develop assessment tasks and learning activities. Therefore, we advocate using the LP as a means of actively designing successful instruction on the part of the teachers rather than an end whereby teachers are expected to use those levels as ready-made recipes without much input on their part.

6.3.3 | Vertical coherence: Dealing with three challenges

There are three challenges in building vertical coherence: assessing complex construct, assessing learning over time, and assessing both breadth and depth of knowledge and practices. In this review, we found only one empirical study dealing with first two challenges. There are, however, several relevant suggestions in other articles for building vertical coherence. Pellegrino and colleagues (NRC, 2014) suggested providing multiple and varied assessment opportunities for students to demonstrate both breadth and depth of their understanding in large-scale assessments. Wilson and Sloane (2000) suggested using LP-based assessments to assess learning over time. At the classroom level, progress variables define the constructs that are aligned with specific learning goals of a curriculum to a level of detail that would allow teachers to monitor student learning at a day-to-day or week-to-week level. Over the course of instruction, students advance along one or more progress variables. Large-scale assessments measure constructs defined by a collection of progress variables. In this way, the classroom assessments and large-scale assessments can provide compatible results across time and context. We hope the above ideas, along with the research of Anderson et al. (2018), will spark future efforts in using LP approaches to build vertical coherence of science assessment systems.

7 | CONCLUSION

While the LP research has clearly advanced in both breadth and depth since Duschl et al.'s (2011) early review, there is still much room for growth. Extensive research has been conducted in developmental coherence. The results suggest that the development of scientific knowledge may take different forms (knowledge enrichment, knowledge integration, and theory transformation) and that scientific practices and knowledge can be integrated in different ways (trends across content, intertwined practices and knowledge, and process-product). However, more research efforts are needed in using LP approaches to promote the development of scientific reasoning across science topics and disciplines. For horizontal coherence, efforts have been distributed unevenly across several subcategories of research. While many empirical LP studies have been conducted to develop and validate LPs, relatively fewer studies have explored LP-based interventions and effective approaches to teachers' learning and use of LPs. We suggest more professional development opportunities for teachers to learn about not only how to use research-developed LPs in classroom teaching but also how to use LP approaches to develop LPs and associated materials that fit their own classroom teaching. Relatively little has been written about using LP approaches to enhance vertical coherence of science education systems. We therefore call for more studies to deal with the three challenges in vertical coherence: assessing complex constructs, assessing learning over time, and assessing both breadth and depth of knowledge and practices.

ORCID

Hui Jin http://orcid.org/0000-0001-7688-703X
Jamie N. Mikeska http://orcid.org/0000-0002-8831-2572

REFERENCES

Adadan, E., & Oner, D. (2014). Exploring the progression in preservice chemistry teachers' pedagogical content knowledge representations: The case of "behavior of gases. *Research in Science Education*, *44*, 829–858. https://doi.org/10.1007/s11165-014-9401-6

Allen, G. E. (2003). Mendel and modern genetcs: The legacy for today. *Endeavour*, *27*, 63–68.

Allchin, D. (2012). Teaching the nature of science through scientific errors. *Science Education*, *96*(5), 904–926. https://doi.org/10.1002/sce.21019

Alonzo, A. C. (2018). Exploring the learning progression-formative assessment hypothesis. *Applied Measurement in Education*, *31*, 101–103. https://doi.org/10.1080/08957347.2017.1408625

Alonzo, A. C., & Gotwals, A. W. (Eds.) (2012). *Learning progressions in science: Current challenges and future directions*. Rotterdam: The Netherlands Sense Publishers. Retrieved from https://doi.org/10.1007/978-94-6091-824-7

Alonzo, A. C., & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education*, *93*, 389–421. https://doi.org/10.1002/sce.20303

American Association for the Advancement of Science (1993). *Benchmarks for science literacy*. New York, NY: Oxford University Press.

Anderson, C. W. (2008). *Conceptual and empirical validation of learning progressions*. East Lansing, MI: Michigan State University.

Anderson, C. W., de los Santos, E. X., Bodbyl, S., Covitt, B. A., Edwards, K. D., Hancock, I. , J. B., & Welch, M. M. (2018). Designing educational systems to support enactment of the Next Generation Science Standards. *Journal of Research in Science Teaching*, *55*, 1026–1052. https://doi.org/10.1002/tea.21484

Bamberger, Y. M., & Davis, E. A. (2013). Middle-school science students' scientific modelling performances across content areas and within a learning progression. *International Journal of Science Education*, *35*, 213–238. https://doi.org/10.1080/09500693.2011.624133

Bennett, R. E. (2010). Cognitively based assessment of, for, and as learning (CBAL): A preliminary theory of action for summative and formative assessment. *Measurement*, *8*, 70–91. https://doi.org/10.1080/15366367.2010.508686

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*, 765–793. https://doi.org/10.1002/sce.20402

Bravo-Torija, B., & Jiménez-Aleixandre, M.-P. (2018). Developing an initial learning progression for the use of evidence in decision-making contexts. *International Journal of Science and Mathematics Education*, *16*, 619–638. https://doi.org/10.1007/s10763-017-9803-9

Breslyn, W., McGinnis, J. R., McDonald, R. C., & Hestness, E. (2016). Developing a learning progression for sea level rise, a major impact of climate change. *Journal of Research in Science Teaching*, *53*, 1471–1499. https://doi.org/10.1002/tea.21333

Cabello-Garrido, A., Espana-Ramos, E., & Blanco-Lopez, A. (2018). Developing a human nutrition learning progression. *International Journal of Science and Mathematics Education*, *16*, 1269–1289. https://doi.org/10.1007/s10763-017-9838-y

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Carey, S. (1986). Cognitive science and science education. *American Psychologist*, *41*, 1123–1130. https://doi.org/10.1037/0003-066X.41.10.1123

Carey, S. (1991). Knowledge acquisition: Enrichment or conceptual change? In S. Carey, & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chen, F., Zhang, S., Guo, Y., & Xin, T. (2017). Applying the rule space model to develop a learning progression for thermochemistry. *Research in Science Education*, *47*, 1357–1378. https://doi.org/10.1007/s11165-016-9553-7

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 61–101). New York, NY: Routledge.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152. https://doi.org/10.1207/s15516709cog0502_2

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Colantonio, A., Galano, S., Leccia, S., Puddu, E., & Testa, I. (2018). Design and development of a learning progression about stellar structure and evolution. *Physical Review Physics Education Research*, *14*(1), 1–23. https://doi.org/10.1103/PhysRevPhysEducRes.14.010143

Corcoran, T., Mosher, F. A., & Rogat, A. D. (2009). *Learning progressions in science: An evidence-based approach to reform*. New York, NY: Columbia University, Teachers College, Center on Continuous Instructional Improvement.

Covitt, B. A., Gunckel, K. L., Bess, C., & Syswerda, S. (2018). Teachers' use of learning progression-based formative assessment in water instruction. *Applied Measurement in Education*, *31*, 128–142. https://doi.org/10.1080/08957347.2017.1408627

DeBoer, G. E. (1991). *A history of ideas in science education: Implications for practice*. New York, NY: Teachers College Press.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, *10*(2–3), 105–225. https://doi.org/10.1080/07370008.1985.9649008

diSessa, A. A., Gillespie, N. M., & Esterly, J. B. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science*, *28*, 843–900. https://doi.org/10.1207/s15516709cog2806_1

Duit, R. (1984). Learning the energy concept in school—Empirical results from the Philippines and West Germany. *Physics Education*, *19*(2), 59–66. https://doi.org/10.1088/0031-9120/19/2/306

Duit, R. (2014). Teaching and learning the physics energy concept. In R. F. Chen, A. Eisenkraft, D. Fortus, J. Krajcik, K. Neumann, J. Nordine, & A. Scheff (Eds.), *Teaching and learning of energy in K–12 education* (pp. 67–85). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-05017-1_5.

Duncan, R. G., Castro-Faix, M., & Choi, J. (2016). Informing a learning progression in genetics: Which should be taught first, Mendelian inheritance of the central dogma of molecular biology. *International Journal of Science and Mathematics Education*, *14*, 445–472. https://doi.org/10.1007/s10763-014-9568-3

Duncan, R. G., Choi, J., Castro-Faix, M., & Cavera, V. L. (2017). A study of two instructional sequences informed by alternative learning progressions in genetics. *Science & Education*, *26*, 1115–1141. https://doi.org/10.1007/s11191-017-9932-0

Duncan, R. G., & Hmelo-Silver, C. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, *46*, 606–609. https://doi.org/10.1002/tea.20316

Duncan, R. G., Rogat, A. D., & Yarden, A. (2009). A learning progression for deepening students' understanding of modern genetics across the 5th-10th grades. *Journal of Research in Science Teaching*, *46*, 655–674. https://doi.org/10.1002/tea.20312

Duschl, R. A., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education*, *47*, 123–182. https://doi.org/10.1080/03057267.2011.604476

Elmesky, R. (2013). Building capacity in understanding foundational biology concepts: A K–12 learning progression in genetics informed by research on children's thinking and learning. *Research in Science Education*, *43*, 1155–1175. https://doi. https://doi.org/10.1007/s11165-012-9286-1

Forbes, C. T., Zangori, L., & Schwartz, C. V. (2015). Empirical validation of integrated learning performances for hydrologic phenomena: 3rd-grade students' model-driven explanation-construction. *Journal of Research in Science Teaching*, *52*, 895–921. https://doi.org/10.1002/tea.21226

Fortus, D., Adams, L. M. S., Krajcik, J., & Reiser, B. (2015). Assessing the role of curriculum coherence in student learning about energy. *Journal of Research in Science Teaching*, *52*, 1408–1425. https://doi.org/10.1002/tea.21261

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

Fulmer, G. W. (2015). Validating proposed learning progressions on force and motion using the force concept inventory: Findings from Singapore secondary schools. *International Journal of Science and Mathematics Education*, *13*, 1235–1254. https://doi.org/10.1007/s10763-014-9553-x

Fulmer, G. W., Liang, L. L., & Liu, X. (2014). Applying a force and motion learning progression over an extended time span using the force concept inventory. *International Journal of Science Education*, *36*, 2918–2936. https://doi.org/10.1080/09500693.2014.939120

Fulmer, G. W., Tanas, J., & Weiss, K. A. (2018). The challenges of alignment for the next generation science standards. *Journal of Research in Science Teaching*, *55*, 1076–1100. https://doi.org/10.1002/tea.21481

Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, *49*, 1181–1210. https://doi.org/10.1002/tea.21054

Furtak, E. M., Bakeman, R., & Buell, J. Y. (2018). Developing knowledge-in-action with a learning progression: Sequential analysis of teachers' questions and responses to student ideas. *Teaching and Teacher Education*, *76*, 267–282. https://doi.org/10.1016/j.tate.2018.06.001

Furtak, E. M., Circi, R., & Heredia, S. C. (2018). Exploring alignment among learning progressions, teacher-designed formative assessment tasks, and student growth: Results of a four-year study. *Applied Measurement in Education*, *31*, 143–156. https://doi.org/10.1080/08957347.2017.1408624

Furtak, E. M., & Heredia, S. C. (2014). Exploring the influence of learning progressions in two teacher communities. *Journal of Research in Science Teaching*, *51*, 982–1020. https://doi.org/10.1002/tea.21156

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de León, V., Morrison, D., & Heredia, S. C. (2016). Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science*, *44*, 267–291. https://doi.org/10.1007/s11251-016-9371-3

Furtak, E. M., Morrison, D., & Kroog, H. (2014). Investigating the link between learning progressions and classroom assessment. *Science Education*, *98*, 640–673. https://doi.org/10.1002/sce.21122

Gabel, C. (2006). Impact of the reform efforts on K–12 science inquiry: A paradigm shift. In D. W. Sunal, & E. L. Wright (Eds.), *The impact of state and national standards on K–12 science teaching* (pp. 215–256). Greenwich, CT: Information Age.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner, & A. L. Stevens (Eds.), *Mental models* (pp. 99–129). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld, & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257–293). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511752902.011.

Gotwals, A. W., & Songer, N. B. (2010). Reasoning up and down a food chain: Using an assessment framework to investigate students' middle knowledge. *Science Education*, *94*, 259–281. https://doi.org/10.1002/sce.20368

Gotwals, A. W., & Songer, N. B. (2013). Validity evidence for learning progression-based assessment items that fuse core disciplinary ideas and science practices. *Journal of Research in Science Teaching*, *50*, 597–626. https://doi.org/10.1002/tea.21083

Gough, D., Oliver, S., & Thomas, J. (2012). *An introduction to systematic review*. London: Sage Publications.

Gunckel, K. L., Covitt, B., Salinas, I., & Anderson, C. W. (2012). A learning progression for water in socio-ecological systems. *Journal of Research in Science Teaching*, *49*, 843–868. https://doi.org/10.1002/tea.21024

Gunckel, K. L., Covitt, B. A., & Salinas, I. (2018). Learning progressions as tools for supporting teacher content knowledge and pedagogical content knowledge about water in environmental systems. *Journal of Research in Science Teaching*, *55*, 1339–1362. https://doi.org/10.1002/tea.21454

Hadenfeldt, J. C., Liu, X., & Neumann, K. (2014). Framing students' progression in understanding matter: A review of previous research. *Studies in Science Education*, *50*, 181–208. https://doi.org/10.1080/03057267.2014.945829

Hadenfeldt, J. C., Neumann, K., Bernholt, S., Liu, X., & Parchmann, I. (2016). Students' progression in understanding the matter concept. *Journal of Research in Science Teaching*, *53*, 683–708. https://doi.org/10.1002/tea.21312

Hammer, D., & Sikorski, T.-R. (2015). Implications of complexity for research on learning progressions. *Science Education*, *99*, 424–431. https://doi.org/10.1002/sce.21165

Hernandez, M. I., Couso, D., & Pintó, R. (2015). Analyzing students' learning progressions throughout a teaching sequence on acoustic properties of materials with a model-based inquiry approach. *Journal of Science Education and Technology*, *24*, 356–377. https://doi.org/10.1007/s10956-014-9503-y

Herrmann-Abell, C. F., & DeBoer, G. E. (2018). Investigating a learning progression for energy ideas from upper elementary through high school. *Journal of Research in Science Teaching*, *55*(1), 68–93. https://doi.org/10.1002/tea.21411

Hirschfeld, L. A., & Gelman, S. A. (1994). Toward a topography of mind: An introduction to domain specificity. In L. A. Hirschfeld, & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511752902.002.

Hokayem, H., & Gotwals, A. W. (2016). Early elementary students' understanding of complex ecosystems: A learning progression approach. *Journal of Research in Science Teaching*, *53*(10), 1524–1545. https://doi.org/10.1002/tea.21336

Hovardas, T. (2016). A learning progression should address regression: Insights from developing non-linear reasoning in ecology. *Journal of Research in Science Teaching*, *53*, 1447–1470. https://doi.org/10.1002/tea.21330

Hsin, M.-C., Chien, S.-P, Hsu, Y.-S., Lin, C.-Y., & Yore, L. D. (2016). Development and validation of a Taiwanese communication progression in science education. *International Journal of Science and Mathematics Education*, *14*, 125–143. https://doi.org/10.1007/s10763-014-9589-y

Inagaki, K., & Hatano, G. (2004). Vitalistic causality in young children's naive biology. *Trends in Cognitive Sciences*, *8*, 356–362. https://doi.org/10.1016/j.tics.2004.06.004

Ioannides, C., & Vosniadou, S. (2002). The changing meanings of force. *Cognitive Science Quarterly*, *2*(1), 5–62.

Jin, H., & Anderson, C. W. (2012). A learning progression for energy in socio-ecological systems. *Journal of Research in Science Teaching*, *49*, 1149–1180. https://doi.org/10.1002/tea.21051

Jin, H., Johnson, M. E., Shin, H. J., & Anderson, C. W. (2017). Promoting student progressions in science classrooms: A video study. *Journal of Research in Science Teaching*, *54*, 852–883. https://doi.org/10.1002/tea.21388

Jin, H., Shin, H. J., Johnson, M. E., Kim, J., & Anderson, C. W. (2015). Developing learning progression-based teacher knowledge measures. *Journal of Research in Science Teaching*, *52*, 1269–1295. https://doi.org/10.1002/tea.21243

Jin, H., & Wei, X. (2014). Using ideas from the history of science and linguistics to develop a learning progression for energy in socio-ecological systems. In R. F. Chen, A. Eisenkraft, D. Fortus, J. Krajcik, K. Neumann, J. Nordine, & A. Scheff (Eds.), *Teaching and learning of energy in K–12 education* (pp. 157–174). New York, NY: Springer. https://doi.org/10.1007/978-3-319-05017-1_9.

Jin, H., Zhan, L., & Anderson, C. W. (2013). Developing a fine-grained learning progression framework for carbon-transforming processes. *International Journal of Science Education*, *35*(10), 1663–1697. https://doi.org/10.1080/09500693.2013.782453

Johnson, P., & Tymms, P. (2011). The emergence of a learning progression in middle school chemistry. *Journal of Research in Science Teaching*, *48*, 849–877. https://doi.org/10.1002/tea.20433

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, *88*, 197–227. https://doi.org/10.1037/0033-295X.88.3.197

Keil, F. C. (1992). The origins of an autonomous biology. In M. R. Gunnar, & M. Maratsos (Eds.), *Modularity and constraints in language and cognition: The Minnesota symposia on child development* Vol.25, pp. 103–137). Hillsdale, NJ: Lawrence Erlbaum Associates.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174. https://doi.org/10.2307/2529310

Lee, H.-S., & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education*, *94*, 665–688. https://doi.org/10.1002/sce.20382

Lehrer, R., & Schauble, L. (2009). Images of learning, images of progress. *Journal of Research in Science Teaching*, *46*, 731–735. https://doi.org/10.1002/tea.20317

Lehrer, R., & Schauble, L. (2012). Seeding evolutionary thinking by engaging children in modeling its foundations. *Science Education*, *96*, 701–724. https://doi.org/10.1002/sce.20475

Lehrer, R., & Schauble, L. (2015). Learning progressions: The whole world is NOT a stage. *Science Education*, *99*, 432–437. https://doi.org/10.1002/sce.21168

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education*, *87*, 517–538. https://doi.org/10.1002/sce.10086

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition*, *99*, 167–204. https://doi.org/10.1016/j.cognition.2004.12.009

Lyon, E. G. (2013). Conceptualizing and exemplifying science teachers' assessment expertise. *International Journal of Science Education*, *35*, 1208–1229. https://doi.org/10.1080/09500693.2013.770180

Makel, M. C., Plucker, J. A., Freeman, J., Lombardi, A., Simonsen, B., & Coyne, M. (2016). Replication of special education research: Necessary but far too rare. *Remedial and Special Education*, *37*, 205–212. https://doi.org/10.1177/0741932516646083

Mayes, R. L., Forrester, J. H., Christus, J. S., Peterson, F. I., Bonilla, R., & Yestness, N. (2014). Quantitative reasoning in environmental science: A learning progression. *International Journal of Science Education*, *36*, 635–658. https://doi.org/10.1080/09500693.2013.819534

Mayr, E. (1982). The growth of biological thought: Diversity, evolution, and inheritance, Cambridge, MA: Belknap Press of Harvard University Press.

McCloskey, M., Washburn, A., & Felch, L. (1983). Intuitive physics: The straight-down belief and its origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*, 636–659. https://doi.org/10.1037/0278-7393.9.4.636

Minstrell, J., & Stimpson, V. (1996). A classroom environment for learning: Guiding students' reconstruction of understanding and reasoning. In L. Schauble, & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 175–202). Mahwah, NJ: Lawrence Erlbaum Associates.

Mohan, L., Chen, J., & Anderson, C. W. (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching*, *46*, 675–698. https://doi.org/10.1002/tea.20314

Moon, A., Courtney, S., Cole, R., & Towns, M. H. (2017). Analysis of inquiry materials to explain complexity of chemical reasoning in physical chemistry students' argumentation. *Journal of Research in Science Teaching*, *54*, 1322–1346. https://doi.org/10.1002/tea.21407

Morell, L., Collier, T., Black, P., & Wilson, M. (2017). A construct-modeling approach to develop a learning progression of how students understand the structure of matter. *Journal of Research in Science Teaching*, *54*, 1024–1048. https://doi.org/10.1002/tea.21397

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academy Press. https://doi.org/10.17226/9596.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press. https://doi.org/10.17226/10019.

National Research Council. (2005). *Systems for state science assessment*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11312.

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K–8*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11625.

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

National Research Council. (2014). *Developing assessments for the Next Generation Science Standards*. Washington, DC: The National Academies Press. https://doi.org/10.17226/18409.

Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, *50*, 162–188. https://doi.org/10.1002/tea.21061

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

Opitz, S. T., Harms, U., Neumann, K., Kowalzik, K., & Frank, A. (2014). Students' energy concepts at the transition between primary and secondary school. *Research in Science Education*, *45*, 691–715. https://doi.org/10.1007/s11165-014-9444-8

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, *328*(23), 463–466. https://doi.org/10.1126/science.1183944

Osborne, J., Henderson, B., MacPherson, A., Szu, E., Wild, A., & Yao, S.-Y. (2016). The development and validation of a learning progression for argumentation. *Journal of Research in Science Teaching*, *53*, 821–846. https://doi.org/10.1002/tea.21316

Osborne, J., Rafanelli, S., & Kind, P. (2018). Toward a more coherent model for science education than the crosscutting concepts of the next generation science standards: The affordances of styles of reasoning. *Journal of Research in Science Teaching*, *55*, 962–981. https://doi.org/10.1002/tea.21460

Paik, S., Song, G., Kim, S., & Ha, M. (2017). Developing a four-level learning progression and assessment for the concept of buoyancy. *Eurasia Journal of Mathematics Science and Technology Education*, *13*, 4965–4986. https://doi.org/10.12973/eurasia.2017.00976a

Pierson, A. E., Clark, D. B., & Sherard, M. K. (2017). Learning progressions in context: Tensions and insights from a semester-long middle school modeling curriculum. *Science Education*, *101*, 1061–1088. https://doi.org/10.1002/sce.21314

Plummer, J. D. (2014). Spatial thinking as the dimension of progress in an astronomy learning progression. *Studies in Science Education*, *50*, 1–45. https://doi.org/10.1080/03057267.2013.869039

Plummer, J. D., & Krajcik, J. (2010). Building a learning progression for celestial motion: Elementary levels from an Earth-based perspective. *Journal of Research in Science Teaching*, *47*, 768–787. https://doi.org/10.1002/tea.20355

Plummer, J. D., & Maynard, L. (2014). Building a learning progression for celestial motion: An exploration of students' reasoning about the seasons. *Journal of Research in Science Teaching*, *51*, 902–929. https://doi.org/10.1002/tea.21151

Plummer, J. D., Palma, C., Flarend, A., Rubin, K., Ong, Y. S., Botzer, B., & Furman, T. (2015). Development of a learning progression for the formation of the solar system. *International Journal of Science Education*, *37*, 1381–1401. https://doi.org/10.1080/09500693.2015.1036386

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*, 211–227. https://doi.org/10.1002/sce.3730660207

Premack, D., & Premack, A. J. (1995). Intention as psychological cause. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate*. Oxford: Clarendon Press.

Reiner, M., Slotta, J. D., Chi, M., & Resnick, L. B. (2000). Naive physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction*, *18*, 1–34. https://doi.org/10.1207/S1532690XCI1801_01

Rivet, A. E., & Kastens, K. A. (2012). Developing a construct-based assessment to examine students' analogical reasoning around physical models in earth science. *Journal of Research in Science Teaching*, *49*, 713–743. https://doi.org/10.1002/tea.21029

Romine, W. L., Todd, A. N., & Clark, T. B. (2016). How do undergraduate students conceptualize acid-base chemistry? Measurement of a concept progression. *Science Education*, *100*, 1150–1183. https://doi.org/10.1002/sce.21240

Sanderson, B. A., & Kratochvil, D. (1971). *Science: A process approach*. Palo Alto, CA: US Department of Health, Education, and Welfare.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*, 90–100. https://doi.org/10.1037/a0015108

Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of U.S. mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, *37*, 525–559. https://doi.org/10.1080/0022027042000294682

Schneider, R. M., & Plasman, K. (2011). Science teacher learning progressions: A review of science teachers' pedagogical content knowledge development. *Review of Educational Research*, *81*, 530–565. https://doi.org/10.3102/0034654311423382

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*, 632–654. https://doi.org/10.1002/tea.20311

Sevian, H., & Talanquer, V. (2014). Rethinking chemistry: A learning progression on chemical thinking. *Chemistry Education Research and Practice*, *15*, 10–23. https://doi.org/10.1039/C3RP00111C

Shea, N. A., & Duncan, R. G. (2013). From theory to data: The process of refining learning progressions. *The Journal of the Learning Sciences*, *22*, 7–32. https://doi.org/10.1080/10508406.2012.691924

Sikorski, T. R., & Hammer, D. (2017). Looking for coherence in science curriculum. *Science Education*, *101*, 929–943. https://doi.org/10.1002/sce.21299

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, *26*, 114–145. https://doi.org/10.2307/749205

Smith, C. L., Carey, S., & Wiser, M. (1985). On differentiation: A case study of the development of the concept of size, weight, and density. *Cognition*, *21*, 177–237. https://doi.org/10.1016/0010-0277(85)90025-3

Smith, C. L., Wiser, M., Anderson, C. W., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic-molecular theory. *Measurement: Interdisciplinary Research and Perspectives*, *4*, 1–98. https://doi.org/10.1080/15366367.2006.9678570

Songer, N. B., Kelcey, B., & Gotwals, A. W. (2009). How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about biodiversity. *Journal of Research in Science Teaching*, *46*, 610–631. https://doi.org/10.1002/tea.20313

Steedle, J. T., & Shavelson, R. J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching*, *46*, 699–715. https://doi.org/10.1002/tea.20308

Stevens, S. Y., Delgado, C., & Krajcik, J. S. (2010). Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal of Research in Science Teaching*, *47*, 687–715. https://doi.org/10.1002/tea.20324

Talanquer, V. (2009). On cognitive constraints and learning progressions: The case of "structure of matter". *International Journal of Science Education*, *31*, 2123–2136. https://doi.org/10.1080/09500690802578025

Testa, I., Galano, S., Leccia, S., & Puddu, E. (2015). Development and validation of a learning progression for change of seasons, solar and lunar eclipses, and moon phases. *Physical Review Special Topics-Physics Education Research*, *11*(2. 020102-1–020102-22. https://doi.org/10.1103/PhysRevSTPER.11.020102

Testa, J. (2018, June). Journal selection process. Retrieved from https://clarivate.com/essays/journal-selection-process/.
Todd, A., & Kenyon, L. (2016). Empirical refinements of a molecular genetics learning progression: The molecular constructs. *Journal of Research in Science Teaching*, *53*, 1385–1418. https://doi.org/10.1002/tea.21262
Todd, A., & Romine, W. L. (2016). Validation of the learning progression-based assessment of modern genetics in a college context. *International Journal of Science Education*, *38*, 1673–1698. https://doi.org/10.1080/09500693.2016.1212425
Todd, A., & Romine, W. L. (2017). Empirical validation of a modern genetics progression web for college biology students. *International Journal of Science Education*, *39*, 488–505. https://doi.org/10.1080/09500693.2017.1296207
Todd, A., Romine, W. L., & Whitt, K. C. (2017). Development and validation of learning progression-based assessment of modern genetics in a high school context. *Science Education*, *101*, 32–65. https://doi.org/10.1002/sce.21252
von Aufschnaiter, C., & Alonzo, A. C. (2018). Foundations of formative assessment: Introducing a learning progression to guide preservice physics teachers' video-based interpretation of student thinking. *Applied Measurement in Education*, *31*, 113–127. https://doi.org/10.1080/08957347.2017.1408629
Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, *4*, 45–69. https://doi.org/10.1016/0959-4752(94)90018-3
Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 3–34). New York, NY: Routledge.
West, P., Rutstein, D. W., Mislevy, R. J., Liu, J., Levy, R., DiCerbo, K. E., & Behrens, J. T. (2012). A Bayesian network approach to modeling learning progressions. In A. Alonzo, & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 257–292). Rotterdam, The Netherlands: Sense Publishers. https://doi.org/10.1007/978-94-6091-824-7_12.
Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching*, *46*, 716–730. https://doi.org/10.1002/tea.20318
Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, *13*, 181–208. https://doi.org/10.1207/S15324818AME1302_4
Wiser, M., & Carey, S. (1983). When heat and temperature were one. In D. Gentner, & A. L. Stevens (Eds.), *Mental models* (pp. 267–297). Hillsdale, NJ: Lawrence Erlbaum Associates.
Wyner, Y., & Doherty, J. H. (2017). Developing a learning progression for three-dimensional learning of the patterns of evolution. *Science Education*, *101*, 787–817. https://doi.org/10.1002/sce.21289
Yao, J.-X., & Guo, Y.-Y. (2017). Validity evidence for a learning progression of scientific explanation. *Journal of Research in Science Teaching*, *55*, 299–317. https://doi.org/10.1002/tea.21420
Yao, J.-X., Guo, Y.-Y., & Neumann, K. (2017). Refining a learning progression of energy. *International Journal of Science Education*, *39*, 2361–2381. https://doi.org/10.1080/09500693.2017.1381356
Zhai, X. M., Li, M., & Guo, Y. Y. (2018). Teachers' use of learning progression-based formative assessment to inform teachers' instructional adjustment: A case study of two physics teachers' instruction. *International Journal of Science Education*, *40*, 1832–1856. https://doi.org/10.1080/09500693.2018.1512772

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Jin H, Mikeska JN, Hokayem H, Mavronikolas E. Toward coherence in curriculum, instruction, and assessment: A review of learning progression literature. *Science Education*. 2019;103: 1206–1234. https://doi.org/10.1002/sce.21525