JOURNAL OF GEOSCIENCE EDUCATION **65**, 219–230 (2017)

Using Systems Thinking in the Design, Implementation, and Evaluation of Complex Educational Innovations, With Examples From the InTeGrate Project

Kim A. Kastens[1] and Cathryn A. Manduca[2]

ABSTRACT

Many geoscience education initiatives now involve cross-departmental or multi-institutional programs. However, the geoscientists who lead such programs typically have little experience or training in program design, leadership, or evaluation. In this commentary, we make the case that geoscientists taking on these ambitious leadership roles can draw on a set of understandings and skills that they already have: the tools and habits of mind of systems thinking. Using examples from the InTeGrate program, we suggest ways to envision, shape, and monitor an educational intervention by thinking of oneself as building and improving a complex system with constituent subsystems. Suggestions for the design phase include the following: Decide on an essential suite of subsystems and plan for them to interact in mutually beneficial ways. Consider using a set of semi-autonomous parallel subsystems that will allow for replication with adaptation as experience accrues. Plan for nonlinear causality chains in which one activity has multiple beneficial outputs, and desired outcomes are supported through multiple influencers. At the implementation phase, leverage feedback loops to nudge actors toward desired behaviors and away from problematic choices. Build technical and social mechanisms to regulate flows of information between and within subsystems, so as to deliver timely, actionable information without overloading the actors. In evaluation, use systems mapping to understand where critical dependencies occur and insert evaluative probes at these locations. Seek out early indicators of emergent phenomena and conceptualize long-term outcomes in terms of modifications to the larger system of higher education. © *2017 National Association of Geoscience Teachers*. [DOI: 10.5408/16-225.1]

Key words: systems thinking, education reform, sustainability, InTeGrate

INTRODUCTION

Geoscience education has reached a pivotal stage in its evolution, where many undergraduate faculty members have adopted some research-based teaching practices, aiming for a more active, more inclusive, more science-like student experience and a deeper level of understanding of Earth processes (Manduca et al., 2017). As a community, we are now being called upon to tackle an even bigger and harder set of problems, challenges that require coordinated action across entire departments, multiple departments, or multiple institutions. The urge to tackle bigger challenges comes from many directions: the realization that improving Earth/human interactions requires interdisciplinary collaboration (Committee on Facilitating Interdisciplinary Research, 2005; ICSU, 2010); the need for more effective workforce preparation in a rapidly evolving economy (Wilson, 2016a, 2016b); the realization that achieving widespread success for underrepresented groups requires supporting the whole student (Seidman, 2005; Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline, 2011); the desire on the part of institutions of higher education to maximize the utility of their assets and to offer distinctive programs (Marcy, 2017); and the availability of funding and recognition for tackling larger-scale educational challenges (National Science Foundation, 2015).

Most college faculty members to do not have training or experience in running large, complicated programs or organizations. Education reform is among a set of challenging problems that have been labeled as "complex" (Holland, 1998; Kania et al., 2014) or "wicked" (Williams and Hummelbrunner, 2010; Ramaley, 2014). Such problems tend to be dynamic, nonlinear, and counterintuitive; are entangled with other issues; involve the interplay of multiple independent factors; and are not amenable to solution by changing one factor. Transforming higher education involves individuals and entities with varied histories and priorities that interact in ways that may inhibit change (Tierney, 1997; Bergquist, 1992; Kezar and Eckel, 2002; Austin, 2013; Fry and the Coalition for Reform of Undergraduate STEM Education, 2014).

Growing from a scientist/educator into a skillful leader of an ambitious intervention in a complex system is hard, and you will need courage, allies, mentors, and much else. In this commentary, we assert that as a geoscientist, you already have a valuable set of thought patterns and habits of mind that you can repurpose in the service of education reform. This way of thinking comes under a variety of labels, including "systems thinking" (Hummelbrunner, 2011; Patton, 2011), "complexity theory" (Hawe et al., 2009), "complex adaptive systems" (North American Primary Care Research Group, 2009), and "system dynamics" (Meadows, 1999). The common thread running through these approaches is that the unit of analysis is a system, a "set of elements whose interconnections determine their behavior" (Climate Interactive, 2016). Interconnections include flows of

---

Received 31 October 2016; revised 6 March 2017 and 30 May 2017; accepted 31 May 2017; published online 7 August 2017.

[1] Lamont-Doherty Earth Observatory of Columbia University, 61 Rt 9W, Palisades, New York 10964, USA, kastens@ldeo.columbia.edu, 1-914-671-2341

[2] Science Education Resource Center, Carleton College, One North College Street, Northfield, Minnesota 55057, USA, cmanduca@carleton.edu

1089-9995/2017/65(3)/219/12

energy, material, and information; contingencies and dependencies; time lags; feedback loops, etc. Your background has conditioned you to regard these as ways to understand natural systems, but in fact many of these concepts have long been used to understand and improve human-built systems, from steam engines to multinational conglomerates (Senge, 2006; Meadows, 2008; Roe, 2009). The goal of this commentary is to offer suggestions on how to apply your systems thinking expertise as you plan and execute your desired education reform and evaluate what you have accomplished. We will not be presenting evidence that systems thinking is superior to other approaches; rather, our assertion is that this approach is particularly accessible to the *Journal of Geoscience Education* readership, many of whom have already invested years or decades in developing systems thinking skills and habits.

We have led or participated in many large and complex programs and institutions, in both science education (Keck Geology Consortium, Digital Library for Earth System Education, the National Science Digital Library, On the Cutting Edge, Sea Education Association, Education Development Center, Project Kaleidoscope) and science (Ocean Drilling Program, RIDGE, EarthCube, Spatial Intelligence & Learning Center, MARGINS). The ideas presented here have emerged from a distillation of all of these experiences, as well as insights from the institutional change literature. However, for the sake of clarity, the examples in this paper are drawn from a single especially large, especially complex project, which is current and familiar to many *Journal of Geoscience Education* readers: the InTeGrate project. InTeGrate is a 6 y effort to improve teaching and learning about Earth in higher education across the United States. "Improve," in this context, connotes both better pedagogy and a closer connection to real-world problems facing society. InTeGrate works at the course scale to change what faculty are prepared to do; at the department and institutional scales to change what faculty are asked to do; and at a national scale to change values, networks, and available resources. InTeGrate's goals, program elements, products, leadership structure, and funding source are detailed at InTeGrate (2017a).

USING SYSTEMS THINKING IN PROGRAM DESIGN

Most geoscientists spend their days trying to understand, describe, and perhaps predict the behaviors of an existing complex system by probing its structures and mechanisms. To become an education reformer requires a shift in role, from observer/explainer to designer/creator. Some readers who have worked in applied fields such as hydrology or soil science may have experienced the process of designing an intervention in the workings of an Earth system. For other readers, this shift in role may be a bigger jump. How do you get past the blank computer screen or blank white board to design a new program? Strategies that we have found useful at the design phase include articulating a shared vision, purposefully selecting the components of the larger system to target, planning for multidirectional interactions among the components, leveraging the power of replication, and designing nonlinear chains of influence.

Articulate a Compelling Vision and Theory of Change

To accomplish transformational change on a substantial scale, within realistic constraints of time, money, and human capacity, highly leveraged interventions are required. Meadows (1999, 2008) identified a hierarchy of 12 leverage points to intervene in a system, stressing that changing components or parameters is less impactful than changing the relationships among components, and that changing the goals or mindset of the system is more powerful still. A compelling vision can attract innovators and early adopters to a nascent program. The goal or vision needs to be sufficiently big and new as to be inspiring, but sufficiently well aligned with the earlier goals of the system actors as to seem achievable. Articulating a theory of change may help the system actors envision the path from first steps toward ultimate goal. As the program matures, the shared vision can attract a broader group of users or participants and then can help to unify this more diverse group in support of the work (Preskill et al., 2014b; Kezar and Gehrke, 2015).

Traditionally, the goals of science education have been to provide skilled professionals and technicians needed for a technology-infused economy (Bush, 1945; National Research Council, 2005), to help students get good jobs in science, technology, engineering, and math (STEM) careers (e.g., Casey, 2012), and to enable students to appreciate and understand the world (e.g., Herschbach, 1996). InTeGrate joined with earlier voices (e.g., MacGregor et al., 2007; Sherman, 2008; Burns, 2010; Ramaley, 2014) in articulating a fourth goal: to prepare problem-solvers with the ability and disposition to tackle profound societal challenges of the 21st century. InTeGrate's vision involves transforming Earth education such that learning occurs in the context of societal issues across the undergraduate curriculum and engages all students. InTeGrate seeks to prepare an Earth-savvy workforce and citizenry able to address urgent, complex problems, such as ensuring access to sufficient energy resources without destroying the environment.

To build this shared vision, InTeGrate planned to provide (1) a compelling articulation of the new vision, (2) attractive and adaptable exemplars of pathways toward the vision, and (3) a supportive, collegial community where the new vision is the norm. Attention to these three attributes is characteristic of STEM reform initiatives that have succeeded in creating new and innovative cultures (Kezar and Gehrke, 2015; Gehrke and Kezar, 2016).

InTeGrate's theory of change conjectures that educating Earth-savvy problem-solvers has the potential to lead to a cascade of desirable changes. As summarized in Fig. 1, teaching about Earth in the context of societal issues would make explicit the role of geoscience in society, interest more students in learning about Earth, and extend opportunities to learn about Earth to new populations. These changes in turn would lead to a citizenry with sufficient Earth literacy to make sustainable choices in their personal lives, a workforce with the skill set to build sustainable structures and processes throughout the economy, and increased capacity throughout society to capitalize on insights from Earth, physical, and social sciences in tackling the environmental challenges of the 21st century.

Although we have shared InTeGrate's mission and theory of change in some detail, the take-home message should be the passion and ambition of InTeGrate's goal, not the specifics of what InTeGrate is trying to accomplish.

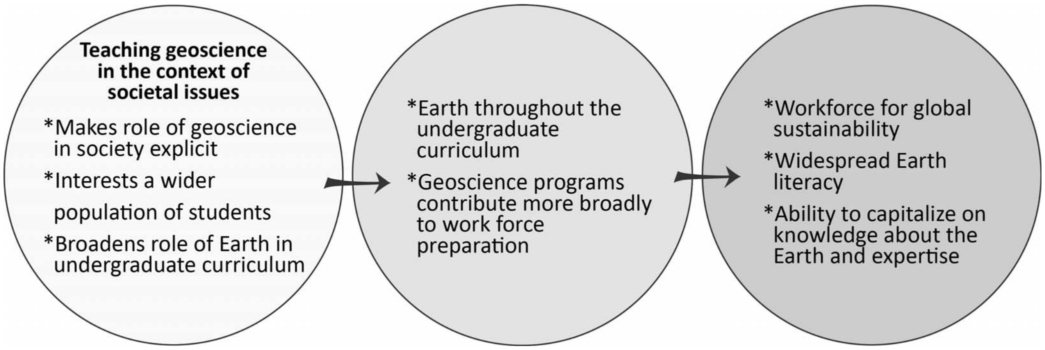


FIGURE 1: Cascade of beneficial consequences envisioned by InTeGrate's theory of change.

Systems thinking can be helpful *regardless* of the goal of an intervention. Whatever the goal, the vision that is articulated will need to be strong enough to attract and motivate actors and clear enough to guide them as they take action.

Select the Components to Target

If a system is "a set of components whose interconnections determine their behavior" (Climate Interactive, 2016), then early design decisions must address those components and interactions to target. Enough components are needed to be able to achieve impact through scale or synergy, but no more than necessary, as each component adds cost and leadership burden. Think carefully about which components to include and your reasons for so doing, as well as the scope of each component. Be equally clear-minded about what to leave out.

InTeGrate targeted three components for intervention: faculty teaching courses; programs or institutions controlling changes at scales larger than a course; and the community of interested educators. This appeared to be the smallest set of interventions that would allow transformation of the system and result in pervasive new types of learning experiences for students that would persist over time.

Component 1: Courses

In American higher education, faculty members have broad powers over the content and approach of their courses, and thus working with faculty to teach Earth-related issues in a societal context was essential. In addition, InTeGrate aimed to improve teaching quality, tapping into a wealth of recent research on effective teaching through student-centered pedagogies (Freeman et al., 2014; Kober, 2015). InTeGrate's materials development activity was designed to create a set of tested materials (InTeGrate, 2017c) and a cadre of faculty who could lead community-wide efforts for adoption. Both the materials and the processes used to create them could then be adapted or adopted by others. To accomplish desired pervasive transformation of teaching and learning about Earth, it was essential that InTeGrate focus not only on geoscience courses but also extend this effort to courses across the entirety of the undergraduate curriculum.

Component 2: Programs and Institutions

Lasting, pervasive change requires work at the department, program, or institutional level (Tobias, 1992; Seymour, 2002), the levels where priorities are set and decisions are made about which courses and degrees to offer. Productive pedagogic or content changes within courses must be endorsed, adopted, or rewarded by others in the institution if they are to transition from research into practice (Fry and the Coalition for Reform of Undergraduate STEM Education, 2014) and expand beyond the work of a single pioneering faculty member. Further, work to broaden participation in STEM fields has demonstrated that increasing the diversity of students requires attention at the programmatic level (Seymour and Hewitt, 1997; Crosling et al., 2008; Engstrom and Tinto, 2008) to promote student motivation and enthusiasm, to provide academic supports, and to cultivate a sense of belonging. InTeGrate's Implementation Program (InTeGrate, 2017b) activity was designed to provide incentives, resources, and coaching for programs to customize activities at their own institution aligned with InTeGrate's overarching goals.

Component 3: Nationwide Community

The third target for InTeGrate intervention was motivated by several intersecting needs: to motivate and support change, to shift values and norms, and to expand the radius of impact. Our belief was that individuals embedded in a community that shared values about the importance of learning about Earth and the role of that knowledge in addressing resource and environmental issues would be empowered and energized to act on those values. As this community became more visible, it would legitimize these values and attract new members, who would, in turn, find support for changing their behavior (Kezar and Gehrke, 2015; Gehrke and Kezar, 2016). While the entirety of the InTeGrate program was designed to promote interaction and learning within networked communities, the professional development program played a key role in establishing and extending this community.

Not Targeted

There are many other components of the education system that impact how teaching and learning about Earth are carried out in higher education. For example, administrators set promotion and tenure criteria, accreditors set cross-institutional requirements, and K–12 educators set students' expectations, all of which impact Earth education. InTeGrate's design did not explicitly address these elements because, although important, they were peripheral to our area of greatest leverage. An early task for a systems thinker in any new context is to draw the boundaries of the system under consideration so as to allow development of a focused and tractable design.

Drive Interaction Between Program Components So As To Be More Than the Sum of the Parts

Systems thinking requires a focus not only on components of a system but also on the interactions between them (Holland, 1998; Meadows, 1999; Hargreaves, 2010; Preskill et al., 2014b). It is these interactions that have the potential to lead to emergent phenomena, like the changes in values and teaching practices that InTeGrate seeks to foster. Examples of interactions between components include: one component creates a body of knowledge or a resource that is used by another component; activities in one component create alliances or working relationships that can be used to launch an activity in a different component; one component identifies needs that inform the trajectory of actions by another component. Mechanisms supporting interactions were designed into the InTeGrate program. For example:

Community Activities Underpin Materials Development

Previous National Science Foundation–funded programs and work by individual faculty had created a wealth of educational resources for teaching about Earth with student-centered pedagogy, as well as a cadre of instructors experienced in deploying reformed teaching practices with undergraduates (Manduca et al., 2010). The first 3 y of InTeGrate's community activities brought this work, and the people who had accomplished it, together with those who were interested in developing new instructional materials supportive of InTeGrate's goals. A series of early workshops (e.g. Gosselin, et al., 2015) gathered resources and allies, and sets of associated Web sites (Narum and Manduca, 2012) were established to capture, organize, and share the gathered insights (InTeGrate, 2017d). The paired workshops and Web sites allowed us to capitalize on the momentum of others without reinventing the wheel (Kania et al., 2014).

Materials Development and Implementation Programs Build Community

InTeGrate's materials development model requires developers to work in teams of three to six faculty members drawn from different institutions and different disciplines. Requiring such diverse teams introduces hurdles in terms of logistics, communication, and conflicting priorities. However, InTeGrate prioritized forming such teams both because it was thought to lead to more widely useable materials, and because a prolonged effort to overcome shared challenges and build a successful product is a way to build enduring collaborative ties across institutions and disciplines (Kezar, 2013). Similarly, Implementation Programs were required to span a department, multiple departments, or multiple institutions, weaving yet more tendrils into the growing network.

The casual observer of InTeGrate or another designed complex system might think it is a matter of luck or happy coincidence that a useful resource (for example, a Web site on teaching geoscience habits of mind) just happens to be available when it is needed. However, it is not a coincidence; it is purposeful consequence of good system design, in which the output of one activity becomes the input for a subsequent activity. Envision the traditional Chinese rice-fish cocultivation system (Xie et al., 2011) in which the foraging activities of the fish benefit the rice plants, and chemical uptakes by the rice improve the water quality for the fish. Keep this vision of mutually beneficial interaction between components in mind as you design your system.

Leverage Replication and Maximize Adaptability by Using Parallel Subsystems

Since at least the time of the Roman legions, leaders responsible for ambitious enterprises have recognized the value of structuring their organizations as a parallel set of subsystems. Such a structure allows for scale-up via replication. Procedures, parameters, information flow pathways, and feedbacks developed and refined for the first few elements can be replicated for multiple parallel elements, without greatly increasing the burden on the infrastructure or leaders. Parallel elements can learn from each others' experiences and mentor newly formed elements, accelerating progress. At the same time, each subsystem can be responsive to its local context, allowing for evolutionary responses to local challenges and opportunities.

Each of the major activities of InTeGrate is structured as a parallel set of subsystems: a set of materials development teams, a set of implementation programs, a set of workshops with companion Web pages. Spinning up the first few instantiations of each subsystem type required vast investments of time, energy, and creativity on the part of leadership and staff, but once each template was well established, subsequent instantiations were much less demanding. Designing InTeGrate as a system of subsystems also provided many leadership opportunities, making it possible to utilize the talents, time, and energy of an army of dispersed, busy individuals.

Design for Nonlinear Causality Chains Rather Than Single Cause→ Single Effect

An older class of education improvement projects tended to focus on a single intervention that would lead toward a single measurable outcome via a fairly linear logic model. Geoscientists will recognize that complex systems rarely work in this linear cause → effect way. In both natural and human-built complex systems, a single action or activity is likely to have multiple consequences, and, conversely, a single desirable outcome is likely to involve multiple contributory causes. For example, decreasing fossil-fuel use leads to less acid rain, less greenhouse warming, and also less employment in extractive industries. Healing the atmosphere's ozone hole involved actions by policymakers, manufacturers, technologists, and consumers. We encourage geoscientists to channel their intuition for complex causality relationships and avoid getting strong-armed into postulating a simplistic cause → effect relationship for an intervention. Instead, we promote the design of a web of actions that push toward desired outcomes.

InTeGrate's design presumes that any given action will have multiple outcomes or consequences. For example, workshops early in the project were designed to achieve four outcomes simultaneously: increased educational capacity of the attendees, increased knowledge base for the project, a more robust community of practice, and recruitment of new allies and leaders into the InTeGrate effort. Codevelopment of instructional materials by three- to four-person teams from different institution types was designed to result in both materials that were not tied to a specific context and enduring collegial relationships within a growing community of practice. Designing activities with multiple beneficial outputs is critical to achieving broad and deep impact from finite resources.

Conversely, InTeGrate's design conjectured that any given desirable outcome would require multiple nudges or influencers rather than a single cause. For example, to create instructors who have internalized InTeGrate values and methods, InTeGrate put in place face-to-face and virtual interactions among teams of materials developers, assessment consultants, and a team leader; provided support through Webinars and on-line materials; and required individual and group reflection on the development process, all in addition to developing and deploying a rubric (InTeGrate, 2016) that articulated and reinforced InTeGrate's pedagogical approaches and priorities.

Advantages of Designing with a Systems Approach

Designing with a systems approach maximizes the chances that an ambitious intervention will be more than the sum of its parts. From the beginning, geoscientists and education reformers can establish the expectation that different parts of the project will both benefit from and contribute to one another, and plan for synergies and efficiencies. Although some missteps are inevitable in any complex undertaking, systems-oriented design can help minimize dead-ends and nonproductive efforts and maximize the run-time available for the most productive activities. Through shared vision and purposeful planning of interactions and influences, the varied components of your program can begin pulling in the same direction.

USING SYSTEMS THINKING IN PROGRAM IMPLEMENTATION

Literature on a systems approach provides strong tactical guidance for implementation (e.g., Senge, 2006; Meadows, 2008). We focus here on strategies for building "system fitness," the ability of a human system to change itself so as to improve performance (Kania et al., 2014). This section describes strategies for managing information flow and monitoring the state of the system, nudging actors toward desired behaviors, and responding to emergent challenges and opportunities.

Manage Flows of Information Between and Within Program Subsystems

A challenge for any large project is figuring out how to move information in ways that allow informed decision making without slowing work to a standstill or overwhelming the capacity of the recipients to take in information. Project leadership needs synoptic information to steer the enterprise and respond to emerging challenges and opportunities. Actors throughout the system need actionable information in a timely manner to make strong decisions as they manage their local responsibilities. Each team or subsystem needs a vigorous internal system of communications, plus access to a strategic subset of the information about what is going on elsewhere in the project. Discoveries, insights, and best practices known in one subsystem can be reused in parallel subsystems—but only if they are shared effectively.

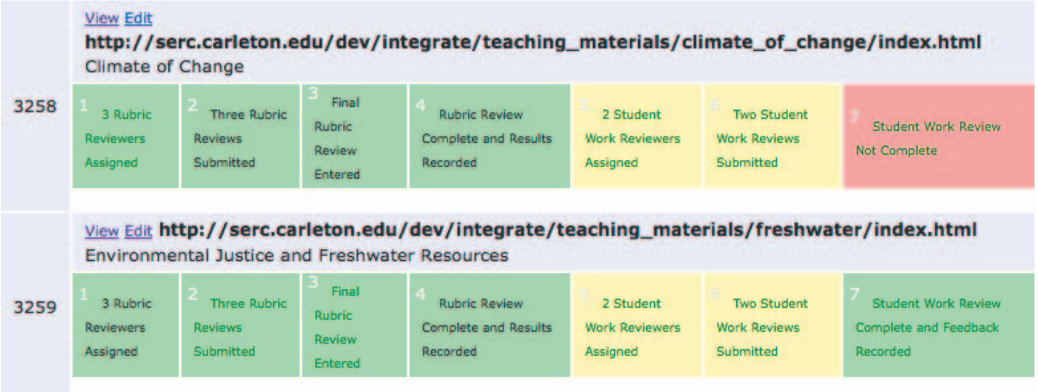
Information can be transmitted by means of artifacts and documents (e.g., rubrics, guidelines, templates, journal articles, Web sites), by human-to-human discussion (face-to-face, virtual-synchronous, virtual-asynchronous), and by moving people (with their embodied expertise) around the system. InTeGrate uses all of these modalities, often in tight combination.

Much information transfer takes place through the InTeGrate Web site (serc.carleton.edu/integrate), which provides both private and public community spaces (Kezar and Gehrke, 2015). All InTeGrate subsystems contribute their accumulating insights and efforts to building this one mega-artifact. For the geographically distributed builders of InTeGrate, the Web site supports the development of a collaborative team by providing a "place" to "convene" and share ideas, plus a visible artifact depicting shared progress.

One always-needed form of information is the state of the system (Kania et al., 2014). System-oriented decision makers continually compare the perceived state of the system with the desired state (goal) and plan their next steps to nudge the system toward the desired state (Meadows, 1999). InTeGrate has many mechanisms to sense the state of its system, including assessments, surveys, interviews, and reflective writings. Going beyond the usual written status reports, InTeGrate has built a suite of Web-based tools, archives, databases, and workspaces that allow team members at different institutions to work collaboratively and keep track of what is going on throughout the system. A dashboard of Web pages that record and update the status of different parts of the system is used to monitor progress, plan workflow, spot problems, and support decision making (Fig. 2).

An artifact can be used to crystallize and communicate a set of decisions or values. For example, the Materials Development and Refinement Rubric (InTeGrate, 2016) embodies InTeGrate's pedagogical values and materials development priorities as crystallized through extensive debate among the leadership team. Assessment Team members use this rubric to communicate expectations to Materials Development teams. A member of the Assessment Team regularly evaluates each set of materials and provides feedback, nudging the materials toward alignment with the rubric. In Meadow's (1999) terminology, a module development team and the associated Assessment Team member form a "self-organizing subsystem" that can create a whole new structure (the module or course) following a set of rules set out by the leadership, but without active involvement of the leadership once the process is established and debugged. There is vigorous detailed communication going on constantly within this self-organizing subsystem, but only a small fraction of that communication bandwidth is used to pass information across the subsystem boundary to and from other subsystems.

InTeGrate's final mode of transferring information from one subsystem to another is by moving a knowledgeable human being. Workshop participants who showed enthusiasm and insight were encouraged to form Materials Development teams, bringing with them the knowledge gathered and created at the workshop. Materials developers who showed leadership potential were encouraged to nucleate new Implementation Programs at their home institutions or coconvene new workshops in their area of expertise. Movement of individuals from one project element to another provides transfer of a level of flexible and nuanced expertise that cannot easily be captured in written documents or short interactions.



3258	View Edit **http://serc.carleton.edu/dev/integrate/teaching_materials/climate_of_change/index.html** Climate of Change						
	1 3 Rubric Reviewers Assigned	2 Three Rubric Reviews Submitted	3 Final Rubric Review Entered	4 Rubric Review Complete and Results Recorded	5 2 Student Work Reviewers Assigned	6 Two Student Work Reviews Submitted	7 Student Work Review Not Complete
3259	View Edit **http://serc.carleton.edu/dev/integrate/teaching_materials/freshwater/index.html** Environmental Justice and Freshwater Resources						
	1 3 Rubric Reviewers Assigned	2 Three Rubric Reviews Submitted	3 Final Rubric Review Entered	4 Rubric Review Complete and Results Recorded	5 2 Student Work Reviewers Assigned	6 Two Student Work Reviews Submitted	7 Student Work Review Complete and Feedback Recorded

FIGURE 2: This screen-shot shows a portion of the webpage that records the progress of instructional modules under development. Green, yellow and red colors indicate the status of the module with respect to a set of seven checkpoints (color available in the online journal). InTeGrate's suite of webpages and web-based tools comprise a "dashboard" that allows distributed actors to coordinate their activities and the leadership team to continuously sense the state of the system.

Leverage Feedback Loops

One of the mechanisms that enables systems to "form a whole that is greater than its parts" (Hargreaves, 2010, 3) is the occurrence of feedback loops. Balancing (i.e., "negative") feedback loops rein in departures from the path toward project goals before they escalate. Reinforcing (i.e., "positive") feedback loops nudge project participants toward actions that align with project values and priorities or that are observed to be effective. Positive feedback loops enable exponential growth and thus enable large impacts from limited resources.

Figure 3 depicts a nested set of two balancing (i.e., negative) feedbacks mechanisms in InTeGrate's materials development program. InTeGrate requires that Materials Development teams include faculty from different types of institutions, which is supposed to result in materials that are suitable for a wide range of students, which in turn is supposed to result in materials being widely adopted/adapted (the straight-line path across the top of Fig. 3). However, this ideal may go astray (dashed line of Fig. 3) if one team member pushes for activities that are only suited for a particular context (e.g., student background, locale, institutional mission). In such a case, as depicted in the inner loop of Fig. 3, the other team members are empowered to pull the development effort back onto a track that will reach more types of students, venues, or institutions. The outer loop of Fig. 3 shows a second feedback that is intended to kick in if overly narrowly targeted materials do make it

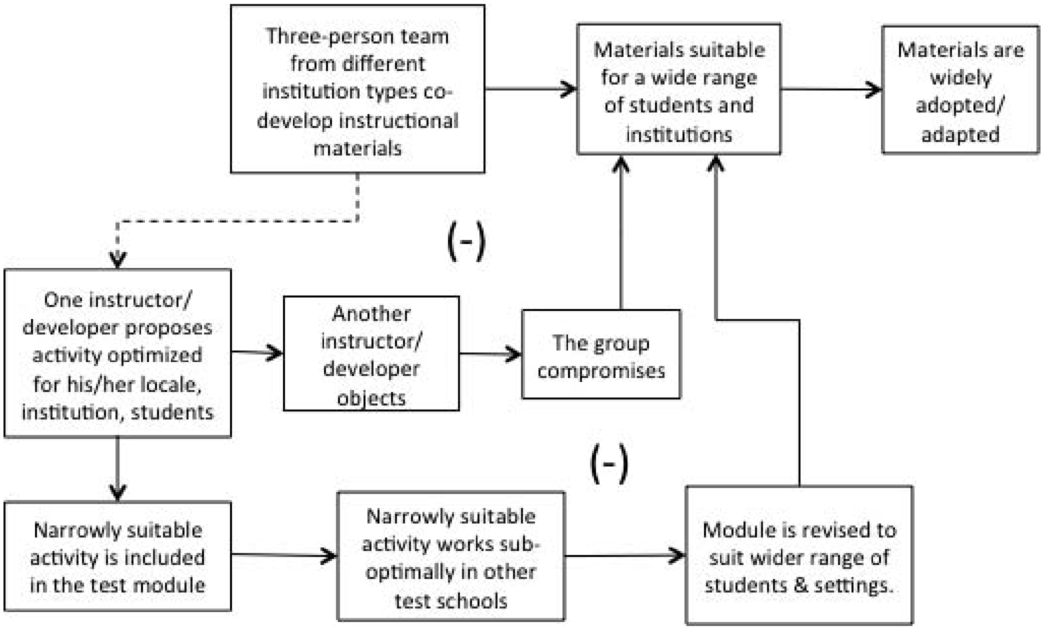


FIGURE 3: Negative (or "balancing") feedbacks can be used to keep a system near a desired state. This systems map depicts a pair of nested feedback mechanisms that are intended to nudge the InTeGrate materials development system toward creating materials suitable for a wide range of students and institutions. The inner loop acts at the design stage, and the outer loop acts at the testing/revision stage of the development process.

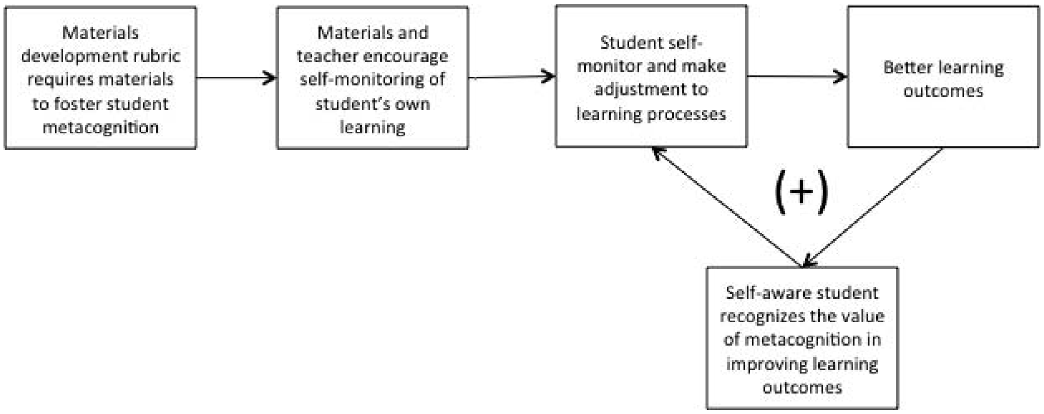


FIGURE 4: Positive (or "reinforcing") feedback loops tend to move a system farther away from its initial state. In this example, both teachers and materials encourage students to self-monitor their own learning, which research shows can lead to better learning outcomes, which in turn can lead to student awareness of the value of metacognition, which in turn can lead to more metacognition and even better learning outcomes. Multiple passes through this reinforcing feedback loop nudge the student toward a more self-aware, self-correcting, self-educating learning style.

through the development stage. All developers are required to test the newly developed materials in their own institutions. During these pilot tests, overly narrow materials will prove to be problematic when tested in the other two institutions, and at the revision stage, the materials will be tweaked back onto the desired path toward widely useable materials.

Figure 4 shows an example of a reinforcing (i.e., positive) feedback loop, working at the level of the student. InTeGrate's materials development rubric (InTeGrate, 2016) places a premium on having students engage in metacognition (Bransford et al., 2000), including having them monitor their own learning process. This is expected to lead to better learning outcomes. The self-monitoring student recognizes that metacognition has been of value in achieving better learning outcomes and thus applies the metacognitive and self-monitoring approaches more widely in his/her studying and learning, which in turn can lead to even better learning outcomes, moving the student away from his/her initial state toward a new state of more effective learner.

Deploy Emergent Strategies to Respond to Challenges and Opportunities

In the division used by Kania et al. (2014) of social interventions into simple, complicated, and complex, one of the attributes of a complex intervention is that it cannot be completely planned in advance using best practices and information that is in hand during the design phase. Instead, information obtained through sensing of the system and its environment must be used in near-real-time to design "emergent strategies" that will capitalize on opportunities and manage challenges. If you have worked in natural resource conservation, you may know this approach as "adaptive management" or "adaptive environmental assessment and management" (Conservation Measures Partnership, n.d.)

As an example of an emergent strategy used to capitalize on an opportunity, the materials development rubric (InTeGrate, 2016) has found use well beyond its intended role of guiding InTeGrate Materials Development teams. It is now being used by other curriculum development projects (e.g., GETSI, 2016) and as a tool for faculty professional development within InTeGrate. In other words, the rubric has become a flow pathway to spread InTeGrate-endorsed pedagogical values beyond the community of individuals directly funded by InTeGrate. Maximizing this flow pathway required sensing the first new uses of the rubric and then adapting project activities to promote this expansion of use.

Similarly, InTeGrate has shown resilience in the face of challenges. For example, the first effort to foster and assess students' mastery of systems thinking showed that instructors were not sure how to teach this topic and that the systems-thinking assessment essay question was yielding only superficial answers. The project responded with a major collaborative effort to revise and test new essay questions, develop a systems thinking Webinar for materials developers, and recruit developers for a dedicated systems-thinking module.

These changes to the plan could be viewed as jury-rigged patches over broken system components. Kania et al. (2014) would encourage us instead to regard them as emergent solutions, redeploying resources after learning by doing, and an essential way of working when tackling a complex problem. Note that all of these invented-in-real-time solutions to emerging challenges are evolutionary in nature—like the solutions arrived at by biological evolution, they repurpose structures and processes that were already in place and adapt them to new purposes. Webinars, module development procedures, and collaborative development of assessment items were already in the InTeGrate toolkit, and one or more of the available tools was pulled forth and adapted to address each emerging opportunity or challenge.

Advantages of Implementing with a Systems Approach

Mindful use of a systems approach during implementation improves the chances that information will flow to where it can be used, and that the local actors in the system will be able to make sense of incoming information and act upon this knowledge constructively. Building "system fitness" by shaping information flows and building local capacities, in turn, increases the chances that the system will be resilient in the face of emerging challenges and proactive when presented with emerging opportunities. Emergent

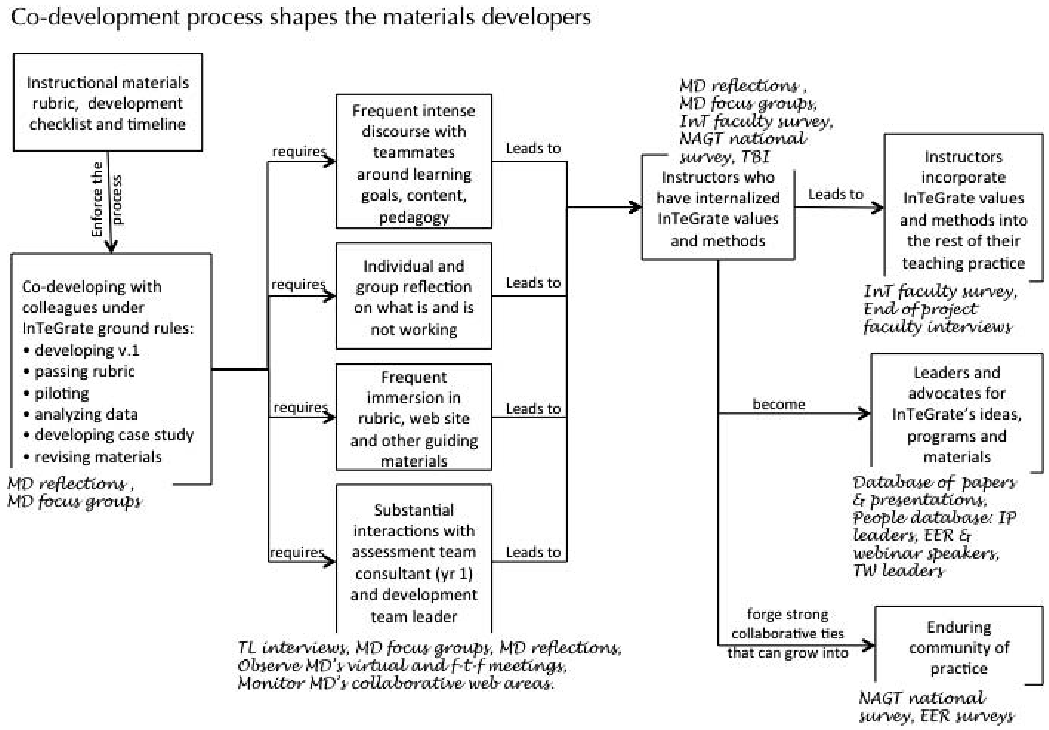


FIGURE 5: A segment of a system map as used for evaluation planning. Italicized lettering marks where the evaluative probes have been inserted. The evaluation moved left to right across the diagram as the program matured. Abbreviations: EER = Earth Educators' Rendezvous; InT = InTeGrate; MD = materials developers; NAGT = National Association of Geoscience Teachers; TBI = Teacher Belief Inventory; TL = Materials Development Team Leaders; TW = travelling workshops.

solutions can develop at all scales, from an individual to an institution. In aggregate, it is the coming together of preplanned capacities plus evolving solutions that enable the emergence of the desired system-level outcomes.

USING SYSTEMS THINKING IN PROGRAM EVALUATION

Using Systems Mapping to Reveal Key Connections, Relationships, and Interdependencies

Evaluators often use a simple kind of tabular "logic model" to summarize the inputs, activities, and expected outcomes of a project. Geoscientists, who are accustomed to reasoning with complicated graphical depictions of natural systems (for example, the Bretherton diagram for climate systems; Earth System Sciences Committee and NASA Advisory Council, 1986), may benefit from using a more complete and intricate form of systems map that includes interactions and dependencies (Preskill et al., 2014a). The type of systems maps we are suggesting differs from the classic logic model in that the systems maps (1) showcase flows and linkages and specify their nature, (2) allow, indeed encourage, depiction of branching, recursion, and bidirectional flows, and (3) are presented as working tools and hypotheses to be tested, rather than as a blueprint that the project should follow to reach a prespecified goal. Evaluators can use such a map to identify the points at which evaluative probes would be most informative, just as a hydrologist would use a map of streams and reservoirs in a watershed to plan out where to insert stream gauges.

Early in the InTeGrate project, the evaluation team and leadership developed overarching system maps for the entire project to depict the design conjectures (Sandoval, 2014) embodied in the proposal. As the project implementation proceeded, we developed more localized systems maps for functional subsystems of the project (e.g., Figs. 3 and 4), and ground-truthed the conjectures with empirical observations. A full set of systems maps can be found at: http://d32ogoqmya1dw8.cloudfront.net/files/integrate/about/integrate_systems_maps.pdf.

Figure 5 shows an example of how system mapping helped to plan and situate InTeGrate's evaluative probes. InTeGrate's materials development effort was initially presented to the community as a mechanism to create pedagogically and scientifically excellent instructional materials and was slated for evaluation through measures of student learning gain. As the system was developed and mapped, however, it became clear that equally important outcomes from the process would be increasing numbers of instructors/developers who incorporated InTeGrate values into their teaching practice, who became advocates for InTeGrate's ideas, programs, and materials, and who became part of an enduring community of practice. We therefore developed a set of four guided reflections for the materials developers to record their experiences and attitudes and a program of interviews with team leaders and focus groups with the materials developers, and we inserted questions probing these aspects of the project into surveys and end of project interviews.

Evaluate for Emergent Phenomena

In systems thinking, emergence is the process by which larger entities, patterns, and regularities arise through interactions among smaller or simpler entities that themselves do not exhibit such properties (Johnson, 2002). The classic example is the coordinated movement of a school of fish or a flock of birds even in the absence of a leader. When intervening in a complex system, important desired outcomes are often emergent phenomena, in that they result from a myriad of interactions and processes among and within smaller subsystems. Even if the contributing interactions and processes cannot be quantified, it can still be valuable to measure the desired emergent phenomenon, while remaining clear-eyed about the complexity of the contributing interactions.

An example within InTeGrate is the desired outcome that more students, nationwide, will flock into STEM college majors and the STEM workforce. InTeGrate's materials development rubric (InTeGrate, 2016) does not require that the materials encourage interest in STEM careers, and the materials developers are not given this as part of their explicit responsibility. Instead, a high-level design conjecture of InTeGrate is that providing students with access to student-centered pedagogy and situating instruction in the context of societal problems of concern to their generation will result in students gravitating toward courses with Earth-related content, geoscience majors, and Earth-related careers. The evaluation team is monitoring the validity of this design conjecture though a pre- and postinstruction survey that asks students about their career interests—even though this is not an explicit learning goal of the instructional materials, and even though we do not necessarily expect to be able to disambiguate the contributions of different factors.

Evaluate for Lasting Changes to the Larger System

Thus far in the paper, we have been considering each newly designed program as a complex system. However, the system of higher education as a whole is also a complex system, within which each program is a subsystem. One way to conceptualize the long-term outcomes of each program is as lasting additions or changes to the structure of this larger system.

For example, the theory of lasting change of InTeGrate and its predecessor On the Cutting Edge (Manduca et al., 2010) emphasizes the development of a community of practice (Lave and Wenger, 1991) of Earth educators that can multiply the effectiveness of funded activities and persist after the initial catalytic funding sunsets. A community of practice is a "group of people who share a concern or passion for something they do and learn how to do it better as they interact regularly" (Wenger-Trayner and Wenger-Trayner, 2015, 1). Development of such communities has been a common thread across other STEM education reform projects that have achieved lasting and widespread change in American higher education (Kezar and Gehrke, 2015; Gehrke and Kezar, 2016). InTeGrate's materials development process, workshops, Webinars, Web site, Earth Educators' Rendezvous, and Implementation Programs are all intended to contribute to the growth of communities and networks of practice, both at small scales and on a national scale.

To evaluate the robustness and effectiveness of InTeGrate's community of practice, the evaluation team developed a conceptual systems dynamics model of how a community of practice can synergistically build the capacity of both individuals and the community through a system of three reinforcing feedback loops (Kastens, 2016b). In this model, individuals accrue both practical and affective value from their involvement in the community. We then used interviews and surveys to query participants at InTeGrate events about new collaborative ties formed at the event, about the balance between what they contributed to and received from the event, and about their sense of belonging to a community of educators who share their interests and concerns (Kastens, 2015, 2016a). We are using these metrics to probe for early indicators that InTeGrate is succeeding in creating a lasting change to the system of high education, an enduring and impactful "community of transformation" (Kezar and Gehrke, 2015).

Advantages of Evaluating with a Systems Approach

Traditional approaches to evaluation do not scale or adapt well to complex projects with interacting parts and emergent outcomes (Hargreaves, 2010). Systems maps allow us to understand and monitor the changing project and adapt the evaluation strategy to most effectively provide formative feedback. Systems thinking enables us to tease out where the desired emergent phenomena should first be detectable, and to probe for early indicators of lasting changes to the larger system.

TRANSFERABLE STRATEGIES, CAVEATS, AND FINAL THOUGHTS

In this paper, we have made the case that systems thinking provides a productive lens for design, implementation, and evaluation of large-scale educational innovations. As geoscientists and educators address complex, wicked problems in education and beyond, the strategies described above can help to initiate and accelerate desirable changes in stubborn systems. Every project is different, and not all strategies will be helpful in a particular program. Systems-inspired approaches that we think have the widest applicability include:

- the division of the effort into subsystems, each with its own goals, processes, and internal communication flows, and each self-organizing within ground rules set by the central project;
- the use of parallel subsystems, with enough commonalities to foster efficiency and yet enough differences to allow adaptability to local circumstances;
- the use of technology-enabled dashboards and communication flows to support distributed decision making;
- the presence of synergistic relationships between workshops and Web sites, or more broadly, between human-mediated and artifact-mediated forms of interactions;
- use of reinforcing and balancing feedback loops to nudge actors toward desired behaviors and away from problematic choices;

- an evaluation that probes lasting structural changes to the larger system.

For additional suggestions, we recommend the writings of Meadows (1999, 2008), Senge (2006), and Kezar and Gehrke (2015), authors who combine pragmatism with inspiration.

A few caveats are in order as well. Systems thinking is a powerful tool, but it is not a panacea. Growing your leadership capacity in educational reform will require a myriad of skills that systems thinking will not help with, such as budgeting, diplomacy, personnel, Institutional Review Boards. The change of vantage point from a scientist seeking to understand a complex system to a change agent seeking to improve a complex system can be jarring. Whereas the scientist's radar is set to hone in on places in the system where there are mysteries to unravel, the change agent's radar seeks out places where there is leverage to shift the system (Meadows, 1999). Human systems involve a new set of drivers—such as goal-seeking toward prestige, acclaim, security, friendship, good grades, or promotion—that differ profoundly from the drivers found in physical systems. The language (both verbal and graphic) of systems thinking may be unfamiliar to collaborators and administrators; if so, using systems reasoning to explain your plans may initially hinder rather than aid communications. You will need to keep in mind not just a lot of different actors and entities, but a lot of different relationships and interactions between actors and entities.

However, you are more prepared than you may think to become an education reformer. As a geoscientist, you are accustomed to dealing with bodies of evidence that combine the quantitative with the qualitative, and that assemble claims by combining multiple lines of evidence (Kastens, 2010; Manduca and Kastens, 2012). It is second nature for you to expect that there are likely to be multiple contributing processes combining to create one observable outcome (Chamberlin, 1890). You do not expect change to happen instantly, but are imbued with the realization that slow processes acting across long time spans can accrue monumental impacts. Relatedly, you accept that systems exhibit time lags, in which minimal change is followed by sudden change, as in an avalanche or a student's mastery of a threshold concept (Stokes et al., 2007). As a geoscientist, you are likely to be an experienced interdisciplinary collaborator (Manduca and Kastens, 2012). Finally, with Earth at risk from environmental degradation and resource depletion, you may find that your motivation to effect change in education about Earth gives you the courage to try new approaches and the endurance to overcome obstacles.

Acknowledgments

This work was funded by the U.S. National Science Foundation STEP Talent Expansion Program through grant DUE-1125331 to Carleton College. The authors thank the InTeGrate Leadership Team and Advisory Board for their insights and suggestions. Thoughtful reviews from Abigail Jurist Levy, Judith Ramaley, Stephanie Pfirman, and the journal editors and peer-reviewers greatly improved the manuscript. Manduca is indebted to the many colleagues who have given her insights, experiences, and opportunities to learn about natural and social complex systems, particularly David Mogk, Heather Macdonald, Michael Mayhew, and Barbara Tewksbury. She in indebted to John Maclaughlin for his introduction to logic models and their use in program design.

REFERENCES CITED

Austin, A.E. 2013. Barriers to change in higher education: Taking a systems approach to transforming undergraduate STEM education. White paper prepared for The Coalition for Reform of Undergraduate STEM Education (CRUSE), p. 18–22. Available at http://www.aacu.org/sites/default/files/u7/PKAL/CRUSE_Foundation_Workshop_2013_White_Papers.pdf (accessed 27 June 2017).

Bergquist, W.H. 1992. The four cultures of the academy. San Francisco, CA: Jossey-Bass Inc., Publishers.

Bransford, J.D., Brown, A.L., and Cocking, R.R. 2000. How people learn: Brain, mind, experience, and school. Washington, DC: National Academies Press.

Burns, W.D. 2010. SENCER in theory and practice. *In* Sheardy, R.D., ed., Science education and civic engagement: The SENCER approach. Washington, DC: ACS Symposium Books, American Chemical Society.

Bush, V. 1945. Science: The endless frontier. *Transactions of the Kansas Academy of Science (1903–)*, 48:231–264.

Casey, B. 2012. STEM education: Preparing for the jobs of the future. *In* U.S. Congress Joint Economic Committee, April. Available at http://www.jec.senate.gov/public/_cache/files/6aaa7e1f-9586-47be-82e7-326f47658320/stem-education—preparing-for-the-jobs-of-the-future-.pdf (accessed 27 June 2017).

Chamberlin, T.C. 1890. The method of multiple working hypotheses. *Science*, 15:92–96.

Climate Interactive. 2016. The climate leader. Available at https://www.climateinteractive.org/programs/the-climate-leader/ (accessed 27 June 2017).

Committee on Facilitating Interdisciplinary Research. 2005. Facilitating interdisciplinary research. Washington, DC: National Academies Press.

Committee on Underrepresented Groups and the Expansion of the Science and Engineering Workforce Pipeline. 2011. Expanding underrepresented minority participation: America's science and technology talent at the crossroads. Washington, DC: National Academies Press.

Conservation Measures Partnership. n.d. The open standards for the practice of conservation: About open standards. Available at http://cmp-openstandards.org/about-os/ (accessed 20 May 2017).

Crosling, G., Thomas, L., and Heagney, M., eds. 2008. Improving student retention in higher education: The role of teaching and learning. London: Routledge.

Earth System Sciences Committee, and NASA Advisory Council. 1986. Earth System Science: Overview: A program for global change. Washington, DC: National Academies Press.

Engstrom, C.M., and Tinto, V. 2008. Learning better together: The impact of learning communities on the persistence of low-income students. *Opportunity Matters* 1:1–21.

Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., and Wenderoth, M.P. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23):8410–8415.

Fry, C.L., and the Coalition for Reform of Undergraduate STEM Education, eds. 2014. Achieving systemic change: A sourcebook for advancing and funding undergraduate STEM education. Washington, DC: Association of American Colleges and Universities.

Gehrke, S., and Kezar, A. 2016. STEM reform outcomes through communities of transformation. *Change*, 48(1):30–38.

Geodesy Tools for Societal Issues (GETSI). 2016. GETSI: Geodesy tools for societal issues. Available at http://serc.carleton.edu/getsi/index.html. (accessed 29 January 2016).

Gosselin, D.C., Burian, S., Lutz, T.M., and Maxson, J. 2015. Integrating geoscience into undergraduate education about environment, society, and sustainability using place-based learning: Three examples. *Journal of Environmental Studies and Sciences*, 3(3):316–330.

Hargreaves, M.B. 2010. Evaluating system change: A planning guide. Available at http://www.innonet.org/resources/files/eval_system_change_methodbr.pdf (accessed 27 June 2017).

Hawe, P., Bond, L., and Butler, H. 2009. Knowledge theories can inform evaluation practice: What can a complexity lens add? *New Directions for Evaluation*, 2009(124):89–100.

Herschbach, D. 1996. Teaching chemistry as a liberal art. *AAC&U Liberal Education*, 82(4):10–17.

Holland, J.H. 1998. Emergence: From chaos to order. Reading, MA: Addison-Wesley Publishing Company, Helix Books.

Hummelbrunner, R. 2011. Systems thinking and evaluation. *Evaluation*, 17(4):395–403.

InTeGrate. 2016. Working with the InTeGrate materials development and refinement rubric. Available at http://serc.carleton.edu/integrate/info_team_members/currdev/rubric.html (accessed 27 February 2017).

InTeGrate. 2017a. About the InTeGrate Project. Available at http://serc.carleton.edu/integrate/about/index.html (accessed 19 May 2017).

InTeGrate. 2017b. Implementation programs. Available at http://serc.carleton.edu/integrate/programs/implementation/index.html (accessed 27 February 2017).

InTeGrate. 2017c. InTeGrate authored modules and courses. Available at http://serc.carleton.edu/integrate/teaching_materials/modules_courses.html (accessed 27 February 2017).

InTeGrate. 2017d. Workshops and Webinars. Available at http://serc.carleton.edu/integrate/workshops/index.html (accessed 27 February 2017).

International Council for Science (ICSU). 2010. Earth System Science for global sustainability: The grand challenges. Paris: International Council for Science. Available at https://www.icsu.org/cms/2017/05/GrandChallenges_Oct2010.pdf (accessed 29 May 2017).

Johnson, S. 2002. Emergence: The connected lives of ants, brains, cities, and software. New York: Simon and Schuster.

Kania, J., Kramer, M., and Russell, P. 2014. Strategic philanthropy for a complex world. *Stanford Social Innovation Review*, 12(3):26–33.

Kastens, K.A. 2010, Multiple lines of reasoning in support of one claim. Earth & Mind: The Blog, Science Education Resource Center. Available at http://serc.carleton.edu/earthandmind/posts/multipledatatyp.html (accessed 27 June 2017).

Kastens, K.A. 2015, Weaving new threads into the GeoEd community of practice. Available at https://d32ogoqmya1dw8.cloudfront.net/files/integrate/about/2015_eer_interviews.pdf (accessed 27 February 2017).

Kastens, K.A. 2016a. Attendees' perceived balance between "getting" and "giving" at the 2016 Earth Educators' Rendezvous: Results from lightning interviews. Available at https://d32ogoqmya1dw8.cloudfront.net/files/integrate/about/lightning_interviews_report_from.pdf (accessed 27 February 2017).

Kastens, K.A. 2016b. Reinforcing feedback loops power effective communities of practice. Earth & Mind: The Blog. Available at http://serc.carleton.edu/earthandmind/posts/commofpract.html (accessed 27 February 2017).

Kezar, A. 2013. Understanding sensemaking/sensegiving in transformational change processes from the bottom up. *Higher Education*, 65(6):761–780.

Kezar, A., and Eckel, P.D. 2002. The effect of institutional culture on change strategies in higher education: Universal principles or culturally responsive concepts? *The Journal of Higher Education*, 73(4):435–460.

Kezar, A., and Gehrke, S. 2015. Communities of transformation and their work scaling STEM reform: Pullias Center for Higher Education, Rossier School of Education, University of Southern California. Available at http://www.uscrossier.org/pullias/wp-content/uploads/2016/01/communities-of-trans.pdf (accessed 3 February 2016).

Kober, N. 2015. Reaching students: What research says about effective instruction in undergraduate science and engineering. Washington, DC: National Academies Press.

Lave, J., and Wenger, E. 1991. Situated learning: Legitimate peripheral participation. Cambridge, UK: Cambridge University Press.

MacGregor, J., Middlecamp, C.H., Millar, S., and Seymour, E. 2007. White Paper: Marshaling the postsecondary STEM education community for significant student learning about global issues: Sketch for a proposed national endeavor. Available at mobilizingstem.wceruw.org/documents/White%20Paper%20National%20Endeavor.doc (accessed 27 February 2016).

Manduca, C.A., Iverson, E., Luxenberg, M., Macdonald, R.H., McConnell, D.A., Mogk, D., and Tewksbury, B. 2017. Improving undergraduate STEM education: The efficacy of discipline-based professional development. *Science Advances*, 3:e1600193.

Manduca, C.A., and Kastens, K.A. 2012. Geoscience and geoscientists: Uniquely equipped to study the Earth.. *In* Kastens K.A., and Manduca, C., eds., Earth & mind II: Synthesis of research on thinking and learning in the geosciences. Boulder, CO: Geological Society of America, Special Paper 486, p. 1–12.

Manduca, C.A., Mogk, D.W., Tewksbury, B., Macdonald, R.H., Fox, S.P., Iverson, E.R., Kirk, K., McDaris, J., Ormand, C., and Bruckner, M. 2010. On the Cutting Edge: Teaching help for geoscience faculty. *Science*, 327(5969):1095–1096.

Marcy, M.B., 2017, The small college imperative: From survival to transformation. White paper from the Association of Governing Boards of Colleges and Universities. Available at https://www.agb.org/sites/default/files/whitepaper_2017_small_college_imperative.pdf (accessed 29 May 2017.)

Meadows, D.H. 1999. Leverage points: Places to intervene in a system. Hartland, VT: Sustainability Institute, p. 1–19.

Meadows, D.H. 2008. Thinking in systems: A primer. White River Junction, VT: Chelsea Green Publishing.

Narum, J., and Manduca, C.A. 2012. Workshops and networks in Brainbridge. *In* Sims, W., ed., Leadership in science and technology, a reference handbook. Thousand Oaks, CA: Sage. p. 443–451.

National Research Council. 2005. Rising above the gathering storm: Energizing and employing America for a brighter economic future. Washington, DC: National Academies Press.

National Science Foundation. 2015. Improving undergraduate STEM education: Education and human resources (IUSE: EHR). Washington, DC: National Science Foundation.

North American Primary Care Research Group. 2009. Learning about complexity science. Available at http://gswong.com/?wpfb_dl=27 (accessed 27 June 2017).

Patton, M.Q. 2011. Developmental evaluation: Applying complexity concepts to enhance innovation and use. New York: Guilford Press.

Preskill, H., Gopal, S., Mack, K., and Cook, J. 2014a. Evaluating complexity: Propositions for improving practice. Available at http://www.fsg.org/publications/evaluating-complexity (accessed 31 October 2016).

Preskill, H., Parkhurst, M., and Juster, J.S. 2014b. Learning and evaluation in the collective impact context. Report from the Collective Impact Collaborative. Available at http://www.fsg.org/publications/guide-evaluating-collective-impact (accessed 28 February 2017).

Ramaley, J.A. 2014. The changing role of higher education: Learning to deal with wicked problems. *Journal of Higher Education Outreach and Engagement*, 18(3):7–22.

Roe, G. 2009. Feedbacks, timescales and seeing red. *Annual Reviews of Earth & Planetary Science Letters*, 37:93–115.

Seidman, A. 2005. Minority student retention: Resources for practitioners. *New Directions for Institutional Research* (special issue on minority retention), 2005:7–24. doi:10.1002/ir.136.

Senge, P.M. 2006. The fifth discipline: The art & practice of the learning organization. New York: Doubleday.

Sandoval, W. 2014. Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, 23(1):18–36.

Seymour, E. 2002. Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology. *Science Education*, 86(1):79–105.

Seymour, E., and Hewitt, N. 1997. Talking about leaving: Why undergraduates leave the sciences. Boulder, CO: Westview Press.

Sherman, D.J. 2008. Sustainability: What's the big idea? A strategy for transforming the higher education curriculum. *Sustainability*, 1(3):188–195.

Stokes, A., King, H., and Libarkin, J.C. 2007. Research in science education: Threshold concepts. *Journal of Geoscience Education*, 55:434–438.

Tierney, W.G. 1997. Organizational socialization in higher education. *Journal of Higher Education*, 68(1):1–16.

Tobias, S. 1992. Revitalizing undergraduate science: Why some things work and most don't. An Occasional Paper on Neglected Problems in Science Education. Tucson, AZ: Research Corporation, Book Dept.

Wenger-Trayner, E., and Wenger-Trayner, B. 2015. Communities of practice: A brief introduction. Available at http://wenger-trayner.com/introduction-to-communities-of-practice/ (accessed February 2016).

Williams, B., and Hummelbrunner, R. 2010. Systems concepts in action: A practitioner's toolkit. Stanford, CA: Stanford University Press.

Wilson, C. 2016a. Status of recent geoscience graduates. Report from the American Geosciences Institute. Available at https://www.americangeosciences.org/sites/default/files/ExitSurvey_2016_final.pdf (accessed 29 May 2017).

Wilson, C. 2016b. Status of the geosciences workforce. Report from the American Geosciences Institute. Available at https://www.americangeosciences.org/workforce/reports (accessed 29 May 2017).

Xie, J., Hu, L., Wu, X., Li, N., Uuan, Y., Yang, H., Zhang, J., Luo, S., and Chen, X. 2011. Ecological mechanisms underlying the sustainability of the agricultural heritage rice-fish coculture system. *Proceedings of the National Academy of Science of the United States of America*, 108: E1381–E1387.