Teaching and Learning About Complex Systems in K–12 Science Education: A Review of Empirical Studies 1995–2015

Susan A. Yoon
University of Pennsylvania

Sao-Ee Goh
Ministry of Education, Singapore

Miyoung Park
University of Pennsylvania

The study of complex systems has been highlighted in recent science education policy in the United States and has been the subject of important real-world scientific investigation. Because of this, research on complex systems in K–12 science education has shown a marked increase over the past two decades. In this systematic review, we analyzed 75 empirical studies to determine whether the research (a) collectively represents the goals of educational policy and real-world science, (b) has considered a variety of settings and populations, and (c) has demonstrated systematic investigation of interventions with a view to scale. Results revealed needs in five areas of research: a need to diversify the knowledge domains within which research is conducted, more research on learning about system states, agreement on the essential features of complex systems content, greater focus on contextual factors that support learning including teacher learning, and a need for more comparative research.

KEYWORDS: complex systems, science education, student and teacher learning

In *Science for All Americans* (American Association for the Advancement of Science, 1989), and the subsequent *Benchmarks for Scientific Literacy* (American Association for the Advancement of Science, 1993), the topic of systems occupied a prominent role as one of the four common themes driving science learning in U.S. K–12 education. Since then, science education researchers have focused on constructing and testing methods for teaching and learning about systems, with an emphasis on complex systems (Hmelo-Silver & Azevedo, 2006; Sabelli, 2006; Wilensky & Jacobson, 2015; Yoon, in press). Although this

research has resulted in a number of new digital tools, curricula, and theoretical frameworks, it has not generated consensus on how best to support student learning. With the recent launch of the *Next Generation Science Standards* (NGSS Lead States, 2013), there is a similar emphasis on systems learning, except that the processes fueling systems, such as energy and matter, are also included as important crosscutting concepts. Such concepts are understood to apply to all the content domains encompassed in the NGSS as core ideas in the Physical, Life, and Earth and Space Sciences as well as Engineering, Technology, and Applications of Science.

As we work to translate the NGSS into curricular experiences for all students, understanding the kind of empirical research that has been done on complex systems in the past would be helpful to determine how best to support classroom learning. Given that systems and their processes are crosscutting concepts, it would be prudent to understand how robust the literature base is in terms of the content domains represented as well as the utility of that content. Similarly, understanding how diverse the research base is with respect to populations it has studied can help us understand whether learning outcomes can be attributed to a majority of students. Finally, for educational resources to be adopted at scale, the reliability of methods should be examined in the research that purports to support improved learning.

We begin this article by defining concepts in the study of complex systems that are central to understanding how they exist and function. The section that follows outlines NGSS teaching and learning goals and recent educational policies that identify systematic frameworks that should be considered when evaluating curricula, resources, and practices that have the potential to be scaled up and that represent real-world scientific inquiry. Next, we describe our methods for narrowing down and reviewing empirical studies of complex systems in science education (CSSE). A review of the empirical studies follows. We conclude with recommendations for classroom learning and directions for future research.

What Are Complex Systems?

The study of complex systems can be summarized as understanding how behavior of phenomena at different scales is related and how larger scale patterns emerge from the interdependent components at lower scales (Bar-Yam, 2016). By studying the patterns that emerge and the interactional processes that lead to these patterns, researchers can better understand how systems adapt, self-organize, fluctuate, and reach and maintain equilibrium. Scientists and engineers have for a long time been interested in investigating structures, dynamics, and states of complex systems that can be applied to real-world problems. For example, in a recent conference, the National Academies amassed a multidisciplinary group of leading American complexity researchers to examine nine pressing global issues. These included ecological robustness (whether the biosphere is sustainable), enhancing robustness through interconnectivity (e.g., power grid structures, disaster relief networks), and how to exert control on the spread of disease (The National Academies, 2009). The goal of this work was to identify the limits, optimal states, and weaknesses within systems such that interventions could be applied to enhance stability in the face of perturbations. Solving these issues, as many of the researchers admitted, is challenging given the nonlinear, stochastic or random, and decentralized nature of complex

systems. Here, and in complexity organizations such as the Santa Fe Institute and the New England Complex Systems Institute, researchers are interested in developing visual and mathematical models (e.g., Servedio et al., 2014) that can help describe and predict various systems states as well as the dynamics that fuel them. Optimization, resilience, and robustness are conditions that complexity researchers study and model for utilitarian purposes (e.g., Allison & Martiny, 2008; Caschera et al., 2011; Levin & Lubchenko, 2008).

In terms of content, complex systems researchers represent a multitude of scientific knowledge domains, including physics (e.g., Fuentes, 2014; Sherrington, 2010), chemistry (e.g., Hordijk, Kauffman, & Steel, 2011), biology (e.g., Sah, Singh, Clauset, & Bansal, 2014), and ecology (e.g., Bettencourt & Brelsford, 2014; Sole, 2015). Complex systems researchers also recognize the universal nature and intersectionality of systems within different domains and seek to highlight this universality and reveal underlying principles that unite them (West, 2014). Although a close cousin, this review does not include studies in the field of systems dynamics that focus on macro-level systems functions (typically of engineered systems) and generally aim to quantitatively track the rate of flow of resources to optimize whole system efficiencies (e.g., Forrester, 1994). Similarly, it does not include studies of dynamic systems, which hail from the field of developmental psychology and focus mainly on identifying the dynamics occurring within and between individual actors that support or constrain mental and behavioral growth (e.g., Lewis, 2000; Thelan & Smith, 1994).

NGSS Teaching and Learning Goals

Emphasizing the utility of science learning and how scientists come to understand and apply their research is a central focus of the *Framework for K12 Science Education* (National Research Council [NRC], 2012) that underpins the NGSS. The vision states,

> The learning experiences provided for students should engage them with fundamental questions about the world and with how scientists have investigated and found answers to those questions.
>
> We anticipate that the insights gained and interests provoked from studying and engaging in the practices of science and engineering during their K12 schooling should help students see how science and engineering are instrumental in addressing major challenges that confront society today, such as generating sufficient energy, preventing and treating diseases, maintaining supplies of clean water and food, and solving the problems of global environmental change. (p. 9)

We can see here that the NGSS vision for science education maps well onto the research that complexity scientists are undertaking in terms of solving problems confronting society and the environment.

With respect to specific content involving systems in the NGSS, arguably all crosscutting concepts bear some relationship to the study of complex systems. To review just three—*concepts of scale, proportion, and quantity* relate to the importance of recognizing what is relevant at different measures of size,

time, and energy and how changes in scale, proportion, or quantity affect a system's structure or performance. For example, to understand the complexity of a particulate system, it is important to recognize that at the component level, particles in a gas system move in straight lines, collide with one another, and alter their speed and direction of motion. Changes in these behaviors such as the number of particles, the energy of particles, and the frequency of collisions can affect the system's outcomes such as energy transfer, random motion of particles, and diffusion.

Concepts of systems and system models include examining the system in detail to determine system interdependencies and recognizing the fact that properties and behaviors that exist at a system level may be different from the lower level interactions from which they emerged. In addition, these system level properties and behaviors are often difficult to predict from knowledge about the components and their interactions (NRC, 2012, p. 92). Concepts in this category illustrate core complex systems concepts of interdependence, emergence, and self-organization, among others (cf. Jacobson, 2001).

Concepts of stability and change emphasize static and dynamic equilibrium over various time scales, and feedback loops. To understand complex systems, one must be able to recognize conditions in which aspects of a system are changing or not, and these can be dependent on the time span during which the observations are made (Booth-Sweeney & Sterman, 2007). The mechanisms of feedback loops in a system also serve to regulate processes and maintain or destabilize the system (Gotwals & Songer, 2010).

Another important challenge articulated in the framework is the need for an inclusive science education that offers equal opportunities for all students to experience high-quality learning experiences. Other policy documents have similarly highlighted goals for a more authentic and current science curriculum in K–12 and instructional practices that reach a diversity of learners in diverse settings (National Science Board, 2010; NRC, 2011; President's Council of Advisors on Science and Technology, 2010). Addressing such details in classroom teaching and learning is not an easy task, however, particularly because the NGSS represents a new vision for science education—one that differs from typical learning activities in U.S. classrooms (Wilson, 2013).

What we aim to determine in this review is whether complex systems research in education similarly reflects the research aims of real-world scientific investigation and the NGSS teaching and learning goals.

Scaling Research and Innovations

The development of classroom resources that can be adopted at the scale intended for NGSS should go through a number of steps and evaluations. Recently, the Institute of Education Sciences (IES) and the National Science Foundation (NSF) in the United States outlined a continuum of methodological levels that influence increased reliability of intervention outcomes at scale (IES/NSF, 2013). The continuum delineates six types of research beginning with foundational and early-stage or exploratory research working with single and ideal populations and ending with scale-up research in which findings can generalize to multiple

populations with attention to situated mediators (a description of the stages of research appears in a later table). Determining how the corpus of educational research in complex systems fares against these categories of research and development can offer information about the reliability of research findings that are intended to support student learning.

Other recent frameworks aimed at producing reliable outcomes that can scale to many populations and educational contexts emphasize a need to consider situated variables such as population and context. Design-based implementation research focuses on designing and testing interventions across different levels and settings of learning (Penuel & Fishman, 2012). This framework suggests that there should be a commitment to an iterative design cycle, attention to particular learner responses, and the development of infrastructures that sustain change. The Carnegie framework for improvement science similarly focuses on building interventions that scale via evidence collected from a variety of educational situations that address variation in performance and system variables that influence it (Bryk, Gomez, & Grunow, 2010; Bryk, Gomez, Grunow, & LeMahieu, 2015). Of central importance to our review here is whether complex systems research in education has tested interventions with consideration to context diversity.

Purpose of the Review

The purpose of this review is to examine the research conducted on CSSE from 1995 to 2015. The year 1995 was selected as the starting point because the bulk of interventions and scholarly work emerged after the publication of the *Benchmarks for Scientific Literacy* (American Association for the Advancement of Science, 1993) where, as previously noted, the concept of systems was prominently featured. We also assumed a need to allow a 2-year period for interventions to be constructed, implemented, and reported.

We are interested in investigating how the corpus of CSSE research addresses real-world complex systems research and applications, fulfils NGSS intended standards and goals, and can be scaled to support all students. The research questions underpinning the review are as follows:

- To what extent does CSSE research represent the state of complex systems research and the NGSS standards and goals?
- To what extent has CSSE research been conducted in a variety of educational settings with a variety of populations?
- To what extent has CSSE research demonstrated systematic investigation in terms of methods toward scaling interventions?

We highlight themes that emerge from this group of studies that illustrate areas of robust research emphasis such as particular content domains, conceptual foci, or methodologies. We also highlight what the research says about concepts that are more or less challenging for students to learn and concepts that the field should investigate in terms of what students know and how to support learning. It is important to note, however, that a meta-analysis of student learning outcomes is outside of the scope of this review.

Method

Literature Search Procedure

This systematic review focused on peer-reviewed empirical articles in the field of science education. Articles for review were selected through the following method. First, we defined the keywords needed to search for complex systems articles. These keywords are: (a) *complex systems*; (b) *system, science education*; and (3) *complexity*. Using these keywords, we performed three different searches with each of the keywords in three different databases: Education Resources Information Center (ERIC), Education Full Text, and PsycINFO. We selected these databases because they are the most commonly used ones for educational studies and previous published review articles with similar content have used these databases (e.g., Gerard, Varma, Corliss, & Linn, 2011). Table 1 shows that from this first-level search, there was a large number of hits from each of the databases. Using the inclusion criteria discussed below, one of the authors studied the titles of the articles as well as skimmed the abstracts, over a 3-month period, to yield the numbers extracted in the last column of the table. This number equaled a total of 234 articles. When the research team examined the results of the first level of extraction, we were concerned that with such large numbers, human error may have caused some articles to be missed that could be germane to the study.

To ensure that all empirical educational articles were selected, we performed another level of systematic review with a finer grained search on ERIC. We chose ERIC as the single database to perform this search because it yielded the largest number of hits in the first level. In this second level, we first searched for the keywords of system AND scien*. We then included another level with the following keywords: (a) *complex**, (b) *decentral**, (c) *emergen**, (d) *causal**, (e) *nonlinear**, and (f) *self-organi**. An asterisk (*) was used in order to include all plural versions of each of the keywords in the searches. The first-level keywords were used in every search, and one keyword from the second level was used for each individual search. For example, the first search was "system AND scien* AND complex*," and the next search was "system AND scien* AND decentral*," and so on. For the second-level search, both the title and abstracts were examined for inclusion criteria, which yielded 130 articles for a total combined first- and second-level search of 364 articles. See Table 1 for the detailed search string results.

Inclusion Criteria

For the literature that resulted from each search, inclusion criteria were also used to narrow the pool of relevant research. The inclusion criteria included studies that (a) encompassed K–12 science education and (b) were focused on empirical studies on teaching and learning about complex systems. For those hits for which the inclusion criteria fit was dubious, the article was examined by one or both of the other authors. The resulting extracted articles were also examined by the other two authors to verify inclusion in the final data set. After checking for duplicates across the two levels of extraction, we identified 203 different studies.

Exclusion Criteria

On more detailed examination of the 203 empirical studies, we recognized that there were some studies that did not focus on the teaching and learning of complexity

TABLE 1

Articles extracted: database, search strings, number of hits, and number extracted after inclusion criteria applied

Level	Database	Search string	No. of hits	No. extracted
Level 1	ERIC	complex systems	247	48
		system, science education	2,273	37
		complexity	7157	20
	Education Full Text	complex systems	255	31
		system, science education	1,288	23
		complexity	5,419	14
	PsycINFO	complex systems	1,901	28
		system, science education	3,097	24
		complexity	3,190	9
		Total		234
Level 2	ERIC	system AND scien* AND complex*	787	106
		system AND scien* AND decentral*	32	2
		system AND scien* AND emergen*	210	10
		system AND scien* AND causal*	79	6
		system AND scien* AND nonlinear*	50	3
		system AND scien* AND self-organi*	39	3
		Total		130
		Total extracted		364

in systems. For example, a number of studies investigated the use of computational tools in learning science (e.g., Azevedo, Cromley, & Winters, 2005; Azevedo, Moos, Greene, Winters, & Cromley, 2008; Azevedo, Winters, & Moos, 2004; Ioannidou et al., 2010; Parnafes, 2010). Other studies focused on instructional strategies such as collaborative learning and self-regulated learning to learn science (e.g., Brady, Holbert, Soylu, Novak, & Wilensky, 2015; Greene, Moos, Azevedo, & Winters, 2008; Randler & Bogner, 2009). Although the topics of these studies were related to systems, for example, exploring differences in student use of self-regulatory strategies when learning about circulatory systems (Greene et al., 2008), since they did not examine the learning and teaching of the complexity associated with the system, they were taken out of our systematic review. Subsequently, we were left with a total of 75 studies (see Table 2 for a list of the studies).

A distribution of the 75 CSSE research studies aggregated in 5-year intervals between 1995 and 2015 is shown in Figure 1. The graph shows an increasing trend in which the publication rate roughly doubled every 5 years. These studies were most commonly featured in *Journal of Research in Science Teaching* (10), *International Journal of Science Education* (9), *Journal of Science Education and Technology* (9), and *Journal of the Learning Sciences* (6).

TABLE 2

Seventy-five CSSE research studies between 1995–2015, included in review

Publication	Subject content	Complex systems characteristics	Target group	Gender	Research setting	School type	Sample size	Research method	Research aim	Educational purpose: dimensions of complex systems in science education
Assaraf and Orion (2005)	Earth science	2, 6, 9	Middle school students	NA	Classroom	NA	50	Mixed methods, nonexperimental	DD and FEE	Student understanding; Student learning
Assaraf and Orion (2010)	Earth science	2, 6, 9	Elementary students	NA	Classroom	NA	40	Mixed methods, nonexperimental	DD	Student learning
Assaraf and Orpaz (2010)	Earth science	3, 9, 10	Middle school students	NA	Classroom	NA	21	Mixed methods, nonexperimental	DD	Student learning
Basu, Sengupta, and Biswas (2015)	Ecology	1, 6, 11	Middle school students	NA	Laboratory	Mainstream public	20	Mixed methods, nonexperimental	FEE	Student understanding; Student learning; Conceptual framework or instrument development
Ben-Zvi-Assaraf, and Orion (2010)	Earth science	1, 6, 9	Middle school students	50% female	Classroom	NA	4	Qualitative, case study	FEE	Student understanding; Conceptual framework or instrument development
Calderon-Canales, Flores-Camacho, and Gallegos-Cazares (2013)	Astronomy	1, 3	Elementary students	46% female	Laboratory	Mainstream public	39	Qualitative, case study	FEE	Student understanding
Carlsson (2002)	Ecology	6, 9	Preservice teachers	NA	Laboratory	NA	10	Qualitative, phenomenology	FEE	Teacher understanding; Conceptual framework or instrument development
Cheng and Gilbert (2015)	Biology	2, 6	High school students	33% female	Laboratory	Mainstream public	3	Qualitative, case study	FEE	Student understanding; Conceptual framework or instrument development
Chi, Roscoe, Slotta, Roy, and Chase (2012)	Chemistry	8, 12, 13	Middle and high school students	NA	Laboratory	Private	41	Mixed methods, quasi-experimental	DD	Student learning
Danish (2014)	Biology, ecology	1, 13	Elementary students	48% female	Classroom	NA	42	Qualitative, nonexperimental	DD	Student learning

TABLE 2 (continued)

Publication	Subject content	Complex systems characteristics	Target group	Gender	Research setting	School type	Sample size	Research method	Research aim	Educational purpose: dimensions of complex systems in science education
Danish, Peppler, Phelps, and Washington (2011)	Ecology	1	Elementary students	NA	Classroom	Mainstream public	40	Mixed methods, nonexperimental	FEE	Student learning; Conceptual framework or instrument development
DeBoer et al. (2014)	Ecology	1, 3, 6	Middle school students	50% female	Classroom	NA	1566	Quantitative, quasi-experimental	FEE	Student understanding; Conceptual framework or instrument development
DeLeo, Weidenhammer, and Wecht (2012)	Physics	8	High school students	NA	Classroom	NA	54	Quantitative, nonexperimental	DD and FEE	Student learning
Eilam (2002)	Ecology	1	High school students	51% female	Classroom	NA	112	Mixed methods, quasi-experimental	DD	Student learning
Eilam (2012)	Ecology	1, 4, 9, 11, 13	High school students	NA	Classroom	NA	50	Mixed methods, nonexperimental	DD	Student learning
Evagorou, Korfiatis, Nicolaou, and Constantinou (2009)	Ecology, complex systems	1, 3	Elementary students	31% female	Informal	NA	13	Qualitative, case study	DD	Student learning
Gazit, Yair, and Chen (2005)	Astronomy	2, 3	High school students	44% female	Laboratory	NA	9	Mixed methods, case study	DD and FEE	Student learning
Ginns, Norton, and Mcrobbie (2005)	Computer science	3, 5, 9	Elementary students	NA	Classroom	NA	30	Qualitative, case studies	FEE	Student understanding
Gotwals and Songer (2010)	Ecology	6, 10	Elementary students	NA	Classroom	Mainstream public	318	Mixed methods, nonexperimental	FEE	Student learning; Conceptual framework or instrument development
Grotzer and Bell Basca (2003)	Ecology	9	Elementary students	NA	Classroom	NA	60	Mixed methods, quasi-experimental	DD	Student learning
Grotzer et al. (2015)	Ecology	9	Elementary students	NA	Classroom	NA	38	Mixed methods, DBR	DD	Student learning

(continued)

TABLE 2 (continued)

Publication	Subject content	Complex systems characteristics	Target group	Gender	Research setting	School type	Sample size	Research method	Research aim	Educational purpose: dimensions of complex systems in science education
Gunckel, Covitt, Salinas, and Anderson (2012)	Ecology	1, 6	Elementary, middle and high school students and teachers	NA	Laboratory	NA	258	Mixed methods, nonexperimental	FEE	Student understanding; Conceptual framework or instrument development
Hmelo, Holton, and Kolodner (2000)	Biology	2, 6	Elementary students	NA	Classroom	NA	42	Mixed methods, quasi-experimental	DD	Student understanding; Student learning
Hmelo-Silver, Liu, Gray, Jordan (2015)	Ecology	2, 3, 6, 11	Middle school students	48%–50% female	Classroom	Mainstream public	145	Mixed methods, nonexperimental case study	DD	Student learning; relationship
Hmelo-Silver, Marathe, and Liu (2007)	Biology, Ecology	2, 6, 11	Middle school students and pre-service teachers	NA	Laboratory	Mainstream public	110	Mixed methods, quasi-experimental	FEE	Student understanding; Teacher understanding
Hmelo-Silver and Pfeffer (2004)	Biology, Ecology	1, 6	Middle school students and preservice teachers	NA	Laboratory	Mainstream public	30	Mixed methods, quasi-experimental	FEE	Student understanding; Teacher understanding
Hogan (2000)	Ecology	3, 9, 10	Elementary students	48% female	Classroom	NA	52	Mixed methods, nonexperimental	FEE	Student understanding; Student learning
Hogan and Thomas (2001)	Ecology	5, 6, 9	High school students	30% female	Laboratory	Mainstream public and private	10	Qualitative, case study	DD	Student learning
Ioannidou, Paraskevopoulos, and Tzionas (2006)	Ecology	3, 9	Elementary students	NA	Laboratory	Not in a school setting	10	Qualitative, case study	FEE	Conceptual framework or instrument development

(continued)

TABLE 2 (continued)

Publication	Subject content	Complex systems characteristics	Target group	Gender	Research setting	School type	Sample size	Research method	Research aim	Educational purpose: dimensions of complex systems in science education
Ioannidou et al. (2010)	Biology	3, 6	High school students and teachers	NA	Classroom	Mainstream public	325	Mixed methods, quasi-experimental	DD	Student learning; Relationship between teaching and learning
Jordan, Brooks, Hmelo-Silver, Eberbach, and Sinha (2014)	Ecology	11	Middle school students	NA	Classroom	Mainstream public	35	Mixed methods, nonexperimental	DD	Student learning
Jordan, Gray, Brooks, Honwad, and Hmelo-Silver (2013)	Ecology	2, 6	Middle school students	NA	Classroom	NA	120	Quantitative, quasi-experimental	DD	Student learning
Jordan, Gray, Demeter, Lui, and Hmelo-Silver (2009)	Ecology	3, 4, 9	Middle school students	NA	Classroom	Mainstream public	45	Mixed methods, nonexperimental	FEE	Student understanding
Keynan, Ben-Zvi Assaraf, and Goldman (2014)	Ecology	4, 9	Middle school students	NA	Informal	NA	20	Qualitative, phenomenology	FEE	Student learning; Conceptual framework or instrument development
Klopfer, Scheintaub, Huang, Wendel, and Roque (2009)	Computer science, physics	5, 6, 13	Middle and high school students	NA	Classroom	Charter and private	47	Qualitative, case study	FEE	Student learning
Klopfer, Yoon, and Perry (2005)	Biology	8	Middle, high school, and pre-service teachers	NA	Informal	Mainstream public and Charter	46	Qualitative, case study	DD	Teacher learning
Klopfer, Yoon, and Um (2005)	Computer science, complex systems	3, 8, 13	Elementary and middle school students	NA	Classroom	Mainstream public	80	Mixed methods, nonexperimental	FEE	Student learning
Levy and Wilensky (2008)	Ecology, complex systems	1, 3, 8	Elementary students	50% female	Classroom	Mainstream public	10	Mixed methods	FEE	Student understanding; Conceptual framework or instrument development

(continued)

295

TABLE 2 (continued)

Publication	Subject content	Complex systems characteristics	Target group	Gender	Research setting	School type	Sample size	Research method	Research aim	Educational purpose: dimensions of complex systems in science education
Levy and Wilensky (2009)	Chemistry	1, 3	High school students	51.1% female	Classroom	NA	904	Qualitative, nonexperimental	DD	Student learning
Levy and Wilensky (2011)	Chemistry, mathematics	1, 3	High school students	51% female	Classroom	NA	81–368	Quantitative, nonexperimental	FEE	Student understanding
Lin and Hu (2003)	Biology	1, 6, 9	Middle school students	NA	Classroom	NA	106	Quantitative, nonexperimental	FEE	Student understanding
Liu and Hmelo-Silver (2009)	Biology	1, 2, 4, 6	Middle school students and preservice teachers	68–71% female	Laboratory	Mainstream public	123	Mixed methods, experimental	DD	Student learning; Teacher learning
McNeal et al. (2014)	Earth science	3, 6	High school students	54% female	Classroom	Mainstream public	205	Mixed methods, DBR	DD	Student learning
Mohan, Chen, and Anderson (2009)	Biology, ecology	1, 3, 8	Elementary, middle and high school students	NA	Laboratory	Mainstream public and NA (Korea)	314	Mixed methods, DBR	FEE	Conceptual framework or instrument development
Pallant and Lee (2015)	Earth science	3, 6, 9	Middle and high school students	50% female	Classroom	NA	512	Mixed methods, nonexperimental	DD	Student learning
Pallant, Lee, and Pryputniewicz (2012)	Earth science, ecology	3, 9	Middle and high school students	NA	Classroom	NA	260	Mixed methods, nonexperimental	DD	Student learning
Penner (2000)	Biology, complex systems	1, 8	Elementary students	50% female	Informal	NA	4	Qualitative, case study	FEE	Student understanding
Peppler, Danish, and Phelps (2013)	Ecology	1, 8, 9, 12, 13	Elementary students	50% female	Classroom	Mainstream public	40	Mixed methods, experimental	FEE	Student learning
Perkins and Grotzer (2005)	Physics	3, 4, 7, 8, 13	Elementary and middle school students	NA	Classroom	NA	163	Quantitative, quasi-experimental	DD	Student learning

TABLE 2 (continued)

Publication	Subject content	Complex systems characteristics	Target group	Gender	Research setting	School type	Sample size	Research method	Research aim	Educational purpose: dimensions of complex systems in science education
Plate (2010)	Ecology, complex systems	4, 9	Middle school students	51% female	Classroom	Mainstream public	147	Quantitative, quasi-experimental	FEE	Student understanding; Student learning; Conceptual framework or instrument development
Puk and Stibbards (2011)	Ecology	7, 8	Preservice teachers	65% female	Classroom	NA	20	Qualitative, case study	FEE	Teacher learning
Reiss and Tunnicliffe (2001)	Biology	2, 6	Elementary, middle and high school students and preservice teachers	61% female	Classroom	Mainstream public and private	158	Mixed methods, nonexperimental	FEE	Student understanding; Conceptual framework or instrument development
Repenning, Ioannidou, Luhn, Daetwyler, and Repenning (2010)	Biology	3, 6, 12	High school students	NA	Laboratory	NA	42	Mixed methods, experimental	DD	Student learning
Resnick (1996)	Ecology, complex systems	12	High school students	33% female	Laboratory	Mainstream public and not in a school setting	12	Qualitative, case study	FEE	Student understanding; student learning; Conceptual framework or instrument development
Richardson and Hari (2009)	Ecology	2, 3, 5, 8	Middle school students	NA	Classroom	NA	82	Mixed methods, nonexperimental	FEE	Student learning
Riess and Mischo (2010)	Biology	6, 8	Elementary students	46% female	Classroom	Mainstream public	424	Quantitative, quasi-experimental	DD	Student learning
Rodriguez, Kohen, and Delval (2015)	Ecology	1, 8	Elementary, middle and high school students	46.3% female	Laboratory	Mainstream public	80	Qualitative	FEE	Student understanding

(continued)

297

TABLE 2 (continued)

Publication	Subject content	Complex systems characteristics	Target group	Gender	Research setting	School type	Sample size	Research method	Research aim	Educational purpose: dimensions of complex systems in science education
Songer, Kelcey, and Gotwals (2009)	Biology	4	Elementary students	29% female	Classroom	Mainstream public	1885	Quantitative, quasi-experimental	DD and FEE	Student understanding; Conceptual framework or instrument development
Stavrou and Duit (2014)	Biology, physics, complex systems	13	High school students	NA	Laboratory	Mainstream public	30	Qualitative, grounded theory	FEE	Student understanding; Student learning
Stavrou, Duit, and Komorek (2008)	Biology, physics	13	High school students	NA	Laboratory	NA	30	Qualitative, case study	FEE	Student understanding; Student learning
Stroup and Wilensky (2014)	Biology, complex systems	1	Elementary and high school students	NA	Classroom	NA	21	Qualitative, case study	FEE	Student understanding; Student learning
Thompson and Reimann (2010)	Ecology	9	High school students	100% female	Laboratory	NA	18	Mixed methods, experimental	FEE	Student learning
Tsurusaki and Anderson (2010)	Ecology, complex systems	2, 3	Elementary, middle and high school students	NA	Classroom	Mainstream public	723	Mixed methods, quasi-experimental	FEE	Student understanding; Conceptual framework or instrument development
Vachliotis, Salta, and Tzougraki (2014)	Chemistry	1, 6	High school students	49% female	Classroom	Mainstream public	91	Quantitative, nonexperimental	FEE	Student understanding; Conceptual framework or instrument development
van Borkulo, van Joolingen, van Savelsbergh, and de Jong (2012)	Ecology	6, 9	High school students	31% female	Laboratory	NA	74	Quantitative, quasi-experimental	DD	Student learning
Varma and Linn (2012)	Ecology	2, 6	Elementary students	48% female	Classroom	NA	190	Mixed methods, DBR	DD	Student learning
Vattam et al. (2011)	Ecology	1, 2, 3, 6	Middle school students	49% female	Classroom	Mainstream public	157	Mixed methods, nonexperimental	DD	Student learning
Verhoeff, Waarlo, and Boersma (2008)	Biology	1	High school students	49% female	Classroom	NA	28	Qualitative, case study	FEE	Student learning
Wilensky and Reisman (2006)	Ecology	1, 3, 8	High school students	50% female	Laboratory	NA	2	Qualitative, case study	FEE	Student learning

TABLE 2 (continued)

Publication	Subject content	Complex systems characteristics	Target group	Gender	Research setting	School type	Sample size	Research method	Research aim	Educational purpose: dimensions of complex systems in science education
Wilensky and Resnick (1999)	Biology, chemistry	1, 3, 6, 8	High school students and teachers (no grade available)	NA	Laboratory	NA	NA	Qualitative, case study	FEE	Student learning
Wilkerson-Jerde and Wilensky (2015)	Ecology, mathematics	1, 6, 8	High school students	45% female	Laboratory	Mainstream public	11	Qualitative, case study and grounded theory	FEE	Student understanding
Wu (2010)	Earth science	2, 3, 4, 6	Study 1: Scientists and high school students; Study 2: High school students	Study 1: NA; Study 2: 100% female	Study 1: Laboratory; Study 2: Classroom	Study 1: NA; Study 2: Mainstream public	Study 1: 18 Study 2: 29	Mixed methods, nonexperimental DBR	DD and FEE	Student learning; Conceptual framework or instrument development
Yoon (2008)	Biology	3, 4, 6, 8, 9, 12, 13	High school students	73% female	Informal	Mainstream public	11	Mixed methods, nonexperimental	DD and FEE	Student learning
Yoon (2011)	Biology	4, 9	Middle school students	54% female	Classroom	Mainstream public	76	Mixed methods, nonexperimental	DD	Student learning
Yoon, Koehler-Yom, Anderson, Lin, and Klopfer (2015)	Biology, complex systems	2, 6	High school teachers	66% female	Informal and Classroom	NA	3	Qualitative, case study	FEE	Teacher learning; Relationship between teaching and learning; Conceptual framework or instrument development

Note. DD = design and development research; FEE = foundational, early-stage, and exploratory research; DBR = design-based research.
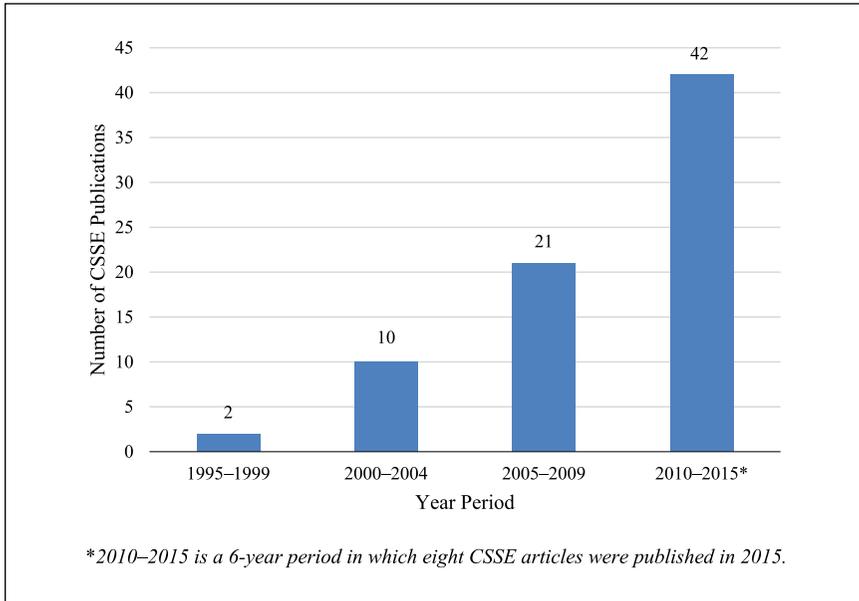
FIGURE 1. Number of CSSE publications over the past 20 years. *CSSE = complex systems in science education.*

Data Analysis

This section describes the analyses performed to address each research question and is therefore divided into three sections. When interrater reliability is reported without extenuating explanatory details, it should be understood that agreement was obtained between the first two authors.

To What Extent Does CSSE Research Represent the State of Complex Systems Research and the NGSS Standards and Goals?

An analysis of the science subjects and the complex systems concepts contained in each study can provide insight into the extent to which CSSE research represents the state of complex systems research in science as well as NGSS standards and goals. Table 3 shows the categorization scheme for science subjects and complex systems concepts.

The CSSE studies focused on the teaching and learning, or understanding, of scientific systems that are complex in nature. These scientific systems are found in myriad knowledge domains. We identified seven common domains or content areas, that is, biology, chemistry, computer science, earth science, ecology, physics/physical science/engineering, and complex systems in general. Some studies were coded for more than one content area due to the interdisciplinary topics they covered. Interrater reliability was completed on 20% (16) of the 75 articles with adequate agreement between the authors (kappa = 0.7; Landis & Koch, 1977).

TABLE 3

Categorization scheme for science content and complex systems concepts

Aspect	Guiding question	Codes and other remarks
Science content	Which content area was the study situated in?	• Biology • Chemistry • Computer science • Earth science • Ecology • Physics/physical science/engineering • Complex systems in general
Complex systems characteristics	What complex systems characteristics were included in the focus of the research?	Structures 1. Levels/aggregates/scale 2. Number and names of variables 3. Connections 4. System organization (linear vs. nonlinear) 5. Inputs/outputs or initial conditions Processes 6. Interdependence/relationships 7. Self-organization 8. Emergence 9. Causality (feedback, cycles) 10. Perturbations States 11. Equilibrium/stability 12. Decentralized 13. Random (order, variations, fluctuations)

Disagreements about the science content of the studies were resolved after discussion.

Constructing a categorization manual to code complex systems characteristics was more challenging due to the different phrasing of processes and structures among the studies. We decided to use a qualitative content analysis approach in which clearly defined categories were developed inductively (Mayring, 2000). The first two authors read and coded 10 studies to identify an initial set of concepts that the authors of the articles themselves used to describe the complex systems characteristics investigated in their respective interventions. The characteristics were then discussed and revised, after which another 10 articles were coded. A preliminary framework of complex systems characteristics was constructed and further elaborated on throughout the process. Studies could be coded as including multiple relevant complex systems characteristics. The two authors then coded the rest of the articles separately and discussed the codes. Any discrepancies in the codes were resolved after several rounds of negotiation.

Finally, a framework was completed comprising three macro categories (i.e., structure, processes, and states) within which 13 complex systems concepts were embedded (see Table 3). These superordinate categories are often referred to in descriptions of differences between complex systems approaches and linear approaches when describing the world. For example, Mitchell (2009) writes that traditionally, science has often adopted a linear and reductionist approach, focusing on "breaking down" these systems to their components for analysis of their behaviors. While this approach simplifies the learning, it does not do justice to the multiple interactions among the components and interrelationships within the system, which give rise to systemic properties and patterns (Capra, 1996). We interpret the notion of 'system components' as referring to system structures. The notions of "multiple interactions" and "interrelationships" refer to processes that fuel systems; giving "rise to systemic properties and patterns" pertains to emergent states.

In our framework, *Structure* characteristics refer to the physical features of the system, including micro and macro levels or different scales, the number and names of variables, how the variables are connected and the number of connections, the system organization (whether it is linear or nonlinear), and inputs and outputs or initial conditions. There are five complex systems characteristics related to the structure category. *Process* characteristics refer to the dynamics and mechanisms that fuel complex systems evolution. These include how the variables in the system are interdependent or form relationships, the processes of self-organization and emergence, the causal nature of processes such as feedback and cycles, and perturbations that trigger shifts in system structures and dynamics. There are also five complex systems characteristics related to processes. *States* characteristics refer to how complex systems exist in the world as a result of shifts or due to existing structures and processes. Three complex systems characteristics of states were identified in the studies.

To What Extent Has CSSE Research Been Conducted in a Variety of Educational Settings With a Variety of Populations?

Analysis of the demographics of research participants in the various studies provided insight into the extent to which CSSE research has been conducted in a variety of educational settings. Three aspects of demographic information commonly provided in peer-reviewed articles were examined. The first was what kind of population participated in the study with respect to students and/or teachers and also the grade range of the students (elementary, middle or high school) and the experience level of the teachers (preservice or in-service). In studies where student ages were given instead, the grade levels were deduced from the given ages. The final two categories included were the gender and the race and ethnicity of the study population. The U.S. basic categorization for races and ethnicities (Office of Management and Budget, 1995) was used to code the race and ethnicity composition in the sample.

While the information was relatively easy to extract with little ambiguity, interrater reliability was still performed on 20% of the articles with very agreement (kappa = 0.85 − 0.9; Landis & Koch, 1977) for all four demographic dimensions. If no information was provided in the article on a particular demographic aspect, it was duly noted. Table 4 describes the details of the coding of the demographic information.

TABLE 4

Categorization scheme for demographics of research participants

Aspect	Guiding question	Codes and other remarks
Target group	What student and/ or teacher group(s) is the study targeting?	Student groups were coded according to the following levels • Elementary (Grades K–6) • Middle school (Grades 7–8) • High school (Grades 9–12) Teacher groups were coded as follows • Preservice (prior to certification) • In-service (after certification)
Gender	What is the gender composition of the study?	• Male • Female
Race and ethnicity	What is the ethnicity of the sample?	The U.S. basic categorization for races and ethnicities was used (Office of Management and Budget, 1995) • American Indian or Alaska Native • Asian or Asian American • Black or African American • Native Hawaiian or Other Pacific Islander • White • Latino/Hispanic Studies that indicated multiple races and ethnicities were coded as "diverse."

To What Extent Has CSSE Research Demonstrated Systematic Investigation in Terms of Methods Toward Scaling Interventions?

Analysis of the research contexts, methodologies, and purposes can provide insight into the extent to which CSSE research has demonstrated systematic investigation in the past 20 years. Specifically, we examined the research setting, school type (if it was conducted in a school), sample size, research methods, and the research and educational purposes of the study.

Research setting refers to where the study was conducted, for example, in a classroom, an informal learning environment (e.g., museum and summer camp), or a laboratory setting.

School type refers to the kind of schools the sample was taken from for studies conducted in a classroom setting. They were coded as public (mainstream), public (charter or special target), or private schools.

Sample size refers to the number of participants involved in the study. If a study included two samples, the larger of the two sample sizes was documented. Such cases are also noted in Table 2.

Research method refers to the sources of information that were used, how the information was sampled, and the types of instruments that were used in data collection. This review largely coded the studies based on the type of data

TABLE 5

Categorization schemes for research context and methodology.

Aspect	Guiding question	Codes and other remarks
Research setting	Is the study conducted in a classroom, informal setting, or laboratory setting?	• Classroom setting: study conducted as part of school curriculum or during school curriculum time • Informal setting: study conducted as part of enrichment programs outside of normal school curriculum or after school hours, for example, summer institutes, museum visits • Laboratory setting: study conducted as an elaborate experimental set-up where there was deliberate control of variables
School type	For studies conducted in classrooms, what type of schools is the sample from?	• Public (mainstream) • Public (charter or special target) • Private • And/or not in school settings
Sample size	What is the sample size used in the study?	If two samples were involved (e.g., different phases of the study), the larger of the two sample sizes was documented. Such cases are noted in the appendix.
Research method	What is/are the research method(s) used in the study?	• Quantitative: only quantitative methods • Qualitative: only qualitative methods • Or mixed methods: both quantitative and qualitative methods used
Research aim	What was the broad aim of the study?	The coding was done according to the three classifications of research by Institute of Education Sciences and National Science Foundation (2013). • *Foundational, Early-Stage, and Exploratory Research:* These studies contribute to core knowledge in education (i.e., basic understandings of teaching and learning). These studies seek to test, develop, or refine theories of teaching or learning and may develop innovations in methodologies and technologies to influence and inform research and development. They may also examine relationships among important constructs in teaching and learning to establish logical connections that form the basis for future interventions or strategies to improve education outcomes.

TABLE 5 (continued)

Aspect	Guiding question	Codes and other remarks
		• *Design and Development Research*: These studies develop solutions to achieve goals or solve problems. These studies draw on existing theory and evidence to design and test interventions or strategies that may include pilot tests of fully developed interventions; testing (i.e., implementation and assessment) of various interventions for comparison.
		• *Efficacy, Effectiveness and Scale-up Research*: These studies contribute to evidence of impact, generating reliable estimates of the ability of a fully developed intervention or strategy to achieve its intended outcomes, which may include testing of strategies or interventions under ideal conditions (efficacy); testing of effectiveness of strategies or interventions under normal circumstances (effectiveness); testing of effectiveness of strategies or interventions in a wide range of populations, contexts, and circumstances (scale-up).
Educational purpose: dimensions of CSSE	What dimension of CSSE did the study investigate?	All studies were categorized into one or more of the following six dimensions of CSSE.
		• *Student understanding*: Study examines students' existing level of complex systems understanding, in the absence of any intervention.
		• *Teacher understanding*: Study examines teachers' existing level of complex systems understanding, in the absence of any intervention.
		• *Student learning*: Study seeks to improve or facilitate student understanding and learning of complex systems. Study typically involves an intervention.
		• *Teacher learning*: Study seeks to improve or facilitate teacher understanding and learning of complex systems. Study typically involves an intervention.
		• *Conceptual framework or instrument development*: Study attempts to develop a conceptual framework or research instrument to measure or interpret complex systems understanding.
		• *Relationship between teaching and learning about complex systems*: Study investigates the relationship between teaching and learning about complex systems.

Note. CSSE = complex systems in science education. 2010–2015 is a 6-year period in which eight CSSE articles were published in 2015.

305

collected—qualitative, quantitative, or mixed methods. Studies coded as qualitative research included case study, ethnography, grounded theory, narrative, and phenomenological research, while studies coded as quantitative included experimental, quasi-experimental, and nonexperimental. Mixed-methods studies were those that include both qualitative and quantitative research methods.

Research aim refers to whether the study was early-stage exploratory research, design and developmental research, or efficacy and scale-up research. This coding was adapted from the IES/NSF (2013) classification of educational research.

Finally, the *educational purpose* refers to the dimension of CSSE investigated in the study. For example, did the study examine student or teacher learning as part of an intervention, or was the research conducted to determine existing levels of complex systems understanding? The six CSSE dimensions were first constructed by one of the authors after reading through half of the articles. The dimensions were subsequently discussed and refined among the three authors. Interrater reliability was also completed on 20% of the articles with respect to this category (that were not reviewed for the construction of the coding scheme) with substantial agreement obtained (kappa = 0.65–0.8; Landis & Koch, 1977).

Further details about the categorization schemes for research contexts, methodologies, and purposes appear in Table 5.

Results

Representations of Complex Systems Research

Science Content Areas

Figure 2 shows the distribution of content areas represented in the group of CSSE studies. Ecology and biology were the most commonly coded content areas, with 38 studies (51%) and 24 studies (32%), respectively. Within these content areas, topics of investigation ranged from environmental issues and ecosystems to human body systems and genetics. The content area of earth science was represented in eight studies (11%) with smaller numbers coded in the remaining content areas. Nineteen studies (25%) were coded in more than one content area.

Complex Systems Characteristics

A total of 185 complex systems characteristics were coded as target learning content in the CSSE studies we reviewed. The distribution of complex systems characteristics coded in the categories and subcategories can be found in Figure 3. Of the total number of complex systems characteristics, 86 (46%) were coded in the *Structures* category. Another 79 (43%) were coded in the *Processes* category, and just 20 (11%) were coded in the *States* category. Among the subcategories, the highest frequency of complex systems concepts appeared in *Interdependence/Relationships*, which was coded in 33 (18%) of studies. The subcategories with the next highest frequencies were *Connections* (coded in 27, or 15%, of studies) and *Levels/Aggregates/Scale* (coded in 25, or 14%, of studies). Relatively low numbers appeared in the subcategories of *Inputs/Outputs or Initial Conditions, Self-Organization, Perturbations, Equilibrium/Stability*, and *Decentralized*. It is

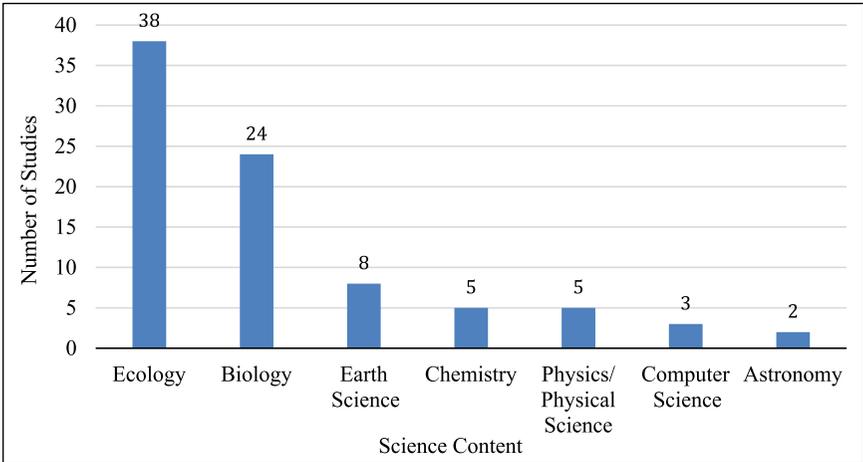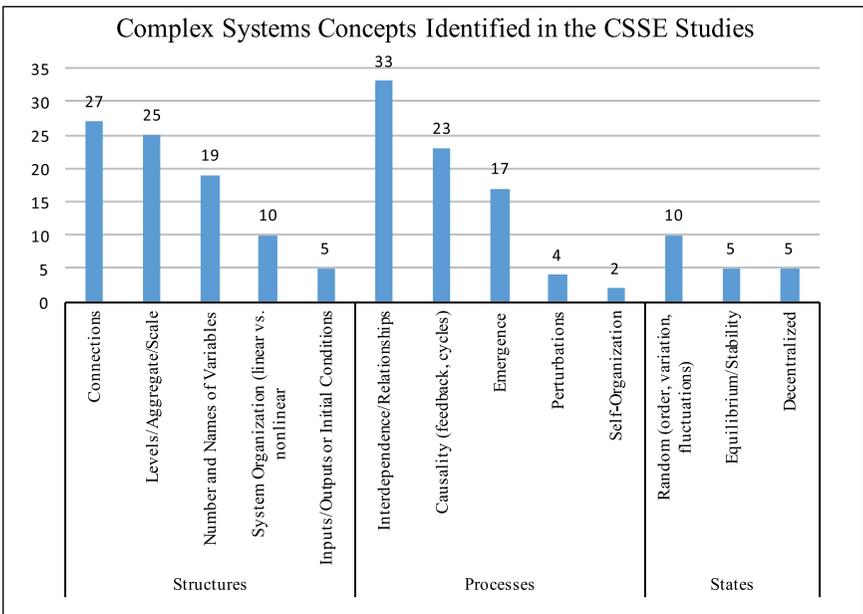FIGURE 2. *Distribution of science content.*

FIGURE 3. *Distribution of complex systems concepts.*

also important to note that some essential complex systems concepts related to real-world scientific investigations were missing in the CSSE studies analyzed. These include the characteristics of robustness and resilience, which would belong in the *States* category.

Educational Settings and Populations

Fifty-nine studies (79%) focused on students as the target groups, with relatively equal distribution across elementary, middle, and high school grades. Twelve studies (16%) included students across elementary, middle, and/or high school grades. In comparison, only 11 studies (15%) focused on teachers as target participants, and 7 studies (9%) included both teachers and students in their samples.

Regarding gender, we were unable to derive or infer the gender composition in 38 (51%) studies as they did not provide sufficient information about their sample. Of the group that did report gender composition, there appeared to be sufficient balance between number of males and females in the study samples.

Similarly, with respect to race and ethnicity, we were unable to distil much information in this category because 59 studies (79%) did not report details about the racial and ethnic composition. While the rest did suggest diverse races and ethnicities in their samples, it was difficult to ascertain the proportion because of the lack of description. Of the 12 studies that provided sufficient information, there were no racial or ethnic groups specifically targeted.

Research Methods

Research Setting and School Type

There were 47 studies (63%) conducted in school settings, 24 studies (32%) conducted in laboratory settings, and 6 studies (8%) conducted in informal settings. One study was conducted in both school and laboratory settings, and one in both school and informal settings. Of the 47 studies conducted in school settings, 27 (52%) did not report sufficient information to distil the types of school the student or teacher samples were from. Of those that did, 19 of the studies (40%) were conducted in mainstream schools while only one study was conducted in a private school.

Sample Size and Methodology

In the category of sample size, forty studies (53%) had sample sizes of 50 or fewer participants. Of this segment, 16 (21%) had sample sizes of 20 or fewer participants. Nine studies (12%) had 51 to 100 participants, and seven studies (9%) had 101 to 150 participants. Two studies were not coded as they did not record their sample sizes (Stroup & Wilensky, 2014; Wilensky & Resnick, 1999).

In the category of methodology, mixed methods were the most common methods used in the CSSE studies. A total of 40 studies (53%) used mixed methods in their investigation, 24 studies (32%) used qualitative methods, and 11 studies (15%) used quantitative methods.

Research Aim and Educational Purpose

In terms of research aim, 41 studies (55%) were coded as Foundational, Early-Stage, and Exploratory (FEE) research, and 28 studies (37%) were coded as Design and Development (DD) research. Six studies (8%) were coded as both FEE and DD. However, no study was coded as Efficacy, Effectiveness, and Scale-up (EES) research.
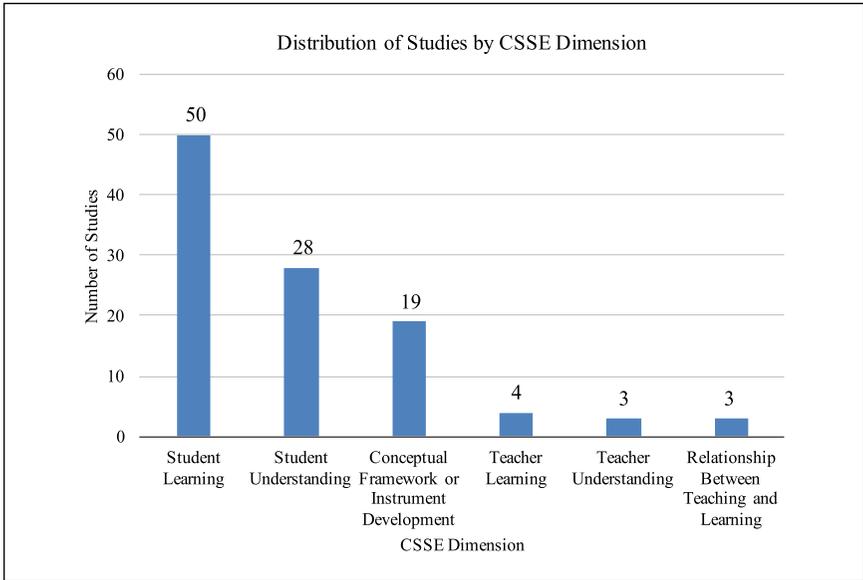
FIGURE 4. Distribution of CSSE dimensions. *CSSE = complex systems in science education.*

For educational purpose, most of the CSSE studies investigated student learning and/or understanding of complex systems. There were 50 studies (67%) that examined how student learning of complex systems can be facilitated or improved, while 28 studies (37%) assessed students' state of understanding. Only seven studies (9%) investigated teacher understanding and/or learning of complex systems. There were 19 studies (25%) in which conceptual frameworks or research instruments were developed to measure or interpret complex systems understanding. Two studies explored the relationship between instructional practice and student learning of complex systems, and another study examined the relationship between teachers' understanding of complex systems and their classroom instruction. Figure 4 shows the distribution of the dimensions of CSSE in the group of studies.

Discussion

In this section, we discuss each of the analyses and highlight areas of strength as well as gaps in the corpus of CSSE research with respect to the research questions. Again, more details about each study can be found in Table 1.

Strong Representation in Biology and Ecology Studies

Complex systems researchers address issues in multiple scientific knowledge domains and search for underlying principles that are universal in nature (West, 2014). We found an abundance of CSSE studies in the domains of biology and ecology. The biological systems in these studies included human body

systems such as the circulatory and respiratory systems (e.g., Cheng & Gilbert, 2015; Hmelo-Silver, Marathe, & Liu, 2007; Ioannidou et al., 2010), cellular systems (Penner, 2000; Verhoeff, Waarlo, & Boersma, 2008), and genetics or genetic engineering (e.g., Klopfer, Yoon, & Perry, 2005; Yoon, 2008). These studies have revealed a number of issues in student reasoning about complex systems. For example, Hmelo-Silver and colleagues have conducted studies with students and teachers to explore differences in how experts and novices understand systems. They found that experts recognize the integrated nature of system structures (components), behaviors (component interactions), and functions (outcomes of interactions) and use the latter two (i.e., behaviors and functions) as deep principles to organize their knowledge of the system (Hmelo-Silver et al., 2007; Hmelo-Silver & Pfeffer, 2004). Novices, on the other hand, only reason about structures and largely ignore behaviors and functions of systems. Verhoeff et al. (2008) found that while students were able to accurately reason about the structures and relationships between levels of cellular organelles, cells, and organs, they had difficulty understanding how the different levels of activity functioned as coherent organ processes. On the topic of genetic engineering, Yoon (2008) found that the process of decentralization was a more challenging concept to grasp than other complex systems concepts. Decentralization is the notion that control of systems (e.g., how genetically modified organisms spread) is often distributed across different components rather than localized in one component.

Studies focused on the domain of ecology, which represent more than half of the total CSSE research in our sample, have likewise highlighted student learning challenges. On the topic of population growth, Wilkerson-Jerde and Wilensky (2015) discussed that students lack inferential reasoning about how individual behaviors generate group-level patterns. They suggested that a greater focus on recognizing mathematical relationships in graphical representations can support learning about exponential growth, which is a typical outcome of complex systems. Identifying essential relationships that fuel complex systems functions has proven to be challenging in various ecological systems. Varma and Linn (2012) found that only a few students in their study understood the relationship between the albedo process and Earth's temperature on the topic of global warming. Other research has investigated what students understand about the impact of human-engineered systems on environmental natural systems. For example, Hogan (2000) found that middle school students tended to reason in a direct fashion about how pollutants affect ecosystems rather than recognizing the importance of indirect effects. Tsurusaki and Anderson (2010) discussed that a sound understanding of the relationship between human activity and ecological systems varies across grade levels.

Need to Diversify Other Knowledge Domains

Collectively, CSSE research in biological and ecological systems has provided a great deal of knowledge about how students reason about complex systems. The plethora of studies have already spawned research in developing learning progressions across grades in topics such as water systems (Gunckel, Covitt, Salinas, & Anderson, 2012), carbon cycling (Mohan, Chen, & Anderson, 2009), and

biodiversity (Songer, Kelcey, & Gotwals, 2009). But our review of the field revealed far fewer studies in the domains of chemistry (Chi, Roscoe, Slotta, Roy, & Chase, 2012; Levy & Wilensky, 2009, 2011; Vachliotis, Salta, & Tzougraki, 2014; Wilensky & Resnick, 1999), physics or physical sciences (DeLeo, Weidenhammer, & Wecht, 2012; Klopfer, Scheintaub, Huang, Wendel, & Roque, 2009; Perkins & Grotzer, 2005; Stavrou & Duit, 2014; Stavrou, Duit, & Komorek, 2008); and astronomy (Calderon-Canales, Flores-Camacho, & Gallegos-Cazares, 2013; Gazit, Yair, & Chen, 2005). Given that an essential goal of K–12 science education is to understand the utility of the scientific research enterprise and recognize how scientists apply their research (NRC, 2012), diversifying the knowledge domains within which students learn about complex systems represents an important area of future CSSE research.

Strong Representation in the Study of Complex Systems Structures and Processes

Our analysis of the complex systems concepts found within the CSSE studies demonstrated relative depth in the areas of systems structures and processes. Within these metalevel categories, a number of the studies attended to learning about connections between systems components, and how they form aggregate levels at different scales. This research is particularly well represented in a series of studies conducted by Levy and Wilensky (2008, 2009, 2011). Using agent-based computational simulations of chemical systems, the authors discuss the importance of developing conceptual knowledge that connects submicrolevel particle behaviors and interactions with emergent structures formed at the macro system-wide level. They further suggest that when students apply reasoning about midlevel structures, such as groups and clusters that can form as intermediate structures between microlevel and macrolevel states, they demonstrate deeper levels of understanding of processes that fuel emergent behaviors and states as well as a deeper level of understanding of the scientific domain.

A preponderance of studies focused on helping students recognize the relationships and interdependencies of system components. Using aquaria, Jordan, Brooks, Hmelo-Silver, Eberbach, and Sinha, (2014) showed that middle school students in their study were more easily able to recognize system interdependencies found in the processes of photosynthesis (plants making food) and limiting factors (e.g., oxygen) but not cellular respiration (converting food to energy) or eutrophication (lack of oxygen). They suggested that a lack of understanding of the reciprocal relationship between photosynthesis and cellular respiration can lead to difficulties in understanding the critical importance of having enough oxygen to maintain healthy aquatic ecosystems. Efforts to enable students to understand these critical relationships and interdependencies in systems have produced learning models and assessment tools that have been examined in multiple lines of research such as the system thinking hierarchical model (e.g., Assaraf & Orion, 2010).

Need for More Research and Interventions on System States

Structures and processes are certainly central features of complex systems and arguably represent a place to start in terms of developing a basic level of understanding. However, there were some gaps in the CSSE literature that warrant

discussion and point to a need to develop a deeper understanding of how systems operate and exist. The relatively low numbers that appeared in the structures and processes categories in the concepts of initial conditions, perturbations, and self-organization (e.g., Assaraf & Orpaz, 2010; Ginns, Norton, & Mcrobbie, 2005; Gotwals & Songer, 2010; Hogan, 2000; Puk & Stibbards, 2011) indicate that few studies examine the critical nature of how complex systems can differ based on initial variables, and how triggers can influence how systems self-organize. Such investigations are important to be able to model phenomena and make predictions, which are two scientific practices at the core of real-world complex systems research and are highlighted the NGSS (NGSS Lead States, 2013). Similarly, the analysis revealed only a few studies that focused on complex systems states such as equilibrium and decentralization (e.g., Basu, Sengupta, & Biswas, 2015; Eilam, 2012; Hmelo-Silver, Liu, Gray, & Jordan, 2015; Peppler, Danish, & Phelps, 2013; Repenning, Ioannidou, Luhn, Daetwyler, & Repenning, 2010; Yoon, 2008) and no CSSE studies that investigated system robustness and resilience. It is clear that more educational interventions are needed that examine system states, which are the focal investigations of scientists concerned with addressing pressing global issues such as biosphere sustainability (The National Academies, 2009).

Strong Representation of Studies About Student Learning

Given the goals of NGSS to provide high-quality K–12 science learning experiences for all students, this review also sought to determine the extent to which the corpus of CSSE studies has conducted research in a variety of educational settings and a variety of populations. The target group analysis showed that research has been conducted on a range of age-groups and grade levels across the K–12 spectrum, although, as we discuss further on, much of this research is non-comparative. Other outcomes of this student learning focus include the importance of using computational tools in instruction, emerging theories about why learning about complex systems is challenging, and theoretical frameworks for assessing complex systems.

A number of studies (e.g., Klopfer, Yoon, & Perry, 2005; Repenning et al., 2015; Vattam et al., 2011; Wilensky & Reisman, 2006; Yoon, Koehler-Yom, Anderson, Lin, & Klopfer, 2015) describe computational tools that have been developed to visualize structures and mechanisms that enable users to view the evolution of systems over time. A particularly robust line of research has been aimed at developing agent-based simulations represented in modeling tools such as NetLogo (Wilensky & Reisman, 2006) and StarLogo (Yoon et al., 2016). These simulations allow students to manipulate and construct facsimiles of scientific systems in which changes in initial conditions, random variation, decentralized interactions, and self-organized emergent behaviors (among other system characteristics) are investigated. Repenning et al. (2015) have developed what they call "collective simulations" that tie together social learning techniques in the classroom of learners with networked computers to engage students themselves in the simulations. Similar earlier efforts to incorporate students as system agents can be found in the concept of participatory simulations (Klopfer, Yoon, & Perry, 2005; Stroup & Wilensky, 2014).

Research on student learning about complex systems has also advanced emerging theories about why learning about complex systems presents challenges. Grotzer and colleagues found that students tend to reason about immediate effects rather than cascading or indirect effects. For example, students fail to realize that a change in one population can have impacts on populations that are not directly linked through domino-like or cyclic complex causal relationships (Grotzer & Bell Basca, 2003; Grotzer et al., 2015). Chi et al. (2012) have hypothesized that learning difficulties about complex causality and nonlinear dynamics may be caused by an inability to recognize the difference between direct schema and emergent schema. Direct schema develop within learners from commonly observed everyday life events. They proceed through linear interactions as simple as needing to shop for food and driving one's car from home to the market. However, emergent schema can be developed only if or when learners realize that complex phenomena exist in a series of nonlinear interactions. For example, ecosystems are weblike structures in which a trigger, like a forest fire in one part of the system, will quickly affect many parts of the forest because they are all interconnected.

CSSE studies of student learning have also generated a few theoretical frameworks that locate learning in different aspects of systems. The Structure–Behavior–Function framework hails from the field of engineering systems design (e.g., Bhatta & Goel, 1997) and is represented in a number of studies (e.g., Danish, 2014; Hmelo-Silver et al., 2007). A systems understanding through the Structure–Behavior–Function lens follows from a hierarchical knowledge of system characteristics. The components or structures (e.g., hybrid or electric motor) and behaviors (e.g., energy consumption) must first be understood in order to work with the system to achieve the desired output or function (e.g., how far a car can travel). Yoon (2008, 2011), based on earlier work by Jacobson (2001), offered a framework encompassing core concepts within complex systems learning that identify "clockwork" versus "complex systems" mental models. A *clockwork orientation* views the world from a Cartesian perspective (Capra, 1982) that generally views the world and its constituents as machines. It is based on a method of analytic thinking that involves breaking up complex phenomena into pieces to understand the behavior of the whole from the properties of its parts. This is in contrast to a *complex systems* view, according to which the essential properties of an organism (or complex system with a constant influx of energy) are properties of the whole—properties that none of the parts have on their own. A system's central properties arise from the interactions and relationships among the parts, which is the dynamic process of *emergence*. This framework investigates the processes that fuel emergence and change in systems from micro to macro levels. The above-mentioned system thinking hierarchical model (e.g., Assaraf & Orion, 2010) represents yet another framework for understanding student learning.

CSSE studies have made significant inroads with respect to what students know about complex systems and how learning can be supported. A prudent next step might be for the field to reach consensus on essential content features of complex systems learning.

Need for More Research on Teacher Learning

Equally important for the field is a greater focus on what teachers know. Our results show that comparatively few studies have worked with teachers as their target population (e.g., Carlsson, 2002; Hmelo-Silver et al., 2007; Klopfer, Yoon, & Perry, 2005; Liu & Hmelo-Silver, 2009; Yoon et al., 2015). We do know from these studies that teachers typically have a weak understanding of complex systems concepts such as behaviors and processes that fuel systems (Hmelo-Silver & Pfeffer, 2004; Liu & Hmelo-Silver, 2009) and that developing teacher expertise in the pedagogies and tools that support complex systems, such as computational modelling tools, is essential to good instruction (Yoon et al., 2015).

Furthermore, the analysis of educational purpose with respect to dimensions of CSSE research showed that a majority of studies focused on improving student learning through specific interventions, which supports the goals articulated in the NGSS. However, few studies have investigated teacher learning needs as well as teachers' states of complex systems understanding. Of particular interest in this category is also a lack of focus on the relationship between teaching or instructional supports and learning, which was the subject of only three studies (Hmelo-Silver et al., 2015; Ioannidou et al., 2010; Yoon et al., 2015). There are well-documented challenges that teachers often experience in adopting new teaching approaches, especially those that are computer supported (Aldunate & Nussbaum, 2013; Ertmer, Ottenbreit-Leftwich, Sendurur, & Sendurur, 2012), as many complex systems interventions are. Furthermore, Gerard et al. (2011) discussed the importance of opportunities for high-quality teacher professional development in order to conduct technology-enhanced inquiry in science classrooms. Thus, it is important for the CSSE field to consider what characteristics of professional development are needed to support teachers in improving students' understanding of complex systems. This is especially essential in light of the fact that the NGSS represents a new vision for science learning that comes with a steep instructional learning curve (Wilson, 2013).

Need for Research on Common Context Features

Overall, the analysis also revealed that a majority of studies did not report on gender composition, race or ethnicity, or the type of school setting where the research was conducted. Almost a third of the studies were also conducted in laboratory settings. Here, it is important to point out that for several decades now, learning research has suggested that populations of learners can be affected by demographic and contextual variables that significantly support or limit learning and participation (Brown, Collins, & Duguid, 1989; Greeno, 1998; Sawyer, 2015). This fact, in part, spawned the development of whole fields of educational research such as the Learning Sciences (Kolodner, 2004; Pea, 2016; Sawyer, 2015, Yoon & Hmelo-Silver, 2017) where the messiness of real-world learning environments and sociocultural activity occupies a central role in understanding how people learn (Bransford, Brown, & Cocking, 1999). To ensure that CSSE research and interventions can impart benefits to all learners, more studies accounting for situated and contextualized factors that mediate learning are needed. Furthermore, it would be helpful to know more about how these factors support or constrain the

development of complex systems understanding, which was not a focus of this review.

Strong Representation of Early-Stage Exploratory Research

With respect to whether CSSE studies have used methodologies and conducted research toward scaling interventions, the analysis showed that the majority of studies were aimed at FEE research. This finding might be expected considering that complex systems research is relatively recent compared to other research in the field of education such as literacy, mathematics, or science. In mathematics education research, for example, a plethora of developmental research has exposed numerous cognitive challenges in the learning of percentage and proportions (Parker & Leinhardt, 1995) that researchers agree is a topic that must be studied. However, the CSSE field still has not come to consensus on frameworks that describe the content of what needs to be learned about complex systems. Furthermore, the aforementioned focus on building computational resources has necessarily required greater emphasis on exploratory research as newer tools continually emerge with affordances that enable different forms of participation and learning that need development and testing. The nature of technological development likewise dictates that some technologies become obsolete before they can be tested at scale. For example, Yoon (2008) and Klopfer, Yoon, and Perry (2005) described participatory simulations that were developed on mobile technologies—Thinking Tags and Palm Pilots—that are not in use today. The relative infancy of the field, in addition to this variability of learning resources, is arguably why we see a preponderance of mixed methodologies to examine implementation outcomes. Such methodologies can provide a more holistic picture of the extent to which populations have learned and why or how they have learned over using only quantitative or qualitative methods (Creswell & Plano Clark, 2011).

Need for More Comparative Research

Our results reveal that some subfields of complex systems research, like student learning, have reached a level of maturity that warrants more systematic investigation of how particular curricular and instructional design choices compare to others. Only 15% of the studies in our data set used quantitative methods; these studies tended to be experimental or quasi-experimental in design (e.g., DeBoer et al., 2014; Liu & Hmelo-Silver, 2009; Peppler et al., 2013; Plate, 2010; Thompson & Reimann, 2010). Thus, CSSE researchers have generally used noncomparative research designs. Noncomparative approaches limit any affordances that can be interpreted as having a significantly positive (or negative) influence on learning compared to other learning activities. Importantly, the noncomparative nature of these studies does not provide adequate information about the added value of learning experiences that can respond to new curricular mandates such as those that are found in the NGSS. Further evidence to support this assertion is suggested by the fact that no studies in our sample were categorized as efficacy, effectiveness, or scale-up research, and the fact that more than half of the studies worked with population sizes of less than 50. Coupled with the aforementioned lack of focus on contextualized variables, it is difficult to determine for whom and under what conditions CSSE interventions can achieve curriculum and

instructional learning goals. More testing across different levels and settings (Penuel & Fishman, 2012) and studies that address variation in performance and system variables (Bryk et al., 2010; Bryk et al., 2015) with larger sample sizes need to be completed in future CSSE research.

Conclusion

Over the past two decades, policy mandates in science education and real-world complex systems research have spurred an interest in helping students learn about complex systems at the K–12 level. This systematic review was aimed at understanding how well the CSSE research reflects three central features of educational interventions. These are the extent to which it represents the goals of real-world complex systems science and the goals of the NGSS; whether CSSE research has been conducted in a variety of settings and populations; and whether CSSE studies have collectively aimed to conduct research to scale interventions. The research questions can also be considered as overarching goals to move the research field forward in achieving high-quality educational experiences in science education for all students. The analysis found critical needs in five areas: (a) a need for more research in different knowledge domains outside of the content areas of biology and ecology, (b) a need for more research on system states as opposed to structures and processes, (c) a need to develop a common understanding of the complex systems content that is essential to be learned, (d) a need to consider contextual factors that will affect the learning environment and population including teacher learning, and (e) more comparative research to determine the value of CSSE interventions over traditional forms of instruction, including an emphasis on what teachers need in professional development activities. Our intention is for this review to mobilize CSSE researchers to work together to address these important needs.

References

References marked with an asterisk indicate studies included in the systematic review.

Aldunate, R., & Nussbaum, M. (2013). Teacher adoption of technology. *Computers in Human Behavior*, *29*, 519–524. doi:10.1016/j.chb.2012.10.017

Allison, S. D., & Martiny, J. B. H. (2008). Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 11512–11519. doi:10.1073/pnas.0801925105

American Association for the Advancement of Science. (1989). *Science for all Americans*. New York, NY: Oxford University Press.

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

*Assaraf, O. B., & Orion, N. (2005). Development of system thinking skills in the context of earth system education. *Journal of Research in Science Teaching*, *42*, 518–560. doi:10.1002/tea.20061

*Assaraf, O. B., & Orion, N. (2010). System thinking skills at the elementary school level. *Journal of Research in Science Teaching*, *47*, 540–563. doi:10.1002/tea.20351

*Assaraf, O. B., & Orpaz, I. (2010). The "life at the poles" study unit: Developing junior high school students' ability to recognize the relations between earth systems. *Research in Science Education*, *40*, 525–549.

Azevedo, R., Cromley, J. G., & Winters, F. I. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional Science*, *33*, 381–412. doi:10.1007/s11251-005-1273-8

Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research & Development*, *56*, 45–72. doi:10.1007/s11423-007-9067-0

Azevedo, R., Winters, F. I., & Moos, D. C. (2004). Can students collaboratively use hypermedia to learn science? The dynamics of self-and other-regulatory processes in an ecology classroom. *Journal of Educational Computing Research*, *31*, 215–245. doi:10.2190/hft6-8eb1-tn99-mjvq

Bar-Yam, Y. (2016). From big data to important information. *Complexity*, *21*(Suppl. 2), 73–98. doi:10.1002/cplx.21785

*Basu, S., Sengupta, P., & Biswas, G. (2015). A scaffolding framework to support learning of emergent phenomena using multi-agent-based simulation environments. *Research in Science Education*, *45*, 293–324. doi:10.1007/s11165-014-9424-z

*Ben-Zvi-Assaraf, O., & Orion, N. (2010). Four case studies, six years later: Developing system thinking skills in junior high school and sustaining them over time. *Journal of Research in Science Teaching*, *47*, 1253–1280. doi:10.1002/tea.20383

Bettencourt, L. M. A., & Brelsford, C. (2014). Industrial ecology the view from complex systems. *Journal of Industrial Ecology*, *19*, 195–197. doi:10.1111/jiec.12243

Bhatta, S. R., & Goel, A. (1997). Learning generic mechanisms for innovative strategies in adaptive design. *Journal of the Learning Sciences*, *6*, 367–396.

Brady, C., Holbert, N., Soylu, F., Novak, M., & Wilensky, U. (2015). Sandboxes for model-based inquiry. *Journal of Science Education and Technology*, *24*, 265–286. doi:10.1007/s10956-014-9506-8

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). How experts differ from novices. In J. D. Bransford, A. L. Brown, & R. R. Cocking (Eds.), *How people learn: Brain, mind, experience, and school* (pp. 19–38). Washington, DC: National Academies Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. doi:10.3102/0013189x018001032

Bryk, A. S., Gomez, L. M., & Grunow, A. (2010). *Getting ideas into action: Building networked improvement communities in education*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved from https://www.carnegiefoundation.org/resources/publications/getting-ideas-action-building-networked-improvement-communities-education/

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

*Calderon-Canales, E., Flores-Camacho, F., & Gallegos-Cazares, L. (2013). Elementary students' mental models of the solar system. *Astronomy Education Review*, *12*(1). doi:10.3847/aer2012044

Capra, F. (1982). *The turning point*. New York, NY: Bantam Books.

*Carlsson, B. (2002). Ecological understanding 2: Transformation—a key to ecological understanding. *International Journal of Science Education*, *24*, 701–715. doi:10.1080/09500690110098877

Caschera, F., Bedau, M. A., Buchanan, A., Cawse, J. de Lucrezia, D., Gazzikam, G., . . . Packard, N. H. (2011). Coping with complexity: Machine learning optimization

of cell-free protein synthesis. *Biotechnology and Bioengineering*, *108*, 2218–2228. doi:10.1002/bit.23178

*Cheng, M. M. W., & Gilbert, J. K. (2015). Students' visualization of diagrams representing the human circulatory system: The use of spatial isomorphism and representational conventions. *International Journal of Science Education*, *37*, 136–161. doi: 10.1080/09500693.2014.969359

*Chi, M. T. H., Roscoe, R., Slotta, J., Roy, M., & Chase, M. (2012). Misconceived causal explanations for "emergent" processes. *Cognitive Science*, *36*, 1–61. doi:10.1111/j.1551-6709.2011.01207.x

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

*Danish, J. A. (2014). Applying an activity theory lens to designing instruction for learning about the structure, behavior, and function of a honeybee system. *Journal of the Learning Sciences*, *23*, 100–148. doi:10.1080/10508406.2013.856793

*Danish, J. A., Peppler, K., Phelps, D., & Washington, D. (2011). Life in the hive: Supporting inquiry into complexity within the zone of proximal development. *Journal of Science Education and Technology*, *20*, 454–467. doi:10.1007/s10956-011-9313-4

*DeBoer, G. E., Quellmalz, E. S., Davenport, J. L., Timms, M. J., Herrmann-Abell, C. F., Buckley, B. C., & Flanagan, J. C. (2014). Comparing three online testing modalities: Using static, active, and interactive online testing modalities to assess middle school students' understanding of fundamental ideas and use of inquiry skills related to ecosystems. *Journal of Research in Science Teaching*, *51*, 523–554. doi:10.1002/tea.21145

*DeLeo, G., Weidenhammer, A., & Wecht, K. (2012). A computer story: Complexity from simplicity. *Science Teacher*, *79*(7), 48–53. doi:10.1057/9781137006226

*Eilam, B. (2002). Strata of comprehending ecology: Looking through the prism of feeding relations. *Science Education*, *86*, 645–671. doi:10.1002/sce.10041

*Eilam, B. (2012). System thinking and feeding relations: Learning with a live ecosystem model. *Instructional Science*, *40*, 213–239. doi:10.1007/s11251-011-9175-4

Ertmer, P. A., Ottenbreit-Leftwich, A. T., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers & Education*, *59*, 423–435. doi:10.1016/j.compedu.2012.02.001

*Evagorou, M., Korfiatis, K., Nicolaou, C., & Constantinou, C. (2009). An investigation of the potential of interactive simulations for developing system thinking skills in elementary school: A case study with fifth-graders and sixth-graders. *International Journal of Science Education*, *31*, 655–674. doi:10.1080/09500690701749313

Forrester, J. W. (1994). Systems dynamics, systems thinking, and soft OR. *Systems Dynamics Review*, *10*(2), 1–14. doi.org/10.1002/sdr.4260100211

Fuentes, M. A. (2014). Complexity and the emergence of physical properties. *Entropy*, *16*, 4489-4496. doi:10.3390/e16084489

*Gazit, E., Yair, Y., & Chen, D. (2005). Emerging conceptual understanding of complex astronomical phenomena by using a virtual solar system. *Journal of Science Education and Technology*, *14*, 459–470. doi:10.1007/s10956-005-0221-3

Gerard, L. F., Varma, K., Corliss, S. B., & Linn, M. (2011). Professional development for technology-enhanced inquiry science. *Review of Educational Research*, *81*, 408–448. doi:10.3102/0034654311415121

*Ginns, I. S., Norton, S. J., & Mcrobbie, C. J. (2005). Adding value to the teaching and learning of design and technology. *International Journal of Technology and Design Education*, *15*, 47–60. doi:10.1007/s10798-004-6193-4

*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*, 259–280. doi:10.1002/sce.20368

Greene, J. A., Moos, D. C., Azevedo, R., & Winters, F. I. (2008). Exploring differences between gifted and grade-level students' use of self-regulatory learning processes with hypermedia. *Computers & Education*, *50*, 1069–1083. doi:10.1016/j.compedu.2006.10.004

Greeno, J. (1998). The situativity of knowing, learning, and research. *American Psychologist*, *53*, 5–26. doi:10.1037/0003-066x.53.1.5

*Grotzer, T. A., & Bell Basca, B. (2003). How does grasping the underlying causal structures of ecosystems impact students' understanding? *Journal of Biological Education*, *38*, 16–29. doi:10.1080/00219266.2003.9655891

*Grotzer, T. A., Powell, M. M., Derbiszewska, K. M., Courter, C. J., Kamarainen, A. M., Metcalf, S. J., & Dede, C. J. (2015). Turning transfer inside out: The affordances of virtual worlds and mobile devices in real world contexts for teaching about causality across time and distance in ecosystems. *Technology, Knowledge and Learning*, *20*, 43–69. doi:10.1007/s10758-014-9241-5

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

*Hmelo, C. E., Holton, D. L., & Kolodner, J. L. (2000). Designing to learn about complex systems. *Journal of the Learning Sciences*, *9* 247–298. doi:10.1207/s15327809jls0903_2

Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: Some core challenges. *Journal of the Learning Sciences*, *15*, 53–61. doi:10.1207/s15327809jls1501_7

*Hmelo-Silver, C., Liu, L., Gray, S., & Jordan, R. (2015). Using representational tools to learn about complex systems: A tale of two classrooms. *Journal of Research in Science Teaching*, *52*, 6–35. doi:10.1002/tea.21187

*Hmelo-Silver, C., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *Journal of the Learning Sciences*, *16*, 307–331. doi:10.1080/10508400701413401

*Hmelo-Silver, C., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, *28*, 127–138. doi:10.1207/s15516709cog2801_7

*Hogan, K. (2000). Assessing students' systems reasoning in ecology. *Journal of Biological Education*, *35*(1), 22–28. doi:10.1080/00219266.2000.9655731

*Hogan, K., & Thomas, D. (2001). Cognitive comparisons of students' systems modeling in ecology. *Journal of Science Education and Technology*, *10*, 319–345. doi:10.1023/a:1012243102249

Hordijk, W., Kauffman, S. A., & Steel, M. (2011). Required levels of catalysis for emergence of autocatalytic sets in models of chemical reaction systems. *International Journal of Molecular Sciences*, *12*, 3085–3101. doi:10.3390/ijms12053085

*Ioannidou, I. A., Paraskevopoulos, S., & Tzionas, P. (2006). An interactive computer graphics interface for the introduction of fuzzy inference in environmental education. *Interacting With Computers*, *18*, 683–708. https://doi.org/10.1016/j.intcom.2005.10.007

*Ioannidou, A., Repenning, A., Webb, D., Keyser, D., Luhn, L., & Daetwyler, C. (2010). Mr. Vetro: A collective simulation for teaching health science. *International*

Journal of Computer-Supported Collaborative Learning, *5*, 141–166. doi:10.1007/s11412-010-9082-8

Institute of Education Sciences & National Science Foundation. (2013). *Common guidelines for education research and development*. Washington, DC: Author.

Jacobson, M. (2001). Problem solving, cognition, and complex systems: Differences between experts and novices. *Complexity*, *6*(3), 41–49. doi:10.1002/cplx.1027

*Jordan, R., Gray, S., Demeter, M., Lui, L., & Hmelo-Silver, C. (2009). An assessment of students' understanding of ecosystem concepts: Conflating ecological systems and cycles. *Applied Environmental Education and Communication*, *8*(1), 40–48. doi:10.1080/15330150902953472

*Jordan, R. C., Brooks, W. R., Hmelo-Silver, C., Eberbach, C., & Sinha, S. (2014). Balancing broad ideas with context: An evaluation of student accuracy in describing ecosystem processes after a system-level intervention. *Journal of Biological Education*, *48*(2), 57–62. doi:10.1080/00219266.2013.821080

*Jordan, R. C., Gray, S. A., Brooks, W. R., Honwad, S., & Hmelo-Silver, C. (2013). Process-based thinking in ecosystem education. *Natural Sciences Education*, *42*, 68–74. doi:10.1080/1533015x.2013.797860

*Keynan, A., Ben-Zvi Assaraf, O., & Goldman, D. (2014). The repertory grid as a tool for evaluating the development of students' ecological system thinking abilities. *Studies in Educational Evaluation*, *41*, 90–105. doi:10.1016/j.stueduc.2013.09.012

*Klopfer, E., Scheintaub, H., Huang, W., Wendel, D., & Roque, R. (2009). The simulation cycle: Combining games, simulations, engineering and science using StarLogo TNG. *E-Learning*, *6*, 71–96. doi:10.2304/elea.2009.6.1.71

*Klopfer, E., Yoon, S., & Perry, J. (2005). Using palm technology in participatory simulations of complex systems: A new take on ubiquitous and accessible mobile computing. *Journal of Science Education and Technology*, *14*, 285–297. doi:10.1007/s10956-005-7194-0

*Klopfer, E., Yoon, S., & Um, T. (2005). Teaching complex dynamic systems to young students with StarLogo. *Journal of Computers in Mathematics and Science Teaching*, *24*, 157–178. doi:10.1016/s0303-2647(03)00115-1

Kolodner, J. (2004). The learning sciences: Past, present and future. *Educational Technology*, *44*(3), 37–42. doi:10.5040/9781474258029.ch-015

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. doi:10.2307/2529310

Levin, S., & Lubchenko, J. (2008). Resilience, robustness, and marine ecosystem-based management. *Bioscience*, *58*(1), 27–32. doi:10.1641/b580107

*Levy, S. T., & Wilensky, U. (2008). Inventing a "mid level" to make ends meet: Reasoning between the levels of complexity. *Cognition and Instruction*, *26*(1), 1–47. doi:10.1080/07370000701798479

*Levy, S. T., & Wilensky, U. (2009). Students' learning with the connected chemistry (CC1) curriculum: Navigating the complexities of the particulate world. *Journal of Science Education and Technology*, *18*, 243–254. doi:10.1007/s10956-009-9145-7

*Levy, S. T., & Wilensky, U. (2011). Mining students' inquiry actions for understanding of complex systems. *Computers & Education*, *56*, 556–573. doi:10.1016/j.compedu.2010.09.015

*Lin, C., & Hu, R. (2003). Students' understanding of energy flow and matter cycling in the context of the food chain, photosynthesis, and respiration. *International Journal of Science Education*, *25*, 1529–1544. doi:10.1080/0950069032000052045

*Liu, L., & Hmelo-Silver, C. (2009). Promoting complex systems learning through the use of conceptual representations in hypermedia. *Journal of Research in Science Teaching*, *46*, 1023–1040. doi:10.1002/tea.20297

Lewis, M. (2000). The promise of dynamic systems approaches for an integrated account of human development. *Child Development*, *71*, 36-43. doi:10.1111/1467-8624.00116

Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, *1*(2). Retrieved from http://217.160.35.246/fqs-texte/2-00/2-00mayring-e.pdf

Mitchell, M. (2009). *Complexity: A guided tour*. New York, NY: Oxford University Press.

*McNeal, K. S., Libarkin, J. C., Ledley, T. S., Bardar, E., Haddad, N., Ellins, K., & Dutta, S. (2014). The role of research in online curriculum development: The case of EarthLabs climate change and earth system modules. *Journal of Geoscience Education*, *62*, 560–577. doi:10.5408/13-060.1

*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. doi:10.1002/tea.20314

The National Academies. (2009). *Keck Futures Initiative: Complex Systems: Task group summaries*. Washington, DC: National Academies Press.

National Research Council. (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. Committee on a Conceptual Framework for New K–12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education*. Washington, DC: National Academies Press.

National Research Council. (2011). *Successful K–12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Washington, DC: National Academies Press.

National Science Board. (2010). *Preparing the next generation of STEM innovators: Identifying and developing our nation's human capital*. Washington, DC: National Science Foundation.

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

Office of Management and Budget. (1995). *Standards for the classification of federal data on race and ethnicity*. Retrieved from https://www.whitehouse.gov/omb/fedreg_race-ethnicity

*Pallant, A., & Lee, H. (2015). Constructing scientific arguments using evidence from dynamic computational climate models. *Journal of Science Education and Technology*, *24*, 378–395. doi:10.1007/s10956-014-9499-3

*Pallant, A., Lee, H., & Pryputniewicz, S. (2012). Modeling earth's climate. *Science Teacher*, *79*(7), 38–42. doi:10.2307/4448063

Parker, M., & Leinhardt, G. (1995). Percent: A privileged proportion. *Review of Educational Research*, *65*, 421–481. doi:10.3102/00346543065004421

Parnafes, O. (2010). When simple harmonic motion is not that simple: Managing epistemological complexity by using computer-based representations. *Journal of Science Education and Technology*, *19*, 565–579. doi:10.1007/s10956-010-9224-9

President's Council of Advisors on Science and Technology. (2010). *Prepare and inspire: K–12 education in science, technology, engineering, and mathematics (STEM) for America's future*. Washington, DC: Author.

Pea, R. D. (2016). The prehistory of the learning sciences. In M. A. Evans, M. J. Packer, & R. K. Sawyer (Eds.), *Reflections on the learning sciences* (pp. 32–58). New York, NY: Cambridge University Press.

*Penner, D. E. (2000). Explaining systems: Investigating middle school students' understanding of emergent phenomena. *Journal of Research in Science Teaching*, *37*, 784–806. doi:10.1002/1098-2736(200010)37:8<784::aid-tea3>3.0.co;2-e

Penuel, W., & Fishman, B. (2012). Large-scale science education intervention research we can use. *Journal of Research in Science Teaching*, *49*, 281–304. doi:10.1002/tea.21001

*Peppler, K., Danish, J. A., & Phelps, D. (2013). Collaborative gaming: Teaching children about complex systems and collective behavior. *Simulation & Gaming*, *44*, 683–705. doi:10.1177/1046878113501462

*Perkins, D. N., & Grotzer, T. A. (2005). Dimensions of causal understanding: The role of complex causal models in students' understanding of science. *Studies in Science Education*, *41*(1), 117–165. doi:10.1080/03057260508560216

*Plate, R. (2010). Assessing individuals' understanding of nonlinear causal structures in complex systems. *System Dynamics Review*, *26*, 19–33. doi:10.1002/sdr.432

*Puk, T., & Stibbards, A. (2011). Growth in ecological concept development and conceptual understanding in teacher education: The discerning teacher. *International Journal of Environmental and Science Education*, *6*, 191–211. doi:10.1080/1533015x.2011.549796

Randler, C., & Bogner, F. X. (2009). Efficacy of two different instructional methods involving complex ecological content. *International Journal of Science and Mathematics Education*, *7*, 315–337. doi:10.1007/s10763-007-9117-4

*Reiss, M. J., & Tunnicliffe, S. D. (2001). Students' understandings of human organs and organ systems. *Research in Science Education*, *31*, 383–399. doi:10.1023/a:1013116228261

*Repenning, A., Ioannidou, A., Luhn, L., Daetwyler, C., & Repenning, N. (2010). Mr. Vetro: Assessing a collective simulation framework. *Journal of Interactive Learning Research*, *21*, 515–537. doi:10.1007/s11412-010-9082-8

*Resnick, M. (1996). Beyond the centralized mindset. *Journal of the Learning Sciences*, *5*(1), 1–22. doi:10.1207/s15327809jls0501_1

*Richardson, M. L., & Hari, J. (2009). Building ecological complexity in the classroom using pea aphids & components of their community. *American Biology Teacher*, *71*, 286–290. doi:10.2307/27669432

*Riess, W., & Mischo, C. (2010). Promoting systems thinking through biology lessons. *International Journal of Science Education*, *32*, 705–725. doi:10.1080/09500690902769946

*Rodríguez, M., Kohen, R., & Delval, J. (2015). Children and adolescents' thoughts on pollution: Cognitive abilities required to understand environmental systems. *Environmental Education Research*, *21*(1), 76–91. doi.10.1080/13504622.2013.862613

Sabelli, N. H. (2006). Complexity, technology, science, and education. *Journal of the Learning Sciences*, *15*, 5–9. doi:10.1207/s15327809jls1501_3

Sah, P., Singh, L. O., Clauset, A., & Bansal, S. (2014). Exploring community structure in biological networks with random graphs. *BMC Bioinformatics*, *15*(220). doi:10.1186/1471-2105-15-220

Sawyer, R. K. (2015). Introduction: The new science of learning. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 1–19). New York, NY: Cambridge University Press.

Servedio, M. R., Brandvain, Y., Dhole, S., Fitzpatrick, C. L., Goldberg, E. E., Stern, C. A., Van Cleve, J., & Yeh, D. J. (2014). Not just a theory: The utility of mathematical models in evolutionary biology. *PLOS Biology*, *12*(12), 1–5. doi:10.1371/journal .pbio.1002017

Sherrington, D. (2010). Physics and complexity. *Philosophical Transactions of the Royal Society of London Series A: Mathematical Physical and Engineering Sciences*, *368*(1914), 1175–1189. doi:10.1098/rsta.2009.0208

Sole, R. (2015). Bioengineering the biosphere? *Ecological Complexity*, *22*, 40–49. doi:10.1016/j.ecocom.2015.01.005

*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*, 610–613. doi:10.1002/tea.20313

*Stavrou, D., & Duit, R. (2014). Teaching and learning the interplay between chance and determinism in nonlinear systems. *International Journal of Science Education*, *36*, 506–530. doi:10.1080/09500693.2013.802056

*Stavrou, D., Duit, R., & Komorek, M. (2008). A teaching and learning sequence about the interplay of chance and determinism in nonlinear systems. *Physics Education*, *43*, 417–422. doi:10.1088/0031-9120/43/4/011

*Stroup, W. M., & Wilensky, U. (2014). On the embedded complementarity of agent-based and aggregate reasoning in students' developing understanding of dynamic systems. *Technology, Knowledge and Learning*, *19*(1-2), 19–52. doi:10.1007/ s10758-014-9218-4

Sweeney, L. B., & Sterman, J. D. (2007). Thinking about systems: Student and teacher conceptions of natural and social systems. *Systems Dynamics Review*, *23*, 285–311. doi.org/10.1002/sdr.366

Thelan, E., & Smith, L. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge: MIT Press.

*Thompson, K., & Reimann, P. (2010). Pattern of use of an agent-based model and a system dynamics model: The application of patterns of use and the impacts on learning outcomes. *Computer & Education*, *54*, 392–403. doi:10.1016/j.compedu .2009.08.020

*Tsurusaki, B. K., & Anderson, C. W. (2010). Students' understanding of connections between human engineered and natural environmental systems. *International Journal of Environmental & Science Education*, *5*, 407–433. doi:10.1002/ 9781118230107.ch14

*Vachliotis, T., Salta, K., & Tzougraki, C. (2014). Meaningful understanding and systems thinking in organic chemistry: Validating measurement and exploring relationships. *Research in Science Education*, *44*, 239–266. doi:10.1007/s11165-013-9382-x

*van Borkulo, S. P., van Joolingen, W. R., Savelsbergh, E. R., & de Jong, T. (2012). What can be learned from computer modeling? Comparing expository and modeling approaches to teaching dynamic systems behavior. *Journal of Science Education and Technology*, *21*, 267–275. doi:10.1007/s10956-011-9314-3

*Varma, K., & Linn, M. C. (2012). Using interactive technology to support students' understanding of the greenhouse effect and global warming. *Journal of Science Education and Technology*, *21*, 453–464. doi:10.1007/s10956-011-9337-9

*Vattam, S. S., Goel, A. K., Rugaber, S., Hmelo-Silver, C., Jordan, R., Gray, S., & Sinha, S. (2011). Understanding complex natural systems by articulating structure-

behavior-function models. *Educational Technology & Society*, *14*, 66–81. doi:10.14786/flr.v1i1.14

*Verhoeff, R. P., Waarlo, A. J., & Boersma, K. T. (2008) Systems modeling and the development of coherent understanding of cell biology. *International Journal of Science Education*, *30*, 543–568. doi:10.1080/09500690701237780

West, G. B. (2014). A theoretical physicists journey into biology: From quarks and strings to cells and whales. *Physical Biology*, *11*(5), 1–6. doi:10.1088/1478-3975/11/5/053013

Wilensky, U., & Jacobson, M. J. (2015). Complex systems and the learning sciences. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 319–338). New York, NY: Cambridge University Press.

*Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3–19. doi:10.1023/a:1009421303064

*Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories—an embodied modeling approach. *Cognition and Instruction*, *24*, 171–209. doi:10.1207/s1532690xci2402_1

*Wilkerson-Jerde, M. H., & Wilensky, U. J. (2015). Patterns, probabilities, and people: Making sense of quantitative change in complex systems. *Journal of the Learning Sciences*, *24*, 204–251. doi:10.1080/10508406.2014.976647

Wilson, S. M. (2013). Professional development for science teachers. *Science*, *340*, 310–313. doi:10.1126/science.1230725

*Wu, H. (2010). Modelling a complex system: Using novice-expert analysis for developing an effective technology-enhanced learning environment. *International Journal of Science Education*, *32*, 195–219. doi:10.1080/09500690802478077

*Yoon, S. A. (2008). An evolutionary approach to harnessing complex systems thinking in the science and technology classroom. *International Journal of Science Education*, *30*(1), 1–32. doi:10.1080/09500690601101672

*Yoon, S. A. (2011). Using social network graphs as visualization tools to influence peer selection decision-making strategies to access information about complex socioscientific issues. *Journal of the Learning Sciences*, *20*, 549–588. doi:10.1080/10508406.2011.563655

Yoon, S. A. (in press). Complex systems and the Learning Sciences: Implications for learning, theory, and methodologies. In *The international handbook of the learning sciences*. New York, NY: Routledge.

Yoon, S. A., & Hmelo-Silver, C. E. (2017). What do learning scientists do? *Journal of the Learning Sciences*, *26*, 167–183. doi:10.1080/10508406.2017.1279546

Yoon, S. A., Klopfer, E., Anderson, E., Koehler-Yom, J., Sheldon, J., Schoenfeld, I., . . .Goh, S. (2016). Designing computer-supported complex systems curricula for the Next Generation Science Standards in high school science classrooms. *Systems*, *4*(4), 38. doi:10.3390/systems4040038

*Yoon, S. A., Koehler-Yom, J., Anderson, E., Lin, J., & Klopfer, E. (2015). Using an adaptive expertise lens to understand the quality of teachers' classroom implementation of computer-supported complex systems curricula in high school science. *Research in Science & Technological Education*, *33*, 237–251. doi:10.1080/02635143.2015.1031099

Authors

SUSAN A. YOON is professor of education in the Teaching, Learning, and Leadership Division at the Graduate School of Education of the University of Pennsylvania; email: *yoonsa@upenn.edu*. Her research interests in the areas of the learning sciences and science education examine how visualization tools and curricula can support the learning of complex scientific phenomena. She has investigated student learning impacts and teacher professional development characteristics including building social networks and social capital to support learning efficacy.

SAO-EE GOH is a researcher with the Academy of Singapore Teachers with the Ministry of Education in Singapore; email: *sedm0253@yahoo.com*. His research interests include professional development and science education.

MIYOUNG PARK is a PhD candidate in the Teaching, Learning, and Teacher Education program at the Graduate School of Education of the University of Pennsylvania; email: *parkmi@gse.upenn.edu*. Her research interests are centered on science education, particularly the preparation, education, and professional development of science teachers. She is also interested in K–12 science curriculum and pedagogy that supports argumentation, and coherence across science topics.