Complexity, Robustness, and Trade-Offs in Evaluating Large Scale STEM Education Programs

Susan A. Yoon, University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA 19104
yoonsa@gse.upenn.edu
Lei Liu, University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA 19104
leil@gse.upenn.edu

Abstract: This study explores the application of the complex systems lenses of robustness and trade-offs to evaluate the success of a STEM education program in terms of sustainability. A mixed methods approach is used to document contextual constraints, trade-offs considered such as modularity vs. one-size-fits-all, and actions taken by the project team to build toward adaptive capacities.

Introduction

The issue of declining participation of America's youth in science, technology, engineering and math (STEM) education and careers has received a great deal of attention over the last few years from industry, government, and education sectors (Business Roundtable, 2005; Domestic Policy Council, 2006; U.S. Department of Education, 2007). As STEM jobs continue to grow, by 2012 the number of positions in science and engineering is estimated to outpace the number of qualified people to fill them by 26% and 15% respectively (NSF, 2006). From a citizenship science perspective, there is a broader issue of a general lack of STEM literacy in society associated with such declining participation. When fewer students are interested in and aware of emerging STEM research such as nanotechnology, the likelihood that they are able to contribute to intelligent decision making about applications that can affect their daily lives decreases (Yoon, 2008). Additional pressures on workforce development exist with the emphasis on developing 21st century and cyberinfrastructure-enabled scientific skills in which learning from and using digital technologies both for developing conceptual knowledge and process skills must now figure prominently in education (NSF, 2007; Partnership for 21st Century Skills, 2007). To respond to these needs researchers in the Learning Sciences and STEM education have focused attention on understanding how to develop, sustain and scale innovative science and technology reform-oriented programs. Investigation and discussion have centered on what mechanisms and variables are required to create the conditions for success and how success may be evaluated (Dede, Honan & Peters, 2005; Yoon & Klopfer, 2006; Yoon et al., 2009). One theme that resonates through much of this research is the challenge reformers face in adapting their programs to specific contexts (Datnow, 2005; Dede & Honan, 2005; Hargreaves & Fink, 2000). Elmore (1996) writes about the difficulties experienced by nested clusters of innovation in educational settings. He states that failures, historically, in generating successful large-scale reforms can be attributed to an "absence of practical theory that takes account of the institutional complexities that operate on changes in practice" (p. 21). Similarly, Goldman (2005) and Fishman et al. (2004) identify the need to account for multiple embedded levels of stakeholders when designing for improvement, who must actively support the reform if it is to succeed. Coburn (2003) reinforces the idea that educational reform and improvement are matters of complexity highlighting the inability of research to address the inherent multidimensionality between and within educational constituents. Building from this reform literature and the calls for research methods and practical theories that take a systems approach, in this paper we introduce the principle of robustness, both as a lens and a tool for understanding issues of sustainability and scale of innovative STEM education programs. We suggest that achieving a robust programmatic state is arguably the sine qua non of success with respect to sustainability, which in turn provides the rationale for going to scale. As we will argue, one of the core measurable mechanisms operating in the service of robustness is trade-offs. We present data, analyses, and results of the first year of a three-year STEM program in which we document the trade-offs experienced and then undertaken in the redesign process to create a more robust program.

Theoretical Framework
Complex Systems

Multidimensionality, nested clusters, and multiple constituents and levels are all characteristics of complex systems. The study of complexity has been a focus of research in many academic disciplines such as biology (Kaufmann, 1995), physical science (Bak, 1996; Prigogine & Stengers, 1984), psychology (Arrow et al., 2000), anthropology (Lansing, 2006), and economics (Stacey, 1996). The ubiquity of complex systems research appears to extend from the commonalities that exist between structures or contexts and behaviors in that they comprise multiple elements that adapt or react to patterns that they create (Arthur, 1999). By studying the patterns that emerge and the interactional dynamics that lead to these patterns, researchers can better understand among other things, how systems adapt, self-organize, fluctuate and either achieve or do not achieve steady

states. For example, in the simple dilemma of the Tragedy of the Commons, one individual may choose to act selfishly and use up resources with no real impact on supplies. However, if many individuals decide to behave in the same way soon the resources will be depleted and the individuals along with the collective community will suffer. Thus the culture or the collective that exists must be considered. It requires one to know the limits of the culture and to adjust behaviors in order to sustain a robust system, which is also in the best interest of individual survival.

Robustness in Complex Systems

Related to understanding the limits of the culture, robustness researchers study the contextual parameters that enable systems to persist or deteriorate in the face of negative often, unpredictable impacts such as the spread of viruses in an ecosystem, continual military conflict between countries, or the collapse of stock markets in economies. As a central complex systems mechanism, various problems across biological and social situations can be investigated through the lens of robustness which has led leading complex systems organizations like the Santa Fe Institute to devote whole research tracks to its study (http://www.santafe.edu/research/topics-innovation-evolutionary-systems.php#4). One of the unifying characteristics that enable its multidisciplinary application is that perturbations impact natural, cultural, and engineered systems. It is hypothesized that studying varying system's responses to these perturbations can provide general rules or principles that can be used to solve long-standing questions such as how evolution that requires variation supports the emergence of phenotypes from genotypes (Wagner, 2005). Others have used notions of robustness to understand how the structure of social networks enables information flow and resource management (Webb & Bodin, 2008). For example, as individual nodes, actors or elements are removed from the network key connections may be broken which may isolate subgroups and thus limit how information gets distributed. Elsewhere we have used social network theories, analyses, and tools to help students and teachers develop better information-seeking strategies to construct knowledge and to gain social capital (Yoon, accepted; Baker-Doyle & Yoon, 2009). In all these studies, creating robust systems requires ease of information flow in order to give and receive resources. Where constraints exist due to cultural or contextual barriers, trade-offs are needed to build adaptive capacities.

Trade-Offs

Perhaps the most well known example of a general application of the idea of trade-offs is the situation of the Prisoner's Dilemma that falls under the general topic of Game Theory. This is a hypothetical case where two prisoners are interrogated separately for having together committed a crime. In the classic scenario, it is assumed that the base prison sentence is 5 years. The pay-offs or decrease of prison time will depend on whether the individuals choose to cooperate with each other or defect. In the context of the game, defecting (claiming innocence and implicating your partner), while your partner cooperates (confesses to the crime), will always provide the largest pay-off to the defector in that no prison time will be served. Similar to the Tragedy of the Commons example, using this simple algorithm, researchers have studied how it is possible that cooperation might have emerged as an evolutionarily adaptive strategy (Axelrod, 1984). While a full blown discussion of the mechanisms that lead to cooperation cannot be made here, one of the key ideas related to trade-offs for our research pertains to the fact that much of the study of strategies that allow populations rather than individuals within systems to benefit are underpinned by the fact that the success of someone or something that includes activities and resources may come at the expense or sacrifice of another person or another thing. For example, Janssen and Anderies (2007) write with respect to the trade-offs that organizations must make in choosing to satisfy short-term or long-term goals, "when all 'low-hanging fruit' is taken to increase robustness cheaply, [systems] will eventually reach a point at which it is no longer possible to generate additional robustness without a cost to performance and/or decreased robustness somewhere else in the system" (p. 44).

Complexity, Robustness, and Trade-Offs in Evaluating STEM Education Programs

Understanding the educational system as a complex system is not new in educational research. In his *Change Forces* series focused on reform, Fullan (1993, 1999, 2003) uses complex systems theory as an organizing framework to reveal core concepts such as non-linearity, unpredictability and multi-level agency that are important issues to contend with in real-world educational systems. However, where states of existence like robustness become goals that drive the reform efforts rather than for description, and where mechanisms like trade-offs become actions, educational research takes a step closer to a "practical theory" that accounts for institutional complexities that Elmore (1996) searches for. The following study documents the multiple constituents, contextual limits, and important trade-offs made toward increasing robustness in a large-scale STEM education reform effort.

Methodology
Context

This work is funded under the National Science Foundation program "Innovative Technology Experiences for Students and Teachers (ITEST)". The ITEST program is designed to increase opportunities for students and teachers in underserved schools to learn and apply information technology concepts and skills in the STEM content areas. Our project is a school district-university partnership between one of the largest school districts in the U.S., a school of education and a leading university nanoscale research center. It aims to achieve the broader ITEST goals by updating standard high school science curricula and training teachers through a curriculum and instruction framework built on five component variables addressing content knowledge, pedagogical content knowledge, and workforce development goals. These variables are: 1) Real world science and engineering applications such as current nanotechnology research and ethical considerations of this research; 2) Educational Technologies to develop content knowledge such as NetLogo computer simulations; 3) Information Technologies for communication, community-building and dissemination such as Google Groups and wikis; 4) Cognitively-rich pedagogical approaches such as problem-based learning; and 5) Investigation and consideration of STEM education and careers.

There are two parts to the scope and sequence of project activities. The first part entails a three-week, 75-hour professional development summer workshop for science teachers, in which they learn to construct and pilot curricular units based on the five component variables. These curricular units are intended to be aligned with school district standards for high school biology and physical science. The summer workshop is followed by the school-year implementation of these units in teachers' classrooms and 5 follow-up Saturday professional development workshops. In addition to the central curriculum and instruction goals, another major goal of the project is to build teaching capacities and self-directed interests through communities of practice structures which include working with groups of teachers from the same school, implementing an online comprehensive professional development database for peer-to-peer interaction, and requiring teachers from different schools to construct units collaboratively. For this study, we focus on the activities and program developments of the first cohort of teachers who participated on the project between August 2008 and May 2009.

Participants

Ten male and six female teachers participated in the workshop from 10 high schools and 1 middle school in the district. The group was racially/ethnically diverse: seven teachers were White, six were African American, and three were Asian. Courses taught ranged from grades 8 to 12 in the content areas of physical science, biology, chemistry, and physics. The average amount of teaching experience was 15.8 years, with a range of 1 to 39 years of experience. Data from 128 students taught by a sample of six teachers were collected with the following racial/ethnic breakdown: 23.4% White, 48.4% African American, 7.8% Hispanic, 14.1% Asian, and 3.9% other races. The study also investigated perceptions and actions of the project team, as the goal was to document and understand the kinds of trade-offs we experienced in the development and implementation of the program. Thus, we include the project team as participants who comprised: a professor, a post-doc and two doctoral students from the school of education; and a professor, two post-docs, and the educational outreach director from the nanoscale research center.

Data Sources and Analyses

Eight data sources were collected for the study and analyzed through a mixed methods approach.

1. A 64-item 5-point Likert-scale survey administered to teachers to measure teacher's self-perceptions of pedagogical practices, beliefs, student participation, and teacher confidence levels with particular emphasis on the project's five variable curriculum and instruction framework. A pre-intervention survey was collected in August 2008 and a post-intervention survey was collected in May 2009. Complete data sets were obtained from 8 participating teachers due to various data collection challenges. A paired-samples t-test was conducted to determine if teachers had statistically significant change in their self-perceptions from the beginning to the end of the 2008-2009 academic year.

2. An 82-item 5-point Likert-scale survey administered to students to measure self-perceptions of attitudes toward science and use of classroom resources and strategies with particular emphasis on the project's five variable curriculum and instruction framework. Pre- and post-intervention surveys were collected from 5 participating teacher's classrooms immediately before and after the ITEST implementation. Times for survey collection were different due to the variability of implementation in the school-year curriculum. For example, some teachers chose to develop units in physical science on the topic of properties of matter, which corresponded with delivery of the standard curriculum at the beginning of the year. Others chose units that corresponded with latter times. A paired-samples t-test on 128 survey responses was conducted to determine if students had statistically significant change in their self-perceptions.

3. An 11-item open-ended research survey administered to teachers that probed in-depth understanding of nanoscale content and STEM education pedagogical beliefs related to the 5 variable project framework. A pre-intervention survey was collected in August 2008 and a post-intervention survey was collected in May 2009. Complete data sets were obtained from 8 participating teachers.

4. Qualitative case studies of 5 project teachers were conducted by the project's external evaluator. Case study classrooms were selected to obtain information from as diverse a sample as possible. The case-study implementation was guided by the following 4 questions: What does the ITEST curriculum look like as it is implemented in case-study classrooms?; To what extent do teachers implement the key elements of ITEST?; To what extent is each enacted unit similar or different from the intended unit?; What supports and challenges are observed and reported in the implementation of the ITEST curricular units? Data collection included: field notes from three classroom observations, informal conversations with teachers, informal conversations and/or focus groups with students, and a formal end-of-unit interview with each teacher.

5. Focus group interview for the entire cohort of teachers conducted during the Saturday workshop held at the end of February 2009. Focus group questions probed for details about affordances and constraints to implementation and suggestions for program redesign for the next cohort of teachers and schools. The cohort was divided into 3 smaller groups. Interviews were transcribed and compared by the external evaluator and the project team in order to identify common themes.

6. Weekly 1.5 hour project management meetings audio-recorded and transcribed.

7. Written and informal verbal reflections of the 4 scientists from the nanoscale science center after specific nano-content meetings and after content modules were implemented during the summer and Saturday professional development workshops.

8. Field notes and informal verbal reflections of the 4 school of education researchers documented throughout the year of implementation as facilitators in participant teachers' classrooms. There were 33 formal classroom observation field notes collected.

For data sources 3, 5-8, data were reviewed systematically by the 4 school of education researchers through a qualitative grounded theory approach (Strauss & Corbin, 1998) in which all open-ended surveys, transcribed meetings, observation notes, and reflections were mined for evidence illustrating various trade-offs that emerged throughout the project year.

Results

With respect to increasing robustness and revealing and using trade-offs that are specific to our ITEST project, we were interested in documenting the emerging themes as they pertained to the 5 variable framework as well as impacts on the goal of building a practitioner community focused on the framework. In this paper, the results of two variables (i.e., real world science and engineering applications and educational technology) are presented and organized into the three sub-categories of: i) Contextual Constraints; ii) Trade-offs; and iii) Actions. We continue to work toward identifying more general categories of constraints and trade-offs that may be applied across similar STEM education programs, however, in this section we list them individually and attempt to apply a meta-level categorization later in the Discussion section.

Real world science and engineering applications

Research in nanotechnology refers to the interdisciplinary study of physical, chemical, and biological phenomena in the 1 – 100 nanometer range. Manipulation of atoms and molecules at this scale have given scientists unprecedented abilities to engineer materials, products and processes in, medicine, the environment, cosmetics, clothing, and computer and automotive technologies (Roco, 2003). Nanotechnology research has undergone rapid growth over the past decade with an estimated annual global research expenditure of $13.9 billion (PCAST, 2008). The pervasiveness of these applications is also immense. The Woodrow Wilson *Project on Emerging Nanotechnologies* states that there are over 800 products currently on the market produced by over 400 companies in 21 countries. Therefore, in our project, incorporating current and cutting edge nanotechnology research and applications including nanoscale content and real world problem solving into the standard science curriculum was a central design element of the teacher constructed PBL units.

Contextual Constraints

Several contextual constraints limited the success of this variable. First from the case studies and focus group interviews, the most widely reported challenge teachers faced was the mismatch in PBL pedagogy and the school district's system of managed instruction, i.e., core curriculum and benchmarks. Due to time challenges, delivering a complete unit was not possible for some teachers choosing instead to deliver discrete lessons throughout the year. Despite significant increases in teacher confidence levels in their ability to use problem-based instructional practices ($t(7)=2.546$, $p=.038$), teacher effectiveness in helping students identify multiple solutions to a problem ($t(7)=2.121$, $p=.078$), and perceptions of increased use of real world applications in the curriculum ($t(7)=3.667$, $p=.008$), no significant differences were found in student perceptions of learning about real world issues, new science innovations or discoveries, teacher's demonstrations of science problems, and student problem-solving. Moreover, while teaching aspects of PBL units was found to occur in classroom observations, even in the best example of one of the case-studies, student final projects had low fidelity with the

original curriculum, which included plans for students to gather, synthesize, and evaluate the various types of information used throughout the unit in order to assess the ethical issues related to their nano-application.

Difficulties also existed in aligning curricular content levels and school contexts with nanoscale applications. Despite significant increases in teacher's perceptions of their ability to incorporate nanoscale science content $(t(7)= 3.130, p=.017)$, scientists frequently expressed concerns about finding the appropriate connections to the topics covered in the high school curriculum. Issues for the scientists also surfaced with respect to empowering teachers to "own" the presented concepts so that they felt compelled to shape it into instruction relevant for their classrooms. From a content knowledge perspective, responses from teachers that probed understanding of nanoscale concepts showed misconceptions at the fundamental level of size and scale as well as in their perceptions of nanoscale science requiring an entirely new understanding of canonical scientific domains. For example, the majority of teachers believed that nanotechnology is a distinctively new science (with its own set of laws), instead of an addition to existing sciences, like biology, chemistry, and physics.

Trade-Offs

The trade-offs we would like to highlight here deal with structural and behavioral factors of social and human systems, which we have discussed elsewhere to be among the most difficult constraints to contend with in the educational system (Yoon & Klopfer, 2006; Yoon et al., 2009). In the first case regarding the implementation of PBL, the expectation was for teachers to develop a coherent unit with the higher-order inquiry skills of evaluation and synthesis however, given district testing pressures, the actualized curriculum looked much different. This is akin to the trade-off of robustness to performance. Essentially, the trade-off parameter can be described in terms of some benefits are better than none in the face of highly uncompromising intransigent constraints. In the second case in aligning curricular content and the existence of fundamental misconceptions, if we were to follow the research on how people learn or knowledge-building theory (Bransford et al. 2000; Scardamalia, 2002), providing scaffolds for teachers to construct their own knowledge would stand a better chance in helping them make curricular connections on their own, promote self-efficacy and self-regulation, and remediate misconceptions. However, the major constraints here are lack of time and resources to undertake the knowledge-building process. A logical trade-off parameter then, given that teachers felt positive about their ability to incorporate nanoscale content into their regular curricula, might be to meet the teachers half-way, i.e., providing instructional support and curricular resources to seed interest and a solid leverage point to explore research and application avenues other than the ones that are directly taught to them.

Actions

Based on the trade-off parameters, a number of changes in project expectations and program delivery occurred for cohort 2 teacher professional development. Instead of mandating a coherent sequential unit, curricula could be constructed drawing on content connections in multiple topics throughout the school year. We identified a teacher from cohort 1 who developed particular expertise in working with project goals in this way and we enlisted her to teach a two-hour module on what we called semi-coherent curricular units where she modeled the process and provided to teachers her own tried and true unit with detailed explanation of the important modifications she made in practice. Actions spawned by the trade-off of meeting the teachers half-way, included constructing a comprehensive curriculum alignment inventory in which 11 nanotechnology and concepts/applications were identified in addition to how they interface with the State Standards and the School District's Core Curriculum for grade 9 Physical Science and grade 10 Biology. Table 1 depicts a sample of the inventory the project team developed. This was used in instruction for teachers when new content was introduced and given to them as a general heuristic to use when building their standards-aligned PBL units.

Table 1: Sample of Curriculum Alignment Inventory.

Nano/Bio Concepts/Applications	Pennsylvania State Standards		SDP Core Curriculum
	Physical Science	**Biology**	
1. What is the nanoscale? How small is nano? 2. Unique properties at the nanoscale	3.4A Explain concepts about the structure and properties of matter 3.1D. Apply scale as a way of relating concepts and ideas to one another by some measure.	3.3A Explain the structural and functional similarities among living things	Physical Science (PS) 3 – Atoms & Periodic Table, & Chemical Bonds BIO 2- Introduction to Chemistry
3. Nanomaterials and surface area	3.4B. Analyze energy sources and transfers of heat. • Determine the efficiency of chemical systems by applying mathematical formulas. • Use knowledge of chemical reactions to generate an electrical current.	3.3B. Describe and explain the chemical and structural basis of living organisms. • Describe relationships between the structure of organic molecules and the function they serve in living organisms. • Identify the specialized structures and regions of the cell and the functions of each.	PS 2 – Matter BIO 3 – Cell Structure and Function BIO 2- Introduction to Chemistry

Educational Technologies to develop content knowledge

Educational technologies (ET) were presented to teachers as distinct in our framework from information technologies (IT) in that ET comprises imaging, simulations, and visualization tools that help scientists and students learn about science content and processes whereas IT is primarily used for communication and interaction. Teachers were instructed on the emerging importance of being able to navigate the cyberinfrastructure for STEM careers, which includes using computational tools for data collection, manipulation, visualization, and prediction. Several digital tools used with teachers included simulations found on the Internet such as The Nano Journey, http://www.nanoreisen.de/english/index.html and NetLogo (Wilensky & Reisman, 2008) software in which a drug delivery system for nanotechnology and cancer was constructed and used with teachers to learn about how to conduct scientific experiments with such simulations.

Contextual Constraints

In comparison to using information technologies, this variable was much less prevalent in teacher's instructional practices. Classroom observation field notes indicated that only 2 out of the total cohort of 16 teachers used NetLogo in their instruction. Where as teachers surveys showed significant increases in the perception of their ability to incorporate IT into their lessons $(t(7)=2.198, p=.064)$, no differences were found with ET. In student surveys, although the item investigating science or other ideas through simulations, images, or animations $(t(125)=2.594, p=.011)$ was found to have positive gains, actual observed practice showed no use of these tools for higher order scientific inquiry skills. Rather, they were used mainly to illustrate or demonstrate a concept. In the case studies, similar results were found along with other contextual constraints. One major challenge was that the teachers did not have administrative privileges to download NetLogo onto their school computers. One teacher did download NetLogo onto his laptop and students individually came up to his computer and tried the simulation for short periods of time. Another teacher said that she was interested in NetLogo, but didn't have a laptop that she could load it onto. Of the other three, one said that he didn't have enough facility to use it, and the other two did not think it would have been feasible to use in their classes due to classroom management issues. All teachers did use other educational technology, the most common application being visualization software that allowed students to move between visualizations of the same item or location at different scales.

Trade-offs

The lack of educational technology use triggered several in-depth discussions during project management meetings about how to facilitate greater applications as well as applications that resembled more closely how scientists really used such software for scientific discovery. NetLogo and similarly StarLogo were simulation tools used by the principal investigator in previous projects with some success in developing computational literacies and 21st century skills (Klopfer & Yoon, 2005; Klopfer et al., 2005). Particularly with respect to the potential for manipulating variables, changing initial conditions, and running multiple experiments, NetLogo/StarLogo modeling software could be extremely useful tools in the science classroom. However, given the difficulties in downloading the software, differences in teachers facility and comfort levels, and the fact that some ET was being used in instruction, the project team explored the trade-off of modularity vs. one size does not fit all. This trade-off is similar to the use and benefits of differentiated instruction. Although it was believed that more focused practice on a few ET applications for all teacher participants would work toward successful ET implementation, again, structural and behavioral realities posed seemingly insurmountable constraints and we needed to differentiate this component.

Actions

Several actions were taken according to the identified trade-off. First, with the easily accessible Internet visualization programs, we developed exemplar lessons to show generically how experimentation could be done with any similar tool. Next, we devoted a good deal of time collecting and vetting ET sites on the Internet that would be appropriate for our project which culminated in a comprehensive Educational Technology Inventory with clickable links. We also planned to deliver parallel sessions that demonstrated different ET applications and associated pedagogies for which teachers could choose to sign up according to their interests and comfort levels. We would offer these iteratively so that when teachers felt that they had developed particular competencies, they could choose to attend increasingly more challenging ET modules to improve their skills.

We have documented and used several more trade-offs as a response to contextual constraints. For example, after realizing how little access we had to guidance counselors in schools and coming to the understanding that they were no longer assigned to just one school, plans to work on a comprehensive nanotechnology education and career package in which teachers and guidance counselors would collaborate to disseminate on a large scale needed to change. Instead we opted for different STEM education and career experiences, one of which was a short, real-world internship where teams of two high school students worked in a modestly paid after school job for two hours a week in the nanoscale professor's research lab. The trade-off and extensive project team

discussions centered around project <u>impact on individual vs. many students</u>. With the variable of cognitively-rich pedagogies, a trade-off that we continually face is how much is enough <u>theory vs. practice</u>. In working toward our goal of building a community of practice one of the major trade-offs we grapple with is whether, how, and how much to build <u>technological, human and/or social capital</u>.

Discussion

With all of these trade-offs, one common theme that organizes around building sustainable or robust programs appears to be negotiating the middle ground. When working with systems with high complexity, that is, many often competing factors that operate under unpredictable conditions, finding the compromise between ideal project fidelity and real-world constraints may be the key to ensuring that at least some parts of the vision survives. From this platform, we can create opportunities for continued growth of activities that have taken a hold in the education system to continue to grow. Another meta-level theme that clearly impacts the project team's decision-making practices is the disposition of flexibility. The adaptive process that these STEM education programs necessarily undergo can work if there is an expectation of flexibility and a willingness to carefully document and analyze the constraints in order to choose another path that may be more successful. Likewise, adopting an ideological stance of pragmatism may be a prudent choice. If we apply from the start, the expectation that ideas or theories such as how people learn best can only be measured or deemed successful in practice, we are forced to take account of the institutional complexities that can exert heavy influence on program success.

There are at least three potential contributions that this study can make. At the level of theory understanding sustainability of innovative STEM education programs through the lenses of complexity, robustness and trade-offs is a promising framework to evaluate the complexity of educational reform. The study also informs practice, operationalizing the framework in terms of documenting contextual constraints, trade-offs considered, and then actions taken. Finally, looking across the trade-offs and identifying the common themes of finding the middle ground, developing dispositions of flexibility, and adopting a pragmatic stance, can be used by other like-minded educational reformers in the design of future STEM education programs.

References

Arrow, H., McGrath, J. & Berhahl, J. (2000). *Small Groups as Complex Systems*. Thousand Oaks, CA: Sage Publications, Inc.

Arthur, W.B. (1999). Complexity and the economy. *Science* (284), 107–109.

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. New York: Copernicus.

Baker-Doyle, K. & Yoon, S. (2009, August). In search of practitioner-based social capital: A social network analysis tool for understanding and facilitating teacher collaboration in a professional development program. Paper to be presented at the annual meeting of the *European Association for Research on Learning and Instruction*, Amsterdam, Netherlands.

Bransford, J., Brown, A., & Cocking, R. (Eds.) (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Business Roundtable (2005). *Tapping America's Potential: The Education for Innovation Initiative*. [Web Resource] Available online. www.businessroundtable.org/publications.

Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3-12.

Datnow, A. (2005). The sustainability of comprehensive school reform models in changing district and state contexts. *Educational Administration Quarterly*, *41*(1), 121-153.

Dede, C., Honan, J. & Peters, L. (2005). Scaling Up Success: Lessons Learned from Technology-Based Educational Improvement. San Francisco, CA: Jossey-Bass.

Dede, C. & Honan, J. (2005). Scaling up success: A synthesis of themes and insights. In, C. Dede, J. Honan & L. Peters (Eds.) *Scaling Up Success: Lessons Learned from Technology-Based Educational Improvement* (pp. 227-239). San Francisco, CA: Jossey-Bass.

Domestic Policy Council (2006). *American Competitiveness Initiative*. Office of Science and Technology Policy: Washington D.C.

Fishman, B., Marx, R., Blumenfeld, P., Krajcik, J. & Soloway, E. (2004). Creating a framework for research on systemic technology innovations. *Journal of the Learning Sciences*, *13*(1): 43-76.

Goldman, S. (2005). Designing for scalable educational improvement. In, C. Dede, J. Honan & L. Peters (Eds.) *Scaling Up Success: Lessons Learned from Technology-Based Educational Improvement* (pp. 67-96). San Francisco, CA: Jossey-Bass.

Hargreaves, A. & Fink, D. (2000). Three dimensions of educational reform. *Educational Leadership*, *57*(7), 30-34.

Janssen, M & Anderies, J. (2007). Robustness trade-offs in social-ecological systems. *International Journal of the Commons*, *1*(1), 43-65.

Johnson, S. (2001). Emergence: The Connected Lives of Ants, Brains, Cities, and Software. New York: Touchstone.

Kauffman, S. (1995). *At Home in the Universe*. New York: Oxford University Press.

Klopfer, E. & Yoon, S. (2005). Developing games and simulations for today and tomorrow's tech savvy youth. *Tech Trends, 49*(3), 33-41.

Klopfer, E., Yoon, S. & Um, T. (2005). Teaching complex dynamic systems to young students with StarLogo. *Journal of Computers in Mathematics and Science Teaching*, 24(2), 157-178.

Lansing, S. & Miller J. (2005). Cooperation Games and Ecological Feedback: Some Insights from Bali. *Current Anthropology 46*(2): 328-334.

National Science Foundation. (2007). Cyberinfrastructure Vision for 21st Century Discovery. Arlington, VA: Author.

President's Council of Advisors on Science and Technology (PCAST). (2008). *The National Nanotechnology Initiative: Second Assessment and Recommendations of the National Nanotechnology Advisory Panel*. Washington, D.C.: Executive Office of the President.

Prigogine, I. & Stengers, I. (1984). *Order Out of Chaos*. New York: Bantam Books.

Roco, M. (2003). Converging science and technology at the nanoscale: Opportunities for education and training. *Nature Biotechnology*, *21*(10), 1247-1249.

Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Smith (Ed.), *Liberal Education in a Knowledge Society* (pp. 67-98). Chicago: Open Court.

Strauss, A. L. & Corbin, J. M. (1998). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (2nd edition), Sage, Newbury Park, CA.

U.S. Department of Education (2007). *Report of the Academic Competitiveness Council*. Washington, DC: Author.

Wagner, A. (2005). *Robustness and Evolvability in Living Systems*. Princeton, NJ: Princeton University Press.

Webb, C. & Bodin, O. (2008). A network perspective on modularity and control of flow in robust systems (pp. 85-118) . In, J. Norberg and G. Cumming (Eds.) *Complexity Theory for a Sustainable Future*. New York: Columbia University Press.

Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep or a firefly: Learning biology through constructing and testing computational theories—an embodied modeling approach. *Cognition and Instruction, 24*(2), 171–209.

Yoon, S. (accepted). Using social network graphs as visualization tools to influence peer selection decision-making strategies to access information about socioscientific issues. *Journal of the Learning Sciences*.

Yoon, S. (2008). An evolutionary approach to harnessing complex systems thinking in the science and technology classroom. *International Journal of Science Education*. 30(1), 1-32.

Yoon, S., Chessler, M., Kumar, R., Williams, D., Flicker, J. & Dunham, S. (2009, April). Understanding emergent challenges and adaptations in the design of an OST STEM project through a Structure, Behavior and Function (SBF) Complexity Lens. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Orange County, CA.

Yoon, S. Liu, L. & Goh, S. (2009). Exploring the process of convergent adaptation in technology-based science curriculum construction. In the proceedings of the *Computer Supported Collaborative Learning annual conference*, Rhodes, Greece.

Yoon, S. & Klopfer, E. (2006). Feedback (F) Fueling Adaptation (A) Network Growth (N) and Self-Organization (S): A Complex Systems Design and Evaluation Approach to Professional Development. *Journal of Science Education and Technology*, *15*(5-6), 353-366.

Acknowledgments