Using systems thinking to design, implement and evaluate the InTeGrate Project

Kim A. Kastens
Lamont-Doherty Earth Observatory of Columbia University
61 Rt 9W, Palisades, NY 10964 U.S.A.
kastens@ldeo.columbia.edu
1-914-671-2341
(no fax)

and

Cathryn A. Manduca
Science Education Resource Center, Carleton College
One North College St.
Northfield, MN 55057 U.S.A.
cmanduca@carleton.edu

Acknowledgements: 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.

InTeGrate White Paper, 8 February 2016
Northfield, MN
Available at http://serc.carleton.edu/integrate/about/pubs.html

Using systems thinking to design, implement and evaluate a complex educational intervention

Abstract

InTeGrate is a 5-year, nation-wide, $10M project to improve undergraduate Earth education, through developing and testing instructional materials, professional development for faculty, and program-level innovations. To envision, shape, and monitor this complex system, we have employed the tools and language of systems thinking. In designing InTeGrate, we planned for a set of semi-autonomous subsystems that would allow for replication with adaption and would interact in mutually-beneficial ways. We laid out non-linear causality chains, in which one activity would have multiple beneficial outputs, and desired outcomes would be supported through multiple influencers. In implementing InTeGrate, we leveraged feedback loops to nudge actors throughout the system towards desired behaviors and away from problematic choices. We built technical and social mechanisms to regulate flows of information between and within subsystems, seeking to deliver timely, actionable information, so as to enable proactive responses to emergent problems and opportunities. In evaluating InTeGrate, we used systems mapping to choose locations to insert evaluative probes and sought out early indicators of emergent phenomena. Although every project differs in detail, InTeGrate's systems approach can serve as a useful model for any education reform effort that seeks widespread and long-lasting impact. [192 words]

Keywords: Systems thinking, geoscience education, sustainability, evaluation, program reform, feedback loops

Introduction

Our society faces an intertwined set of challenging problems that have been labeled as "complex" (Holland, 1998; Kania, Kramer & Russel, 2014) or "wicked" (Williams and Hummelbrunner, 2009; Ramaley, 2014). Such problems tend to be dynamic, non-linear, and counter-intuitive; 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 is one such complex problem, because it involves individuals and entities with varied histories and priorities that interact in ways that may retard change (Austin, 2013; Bergquist,1992; Fry, 2014; Kezar and Eckel, 2002; Tierney, 1977).

Scholars of natural and human-built systems have developed a tool kit and a vocabulary for dealing with complex, non-linear dynamic situations. This body of thought comes under a variety of labels including "systems thinking" (Hummelbrunner, 2011), "complexity theory" (Hawe, Bond & Butler, 2009), "complex adaptive systems" (North American Primary Care Research Group, 2008) and "system dynamics" (Meadows, 1999). The common thread running through these approaches is that the unit of analysis is a *system*, which may be defined as a "set of elements whose interconnections determine their behavior" (Climate Interactive, 2016). Controlled by interactions rather than by simple attributes, a system can exhibit behavior that is adaptive, dynamic, goal-seeking, self-preserving, evolutionary—and often unexpected. Much of the systems literature strives merely to understand how complex systems work, but there is also an applied literature that uses systems thinking to design and evaluate interventions aimed at complex problems (Kania et al, 2014; Meadows, 1999; North American Primary Care Research Group, 2008; Patrizi 2013; Quinn-Patton, 2015; Williams and Hummelbrunner, 2009).

In this paper, we describe how systems thinking has been used in the design, implementation, and evaluation of the InTeGrate project, a five-year, $10 million effort to transform learning about Earth in higher education across the United States (InTeGrate Program, 2015b). InTeGrate's project design combines interventions and activities at multiple scales and leverage points, with a conscious focus on maximizing productive interactions between activities. Specifically, InTeGrate facilitates work at the course scale to change what faculty are ready to teach; work at the department and institutional scales to change what faculty are asked to teach and do; and work at a national scale to change values, networks, and available resources. InTeGrate's use of systems thinking to promote widespread change in higher education provides a model that can be adapted for use in other disciplines or other nations.

Using system thinking to *design* InTeGrate

InTeGrate was prompted by a solicitation from the U.S. National Science Foundation to transform higher education within the geosciences, with an emphasis on environmental and resource grand challenges facing the nation. Designing InTeGrate required a shift away from single interventions that would cause a desired change. Instead, transformative change required identifying a shared vision, defining key elements of the system, selecting the elements to target, planning for multi-directional interactions among these elements, anticipating that a given action would have multiple outcomes, and attending to the goals and priorities of multiple types of actors.

Shared Vision

To accomplish transformational change on a national scale, within realistic constraints of time, money, and human capacity, requires highly leveraged interventions. Meadows (1999, 2008) identifies 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. 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 STEM careers (e.g. Casey & U. S. Joint Economic Committee, 2012), and to enable students to appreciate and understand the world (e.g. Herschbach, 1996). InTeGrate joined with earlier voices (e.g. Burns, 2010; MacGregor, Middlecamp, Millar & Seymour, 2007; Ramaley, 2014; Sherman, 2008) in building towards a fourth goal: to prepare problem-solvers with the ability and disposition to tackle profound societal challenges of the 21st century.

InTeGrate aims to transform Earth education such that learning occurs in the context of societal issues across the undergraduate curriculum and engages all students. InTeGrate's founders considered these changes to be essential for preparing 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 strong assertion of the new vision, (2) attractive and easily adaptable exemplars of the new vision, and (3) a supportive, collegial community where the new vision is the norm.

InTeGrate's theory of change conjectured that educating Earth-savvy problem-solvers had the potential to lead to a cascade of desirable changes. As summarized in figure 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.

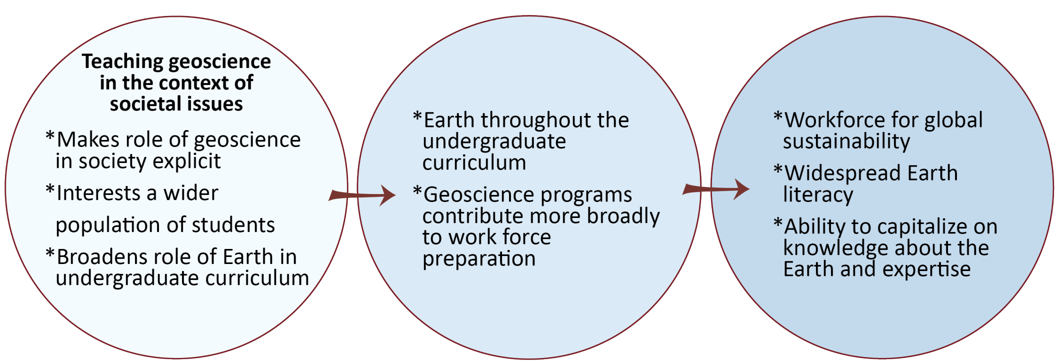

Figure 1: Hypothesized impact of shifting goals and strategies on grand challenge of living sustainably and justly on Earth.

Selecting 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 which components and interactions to target. 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. Although other components are also influential (Austin, 2013), these three were selected as the smallest and most tractable set of components that could plausibly result in pervasive and lasting changes.

Component 1: Courses: In American colleges and universities, faculty have broad powers over the content and context of their courses. Thus, working with faculty to include Earth-related issues in a societal context in their courses was fundamental to meeting the project's goals. In addition, InTeGrate aimed to improve teaching quality, tapping into a wealth of recent research on effective teaching through student-centered pedagogies (Kober, 2014; Freeman et al., 2014). Our design anticipated barriers to change (Elder & MacGregor, 2009) including lack of examples of teaching about the Earth in the context of societal issues, the time needed to revise a course, and resistance to taking time away from traditional content.

InTeGrate's materials development activity was designed to create a cadre of faculty who could demonstrate the path forward with exemplary materials tested in a variety of instructional settings. Both the materials and the processes used to create them could then be adapted or adopted by others. Over 100 faculty were involved in developing and testing materials. Materials developers worked in teams of 3 to 6 faculty members, drawn from different types of institutions. They were selected through a competitive proposal system, and received a stipend for their work. Materials were required to conform to the InTeGrate Materials Development and Refinement Rubric (Steer et al., in review; InTeGrate Program, 2015c), which requires that materials use research-tested pedagogical approaches, connect geoscience to grand challenges facing society, make use of authentic geoscience data, and foster interdisciplinary problem solving, systems thinking, and geoscientific habits of mind. The process of recruiting and selecting materials development teams was actively managed to yield a balanced portfolio of materials suitable for general education, teacher preparation, geoscience majors, and other majors (InTeGrate Program, 2016). A group with expertise in student-centered pedagogy and learning assessment (the Assessment Team) mentored the development process and enforced the requirements of the rubric.

Component 2: Programs & Institutions: There are limits to what can be accomplished by an individual faculty member. Lasting change requires work at the department, program or institutional level (Seymour, 2002; Tobias, 1992). Productive pedagogic or content changes within courses must be endorsed, adopted or rewarded by others in the department or institution if they are to transition from research into practice (Fry, 2014) and expand beyond the work of a single pioneering faculty member. Which courses are offered within a degree program are decisions that are made by departments or institutions. Further, work to broaden participation in STEM fields has demonstrated that increasing the diversity of students requires attention at the programmatic level (Crosling, Thomas & Heagney, 2008; Engstrom & Tinto, 2008; Seymour and Hewitt, 1997) to promote student motivation and enthusiasm, to provide academic supports, and to cultivate a sense of belonging. We designed the proposal-based implementation program (IP) activity to provide incentives, resources, and coaching for programs to customize activities at their own institution aligned with InTeGrate's overarching goals. Robust evaluation and reporting would allow others to learn from these initial examples.

InTeGrated includes sixteen IPs (InTeGrate Program, 2015a). Teams were chosen through competitive proposals, and the process of recruiting, selecting, and mentoring IP teams was actively managed to yield a diverse, nation-wide, portfolio of institution types and program models. The IPs span multiple courses and programmatic activities within a single department, multiple departments within an institution, or multiple institutions. Two focus on teacher preparation, and approximately half target non-traditional students or groups underrepresented in science. The majority seek to broaden Earth literacy across their campuses, reaching students in departments ranging as far afield as political science, nursing, and business.

Component 3: Nationwide community: The third target for intervention was motivated by several intersecting needs: to motivate and support change, to shift values and norms, and to expand the radius of impact. The inspiration and learning required for successful change operate on a scale larger than an individual institution: we know this from our experience with progress in science, which is supported through a plethora of workshops, conferences, professional groups, and inter-institutional collaborations: "Individuals do not produce scientific knowledge—communities do" (Ford, 2008, p. 410). 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.

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 by combining face-to-face and virtual gatherings with development and use of associated on-line resources. Synchronous learning opportunities included webinars that targeted specific problems, and multi-day workshops. Early workshops focused on gathering and sharing resources, insights, and best practices in both pedagogy and program development (e.g. Gosselin, Burian, Lutz & Maxson, 2015), while later workshops shifted focus towards dissemination of InTeGrate materials and approaches. On-line resources include collections of contributed educational materials, strategies for teaching for a sustainable future, information about workforce needs, and strategies for recruiting and supporting students from underrepresented minorities. The website and workshop/webinar programs are tightly coupled, in that much of the web content was initially developed to support workshops and then expanded using materials contributed and created by workshop participants.

More than the sum of the parts: driving interaction between program components

Systems thinking requires a focus not only on components of a system but also on the interactions between them (Hargreaves, 2010; Holland, 1998; Meadows, 1999; Preskill, Gopal, Mack, & Cook, 2014). It is these interactions between components that lead to emergent phenomena like the changes in values and teaching practices that InTeGrate seeks to foster. We designed several interactions to create strong linkages between the different components of the system, for example:

Community activities underpin Materials Development: Previous NSF-funded programs and work by individual faculty had created a wealth of educational resources for teaching about the 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 three years 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 in-gathering workshops gathered resources and allies, and a set of associated web sites (Narum & Manduca, 2012) were established to capture, organize and share the gathered insights. The paired workshops and websites allowed us to capitalize on the momentum of others while avoiding reinventing the wheel (Kania, et al., 2014).

Community activities lay a foundation for Implementation Programs: A second set of early workshops was designed to pull together community wisdom about challenges that manifest at the programmatic level. These program-level workshops addressed interdisciplinary learning; pathways to the workforce; and how to support a diverse student body. Representatives from programs that had made progress in these areas came together to share experiences, synthesize recommendations, and create on-line resources for use by implementation program developers, laying a foundation of knowledge that could be adapted to local conditions. IP proposals were asked to build upon this body of knowledge.

Materials development and implementation programs build community: InTeGrate's material development model requires developers to work in teams drawn from across institutions and in many cases across 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 are required to span a department, multiple departments or multiple institutions, weaving yet more tendrils into the growing network.

Building non-linear causality chains

The InTeGrate program design sought to create a web of actions that would push toward desired outcomes. Rather than a linear cause→effect chain, InTeGrate's design presumes that any given action will have multiple outcomes or consequences. For example, in-gathering workshops were designed to achieve at least four outcomes simultaneously: increased 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. Co-development of instructional materials by 3- to 4-person teams from different institution types was designed to result in both materials that are 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.

Likewise, 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; required individual and group reflection on the development process; all in addition to developing and deploying a rubric that articulates and reinforces InTeGrate's pedagogical approaches and priorities. InTeGrate makes no effort to disambiguate the individual impact of these support structures; rather the aggregate package of supports is viewed as synergistically generating a set of influences and feedbacks that collectively shift the system in the direction of the desired outcome.

Advantages of Designing with a Systems Approach

Thinking about the needs of individuals and communities, courses and programs, all at the same time, allowed us to optimize for impact at multiple scales. By putting interactions among components in place as part of the initial design, we were able to plan for synergies and efficiencies, and establish the expectation that different parts of the project would both benefit from and contribute to one another. Systems thinking was present in InTeGrate's design from the earliest planning discussions, but as the project has progressed we have found it useful to be more explicit and purposeful in our use of the theory and language of systems thinking for understanding and explaining how to design for change.

Using systems thinking to *implement* InTeGrate

Literature on a systems approach provides strong tactical guidance for implementation. We focus here on strategies that we have used to build system fitness, the ability of a human system to change itself so as to improve performance (Kania et al., 2014). System fitness is determined in large part by the flow of information through the system and the capacity for actors in the system to make sense of and act upon this information. This section describes our strategies for managing information flow, monitoring the state of the system, nudging actors towards desired behaviors, and responding to emergent challenges and opportunities.

Flows of information between 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. Actors throughout the system need actionable information in a timely manner to make strong decisions as they manage their local responsibilities. Project leadership needs synoptic information to steer the enterprise and respond to emerging challenges and opportunities.

Most of InTeGrate's work is accomplished by collaborative teams of 3 to 15 people: materials development, implementation program, workshop convening, and assessment, evaluation, and leadership teams. From a systems perspective, these teams can be seen as subsystems within the larger InTeGrate system, each with its own purpose or goal, its own elements, its own internal relations and structures, and its own vigorous internal information flows. Because InTeGrate's design is modular, discoveries, insights and best practices known in one subsystem can be re-used in parallel subsystems, but only if they are shared effectively. To manage information flow within and between subsystems, InTeGrate combines human interactions, web-based information sharing, and movement of individuals with expertise.

InTeGrate provides numerous formal opportunities for transfer of information via direct interactions. Materials development teams meet face-to-face in meetings with other teams from their cohort, and members of earlier cohorts lead professional development webinars for new cohorts. The Assessment Team meets with each materials development team twice. Implementation programs attend virtual meetings and have professional development webinars in response to their emerging needs. Quarterly leadership team meetings and the annual Advisory Board meeting are convergence points for flows of information from and to all of the subsystems.

Complex information sometimes requires a combination of human and artifact-based information flows. For example, the Materials Development and Refinement Rubric embodies InTeGrate's values and priorities into an artifact, which the Assessment Team uses 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 towards 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 de-bugged.

Much information transfer takes place through the InTeGrate website (http://serc.carleton.edu/integrate), which supports both internal communication among team members and dissemination of project results, findings, and resources to others. All InTeGrate subsystems contribute their accumulating insights and efforts to building this one mega-artifact. For the geographically distributed builders of InTeGrate, the website supports the development of a collaborative team by providing a "place" to "convene" and share ideas, plus a visible artifact depicting shared progress.

One important form of information is the state of the system. Kania et al. (2014) stress that using systems thinking to implement a social change requires a constant process of "sensing" the system and its environment to ensure that resources are applied where opportunities are greatest. System-oriented decision-makers continually compare the perceived state of the system with the desired state (goal) and plan their next steps so as to nudge the system towards 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 for use by those actively engaged in building InTeGrate. Purpose-built, web-accessed tools allow team members at different institutions to work collaboratively and keep track of what is going on in the system. These technology-based tools create a kind of a dashboard, comprising webpages that record and update the status of different parts of the system (figure 2). The dashboard is used to monitor progress, plan workflow, spot problems, and support decision-making.

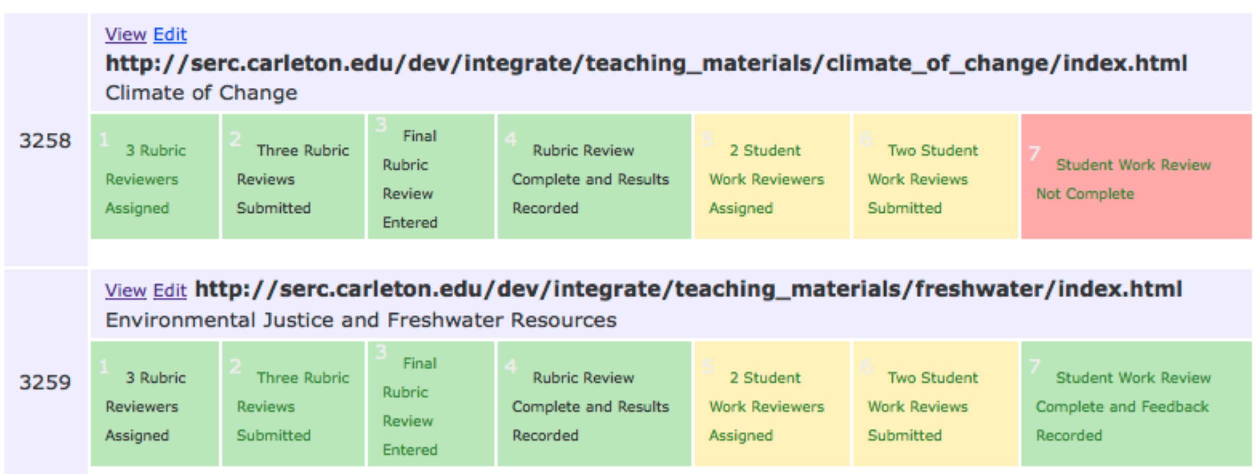

Figure 2: This screen-shot shows a portion of the page that records the progress of in-development instructional modules and courses as they move through a series of seven checkpoints. InTeGrate's suite of webpages and web-based tools comprise a "dashboard" that allows the leadership team to continuously monitor the state of the system.

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 or Implementation Program teams, bringing with them the knowledge gathered and created at the workshop. Materials developers who show leadership potential were encouraged to nucleate new IP's at their home institutions, or co-convene new workshops in their area of expertise.

This system successfully supported the spin up of the project to its current scale. The use of all three information flow strategies has been essential. The website provides the information needed for project members to stay informed in a format that allows them to find what they need when they need it. However, information does not by itself convey insight, nor does it support synthesis across project elements. Meetings, working groups, and other face-to-face and virtual interactions address these needs. Movement of individuals from one project element to another provides transfer of a level of nuanced expertise that cannot easily be captured in written documents or short interactions.

Leveraging Feedback Loops:

One of the mechanisms that enables systems to "form a whole that is greater than its parts" (Hargreaves, 2010, p. 3) is the occurrence of feedback loops. In a feedback loop, A causes or influences B, and then B causes or influences C, which in turn loops back around and influences A, directly or indirectly. In some circumstances, influence can continue to flow around the loop multiple times.

A negative (or "balancing" feedback) is one in which the influence coming from the loop tends to moderate the initial direction of change, tending to pull the system back towards an equilibrium position. Negative feedback loops rein in departures from InTeGrate's values before they escalate. Figure 3 depicts a nested set of two balancing feedback loops in InTeGrate's materials development program. InTeGrate requires that materials development teams comprise faculty from different types of institutions, who must all pilot test the newly developed materials. As depicted in the inner loop of figure 3, if one developer begins to develop materials that are only well-

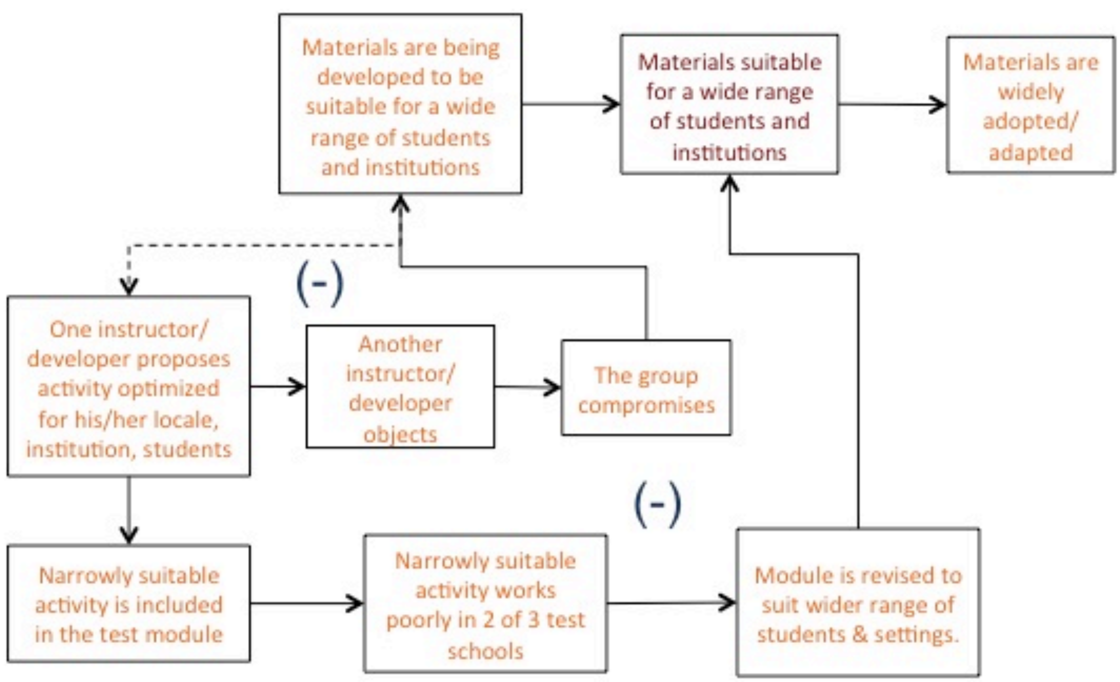

Figure 3: Negative/balancing feedback loops can be used to keep a system near a desired equilibrium state. This systems map depicts a pair of nested feedback loops that tend to keep the InTeGrate system moving towards development of materials that will be useable with a wide variety of students in a wide variety of geographic and institutional settings.

suited to his or her own particular students, or teaching style, or geographic venue, or institution type, the other team members are expected and empowered to pull the development effort back onto a track that will reach more types of students, venues, institutions. As depicted in the outer loop, if overly-narrowly targeted materials make it through the development stage into the piloting tests, they prove to be problematic when tested in the other two institutions, and at the revision stage the materials are tweaked back onto the desired path towards widely useable materials.

In contrast, a positive (or "reinforcing") feedback loop is one in which the influence coming back from the loop tends to push the system even farther in the direction of the initial impetus. In InTeGrate, reinforcing feedback loops nudge InTeGrate participants towards actions that align with InTeGrate's values and priorities or that are observed to be effective. For example, InTeGrate's materials development rubric places a premium on having students engage in metacognition (Bransford, et al, 2000), including having them monitor their own learning process (figure 4). 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 towards a new state of more effective learner.

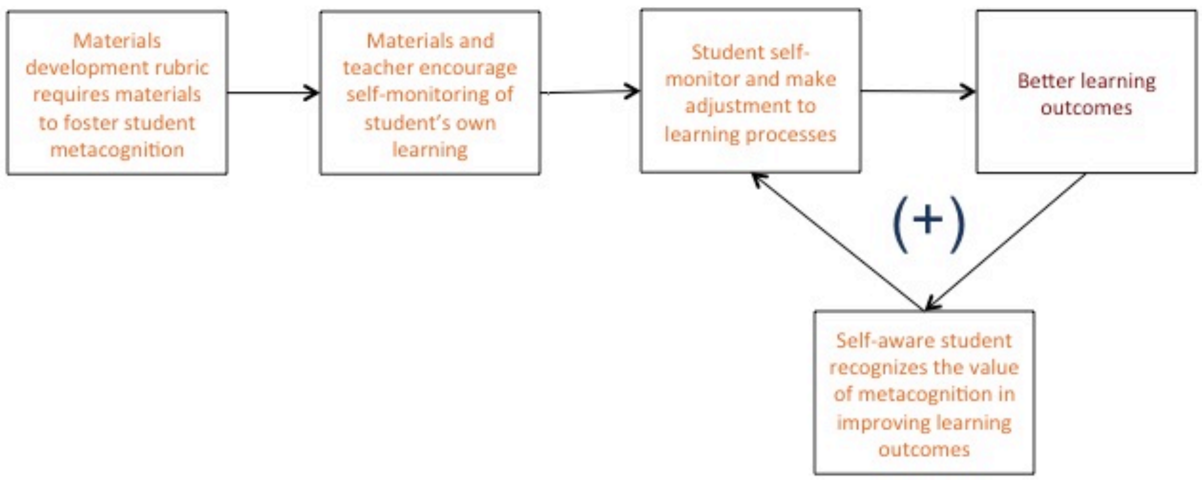

Figure 4: Positive (or "reinforcing") feedback loops tend to move a system farther away from its initial state. In this example, both teachers and materials encourage students to self-monitor their own learning, which research shows can lead to better learning outcomes, which in turn can lead to student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

Deploying Emergent Strategies to Respond to Challenges and Opportunities

In Kania et al's (2014) division 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, the information obtained through continuous 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.

As an example of an emergent strategy to capitalize on an opportunity, the materials development rubric 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 the InTeGrate project. 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 to create curriculum. 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 for classroom use. The first

Implementation Program solicitation did not yield strong proposals that would bring substantial numbers of underrepresented minorities into the geoscience education pipeline. The project responded by inviting two proposing teams with the desired goals but immature work plans to engage in a mentoring process to improve their proposals, under the guidance of an advisory board member.

These changes and additions 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, 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, in that—like the solutions arrived at by biological evolution—they re-purpose structures and processes that were already in place and adapt them to new purposes. Webinars, module development procedures, skilled advisory board members, and collaborative development of assessment items were already in the InTeGrate toolkit, and one or more such tools were pulled forth and adapted to address each emerging challenge.

Advantages of implementing with a systems approach:

You can't anticipate and plan for everything. 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 solutions can develop at all scales from an individual actor in a local knowledge network, to projects within departments or across institutions. In aggregate, it is the coming together of pre-planned capacities plus evolving solutions that enable the emergence of the desired system level outcomes.

Using systems thinking to *evaluate* InTeGrate

Using systems mapping to reveal key connections, relationships and interdependencies:

InTeGrate is large and complex, involving many moving parts: dozens of modules and courses, hundreds of faculty, thousands of web pages, and potentially tens of thousands of students. To stretch our human minds around this complexity, we have engaged in systems mapping. If a "system" is "a configuration of interacting, interdependent parts that are connected through a web of relationships…." (Hargreaves, p. 3), a system map is a visual depiction of a portion of that configuration, naming key parts and showing key interactions and dependencies. Early in the project, the evaluation team and leadership developed overarching system maps for the entire project. As the project has been built out, we developed more localized systems maps for functional sub-systems of the project (figures 3 and 4). [1]

InTeGrate systems maps differ from the classic logic models used in intervention planning and evaluation in that the systems maps (a) showcase flows and linkages and specify their nature, (b) allow, indeed encourage, depiction of branching, recursion, and bi-directional flows, and (c) are presented as working tools and hypotheses to be tested, rather than as a blueprint that the project must follow to reach a pre-specified goal. The earliest InTeGrate system map was a conjecture map, in the sense of Sandoval (2013): a diagram depicting the design conjectures that the Principal Investigators had made in the proposal At this point, 80% of the way through the project, the systems maps are best viewed as a hybrid of conjectures about how InTeGrate is intended to work and empirically grounded observations of how InTeGrate is in fact working, with additional empirical evidence accruing as the project proceeds.

As one of their nine propositions for evaluating complexity, Preskill, et al. (2014) recommend "Seek to understand and describe the whole system, including components and connections." Systems mapping is one of the tools they recommend to enact this proposition. For the InTeGrate evaluators, the systems maps have allowed us to inventory where evaluative probes have already been introduced into the system and where additional evaluation effort would be valuable. For example, the materials development effort was initially presented to the community as a mechanism to create pedagogically and scientifically excellent instructional materials. Early on, though, it

---

[1] The full set of InTeGrate systems maps can be found at: http://d32ogoqmya1dw8.cloudfront.net/files/integrate/about/integrate_systems_maps.pdf

became clear that equally important outcomes from the process would be instructors/developers who incorporate InTeGrate values into their teaching practice, who become advocates for InTeGrate's ideas, programs, and materials, and who become 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 and focus groups with the materials developers.

Evaluating for emergent phenomena and for lasting changes to the larger system

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, 2001). The classic examples are flocks of birds or schools of fish that move together, despite their lack 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 sub-systems. An example within InTeGrate is the desired outcome that more students, nationwide, will move into STEM college majors and the STEM workforce. InTeGrate's materials development rubric 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 towards 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 post-instruction survey that asks students about their career interests--even though this is not an explicit learning goal of the instructional materials.

InTeGrate is embedded within the larger system of higher education. One way to conceptualize InTeGrate's long-term outcomes 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 & Wenger, 1991) of Earth educators, which can multiply the effectiveness of funded activities and be capable of persisting 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" (Wegner-Trayner & Wegner-Trayner, 2015). Communities of practice are committed to a shared domain of interest, value their collective competence, engage in joint activities, share information, build relationships that enable them to learn from each other, care about their standing with each other, and develop a shared repertoire of resources (experiences, stories, tools, ways of addressing recurring problems). InTeGrate's materials development process, workshops, webinars, website, 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. The expectation is that once a geo-ed community of practice has been established as a well-functioning component of the system of higher education, it can be maintained at lower cost than the heavy lift of building it in the first place.

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, 2016). In this model, individuals accrue both practical and affective value from their involvement in the community. We draw on the Wenger-Trayner framework for evaluating communities of practice (Wenger, Trayner, & de Laat, 2011; Wenger-Trayner, 2014), which combines quantitative and qualitative indicators of immediate value, potential value, applied value, and realized value, together with narrative "value-creation stories" that show how individuals traversed across the boundaries from one type of value to the next. For indicators of value gained, the evaluation team has access to surveys at the end of events, numbers and demographics of new and repeat participants, interviews with event participants, a survey of IP faculty, responses to online questions posed to materials downloaders, and a national survey of geoscience educators. As sources for value creation stories, we have written reflections from materials developers and testers, web-published Instructor Stories, interviews with the Leadership Team and notes from Leadership Team meetings, and case studies of the Implementation Programs. We are using these metrics and value-creation stories 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 & 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 (Hargraves, 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.

Final Thoughts & Transferable Strategies

In the work described in this paper, we drew on our experience as geoscientists studying the Earth system (Manduca & Kastens, 2012) plus systems ideas developed in other disciplines, and found that that systems thinking provided a productive lens for design, implementation and evaluation.

Our *design* approach differed from the traditional linear model of single actions causing single effects. Instead, we found that the diversity of interactions and feedbacks within the complex system ensured that single actions would have multiple outcomes and multiple actions were needed to cause the system to change its behavior. Critical to our *design* was a focus on shifting the behaviors of actors at all levels, by addressing their values or goals and by reducing barriers to change.

During *implementation*, the five year time scale for this project required that we scale up quickly, developing management strategies and information flows that could allow the project to grow rapidly in scope and impact. The systems lens, with its focus on flows, feedbacks and adaptations was ideal for this. Building "system fitness" by shaping information flows and building local capacities increased the chances that the system would be resilient in the face of emerging challenges and proactive when presented with emerging opportunities.

In seeking to *evaluate* the impact of a short project on a long-term problem, the systems lens helped us focus on what could emerge and endure from our actions. Systems mapping and conceptualizing certain desired outcomes as emergent phenomena allowed us to identify and probe for leading indicators of progress towards systemic change.

Thinking back over the totality of our InTeGrate experience, we are struck by the power of replicability and adaptability. Constructing an ambitious intervention as a set of subsystems (e.g. the set of materials development teams, the set of implementation programs, the set of workshops with companion webpages) has allowed for scale-up via *replication*. The set of procedures, parameters, information flow pathways, feedbacks and dashboard components developed and refined for the first few elements were relatively easily replicated for multiple parallel elements, without greatly increasing the burden on the infrastructure or leaders. Moreover, the parallel elements could then learn from each others' experiences and mentor the newly-formed elements, accelerating progress. The same subsystem strategy allowed for *adaptability*. Structures or strategies developed and tested in one context could be repurposed for use in another part of the project, allowing for evolutionary responses to challenges and opportunities. Finally, designing and constructing InTeGrate as a system of subsystems made it possible to utilize the talents, time and energy of a vast army of dispersed, busy individuals. Having many subsystems allows for many leadership opportunities, each of a magnitude that can be accommodated alongside a full teaching load and research agenda. InTeGrate's locally- and thematically-adapted subsystems provide many attachments points for faculty members to get involved, regardless of course taught or home discipline.

Some of InTeGrate's processes and structures were specifically tailored for the idiosyncratic context of geoscience education. But others have the potential to be adopted or adapted by new projects and perhaps become the new norm. Systems-inspired approaches that we think have the most transfer potential include:

- the synergistic relationship between workshops and website;
- the use of technology-enabled dashboards and communication flows to support distributed decision-making;
- structuring of reinforcing and balancing feedback loops to nudge actors towards desired behaviors;
- division of the effort into subsystems, each with its own goals, processes, and internal communication flows, 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.

As we address complex, wicked problems within higher education and beyond, these strategies may be of use in managing large projects, collaborations, or networks that can accelerate the rate of change in stubborn systems.

References Cited

AAC&U (2011). The LEAP Vision for Learning: Outcomes, Practices, Impact, and Employers' Views. Association of American Colleges and Universities: 28.

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

Bergquist, W. H. (1992). *The Four Cultures of the Academy*. Jossey-Bass Inc., Publishers, 350 Sansome Street, San Francisco, CA 94104-1310.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). How people learn*: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.

Burns, W.D. (2010) SENCER in Theory and Practice. In Sheardy, R.D. ed, *Science Education and Civic Engagement: The SENCER Approach,* American Chemical Society.

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

Casey, B. (2012, April). STEM Education: Preparing for the Jobs of the Future. In *US Congress Joint Economic Committee, April*.

Climate Interactive. (2016). The Climate Leader. Retrieved from https://www.climateinteractive.org/programs/the-climate-leader/

Crosling, G., Thomas, L., & Heagney, M. (Eds.). (2008). *Improving student retention in higher education: the role of teaching and learning*. Routledge.

Elder, J. L., & MacGregor, J. (2008). *The sustainability movement in higher education: An overview*. Essay for January 2009 meeting on Mobilizing STEM education for a sustainable future. Retrieved from: http://mobilizingstem.wceruw.org/documents/jan09meet.html

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

Ford, M. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education*, *92*(3), 404-423.

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

Fry, C. L., & The Coalition for Reform of Undergraduate STEM Education (Eds.). (2014). *Achieving systemic change: A sourcebook for advancing and funding undergraduate STEM education*: Association of American Colleges and Universities.

GETSI (2016). GETSI: Geodesy Tools for Societal Issues. Retrieved Jan. 29, 2016, from http://serc.carleton.edu/getsi/index.html.

Gosselin, D. C., Burian, S., Lutz, T. M., & 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.

Hawe, P., Bond, L., & 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 Eduction, 82(4)*, 10-17.

Holland, John 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 Program. (2015a). Implementation Programs. Retreived Feb. 1, 2016, from http://serc.carleton.edu/integrate/programs/implementation/index.html

InTeGrate Program. (2015b). InTeGrate: Interdisciplinary Teaching about the Earth for a Sustainable Future. Retrieved Feb. 1, 2016, from serc.carleton.edu/integrate.

InTeGrate Program. (2015c). Working with the InTeGrate Materials Development and Refinement Rubric. Retrieved Feb. 1, 2016, from http://serc.carleton.edu/integrate/info_team_members/currdev/rubric.html.

InTeGrate Program. (2016) InTeGrate Authored Modules and Courses. Retrieved Feb. 1, 2016, from http://serc.carleton.edu/integrate/teaching_materials/modules_courses.html

Johnson, S. (2002). *Emergence: The connected lives of ants, brains, cities, and software*. Simon and Schuster.

Kastens, K. (2016). Reinforcing feedback loops power effective communities of practice. *Earth & Mind: The Blog.* Retrieved from http://serc.carleton.edu/earthandmind/posts/commofpract.html

Kania, J., Kramer, M., & Russell, P. (2014). Strategic philanthropy for a complex world. *Stanford Social Innovations Review, 12*(3).

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

Kober, N. (2015). *Reaching Students: What Research Says About Effective Instruction in Undergraduate Science and Engineering*. National Academies Press.

Kezar, A., & 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. Retrieved February 3, 2016 from http://www.uscrossier.org/pullias/wp-content/uploads/2016/01/communities-of-trans.pdf.

Kuh, G. D. (2008). High-Impact Educational Practices. Washington, DC, Association of American Colleges and Universities**:** 35.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge university press.

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

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

Manduca, C. A., & Kastens, K. A. (2012). Mapping the domain of complex earth systems in the geosciences. In K. A. Kastens & C. Manduca (Eds.), *Earth & Mind II: Synthesis of Research on Thinking and Learning in the Geosciences, Geological Society of America Special Publication* (pp. 91-96). Boulder: Geological Society of America.

Meadows, D. H. (1999). *Leverage points: Places to intervene in a system* (pp. 1-19). Hartland, VT: Sustainability Institute.

Meadows, D. H. (2008). *Thinking in Systems: A Primer*. White River Junction, VT: Chelsea Green Publishing.

Narum, Jeanne and Manduca, Cathy (2012) Workshops and Networks in Brainbridge, William Sims, editor, *Leadership in Science and Technology, A Reference* Handbook. *ISBN 978-1-4129-7688-6* p. 443-451

North American Primary Care Research Group: Adapted from Zimmerman, B., Lindberg, C., Plsek, B. (2009). "A Complexity Science primer: What is Complexity Science and Why Should I Learn About It?". Retrieved December 5, 2015.

Patrizi, P., Heid Thompson, E., Coffman, J., & Beer, T. (2013). Eyes wide open: Learning as strategy under conditions of complexity and uncertainty. *The Foundation Review*, *5*(3), 7.

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

Preskill, H., Gopal, S., Mack, K., & Cook, J. (2014). Evaluating complexity: Propositions for improving practice.

Ramley, 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.

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.

Steer, D. N., Iverson, E., Egger, A., Kastens, K. A., Manduca, C. A., & McConnell, D. (2016 in review). Developing student-centered curricula for outcomes-based education: An example melding Geoscience and societal issues. *Curriculum Journal*.

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*. Research Corporation, Book Dept., 6840 East Broadway Boulevard, Tucson, AZ 85710-2815.

Wenger, E., Trayner, B., & de Laat, M. (2011). Promoting and assessing value creation in communities and networks: A conceptual framework. Heerlen, the Netherlands: Ruud de Moor Centrum, Open University.

Wenger-Trayner, B. (2014). World Bank report and recommendations: Learning partnerships in the program for capacity building to strengthen good financial governance in Southern and Eastern Africa: 2010-2014.

Wenger-Trayner, E., & Wenger-Trayner, B. (2015). Communities of practice: A brief introduction. Retrieved from http://wenger-trayner.com/introduction-to-communities-of-practice/

Williams, B., & Hummelbrunner, R. (2010). *Systems concepts in action: a practitioner's toolkit*. Stanford University Press.