

## **Determining Progress in Improving Undergraduate STEM Education: The Reformers' Tale**

*A White Paper commissioned for the Forum, "Characterizing the Impact and Diffusion of Engineering Education Innovations", February 7-8, 2011.*

Elaine Seymour, Kris DeWelde, and Catherine Fry

We set out to answer four related questions about change in science, technology, engineering, and mathematics (STEM) education at the undergraduate level. They are:

1. What do we know about the extent and nature of uptake of proven STEM education practices?
2. What indicators or measures of "uptake" or "scale-up" have been used to assess these?
3. What are the sources and nature of constraints and resistance to change?
4. What leverages change and what sustains it?

The second two questions were more familiar and we draw on some of co-author Seymour's work in addressing them. However, neither author had looked carefully at published sources estimating the extent of change in STEM education or the indicators by which these were arrived at. Indeed, we had somewhat taken on trust the commonly-offered assessment that only marginal inroads had been made into the traditional undergraduate curriculum and the ways in which it is taught. We therefore began by reviewing available published work. As we describe, the results were not quite what we expected to find.

Since the mid-1980's researchers have pointed to problematic outcomes in undergraduate STEM education in terms of declining enrollment, drop-outs, and field-switching (reviewed in Seymour and Hewitt, 1997 and Seymour, 2001). By the late 1990's, the causes of the problem had been firmly located in inadequate teaching methods and curriculum content issues. This diagnosis was reflected in a series of reports from national commissions and panels, each of which called for fundamental changes in STEM teaching and learning. These reports drew on both U.S. and international comparative data sources in arguing that, unless the serious problems evident in STEM education were addressed, the U.S. would lose its world leadership in science and technology (*e.g.*, National Science Foundation, 1996; Boyer Commission on Educating Undergraduates in the Research University, 1998; National Research Council, 2003a, 2003b; National Research Council, 2007). Follow-up studies reported either modest or disappointing progress (The Reinvention Center, 2002; National Research Council, 2010). The most recent reports have strengthened the urgency of their calls for improvement in STEM education at all levels (President's Council of Advisors on Science and Technology, 2010; National Science Board, 2010).

Articles by researchers echo the impression of unsatisfactory progress reflected in national reports (Handelsman et al., 2004; DeHaan, 2005; Dancy and Henderson, 2008; Fairweather, 2008; Baldwin, 2009; Henderson et al., 2010). While they convey a consensus that promising teaching and learning innovations are not being taken up at a satisfactory level or pace, the evidentiary base for these claims is not always clear. For example, an often-cited source for assertions of low uptake is the National Research Council's report, "Improving Undergraduate Instruction in Science, Technology, Engineering, and Mathematics: Report of A Workshop" (2003b). On page one, this report asks, "why introductory science courses in many colleges and universities still

rely primarily on lectures and recipe-based laboratory sessions where students memorize facts and concepts, but have little opportunity for reflection, discussion, or testing of ideas?” However, no data are offered or other evidence cited to support this claim.

We therefore surveyed the available literature and extracted what could be said from evidence about the nature and extent of “uptake” or “scale-up” of research-grounded teaching materials and methods. Table 1 summarizes the results of our search and categorizes evidence of implementation of reforms by their institutional reach.

### **Table 1: Evidence of Uptake or Scale-up of Research-Grounded Improvements in STEM Education by Extent of Institutional Reach**

#### **Whole institution:**

- A medical school overhauled the entire curriculum to incorporate active learning and other research-based pedagogy (Elizondo-Montemayor et al., 2008).

#### **Within a single university, either the whole department or whole college:**

- There are multiple examples of entire departments or colleges changing their curriculum to incorporate active learning and other research-based innovations. These include departments of biology (Ono et al., 2007), engineering, (Pundak and Rozner, 2008; Merton et al., 2009), chemistry (Ege et al., 1997; Coppola et al., 1997), and physics (Dori and Belcher, 2005).

#### **Across universities, sub-departmental:**

- The SCALE-UP (Student-centered Active Learning for Undergraduate Programs) approach has been adopted at more than 50 new institutions (Beichner and Saul, 2003; Beichner et al., 2007; Beichner, 2008) and in new disciplines (Oliver-Hoyo et al., 2004; Beichner et al. 2007).
- A 2003 meta-analysis of 110 single-institution assessments reported that over 1,000 institutions offer some form of learning community program, many of which are longstanding and regularly offered. They are established in all types of institutions (research universities, engineering schools, regional public institutions, liberal arts colleges, and community colleges) (Taylor et al., 2003).
- The SCALE (System-wide Change for all Learners and Educators) program has developed networks of STEM and education faculty who successfully use collaborative approaches for teacher preparation (Hora 2007).
- Concept inventories are now used in a variety of institutions. There are 23 distinct concept inventories developed or under development (Libarkin 2008).
- Surveys of faculty across institutions report higher levels of faculty awareness of teaching alternatives, some increased inclination to use them, but mixed results for uptake. For example:
  - For engineering departments, there are high levels of awareness of innovations (82%), but much lower adoption levels (47%) (Borrego et al., 2010).
  - A survey of faculty in Louisiana shows that few faculty have been trained in teaching methods, but those who have are more likely to use that information to improve teaching and to consider teaching an important part of their professional identities (Walczyk et al., 2007).

- For physics faculty, survey data indicate that despite awareness of innovations and a willingness to use those ideas in their teaching, to a large extent, actual teaching practices are still traditional and not consistent with research-based practices (Dancy and Henderson, 2010).
- The UCLA-HERI faculty survey reports an increase in faculty using student-centered pedagogies in recent years (Higher Education Research Institute at UCLA, 2009).

**Within a single university, sub-departmental:**

- Evidence from two faculty development programs at two different universities shows courses have been modified or developed as a result, evidence of strengthened faculty networks, collaborations on grant proposals, and increases in interdisciplinary teaching. Both programs were aimed at integrating environmental and sustainability issues across the curriculum (Barlett and Rappaport, 2009).
- A study of transforming a single course at one university reports that the implementations were made and sustained with a high degree of fidelity over time and across instructors (Chasteen et al., 2010).

The scatter of clues from this small body of research broadly indicates some uptake by departments, some success in getting usage of particular teaching methods that are supported by institutions, and their spread at sub-departmental levels and across samples of individual faculty. They also suggest increased levels of faculty awareness of alternative ways to teach and some inclination to try them. While this does not indicate a strong national response to the call for change, it does not seem to justify the consensus expressed in post-Millennium reports and articles that little or nothing has happened to improve undergraduate STEM education. Rather, we get a sense that we simply do not have enough information to reach such a judgment.

*Inquiry Methods*

In light of the incomplete picture of the extent of change that emerged from our review of relevant literature, we decided to pose our set of questions to a sample of people who, over the last three decades, have been highly involved in efforts to improve quality and access in STEM undergraduate education. We chose a panel of expert witnesses, comprising project directors, principal investigators (PIs), co-PIs, evaluators, and funding officers of multiple (current and former) STEM education reform grants. In the limited time available, we restricted our panel to ten members, with the intention of subsequently extending it for a fuller inquiry. Panel members have been engaged with STEM education reform projects between eleven and thirty years; some also have extensive experience as scholars of STEM education, workshop and institute organizers, and advisors to other projects.

The resulting sample of projects represented by the panel and referenced during interviews included: Center for the Advancement of Engineering Education (CAEE), Center for the Integration of Research, Teaching, and Learning (CIRTL), Process-Oriented Guided Inquiry Learning (POGIL), Peer-Led Team Learning (PLTL), Science Education for New Civic Engagements and Responsibilities (SENCER), The Reinvention Center, two National Learning Communities projects, multiple Cooperative Learning initiatives and workshop programs, Women in Science and Engineering (WISE), Recognition Awards for the Integration of Research and Education (RAIRE), the EXCEL Engineering Coalition, the Systemic Reform Projects in

Chemistry, and Multi-Initiative National Dissemination (MID) workshops, Curriculum for the Bioregion, Mobilizing STEM Education for a Sustainable Future (MSE), the Student Assessment of their Learning Gains online instruments (SALG), Doctoral programs in STEM education at Purdue University and the University of Michigan, the Washington Center for Improving the Quality of Undergraduate Education, and several other university-based Teaching and Learning (T&L) Centers.

We sent the research questions to our panel members in advance, and asked them to discuss their responses in an open-ended telephone interview of about one hour. The questions were the sole protocol, but, in the ethnographic tradition, related lines of thought, observations, and examples offered by panel members were followed up. The interviewer (Seymour) took shorthand notes of the interviews which were then transcribed. The resulting text data were coded with a code word system focused on the four questions and further shaped by panel members' responses. Definitions encapsulating code content were logged into a codebook. Codes were grouped into themes and searched across all text data files, along with text segments of particular interest. Code frequencies were run and descriptive tables built from clustered items. These processes were enabled by The Ethnograph V.6.0 1.0 (2010), a set of software programs designed to assist with text data analysis.

The panel's answers to Questions 1 and 2 are presented in four descriptive tables, and answers to Questions 3 and 4 are offered in a narrative discussion. All counts offered in the tables are of the numbers of panel members offering any observation. The maximum response in some categories exceeds ten where panel members offered different but related observations.

### *Definitional Issues*

We asked panel members to clarify their understanding of the terms "innovation", "uptake", and "scale-up". On reflection, they saw "innovations" as methods or materials that were still being tested, and used a variety of qualifying adjectives (e.g., "research-grounded," "scientific", "proven," "quality," "good") to reference classroom resources that were proven and well-used. However, they agreed that there was no good, generic noun to reference all non-traditional teaching methods and materials and understood why faculty at large might use "innovation" to describe anything unfamiliar.

They took "uptake" to mean use or adaptation of any classroom resource by individuals or groups of faculty, usually in a single discipline, and in a few courses or sections. This might occur in single schools or spread to users in several institutions by the efforts of a funded project. They saw uptake as essentially informal--the result of individual faculty choices--and the product of promotional efforts by project networks, or the efforts of leaders in particular institutions. Host institutions might not know this was happening.

"Scale-up" was interpreted as adoption of a resource into many courses in one or several departments or colleges (e.g. engineering), across one or several institutions, or in whole divisions or institutions. "Scale-up" implied formal decisions to adopt something as "a reform," following the promotional efforts of both formal and informal leaders. For whole institutions this implied policy decisions. Reforms, or scaled-up "innovations" were essentially public, formal, and durable. One project director described the resulting normative shift as the feeling that "everyone is doing this."

**Question 1: How did panel members assess the extent and nature of uptake of proven STEM education methods and materials?**

We asked the panel to offer their estimates of both uptake and scale-up, whether from direct experience of projects in which they were (or had been) involved, or from their wider experience of national reform efforts. In their responses, summarized in Table 2, we noted some blurring of the distinction between high levels of “uptake” and “scale-up”.

**Table 2. Panel Members’ Estimates of Nature and Extent of Uptake and Scale-up**

**More positive overall estimates, largely defined as “uptake” N=6**

- Many combinations of methods are in widespread use--POGIL, PLTL, just-in-time teaching, calibrated peer review, etc. These and other innovations (*e.g.*, learning communities, small group learning) as have gained a foothold in many institutions and disciplines. (Further growth was thought to depend on good professional development.)
- Greater awareness of T&L research and its implications for teaching.
- More disciplinary discussion of education issues, notably at meetings.
- More publications available on what works better or worse.
- An overall sense of an upward trend: "We haven't slipped back as far as we used to. The high energy is still there."

**Observations on the nature of uptake N=14**

- Most change is in lower level classes.
- Most uptake is at the individual level which is easier than at the institutional level. We lack strategies for institutional change.
- Pedagogical innovations with most uptake are marginal, notably, easy-to-incorporate technical additions to traditional pedagogy that do not require fundamental change in how faculty think about their pedagogy. (cf. Note 1)
- Also more often adopted are innovations that are “concrete” and those that can be grasped without much explanation. (cf. Note 2)

**Examples of high level uptake evident in many institutions N=8**

- Over 1,000 institutions offer some form of learning community program, many of which are longstanding and regularly offered. They are established in all types of institutions (research universities, engineering schools, regional public institutions, liberal arts colleges, and community colleges). In contrast with individual class uptake, they require teams of faculty and administrators to set up and run them. (Taylor et al., 2003)
- Other resources that are widely used in many institutions: POGIL (In 2009: 800-1,000 people "were doing something they identify as POGIL"); PLTL (2008 Director’s survey yielded 300 users); various forms of small group learning were reported as widely used, as also were SENCER courses, and the SALG online instruments.

- Growth of professional development programs in many institutions *e.g.*, DELTA program (University of Wisconsin) is offered annually and attracts good participation levels; CAEE engineering teaching portfolio program for graduate students.
- Growth of T&L Centers and Engineering Education Centers. Estimates of faculty usage however vary. CAEE report their “CELT” model is successfully used in other centers and departments. Others report some good T&L Centers with low faculty use of their services.

### **Departmental, Divisional, and Whole Institution Scale-up N=12**

- All project directors offered examples of departmental, sub-departmental (*e.g.*, all introductory courses), and institutional uptake of curriculum or pedagogy reforms, some over long periods. Examples offered in interviews were the iceberg tip. They (and others informants whom they cited) had many more examples. They pointed to the value of collecting and categorizing these in a public national registry that was regularly updated. (cf., Note 3)
- The strongest example of success in a nationwide reform effort was agreed to be achieving higher rates of enrollment, persistence to graduation, and entry to doctoral programs, of women in STEM disciplines. Success was attributed to sustained effort and funding, good publicity, and to both on-campus and nationally-organized women’s groups (*e.g.*, WISE, WEPAN, SWE, NCWIT). However, this major reform was achieved outside of the classroom rather than by pedagogical or curriculum changes.

### **Factors affecting progress N=8**

- Uptake nationwide is seen as variable and uneven. Some institutions (*e.g.*, community colleges and liberal arts colleges) are more flexible in their teaching structures and have cultures that are more open to change.
- Some saw greater uptake in both engineering and physics. (Not a unanimous view). Is engineering inherently more geared to change and responsive to outside influences?
- “Big ideas” like adoption of the scholarship of teaching and learning or, learning through diversity are more appealing to senior administrators than to faculty who value resources that they can immediately use. (cf., Note 4)
- Observes a time lag for uptake. (*e.g.*, PLTL director sees progress in getting PLTL into department curricula as “the fruit of seeds sown 8-10 years ago.”)

### **Estimates of little or no change N=8**

- That STEM faculty teaching is unchanged is seen as the dominant NSF view: “that there is little or no connection between research on student learning and what’s happening in the classroom.” (cf., Note 3)
- Some long-time project directors report no signs of uptake of their own resource, or value shifts towards other reforms, in their own institutions. “Without institutional framework and college-level support, innovation all falls on the individual professor.”
- In presentations, some directors reported getting the same questions about their initiative that they got 10-12 years ago. They ascribed this to faculty being unaware of evidence of efficacy in the research literature that they and others have produced.

- Notwithstanding the examples of uptake they offered, for many initiatives, the impact of particular initiatives has not spread much beyond the immediate group of innovators. (This pattern was first noted in the EXCEL coalitions in 1980s).

*Table 2 Notes*

1. “People pick up the stuff they can easily use without too much work on their own part.” “We get most frequent hits on our web-site to find out how to use teaching portfolios and download syntheses of research papers.” The most common example of this kind of uptake was what one panelist called, “the clicker explosion.” Uptake of this resource was thought to be widespread because it took little effort and gave the appearance of making a change without actually disturbing the teaching status quo. “It makes faculty look good in a climate favoring change.” Panelists did not decry the value of clickers, but to use them meaningfully took some work on the teacher’s part: “If used effectively, personal response systems are very effective. But they are often used in trivial ways—as a ‘fancy thing,’ something that’s ‘really cool.’ The issue is the quality of the questions for which clicker responses are solicited.”
2. The CAAE director, Cindy Atman, offered her observations that the classroom innovations most likely to get adopted are those that are “concrete” and/or “sticky.” She recounted how the application of misconception research took off among engineers because it is inherently concrete: the basic idea is straightforward; the research results match what teachers experience in class; student misconceptions can be identified; and there are available strategies to help address them. Concept inventories have a similar concrete appeal. “Sticky” innovations are those that many people relate to and seem to “get” with very little explanation. “They seem to speak to people.” She cited widespread uptake by engineers of visual representations of design process data as an example.
3. Panelists debated how much NSF officers knew about the extent of uptake or scale-up. NSF officers feel they know what is happening because they get constant feedback from the field. However, some project directors thought individual NSF officers have only a partial view of how much change is happening. They proposed a national inquiry to establish the extent of uptake of established resources, and of formal adoption (scale-up) in particular institutions. Every project director cited examples of institutional adoption of their own resources and pointed to others who could do the same. We invited all the project directors to offer a list of institutions, schools, and departments that had formally adopted their projects’ products. Two directors (Learning Communities and SENCER) were able to provide provisional lists in the limited time available.<sup>1</sup> All directors indicated that, given more time, and (in two cases) some resources, they could generate such a list. They were unanimous that a formal effort to collect, categorize, and publish this information would be valuable for many purposes.
4. One evaluator offered the view that discouragement about the extent or pace of change may arise because program organizers miss the distinction between what they think should be taken

---

<sup>1</sup> Over 1,000 institutions offer some form of learning community program. Those that enroll large numbers of students each year in LC programs include 7 research universities, 5 community colleges, 4 regional public institutions, 3 liberal arts colleges, and an unspecified number of engineering schools. Clusters of SENCER courses have been adopted in 2 private universities, 1 research university, 2 community colleges, and 1 public liberal arts college. Disciplines known to use SENCER courses include, biology, chemistry, mathematics, anatomy and physiology, natural sciences, and integrated linked courses in the sciences, social science, and humanities.

up and valued and what is actually taken up and valued. Gains to participants that appear in evaluation data tends to be of the latter kind. “The logic of this is to rethink program goals.” The same evaluator questioned why project directors and funders want innovations to scale-up, as opposed to spreading among faculty more organically. “It’s a distributive approach whose traditional indicators are scale-up by geography, university systems, disciplinary majors, or non-majors. It’s a widgets approach to innovation. The more we focus on the broader impacts of the project, the more we are apt to lose the intellectual merit.”

## **Question 2. What indicators or measures did panel members use to reach their assessments?**

We asked the panel members to explain how they had reached their assessments of progress that are reflected in Table 2. What formal or informal indicators had they, or their project evaluators, used to get a sense of progress, both for the resources they had developed, and for the STEM education reform effort overall? Although all of their projects were formally evaluated, and examples of their measurement methods were offered, the relative volume and detail of their answers indicated that panel members relied far more on what they had directly observed than on formal change measures. That said, they were acutely aware of the need for good measures of change, described some of their problems with existing measures, and explained what kinds of measures they would like to have. These observations are gathered into Tables 3, 4, and 5. Table 3 summarizes the informal, largely observational, indicators used by panel members to monitor progress in STEM education reform, Table 4 offers samples of outcomes found by formal measures, and Table 5 outlines problems with current measures and unmet evaluation needs.

### **Table 3. Change Indicators used by Panel Members to assess Progress**

#### **Indicators of institutional uptake of innovations N=10**

- Gave examples of institutions that have provided money and resources to keep innovations going.
- Central resources are created (some with endowments), notably, teaching and learning centers, creation of new staff positions to service them.
- Examples of institutions that continue staff positions required to service innovations after external grant ceases.
- Development in some institutions of professional development programs for faculty, post-docs, and/or TAs.
- Development of a national directory of participating institutions (*e.g.*, All directors of learning communities are asked to register).

#### **Within larger initiatives, communities and networks have developed, and are attracting new recruits N=8**

- Community activity in larger initiatives (*e.g.*, regional networks of learning communities). Consortia of participating institutions hold meetings and conferences for faculty and administrators; deans and provosts become engaged through these meetings. Their national conferences’ attendance figures are also indicators of growth
- Summer institutes that have run for multiple years promote ideas, research, and know-how, build networks and work teams (*e.g.*, cooperative learning, learning communities, and SENCER)
- Reform communities are created and maintained by list serves and online networks.

- Attraction and engagement of recruits who then become part of projects' dissemination efforts.
- Growth of international circles that discuss student-centered education issues. Project directors and researchers get many invitations to talk about their own work and education change in the US. Builds awareness and prestige.

#### **Indicators in higher education program growth and published scholarship N=7**

- Articles, monographs, and books appear in increasing numbers on educational scholarship authored by STEM education practitioners and researchers.
- Ph.D. programs have developed in engineering, chemistry, and physics education, and an increasing number of doctoral theses focus on STEM education topics.
- More journals publish articles on STEM education issues, and new (including online) journals have arisen to carry them. Some reform groups have their own journal.

#### **Greater awareness and acceptance of T&L research and teaching methods N=4**

- Encounter more awareness of T&L research at all levels of academe.
- Aggressive expressions of skepticism about student-centered forms of instruction that were common in presentations and workshops 15 years ago have disappeared. Even if this reflects political correctness, such behavior is no longer legitimated.
- Meet more college administrators who are aware of the kinds of change that are needed to improve student learning.
- Institutional documents (*e.g.*, mission statements), and the language they employ, indicate some reform ideas have become "givens".

#### **Spontaneously-offered reports of use and promotion of innovations N=3**

- Directors (*e.g.*, CAEE, POGIL and PLTL) learn about uptake of aspects of their innovations when people spontaneously contact them. They encounter strangers at meetings who describe uptake of their innovation.
- These encounters also indicate that adapters are enabling further uptake:

"I no longer feel have to be out there convincing people because there's enough knowledge of it out there, and it's other people who do the convincing on the basis of their own practical experiences."

#### **Institutional shifts in formal criteria for hiring, promotion, and tenure N=3**

- In some institutions, recruitment criteria now include familiarity with particular teaching achievements, and scholarship of teaching and learning, Caveat: Traditional criteria may persist; panel cited examples of administrator interventions to correct decisions made without reference to new formal criteria.

#### **Publication of articles on effective use of an innovation; their citation in the presentations and proposals of others N=2**

- Inclusion of these articles and citations in faculty portfolios and their acceptance for promotion and tenure purposes.

### **Positive shifts observed in faculty attitudes toward their teaching role N=2**

- Observes shifts in orientation towards students as learners in faculty stories told in summer institute presentations. Beyond such "conversions" or "moral shifts" faculty do not return to traditional teaching methods.
- Evaluators found that the DELTA (professional development) program "is helping to correct for the anti-teaching element in professional academic socialization: it is helping young faculty to find their way back to the ideal of becoming good teachers. It enables them to embrace teaching both as a role and as an identity. It achieves their rediscovery of teaching. The logic of these findings is that this should be a central program goal."

### **Behavioral shifts noted include faculty attention to their teaching outcomes N=2**

- See faculty adapting or developing learning assessment instruments that reflect their changed content or pedagogy.
- Faculty make more use of student feedback tools (*e.g.*, the SALG online instrument) to solicit student feedback "in real time," as well as at course end.

### **Innovation spreads into new disciplines or from the sciences into non-sciences. N=2**

- POGIL was recently taken up in mathematics and biology; Undergraduate research has spread into social science and humanities in multiple institutions.

Taking the findings offered in Tables 2 and 3 together, the panel members offered a more comprehensive, detailed, and nuanced picture of the extent and nature of progress in improving the quality of STEM undergraduate education than does the existing body of published work. Notwithstanding their commitment to the reform effort in broad, and their projects' contributions to it, their assessments are not rosy-tinted. They describe a wide spectrum of uptake and scale-up with both encouraging and discouraging dimensions. They also reference a process of faculty development that includes stages of increased awareness of teaching alternatives, and shifts in willingness to try them, as precursors to change. Early in the change process, they record the tendency to pick up what is easiest and disturbs traditional teaching methods least. However, they also observe (with examples and some documented data) "widespread" use of research-grounded methods, both singly and in combination. Exactly how extensive this usage is nationwide remains, as they point out, to be documented.

The panel members also identified project directors and evaluators as primary sources of more information about the uptake of project-developed products either than is currently available in the public record, or than has been collected from them—singly, or, more significantly, across funded projects. Some are clearly aware of, and some could provide, categorized listings of institutional and departmental implementation of reformed curricula and pedagogy, and also instances of high levels of implementation that fall short of formal adoption. Others indicated that they would need additional funding to conduct such inquiries, but see themselves as well-placed and motivated to undertake this work. From their examples listed in Table 2, the category of "high-level uptake" of any resource both within an institution and across institutions (*i.e.*, without formal adoption) looks like a worthwhile target for inquiry its own right. Finally, the array of change indicators offered includes more dimensions than are typically included in a formal evaluation study. Although some indicators, such as information acquired in spontaneous

encounters, could not easily be formalized, many others could form the basis for a set of evaluation questions to use singly or in combination.

All of the projects represented in the sample also used formal evaluation methods. However, panel members' accounts of evaluation questions they had asked, the measures that they had used, and the results these had yielded (summarized in Table 4), are clearly only a selection of all the evaluation methods deployed in their projects. What panelists raised appeared to be examples of what stood out as having particular importance.

**Table 4. Evaluation Questions, Measures, and Findings Reported by Panel Members**

**Evaluation questions explored by surveys N=5**

- POGIL: What was the extent and nature of current and planned usage (annual self-report survey)?
- The Reinvention Center: How much T&L Center participants had developed along lines recommended by Boyer Commission?
- CAEE: How many engineering faculty were doing education research?
- CIRTL: Which aspects of the DELTA professional development program were of shorter and longer term use to participants?
- RAIRE: How much did the introduction of undergraduate research experiences increase?

**Findings from intermediate and longer-term data collection N=3**

- PLTL: 10-12 years out, young faculty were using classroom methods learned as undergraduate peer leaders.
- CIRTL/DELTA: Immediate gains morphed into other gains with time; Gains in practical classroom know-how were far stronger than development of teaching-as-scholarship.
- SENCER: Reformed courses got into the curriculum and stayed there.

**Measures of student learning and learning retention N=2**

- SENCER: Use of Student Assessment of their Learning Gains (SALG)<sup>2</sup> instruments as part of project evaluation and student feedback to individual faculty.
- CAEE: Development of instruments to gain insight into student learning gains.

**Meta-analysis of evaluation report data N= 2**

- The National Learning Communities Project (Taylor et al., 2003) used meta-analysis of 17 project evaluation reports to establish extent of LC use, and outcomes for students and faculty.
- NSF uses evaluation reports for internal assessment of individual innovation uptake and extent of overall change.

**National survey includes question(s) about institutional use of an innovation. N=1**

- The National Survey of Student Engagement (NSSE) includes questions about institutions' use of learning communities.

---

<sup>2</sup> [www.salgsite.org](http://www.salgsite.org)

Panel members were concerned that the STEM education reform communities of which they were a part were experiencing difficulty in documenting the nature and extent of the gains that they observed. They discussed some of the problems that make evaluation difficult and explained their evaluation needs. These observations (which are, again, unlikely to be other than a sample) are summarized in Table 5.

### **Table 5. Problems and Needs in Measuring Change**

#### **Project evaluation data needs to be synthesized, made public, and regularly updated N=6**

An overall, national picture of change is needed. Evaluation reports do not circulate more widely than the initiative leaders and funding agency staff. Current and prospective PIs need this information. Meta-analysis of evaluation data can address this. The NSF is encouraged to develop the capacity to analyze multiple evaluation data sets. Data mining is possible: using NSF data, the PLTL director found 300 projects that were using PLTL.

#### **Most needed are of measures of student learning that match course learning objectives, and of learning retention over time N=6**

Some instruments have been developed for individual faculty (by T&L Centers), for particular disciplines (concept inventories), for project users (*e.g.*, CAEE), and for student feedback purposes (*e.g.*, SALG), but there is little available to measure what students remember and carry with them into subsequent courses. More of both are needed to assess student learning outcomes. Development of institutional research capacity is needed to track (for large samples) student learning outcomes from the introduction of proven practices or improved course content.

#### **Cultural barriers to developing common implementation measures N=5**

There are problems comparing like with like. Project directors and evaluators experience wide variation in the way that proven practices are adapted, including implementations that violate the underlying principle or omit essential elements. Only innovations that are very concrete appear to avoid this problem (*e.g.*, use of concept inventories).

Some program directors do annual surveys of current and planned use of their resource. The results are estimates because return rates are modest, defining what counts as instances of implementation makes comparison hard, and answers are self-reported.

Questions about overall progress are particularly hard to answer for large, multi-institutional projects, partly because of sheer size; partly because such collaborations seem inherently “unnatural” and meaningful commonalities are hard to find.

#### **Longitudinal data of most kinds are lacking N=4**

Original project funding often expires before this is feasible: innovations are still evolving and not in final form. Directors report that new funding for evaluation purposes is hard to secure. One project director with full-time assessment staff saw these positions as crucial to long-term tracking.

#### **Evaluation findings (including for meta-analysis) are hard to find in journals N=2**

Many journals do not accept articles based on evaluation studies. Implementation accounts are hard to track because they appear in a wide array of education journals.

### **Publication of articles on the effective use of an innovation, and their citation in presentations and proposals N=2**

This is a normal measure for judging the quality of a research paper. It should be applied to peer-reviewed education research articles for promotion and tenure purposes.

### **Younger faculty members are active in reform efforts. Uptake on innovations is also thought to be higher among younger faculty, but this is hard to measure N=2**

The extent of pre-tenured uptake of proven methods and materials is hard to establish because they “stay in the closet until after tenure. Many may do it but they don't talk about it.”

### **Fundamental problems of studying “scale-up” N=1**

"What we are seeing as evaluators is that the current scale-up model is not working. However, we don't yet have the metrics, methods, or a base in the literature to study dissemination based on ground-up observations of participant gains. That is, we don't yet have ways to study the problem."

Taken together, the questions raised and issues addressed in Tables 4 and 5 reflect a sophisticated understanding of what evaluation data it is important to collect, why this may be difficult, and what not-yet-available measures are needed. The panel also point to opportunities for new forms of data collection and for further analysis, synthesis, and publication of evaluation data already collected.

### **Questions 3 and 4: What are the sources and nature of constraints and resistance to change? What leverages change and what sustains it?**

#### **Conditions for Change: Resistance, Leverage, and Sustaining Reform**

The panel members described the conditions for change in highly inter-connected ways. They identified major sources of structural, cultural, and personal-professional resistance to fundamental changes in the ways in which STEM education is organized and conducted. They also discussed what kinds of leverage can address these and offered examples of where these have been used with success. Finally, they discussed the factors that sustain or fail to sustain otherwise successful change initiatives. Rather than offering this information in tabular form, we present our informants' observations on these three attributes of change by topic. These are augmented by contributions from our own work and that of other scholars.

#### **Changing institutions: possibilities and limitations**

Academic institutions are inherently less nimble than business organizations. They have fewer external drivers, and respond less directly to their customers and markets. Academic institutions are geared to preserving customary ways of carrying out formal tasks even when some outcomes are dysfunctional. Institutional structures service older teaching methods and do not easily accommodate new ones. This inertia is hard to break, involving as it does both structural and cultural change. Indeed, resistance to change may be seen as normal in academic institutions.

Aware of an unsuccessful history of experiments with both grass-roots-only initiatives and top-down led or mandated changes, the panel members were unanimous that successful reform requires combined top-down and bottom-up effort. It also requires institutional commitment to

the value of teaching. Collectively, they and others have expended considerable effort in convincing deans, provosts, and college presidents of the importance to their missions of high quality teaching and teaching scholarship: “The idea that research excellence isn’t enough anymore has taken hold.”

Panel members cited instances where senior administrators came to see changes as in their own best interests—as something for the President to brag about. Examples included the success of particular teaching and learning centers, undergraduate research programs, and women-in-science programs. They proposed that awareness of both successes and failures can motivate an administrative response: innovations shown to improve retention and completion rates can prompt administrative leadership of a top-down/bottom-up effort. A dismal pass rate (30-40%) in inorganic chemistry prompted uptake of a senior department member’s suggestion to try the introduction of discussion groups into all courses for one experimental semester. That this was a purely departmental decision evokes the panel members’ caveat that institutions *per se* lack the power to take on the forces that determine curriculum and pedagogy. In these matters, institutional leverage was seen as marginal, not central. What then do institutions have the power to do?

### *Establishing Official Criteria for the Faculty Rewards System*

All project directors identified formal institutional rewards systems as the main structural deterrent to faculty who are otherwise disposed to revise their teaching. As one observed:

“What is not true is that faculty don’t care about students. Rather, the strategies that would improve their pedagogy are not as yet embedded in faculty positions and rewards.”

As an important indicator of change, panel members cited examples of institutions that have changed their criteria for hiring, promotion, and tenure to include evidence of teaching skills and scholarship. These changes reflect Boyer’s (1990) proposal that research and teaching achievements be judged by parallel criteria. However, they also gave examples of resistance to revised official criteria by tenure committees who continued to operate by traditional criteria that exclude teaching excellence or scholarship. Project directors cited instances where they and others had intervened on behalf of a candidate to insist that the new official criteria be followed. One project director observed, “Some tenure processes have stepped up to the 21<sup>st</sup> century; others have a firm grasp of their rear-view mirror.”

What has not been documented is the extent of institutional changes in their rewards criteria, the numbers of faculty who have benefitted from these, or the extent of losses of faculty denied positions, promotions, or tenure because their teaching achievements and scholarship were discounted. Based on their direct knowledge, the panel was broadly optimistic that the situation for young classroom reformers was improving. One program director cited several tenure successes among his program participants, although he had no information about losses. Others thought it now more possible for young faculty to negotiate career paths that emphasize teaching. They suggested that such changes are prompted by institutional concerns to meet the criticisms of parents and state legislatures about increasing tuition in the face of complaints about poor quality teaching. However, they saw this as a stronger imperative for administrators than for departments whose criteria for hiring and approved career development remain more traditional. Again, we lack a nationwide picture of these significant indicators of change.

Aside from institutional progress in restructuring the formal rewards system, project directors described the success of using modest financial rewards to encourage and reward faculty. Several

initiatives (and some institutions) offer individual faculty small grants on a competitive basis for design and implementation of classroom change experiments. Faculty respond well even to modest rewards because these acknowledge good work and confer status. Examples were also given of administrators who offered financial awards to whole departments for undertaking sustained teaching improvements that resulted in desired (and documented) student outcomes. They saw this as a promising way to leverage departmental changes in curriculum and pedagogy.

#### *Other institutional contributions to change*

A major contribution to both faculty uptake of an innovation and to sustaining it after external funding has ceased, is to create new structures and add faculty lines and/or staff positions that service implementation. As one panelist observed, “that’s what ‘institutionalization’ means.” Additions to faculty positions have to come from institutions, not departments. Project directors cited such additions as critical elements in institutionalizing successful projects. New administrative (and physical) structures may be needed to accommodate new ways of teaching. Panelists offered examples of insistence on conformity exercised by senior administrators where a revised set of curriculum criteria has been agreed by departments. Strategic use of central resources also plays a role: for example, institutional research capacity may be extended to monitor the outcomes of an institutionally-adopted project over time.

An institutional initiative with considerable power to encourage faculty uptake of new teaching methods and materials is changing policies on faculty time allocation. Institutional policy and provision for released time to individual innovators to do their work is critical both to their motivation and their chances of success. It is especially critical for the principal investigators (PIs) of reform efforts who are also working faculty members. As several project directors pointed out, very few of their number have either full or part-time appointments that allow them to give sufficient attention to organizing reform efforts that may be multi-institutional in scope. The expectation that faculty PIs can lead a major project without appropriate time allowances from their institution undermines their projects’ administrative efficiency. The common situation of part-time PIs may be an inappropriate transposition to education reform of a traditional research funding model.

Finally, the panelists pointed to the significance for change of the power and influence of senior members both of college administration and in particular disciplines, who champion education reform. In Seymour (2001), we pointed to the importance of what we dubbed, “radicalized seniors” in publicly promoting educational improvements, legitimating their uptake, protecting younger faculty reformers from negative consequences of their work, and using their power and influence to leverage change at the national, institutional, departmental, and disciplinary levels. Panelists cited examples of Nobel Laureates, college presidents, provosts, and deans who have spear-headed reforms in particular institutions or nationwide. The panelists were unanimous in their view that getting proactive administrators and disciplinary leaders was essential to durable within-institution and nationwide change. They cited examples of such leadership: the seven engineering deans who put the EXCEL coalition together; and a group of senior women who insisted their institution reverse its decision not to continue to fund its WISE program. In one ADVANCE program, senior women administrators insisted that departments include suitably qualified and experienced, locally-available women in their hiring searches, and refused to consider its hiring proposals without evidence of this. Some program directors have risen to senior positions in their institutions, and now hold seats on key committees such as finance, hiring, promotion and tenure—a development that gives change advocates more direct leverage.

### *Role of Professional Development*

Administrators can also provide the resources for teaching and learning centers and other forms of professional development for faculty, graduate students, post-docs and graduate teaching assistants. They need also to actively leverage their use. Panelists urged that improving the teaching capacity of young and future faculty is crucial for nationwide reform: institutions need to establish professional development for current and future teachers so that they do not have to invent them from scratch. Often neglected, they urged that development of institution-wide teaching assistant education programs grounded in sound teaching methods to support the work of faculty reformers with larger classes and as part of graduate professional socialization (see also Seymour, 2005).

Panelists commented on the growth of teaching and learning centers and advocated making them the spearhead for change. Institutions need to set up and support centers that provide practical assistance (notably with learning assessment) to all teachers (including post-docs, instructors, and graduate students). However, they argued that creating centers is not enough, citing examples of well-run centers that faculty do not use. They urged that institutions make them accessible, and promote and reward their use: "We can't wait for faculty to discover them. You have to get institutions to push their use." A common problem is that centers are under-used because their staff have education, but not disciplinary degrees, or are post-doctoral appointees. To be respected, center staff need high skill sets and status derived from research engagement. CAEE's "CELT" center at the University of Washington School of Engineering claims success partly because its director had status as an engineering faculty member, a Ph.D. in education, and "looked credible" as an older faculty member.

Outside of institutions, and predating teaching and learning centers, regional, disciplinary based workshops have, for two decades, been the main carriers to colleagues of knowledge and know-how developed by larger initiatives (*e.g.* the chemistry coalitions' MID workshops). Project Kaleidoscope is probably the best known example. There is strong evidence that workshops grow cross-institutional reform by drawing in, educating, enabling, and incorporating recruits into the change effort (Andrews, 1997; Hilsen and Wadsworth, 2002; Connolly and Millar, 2006). Project directors who had a good deal of experience with workshop organization observed that, to work optimally, they must be of sufficient duration, offer repeated exposure in a progressive sequence, provide support for new reformers back in their home departments, use "old hands" as facilitators, and build facilitator capacity among newer recruits:

"Workshops give strength, motivation, and skills to people who don't feel a connection to their department colleagues."

"They offer portals to people who are like-minded, people that they wouldn't ordinarily meet—people in different disciplines, senior people."

Panel members observed that the continuing role of reform-associated workshops has been thrown into doubt because of difficulties in finding adequate funding. In recent years, professional and disciplinary societies have stepped into this role by mounting workshops as part of annual meetings.

## Getting Faculty Buy-in

### *Role of internalized disciplinary and academic cultures in resistance to change*

A central question on the minds of STEM education reformers and of their funders, is, notwithstanding greater awareness of research-grounded teaching methods, why do faculty not use more of the sound instructional practices that are available to them? To which one project director added: “Why don’t faculty apply to their teaching the same standards and care that they use in their own research?” We cannot attempt a comprehensive answer to either question here. However, we offer some insights from panel members who have longstanding experience in seeking practical ways to address them.

Several strong themes are evident in the panelists’ analysis. First, one cannot assume that faculty members are free agents. Their professional choices are shaped by their professional socialization beginning in graduate school or earlier. These are reinforced by cultural values, norms, and beliefs that are widespread in their discipline and reinforced in their departments—norms that are often strong enough to withstand institutional change efforts (see also Seymour, 2001). Added to this culture over the later 20<sup>th</sup> century is loss of respect for teaching, which is also widely echoed in the general population (Seymour, 2006). One panelist cited Barzun’s (1991) observation of academe’s loss of a sense of teaching as “a calling.” Another pointed to the academy’s de-professionalization: unlike other professions, academe no longer offers its entrants an education for one important part of its professional role. Learning how to teach well has, thus, to be retrofitted. Meanwhile, the dominant cultural norm has been that teaching is a “talent”: thus, training may be perceived as ineffective for the untalented and unnecessary for the rest. It is this culture that prompted panelists to press enlightened administrators to provide adequate systems of professional teaching and learning education and push for their use.

The panel also discussed the extension of the norm of academic freedom to include the beliefs that every classroom is unique and that faculty are the best judges of how to teach any class. Many of the panel themselves subscribed to this view of professional “rights”. However, they also saw this as the main source of a dominant tendency to adapt rather than adopt innovations. In journal accounts, presentations at meetings, and live encounters with adapters, they commonly find usages of their pedagogical methods that miss essential elements or violate central principles. Their observation is supported by Dancy and Henderson’s (2010) findings that, in implementing innovations, faculty altered them in ways that made them more ‘traditional’ and were significantly different from the original research-grounded techniques. The panel debated whether this widespread adapter behavior is explicable in terms of the inevitable variability of teaching situations, or in terms of this belief alone. One panelist defined the presumed uniqueness of all classroom settings as a myth, pointing to the dominant pattern in research universities of very large classes taught in a uniform manner. In either case, it creates hybrids that panelists described as SENCER-like, POGIL-like, PLTL-like, SALG-like, etcetera, that make evaluation and meta-analysis very difficult (cf., Century et al., 2010).

### *Student Resistance*

The panelists observed that the culture locks faculty into some dysfunctional teaching norms and practices. These are highly evident in their tolerance of high rates of student failure and of the “normative resistance” (Seymour, 2005) found among students in many traditionally-taught (especially large) classes. Students exhibit an array of disquieting behaviors that express disinterest, disengagement, and disrespect (of each other and their professors). Students distance

themselves from their classes by “arriving late, sitting at the back, talking, reading non-class material, doing work for other classes, eating, sleeping, and failing to ask or answer questions” (Seymour, 2005 p. 128). These behaviors are well-known to faculty, and, for those who are hesitant to change their teaching methods, may be a rationale for ‘letting sleeping dogs lie.’ Some panelists noted, (and we and others have discussed) culturally-supported explanations for dysfunctional behaviors that deflect blame away from teaching methods and onto (presumed) inherent students flaws (Seymour and Hewitt, 1997; Campbell et al., 2001; Seymour, 2005). Tolerance of bad student responses to their teaching does not necessarily prompt faculty to try other methods because to do so carries a high emotional load of risks, fears, and potential losses. By using unfamiliar materials and methods, faculty fear exposure of what they do not (yet) know, loss of control over class activities, collegial accusations of reducing the “rigor” of their courses, or failing to cover a normatively defined canon. They also fear student resistance that will be reflected in their institution’s course evaluation surveys that (albeit woefully irrelevant) are nevertheless included in their tenure portfolios as evidence of teaching competence.

Notwithstanding reforms in K-12 teaching methods, many students enter college having learned ways of navigating course material that focus on collection and memorization of information. Little may have been demanded of them by way of inquiry, independent reflection, or application of knowledge. Early college classes, traditionally taught, do not disrupt learning techniques focused on passive absorption and short-term memorization. However, students find that their ‘learning’ methods do not work in courses where teachers demand active engagement with ideas, independence of thought, application of ideas in new situations, and responsibility for learning. Such approaches expose inadequacies in both students’ traditional learning practices and their depth of conceptual understanding. In essence, an implicit “contract” between students and teachers about what each expects of the other in terms of teaching and learning behavior is suspended. The resulting student discomfort commonly provokes expressions of anxiety, complaints, withdrawal, and both passive and active resistance. Innovators can expect student resistance to be most acute in early stages of reform. Seymour and Laursen (2005) report that innovating faculty can typically expect two semesters of resistance from the inception of a course requiring active, interactive, or inquiry-based learning of students. Acceptance and appreciation develop as faculty gain experience and confidence in using their new materials and methods, as students and TAs become aware of increased learning gains, and the approach comes to be seen as “normal” for this course (Seymour, 2005).

#### *Problematic departmental responses*

The panel was agreed that discovering how to leverage change in departments is critical, and “that we have never really known how to do it.” Routine resistance to teaching changes is, as panelists noted, so strong in some sub-disciplines (*e.g.*, organic chemistry), it can hardly be coincidental, and is best explained in terms of deeply internalized cultural norms about appropriate teaching methods. Faculty who choose to come to a summer institute or workshop have to struggle to break loose from these internalized restraints. As one organizer observed, “although our whole program is built on a critique of traditional practice, we avoid direct attacks which only provoke resistance. Instead, we find ways to get participants to see the critique and its personal consequences for themselves.” The power of departmental culture and customary practices makes it difficult to see “when and how STEM faculty get to exercise choice.”

That said, panel members observed that the nature of resistance has shifted because it has become harder to ignore widely circulating knowledge of improved teaching methods and their results. Many more faculty now know about research on how students learn best and why lecturing is less

effective than more active and interactive forms of pedagogy. They also hear directly from colleagues at disciplinary meetings and elsewhere about good results from research-grounded teaching. The panel acknowledged the growing role of disciplinary societies in spreading and thereby legitimating uptake of good practices. As one panelist observed, it had, thus, become a more conscious "moral" choice not to try them. Others added the observation that the ground has shifted from arguing that the evidence of efficacy is insufficient, to the rationale that to make changes would take time and effort that they simply cannot afford. Beichner and colleagues (2007), Dancy and Henderson (2008, 2010), and Fairweather (2008) all present evidence of faculty claiming that the time commitment involved deters them from making changes in their teaching methods. Panel members included among the contributions that institutions can make, providing faculty with time for educational development work. However, where faculty resist trying something already developed, they treat "it will take too much time" as a culturally-legitimated fend-off.

They offered some partial strategies for departmental change from their experience. Securing the buy-in of chairs is critical because improving basic courses requires their help, and it is largely chairs that decide who teaches what courses. Chairs can also protect the early years of the curriculum from pressures by serviced majors (largely engineering and premed). As discussed, several panelists cited the effectiveness of institutions offering annual grants to departments that demonstrate they are improving student learning and persistence, and reducing drop-out rates.

In a situation of greater understanding that something has to be done, particularly to address problems in introductory courses, panelists lamented a common departmental response, namely to side-step problems with questionable teaching methods by employing adjuncts, lecturers, and other specially-appointed faculty to teach basic courses, including non-majors and serviced majors. One panelist labeled this as a form of "unbundling", that is, the corporate practice of separating and redistributing its component parts. Others commented on the lower departmental status and salaries of faculty who are hired "just to teach" (cf., Fairweather, 2008).

*Under these conditions, what may best secure faculty engagement?*

The panel members focused on four main strategies. The first, establishment and active promotion of professional development, has been discussed. Three others are discussed below.

#### Building communities and networks

Panel members pointed to many examples of communities and networks that function to keep like-minded activists supported, motivated, and engaged. Historically, these have developed in larger initiatives and are sometimes referred to as "communities of practice." They provide empowerment for faculty who might otherwise be isolated, accumulate a reservoir of know-how and resources for members to tap into, and offer the satisfactions of working together on concrete projects, often with people they might not otherwise meet. Much of the daily life of such communities and networks goes on by electronic means. However, panelists observed from experience that finding funds for meetings, institutes, and workshops is money well spent. Face-to-face occasions bring together experienced classroom innovators with interested newcomers. Workshops and active working collaborations build the reform effort by attracting, training, motivating, and supporting recruits who become the next activists and carry the reform effort forward. Engagement in community conversations also enables faculty to deepen their understanding of the underlying principles of an innovation and develop a sense of ownership of the adaptations they develop. Ownership was identified as one of the factors that sustain

individual reformers over the long term, and was also identified in the literature as a key component of successful institution-wide change (Elizondo-Montemayor et al., 2008). Increasingly, disciplinary and professional societies play an active role in developing communities of education research and practice and providing the venues to sustain working collaborations.

Several panelists offered examples in a single institution where development of a cross-department community of faculty and administrators, and, in some instances, students, advisors, and staff, was sufficient to instigate and sustain a reform effort. These included effectively addressing the campus concerns of women via the ADVANCE initiatives. The diverse composition of such communities was seen to contribute to their effectiveness.

#### Offering faculty resources they can find and use

Putting their observations together, the panelists' collective theory of change was that faculty are more likely to try out high quality materials and methods if they are introduced to them by people they respect, if they can easily find them, and if they are something they can easily use. Because adaptation rather than adoption is the norm, the core elements and underlying principles of an innovation also need to "make sense" to faculty. Fairweather (2008) also argues that faculty need good, practical information on implementation strategies.

On the topic of what faculty can use, one director reported that pro-actively disseminating the results of using learning assessments developed by project was a catalyst for their wider use. As referenced in Table 2, one project evaluator reported that what early career participants took away and used (both immediately and over time) from professional development experiences was practical know-how in teaching methods and classroom management rather than more complex professional activities such as education scholarship. Also in Table 2, Note 2, we offered CAEE's project director's observation that the classroom innovations most likely to get adopted are those that are "concrete" and/or "sticky," that is, easy to relate to.

What faculty can use is also related to the time and effort it is perceived to require. The panelists argued that, for reforms to be taken up and get established, they must not demand more than faculty can give over time. Borrego and colleagues (2010) also found that, among engineering faculty, the reforms requiring the least amount of coordination and time commitment were the most likely to be used.

The panel also urged that getting knowledge and know-how out to faculty requires more proactive forms of dissemination. "It is critical to synthesize what we know for faculty to find easily and use." This will require more than placing materials in National Digital Library collections which are far less searched for materials than popular search engines (Manduca et al., 2006; McMartin et al., 2008). In addition to published accounts of successful innovations, "the NSF has a lot of evaluation data that should be synthesized and published in a single, easily-found place with regular updates." Several panelists mentioned current Congressional and Presidential pressures on the National Science Foundation to compile and use what they know works to improve graduation and address drop-out rates. The panel strongly encouraged the NSF to use this situation to "give themselves permission" to become more actively engaged both in synthesizing and disseminating what they know, and directly promoting proven practices. The recently-released President's Council of Advisors on Science and Technology (2010) report goes further: "relatively little Federal funding has historically been targeted toward catalytic efforts with the potential to transform STEM education, too little attention has been paid to replication

and scale-up to disseminate proven programs widely, and too little capacity at key agencies has been devoted to strategy and coordination.”

### The role of evidence

Panelists held the view that proving that a resource works is necessary, but it is not sufficient by itself to persuade faculty to use it. This view is also taken by Seymour (2001), Dancy and Henderson (2008, 2010), Fairweather (2008), Borrego and colleagues (2010), and others. We have long known that validation by disciplinary colleagues with high research prestige is a persuasive source of validation for an educational innovation (Foertsch et al., 1997). Providing such validation is one of several contributions that “radicalized seniors” make to the education reform effort. However, project directors who offer summer institutes offered another effective source of persuasion, namely the personal testimony of experienced innovators offered as stories at these meetings. They recounted the combined effect on newcomers of engaging in hands-on experiences of new methods and materials, hearing personal accounts from people who use them, and being able to raise issues openly in a safe and supportive environment. All of these experiences in combination contributed to the many “conversion experiences” that they had witnessed. Evaluators concurred that, once faculty had reached an intellectual and personal understanding of what they needed to do in their own classrooms, and had experienced using it with their students, they found that they could not return to how they had previously taught. This observation, noted also by project evaluators, underscores project directors’ strong belief in the value of live community engagement made possible at meetings, workshops, and institutes.

What, then, did the panel members see as the role of evidence? They argued that, strong evidence of success proactively disseminated in disciplinary and professional meetings or project meetings may be a good driver of uptake. As indicated in Table 2, they pointed to widespread awareness of evidence in support of many innovations. However, like Borrego and colleagues (2010) and Dancy and Henderson (2010), they observed that awareness levels far outpaced adoption levels. Beyond mere awareness of the evidence, greater faculty acceptance and understanding of the evidence in support of reforming teaching practice is a necessary precursor to change in teaching behavior. However, it does not necessarily prompt it. Both Henderson and Dancy (2007) and Norton and colleagues (2005) found that faculty conceptions and beliefs about teaching were more “reformed” than their actual practice. Finally, the panel were doubtful whether any amount of evidence would sway those most opposed to change: “We won’t get more change by producing more work that’s focused on becoming more credible with disciplinary conservatives.”

Several panelists pointed out that lack of evidence of the efficacy of a practice did not undermine the use of lecture-only teaching. Nor did it curb uptake of practices that are widely believed to be inherently good. They pointed to the major expansion of undergraduate research programs funded by large grants. It was clear in both published articles and in proposal narratives that both UR funders and directors were looking for evidence of the student gains they perceived in these experiences (and of ways to measure them). However, they went ahead without either. Their confidence in the method was based on their own direct observation of student responses and outcomes—a form of validation that the panelists broadly shared. Research-based evidence for such confidence is now available, and Laursen and colleagues (2010) offer a comprehensive review of their own and other studies of the benefits of undergraduate research. Clearly, academics do not necessarily require evidence before making pedagogical choices, but they do have to believe in their value. One panel member offered the example of the growth of Charter

schools to suggest that action-before-evidence may be normal in public as well as individual education choices.

### **Larger influences on change**

Finally, the panel discussed four types of leverage that have the capacity both to instigate and to sustain change: external pressures, creating a demand for change, transcendent concerns that overcome resistance, and the role of adequate funding.

*The significance of external leverage* has been noted since the early days of the STEM reform effort (Seymour, 2001). Observers pointed in particular to the impact on engineering education reform of its accreditation agency, the Accreditation Board for Engineering and Technology (ABET). The earliest of the major coalitions, EXCEL, was grounded in ABET's concern to restore design to the engineering curriculum, a goal which the EXCEL coalition achieved. Discipline-based coalitions can also leverage change within departments and schools by their active dissemination of know-how. The power of accrediting agencies to encourage and endorse change is beginning to be explored by current national change initiatives (*e.g.*, Mobilizing STEM Education for a Sustainable Future).

Also explored more recently is the proposition that *change can be generated by creating a demand for it*, for example, by students or employers (Zemsky, 2009). The director of SENCER described an example in which students had successfully lobbied for introduction of SENCER courses in the curriculum.

We noted in our literature review examples of widespread change within particular institutions that were a response to cross-disciplinary concern to prepare students to address serious environmental challenges (Barlett and Rappaport, 2009). Our panelists offered further examples of *transcendent concerns that had overcome administrative, departmental, and individual resistance* and had generated cross-institutional efforts to address a widely-shared concern. They offered examples of institutional responses to global and regional environmental problems. Engineering panel members cited positive school of engineering responses to the "Fourteen Grand Challenges awaiting engineering solutions in the 21st century" that were delineated in 2008 by the National Academy of Engineering. Other issues with the ability to rally administrators and faculty into concerted action were regional concerns over job shortages for graduates, reported in both "Rust Belt" schools and in California where environmentally-focused industries are developing:

"The action around here has shifted towards producing students who are globally competitive. Here institutions and departments are open to change. It's a matter of regional survival. We can't go on preparing students for jobs that don't exist anymore."

A shared concern to break the patterns of disadvantage that have historically limited entry to STEM education and professions to students of color and women of all races and ethnicities was also cited as the genesis of institution-wide change efforts.

In documenting the adoption of inter-disciplinary, environmental problem-focused curricula at two universities (Emory and Tufts), Barlett and Rappaport (2009) did not draw the broader inference that institutional and personal resistance to innovation in STEM education may be surmounted by a rallying call with sufficient power to command widespread response. However, this proposition was a topic of lively conversation with panel members. We also discussed the failure of national reports to evoke such a response by pointing to the poor performances of U.S.

students indicated in comparative international data and their dismal economic consequences for the country. Clearly, appeals to the patriotism of STEM educators are less effective than other rallying calls identified by our panel members.

All panel members discussed *the role of external, institutional, and other sources of funding in both enabling and restricting desired changes* on the premise that “money enables change.” Project directors of initiatives with uptake in many institutions were particularly concerned about how they would survive, let alone maintain growth, in projects that had established their value but were not viewed by their PI’s institution as their responsibility. On no other issue was the distinction between uptake and scale-up/reform more sharply made: “We don’t currently have any funds to run the project. If there’s any funding, it has shifted to people who are using what we developed.” There was consensus that “funding must be adequate to the size of the task. It’s not enough just to grow communities; you need to support them if you want national outcomes.” They pointed to a long-standing, but unrealistic, expectation by funders that, by means unspecified, projects will become “institutionalized,” or otherwise sustain themselves, and observed (from experience) that “the typical PI does not have the skills to invent a self-sustaining structure for a project that is being used in lots of institutions.”

As veterans of many grant proposals, they noted patterns in the solicitations and awards behavior of funders that they assessed as counterproductive for scale-up. Among these, they observed an unstated but normative timeline for achieving project goals, including the undefined expectation of “institutionalization.” Refusal of further funding after a program, however successful, is “thought to have received enough,” effectively dismantles what has been built up. One observed, “reluctance to keep a good thing going is built into NSF’s distribution of funds.” They understood that the NSF in particular was struggling with the restraints built into its original mission to create new knowledge rather than to disseminate and promote proven practices. They also understood this imperative, while part of its formal remit, might also be reinforced by fears of interfering with either states’ rights or faculty autonomy. They also wondered “if the NSF had a model for scale-up or were expecting grantees to discover it by trial and error.” Whatever the case, they stressed that the focus on innovation at the expense of consolidation and growth was self-defeating: “They never get to see what could be achieved if successful innovations were promoted and supported.”

The “mixed messages” they noted in patterns of NSF funding confused program directors. “Some projects with demonstrated success continue to get funding, but they starve and stunt the growth of others.” They tried to make sense of this. The composition of review panels, or the guidelines they were instructed to use, were thought to reduce chances of sustaining successful programs:

“I question the criteria by which a decent proposal from a project such as this with widely acknowledged teaching methods and growing faculty uptake gets rated only as ‘fair’.”

Another panelist (with rotator experience) wondered if “the rotators system creates loss of institutional memory.” Whatever their causes, they described the damaging effects of what seemed irrational withdrawals of funds:

“Cutting off funding from workshop-based projects that have developed expertise in how to do them well, that are getting good response, attracting new facilitators, and growing the movement, is disastrous. It demoralizes people who have invested a lot and have taken professional risks to do this work. It is important to reward people. It’s also important not to punish them for doing a good job.”

“Innovations are now intrinsically sustained by the concern of people to make a difference. I question whether this is enough.”

All the directors were expecting more difficult financial days ahead: “I am afraid that large budget cuts will choke off progress.” They were trying in various ways to prepare for this. For example, CAEE focused on building faculty capacity in research and scholarship in the expectation that external funds will run out and that individuals and small groups will have to continue the work unaided. Some multi-institutional projects were looking for ways to become financially independent from original host institutions that had accepted their grant “overhead” but did not support them. They were exploring endowment funding, establishment as independent “not-for-profit” entities, seeking host institutions other than universities, copyrighting their products, and exploring the option of charging fees for service. One project had shifted to a business model that transfers some of direct costs of summer institutes to the institutions that send participants.

We were surprised to discover that these long-time directors of large projects who share the same problems of survival and growth did not talk to each other. They described opportunities to discuss shared concerns and ideas with other project directors as rare. However, they thought the time ripe for a series of meetings among project directors, evaluators, and seasoned NSF officers that would offer opportunities to brainstorm ideas for uptake and scale-up grounded in their experiences. They also identified the issues of how to get wider uptake of good methods and materials and how to take successful programs, develop them to scale and sustain them, as research questions deserving of an awards category in their own right.

## **Conclusions**

On the basis of the information that we derived from published sources, directly from our panel members, and indirectly from their sources of formally and informally gathered evidence, it is at least plausible to argue that the widely cited view of the National Research Council’s report (2003b)--that little change has been made in the learning experiences of undergraduates in introductory science courses--may be insufficiently grounded. Although this report offers no evidence in support of its claim, we note that it is repeatedly cited in other work as if it were a solid source of evidence. The claim (or some version of it) may or may not be correct. However, the unfortunate practice of not checking the nature of the data supporting a claim, and the credence given to a highly respected source, may have created a research myth. Alternatively, the actual extent of progress may be only partially documented by an overly narrow set of indicators. In either case, in so far as funders view this claim as an accurate estimate of change, it has consequences for decisions affecting the reform effort.

What are the conclusions that the panel members have collectively reached about progress in STEM education reform? They offer many kinds of evidence--both documented and observed--that change is happening: some as outcomes, some as shifts in professional attitudes and values, some in the distribution of awareness and knowledge, and some as identification of stages in change as a process. They imply that what they and co-reformers “see” could (given the imagination, tools, and funding) also be captured by more formal inquiries to build a many-faceted picture of the processes and outcomes of change. They argue strongly that data already gathered by many initiatives could, and should, be mined, synthesized, and made available. They suggest new lines of research needed to establish the extent of uptake and scale-up on a national basis, and offer themselves and their fellow PIs, project directors, and project evaluators as a collective resource in developing research questions and methods. They understand many

dimensions of structural, cultural, and personal-professional inertia and resistance to change; and they offer some ways to address these that are already working but need stronger, nationwide promotion. They also understand the internal struggles of their primary funder, the National Science Foundation, that inhibit its direct engagement in promoting the proven practices that they have helped to develop, and they encourage the NSF to become an active partner in finding practical answers to the question, “How do you take successful programs, develop them to scale, and sustain them?”.

This is not the article we had expected to write when we undertook this commission. However, even with an admitted under-sampling of the available experts, and perhaps of available literature, we hope that our panel members have collectively offered enough to lift the reform community and its funders out of the slough of despond into which it may recently have fallen for want of some well-founded good news. That said, we clearly need a series of inquiries that are both multi-dimensional and national in scope in order to ground the actual extent and nature of what has and has not been achieved.

### **Acknowledgments**

The authors offer their thanks to the following colleagues for their help and insights in preparing this paper. We hope that we have correctly represented their observations.

Cindy Atman, Director, Center for the Advancement of Engineering Education (CAEE), Director, UW Center for Engineering Learning and Teaching, Mitchell T. and Lella Blanche Bowie Endowed Chair, College of Engineering, University of Washington.

Myles Boylan, Program Director, Division of Undergraduate Education, the National Science Foundation.

William David Burns, Executive Director, National Center for Science and Civic Engagement (NCSCE), PI, Science Education for New Civic Engagements and Responsibilities (SENCER).

Mark Connolly, Assistant Scientist, Wisconsin Center for Education Research (WCER), University of Wisconsin at Madison.

James S. Fairweather, Professor, Education Administration, Michigan State University.

Wendy Katkin, Founder and Director Emerita, The Reinvention Center, University of Miami.

Richard Moog, Project Director and PI for the POGIL Project, Project Coordinator for the Middle Atlantic Discovery Chemistry Project (MADCP), Professor of Chemistry, Franklin & Marshall College.

Jean MacGregor, Director, Curriculum for the Bioregion Initiative, Director, National Learning Communities Dissemination Project, Co-director, National Learning Communities Project, Senior Scholar, Washington Center for Improving the Quality of Undergraduate Education, the Evergreen State College.

Karl A. Smith, Professor of Cooperative Learning in Engineering Education, Fellow of Discovery Center, College of Engineering, Purdue University, Morse–Alumni Distinguished Teaching Professor and Professor of Civil Engineering at the University of Minnesota.

Pratibha Varma-Nelson, Professor of Chemistry; Executive Director, Center for Teaching and Learning, Indiana University-Purdue University, Indianapolis; Co-PI, Peer-Led Team Learning (PLTL) development and dissemination project.

## References

- Andrews, G. J. (1997). Workshop evaluation: Old myths and new wisdom. In Fleming, J. A. (Ed.), *New Perspectives on Designing and Implementing Effective Workshops* (pp. 71-85). San Francisco: Jossey-Bass.
- Baldwin, R. G. (2009). The climate for undergraduate teaching and learning in STEM fields. *New Directions for Teaching and Learning*, (117), 9-18.
- Barlett, P. F., and Rappaport, A. (2009). Long-term impacts of faculty development programs: The experience of TELI and Piedmont. *College Teaching*, 57(2), 73-82.
- Barzun, J. (1991). *Begin Here: Forgotten Conditions of Teaching and Learning*. Chicago: University of Chicago Press.
- Beichner, R. J. (2008). The SCALE-UP project: A student-centered active learning environment for undergraduate programs. *Commissioned Paper for the National Academies Workshop: Evidence on Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education*. Retrieved Nov. 2, 2010 from [http://www7.nationalacademies.org/bose/PP\\_Commissioned\\_Papers.html](http://www7.nationalacademies.org/bose/PP_Commissioned_Papers.html)
- Beichner, R. J., and Saul, J. M. (2003). Introduction to the SCALE-UP (Student-centered activities for large enrollment undergraduate programs) project. In *Proceedings of the International School of Physics "Enrico Fermi"* (pp. 1-17). Varenna, Italy.
- Beichner, R. J., Saul, J. M., Abbott, D. S., Morse, J. J., Deardorff, D. L., Allain, R. J., Bonham, S. W., Dancy, M. H., and Risley, J. S. (2007). The student-centered activities for large enrollment undergraduate programs (SCALE-UP). In Redish, E. F., and Cooney, P. J. (Eds.), *Research-based Reform of University Physics*. Retrieved Nov. 2, 2010 from [http://www.compadre.org/PER/per\\_reviews/volume1.cfm](http://www.compadre.org/PER/per_reviews/volume1.cfm)
- Borrego, M., Froyd, J. E., and Hall, T. S. (2010). Diffusion of engineering education innovations: A survey of awareness and adoption rates in U. S. engineering departments. *Journal of Engineering Education*, 99(3), 185-207.
- Boyer, E. L. (1990). *Scholarship Reconsidered: Priorities of the Professorate*. San Francisco: Jossey-Bass.
- Boyer Commission on Educating Undergraduates in the Research University. (1998). *Reinventing Undergraduate Education: A Blueprint for America's Research Universities*, Carnegie Foundation for the Advancement of Teaching, Menlo Park, CA. Retrieved November 2, 2010 from: <http://naples.cc.sunysb.edu/Pres/boyer.nsf/>
- Campbell, J., Smith, D., Boulton-Lewis, G., Brownlee, J., Burnett, P. C., Carrington, S., and Purdie, N. (2001). Students' perceptions of teaching and learning: The influence of students'

- approaches to learning and teachers' approaches to teaching. *Teachers and Teaching: Theory and Practice* 7(2): 173-187.
- Century, J., Rudnick, M., and Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *American Journal of Evaluation*, 31(2), 199-218.
- Chasteen, S. V., Pepper, R. E., Pollock, S. J., and Perkins, K. K. (2010). But does it last? Sustaining a research-based curriculum in upper-division electricity & magnetism. In *Physics Education Research at Colorado (PERC) Proceedings*. Retrieved Nov. 3, 2010 from <http://www.compadre.org/PER/items/detail.cfm?ID=10321>
- Connolly, M. R., and Millar, S. B. (2006). Using educational workshops to improve instruction in STEM courses. *Metropolitan Universities*, 17, 53-65.
- Coppola, B. P., Ege, S. N., and Lawton, R. G. (1997). The University of Michigan undergraduate chemistry curriculum 2. Instructional strategies and assessment. *Journal of Chemical Education*, 74(1), 84-94.
- Dancy, M. H., and Henderson, C. (2008). Barriers and promises in STEM Reform. *Commissioned Paper for the National Academies Workshop: Evidence on Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education*. Retrieved Nov. 2, 2010 from [http://www7.nationalacademies.org/bose/PP\\_Commissioned\\_Papers.html](http://www7.nationalacademies.org/bose/PP_Commissioned_Papers.html)
- Dancy, M., and Henderson, C. (2010). Pedagogical practices and instructional change of physics faculty. *American Journal of Physics*, 78(10), 1056-1063.
- DeHaan, R. L. (2005). The impending revolution in undergraduate science education. *Journal of Science Education and Technology*, 14(2), 253-269.
- Dori, Y. J., and Belcher, J. (2005). How does technology-enabled active learning affect undergraduate students' understanding of electromagnetism concepts? *Technology*, 14(2), 243-279.
- Ege, S. N., Coppola, B. P., and Lawton, R. G. (1997). The University of Michigan undergraduate chemistry curriculum 1. Philosophy, curriculum, and the nature of change. *Journal of Chemical Education*, 74(1), 74-83.
- Elizondo-Montemayor, L., Hernández-Escobar, C., Ayala-Aguirre, F., and Aguilar, G. M. (2008). Building a sense of ownership to facilitate change: the new curriculum. *International Journal of Leadership in Education*, 11(1), 83-102.
- Fairweather, J. (2008). Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education: A status report for the National Academies National Research Council Board on Science Education. *Commissioned Paper for the National Academies Workshop: Evidence on Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education*. Retrieved Nov. 2, 2010 from [http://www7.nationalacademies.org/bose/PP\\_Commissioned\\_Papers.html](http://www7.nationalacademies.org/bose/PP_Commissioned_Papers.html)
- Foertsch, J. A., Millar, S. B., Squire, L. L., and Gunter, R. L. (1997). Persuading professors: A study of the dissemination of educational reform in research institutions (Vol. 5). Madison: University of Wisconsin-Madison, LEAD Center.

- Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R. L., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S.M., and Wood, W. B. (2004). Scientific teaching. *Science* (304), 521–522.
- Henderson, C., and Dancy, M. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics - Physics Education Research*, 3(2), 1-14.
- Henderson, B. C., Finkelstein, N., and Beach, A. (2010). Beyond dissemination in college science teaching: An introduction to four core change strategies. *Journal of College Science Teaching*, 39(5), 18-25.
- Higher Education Research Institute at UCLA. (2009). The American College Teacher: National Norms for 2007-2008. Los Angeles, CA: University of California, Los Angeles.
- Hilsen, L. R., and Wadsworth, E. C. (2002). Staging successful workshops. In Gillespie, K. H. (Ed.), *A Guide to Faculty Development: Practical Advice, Examples, and Resources*. Bolton, MA: Anker.
- Hora, M. T. (2007). Analyzing cultural processes in higher education: STEM and education faculty collaboration in teacher education. *Presented at the Annual Meeting of the American Educational Research Association*. Chicago, Illinois. Retrieved Nov. 2, 2010 from <http://scale.mspnet.org/index.cfm/19915>
- Laursen, S. L., Hunter, A.-B., Seymour, E., Thiry, H., and Melton, G., (2010). *Undergraduate Research in the Sciences: Engaging Students in Real Science*. San Francisco, CA: Jossey-Bass.
- Libarkin, J. (2008). Concept inventories in higher education science. *Commissioned Paper for the National Academies Workshop: Evidence on Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education*. Retrieved Nov. 2, 2010 from [http://www7.nationalacademies.org/bose/PP\\_Commissioned\\_Papers.html](http://www7.nationalacademies.org/bose/PP_Commissioned_Papers.html)
- Manduca, C. A., Fox, S., & Iverson, E. R. (2006). Digital library as network and community center: A successful model for contribution and use. *D-Lib Magazine*, 12(12). <doi:10.1045/december2006-manduca>
- McMartin, F., Iverson, E., Wolf, A., Morrill, J., Morgan, G., and Manduca, C. (2008). The use of online digital resources and educational digital libraries in higher education. *International Journal of Digital Libraries* 9, 65-79.
- Merton, P., Froyd, J. E., Clark, M. C., and Richardson, J. (2009). A case study of relationships between organizational culture and curricular change in engineering education. *Innovative Higher Education*, 34(4), 219-233.
- National Research Council. (2003a). Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics. In Fox, M. A., and Hackerman, N. (Eds.), *Committee on Recognizing, Evaluating, Rewarding, and Developing Excellence in Teaching of Undergraduate Science, Mathematics, Engineering, and Technology*, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.
- National Research Council. (2003b). Improving Undergraduate Instruction in Science, Technology, Engineering, and Mathematics: Report of a Workshop. In McCray, R. A., DeHaan, R. L., and Schuck, J. A. (Eds.), *Center for Education, Division of Behavioral and Social Sciences and Education*. Washington, DC: National Academies Press.

- National Research Council. (2007). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Committee on Science, Engineering, and Public Policy (COSEPUP). Washington, DC: National Academies Press.
- National Research Council. (2010). *Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category 5*. By Members of the 2005 "Rising Above the Gathering Storm" Committee; Prepared for the Presidents of the National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. Washington, DC: National Academies Press.
- National Science Board. (2010). *Preparing the Next Generation of STEM Innovators: Identifying and Developing our Nation's Human Capital*. NSF Publication NSB1033. Arlington, VA.
- National Science Foundation. (1996). *Shaping the Future: New Expectations for Undergraduate Education in Science, Mathematics, Engineering, and Technology*. A report on its review of undergraduate education by the advisory committee to the National Science Foundation Directorate for Education and Human Resources. NSF Publication 96-139. Arlington, VA.
- Norton, L., Richardson, J. T. E., Hartley, J., Newstead, S., and Mayes, J. (2005). Teachers' beliefs and intentions concerning teaching in higher education. *Higher Education*, (50), 537-571.
- Oliver-Hoyo, M. T., Allen, D., Hunt, W. F., Hutson, J., and Pitts, A. (2004). Effects of an active learning environment: Teaching innovations at a research I institution. *Journal of Chemical Education*, 81(3), 441-448.
- Ono, J. K., Casem, M. L., Hoese, B., Houtman, A., Kandel, J., and McClanahan, E. (2007). Development of faculty collaboratives to assess achievement of student learning outcomes in critical thinking in biology core courses. In Deeds, D. and Callen, B. (Eds.), *Proceedings of the National STEM Assessment Conference, October 12-21, 2006* (Vol. 104, pp. 209-218). Open Water Media, Inc.
- President's Council of Advisors on Science and Technology. (2010). *Prepare and Inspire: K-12 Education in Science, Technology, Engineering, and Mathematics (STEM) for America's Future*. Retrieved Nov. 3, 2010 from <http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stem-ed-final.pdf>
- Pundak, D., and Rozner, S. (2008). Empowering engineering college staff to adopt active learning methods. *Journal of Science Education and Technology*, 17(2), 152-163.
- Seymour E. and Hewitt, N. M. (1997) *Talking about Leaving: Why Undergraduates Leave the Sciences*. Boulder, CO: Westview Press.
- Seymour, E. (2001). Tracking the process of change in U.S. undergraduate education in science, mathematics, engineering, and technology. *Science Education* (86): 79-105.
- Seymour, E. and Laursen, S. (2005). Student resistance and student learning in undergraduate science classes using active learning pedagogies: A white paper. *Ethnography & Evaluation Research*, University of Colorado at Boulder, February 2005.
- Seymour, E., with Melton, G., Wiese, D. J., and L. Pedersen-Gallegos. (2005) *Partners in Innovation: Teaching Assistants in College Science Courses*. Lanham, MD: Rowman and Littlefield Publishers, Inc.
- Seymour, E. (2006). Testimony to the Research Subcommittee of the Committee on Science of the U.S. House of Representatives Hearing on Undergraduate Science, Mathematics and Engineering Education: What's Working? March 15, 2006.

- Seymour, E. (2007). The U.S. experience of reform in science, technology, engineering and mathematics (STEM) undergraduate education. Presented at *Policies and Practices for Academic Enquiry: an International Colloquium*, held at the Marwell Conference Centre, Winchester, UK, 19/21 April, 2007. Retrieved Nov. 3, 2010 from [http://portal-live.solent.ac.uk/university/rconference/rccolloquium\\_home.aspx](http://portal-live.solent.ac.uk/university/rconference/rccolloquium_home.aspx).
- Taylor, K., with Moore, W. S., MacGregor, J. and Lindblad, J. (2003). Learning Community Research and Assessment: What We Know Now. National Learning Communities Project Monograph Series. Olympia, WA: The Evergreen State College, Washington Center for Improving the Quality of Undergraduate Education, in cooperation with the American Association for Higher Education.
- The Ethnograph, Qualitative Data Analysis Software, Version 6.0 1.0 (2010) John Seidel, Qualis Research Associates: Colorado Springs, CO.
- The Reinvention Center. (2002). Reinventing Undergraduate Education: Three Years After the Boyer Report. The University of Miami.
- Walczyk, J. J., Ramsey, L. L., and Zha, P. (2007). Obstacles to instructional innovation according to college science and mathematics faculty. *Journal of Research in Science Teaching*, 44(1), 85-106.
- Zemsky, R. (2009). *Making Reform Work: The Case for Transforming American Higher Education*. Piscataway, NJ: Rutgers University Press.