JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 48, NO. 8, PP. 952–984 (2011)

Facilitating Change in Undergraduate STEM Instructional Practices: An Analytic Review of the Literature

Charles Henderson,[1] Andrea Beach,[2] and Noah Finkelstein[3]

[1]*Department of Physics and Mallinson Institute for Science Education, Western Michigan University, 1903 W. Michigan Ave., Kalamazoo, Michigan 49008-5252*
[2]*Director of Faculty Development, Western Michigan University, 3430 Sangren Hall, Kalamazoo, Michigan 49008*
[3]*Department of Physics, University of Colorado Boulder, UCB-390, Boulder, Colorado 80309-0390*

Received 6 June 2010; Accepted 3 August 2011

Abstract: This article reviews current scholarship about how to promote change in instructional practices used in undergraduate science, technology, engineering, and mathematics (STEM) courses. The review is based on 191 conceptual and empirical journal articles published between 1995 and 2008. Four broad categories of change strategies were developed to capture core differences within this body of literature: disseminating curriculum and pedagogy, developing reflective teachers, enacting policy, and developing shared vision. STEM education researchers largely write about change in terms of disseminating curriculum and pedagogy. Faculty development researchers largely write about change in terms of developing reflective teachers. Higher education researchers largely write about change in terms of enacting policy. New work often does not build on prior empirical or theoretical work. Although most articles claim success of the change strategy studied, evidence presented to support these claims is typically not strong. For example, only 21% of articles that studied implementation of a change strategy were categorized as presenting strong evidence to support claims of success or failure of the strategy. These analyses suggest that the state of change strategies and the study of change strategies are weak, and that research communities that study and enact change are largely isolated from one-another. In spite of the weak state of the literature, some conclusions related to the design of change strategies can be drawn from this review. Two commonly used change strategies are clearly not effective: developing and testing ''best practice'' curricular materials and then making these materials available to other faculty and ''top-down'' policy-making meant to influence instructional practices. Effective change strategies: are aligned with or seek to change the beliefs of the individuals involved; involve long-term interventions, lasting at least one semester; require understanding a college or university as a complex system and designing a strategy that is compatible with this system. © 2011 Wiley Periodicals, Inc. J Res Sci Teach 48: 952–984, 2011

Keywords: teacher change; science education; higher education; pedagogy; professional development

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: National Science Foundation; Contract grant numbers: DRL-0723699. SES-0623009.

Correspondence to: C. Henderson; E-mail: charles.henderson@wmich.edu

DOI 10.1002/tea.20439
Published online 2 September 2011 in Wiley Online Library (wileyonlinelibrary.com).

Recent decades have seen increasing calls for fundamental change in the teaching of science, technology, engineering, and mathematics (STEM). National commissions, state panels, university administrators, and individual researchers have expressed concern that the United States will lose its role as a leader in science and technology fields due to outdated and inappropriate instruction practices (e.g., Center for Science, Mathematics, and Engineering Education, 1999; The White House, 2009). The broader conversations in higher education have been dominated by the need to fundamentally shift the entire undergraduate education paradigm from instruction- or teacher-centered to learning- or student-centered (e.g., Barr & Tagg, 1995). These concerns have led to significant expenditures of time and money on research and development to improve teaching and learning. Yet, these efforts have met with only modest success (Bok, 2006; Boyer Commission on Undergraduates in the Research Universities, 1998; Handelsman et al., 2004; Kezar, 2001; Seymour, 2001).

As we began to investigate approaches to solving this problem of impact, we quickly noticed that there are at least three distinct research communities involved in efforts designed to improve undergraduate instruction in STEM disciplines. *Disciplinary-based STEM education researchers* (SER) are typically situated in STEM-related departments, frequently in a college of arts and sciences, but sometimes in a college of engineering or as disciplinary STEM specialists in a college of education. SER researchers are particularly focused on studying student learning within their disciplines. This knowledge is used to develop and disseminate discipline-specific curricular materials to improve student learning. *Faculty development researchers* (FDR) are typically situated in centers for teaching and learning. The mission of these centers is commonly to provide professional development for all instructors at an institution. Therefore, FDR researchers often focus on providing faculty with more general pedagogical skills or motivation and tools for self-improvement. *Higher education researchers* (HER) are typically situated in a departments of educational leadership in a college of education and, sometimes, in university administration. HER researchers often study how cultural norms, organizational structures, and state and national environments and policy influence higher education practices. In contrast to the other groups, HER frequently focus their efforts beyond individuals to study and impact institutional or national level issues.

In addition to unique disciplinary interests and histories, each of these groups has its own distinct professional societies, journals, and disciplinary norms. The work of these groups is deeply complementary. One would hope that these three research communities, each of which has compelling findings in STEM education, regularly share their findings with each other, since doing so would yield better understanding and theory regarding change in teaching and learning. However, as we discuss later, connections between these research communities are rare. Thus, we see the need to provide an overview of the current thinking from these and related communities that study how to promote change in instructional practices used in undergraduate STEM. We see the promise of interdisciplinary research to address this problem and believe that interdisciplinary research will be most successfully built on the foundations developed by these currently isolated research communities. The framework for categorizing change strategies presented in this manuscript provides a useful tool for facilitating such interdisciplinary research.

We define several key terms related to change as follows:

Instructional change: Alterations in classroom practices used by individual STEM faculty.

Change agent: An individual or group attempting to create instructional change among one or more instructors. Although the ultimate goal of change agents is typically to

change the instructional practices of large numbers of individual faculty members, a change agent can approach this goal from a variety of levels (e.g., from individual faculty practices to national policy).

Change practices: The specific tactics that change agents use to promote instructional change (e.g., dissemination of textbooks or other materials, or the specific nature, duration, and content of workshops, etc.).

Change strategy: Coherent plan of action that guides a change agent's choice of practices (e.g., provide tested, ready-to-use materials for teaching a specific course).

Change theory: An analytic construction that describes a generalizable set of observations in a descriptive and likely predictive sense (e.g., Rogers' (2003) theory of diffusion of innovations).

Although the emphasis of this review is on change strategies relevant for changing instructional practices used in undergraduate STEM courses, we do not think that change strategies and related issues discussed here are unique to the STEM disciplines. Rather, we suspect that the results can be easily generalized to undergraduate instruction in other fields and can also help to inform thinking in other levels of the educational system.

Studying Change Strategies

The goal of this article it to provide a review of the conceptual and empirical literature related to creating instructional change in undergraduate STEM courses. The research questions that guided this review are:

(1) What strategies are commonly used by change agents to promote instructional change in undergraduate STEM courses?
(2) How do change strategies described by articles relate to the disciplinary background of the authors (i.e., SER, FDR, HER)?
(3) To what extent do articles build on or make connections to other work on change?
(4) How well does the evidence presented in an article support the claimed effectiveness of the change strategy studied?
(5) What research-based assertions can be made about best practices for creating change?

Collectively, these research questions seek to elucidate the landscape of educational change in undergraduate STEM courses. The first two of these questions help us identify categories of change and communities of change. The second two seek to identify the degree to which these efforts follow standard scientific research practices. The final question identifies the implications that can currently be drawn from this body of research.

Previous Reviews

Understanding how to promote change in higher education has been an ongoing topic of interest to each of the three research communities described above. Each community, to a greater or lesser extent, has also made some efforts to summarize the knowledge about how and why change occurs, and strategies to promote change. In this section, we identify relevant prior work and discuss how it differs from the literature review presented in this article. Later, in the results section, we will highlight how the findings from these previous works compare with our findings.

The FDR community has the strongest history of conducting systematic reviews of the literature on creating changes in faculty teaching practices (Amundsen et al., 2005; Emerson

& Mosteller, 2000; Levinson-Rose & Menges, 1981; Weimer & Lenze, 1991). The goals of change strategies within the FDR community are most commonly focused on helping faculty improve their teaching practices via self-reflection or on helping faculty learn how to use a new piece of teaching technology that is applicable in a wide variety of disciplines. Each of the four reviews builds on the previous review. Emerson and Mosteller (2000), in particular, do a nice job of summarizing changes in thinking about change strategies within the FDR community. As an example of the types of change strategies studied by the FDR community, Emerson and Mosteller (2000) group the reviewed literature into four categories: (1) interventions by professional consultants and facilitators, (2) workshops, seminars, and courses, (3) mentoring programs, and (4) action research (p. 20). They conclude that effective change strategies should be collegial, extend at least one full semester, and be focused and concrete (Emerson & Mosteller, 2000, p. 20).

The HER community also has a strong history of studying change and has produced at least two significant reviews of the literature (Dill & Friedman, 1979; Kezar, 2001). However, unlike the FDR community, the HER community is much more interested in developing perspectives and methodologies for studying how change happens rather than proposing strategies for making change happen. Thus, many of the frameworks for understanding change from the HER perspective provide little or no guidance for the change agent. For example, Kezar (2001) describes six categories of change theories: evolutionary, teleogical, life cycle, dialectical, social cognition, and cultural. Only the teleogical and social cognition theories provide guidance for the change agent. The types of change that interest the HER community are quite broad, and are related to all aspects of higher education institutions, not just teaching. Thus, the types of changes envisioned by the HER community are much more widely conceived than the current review.

The SER community has a much weaker history of studying change in instructional practices used in undergraduate instruction. This community does have a very strong history of studying effective instructional practices and has produced a number of works identifying ''best practices'' for undergraduate instruction (e.g., AAAS, 2004; Handelsman et al., 2004; Labov, Singer, George, Schweingruber, & Hilton, 2009; National Research Council, 1999; Redish, 2003; Wood, 2009). However, although there have been conferences (e.g., Spalter-Roth, Fortenberry, & Lovitts, 2007) and journal volumes (e.g., Larson, 2006) devoted to this issue, we are not aware of any systematic attempts to synthesize the literature related to strategies for spreading these best practices to other faculty. There have been attempts to provide overviews of change strategies from the SER perspective (Millar, 1999; Seymour, 2001; Smith, Linse, Turns, & Atman, 2004). For example, Seymour (2001) undertook an ethnographic analysis of work presented at a 1998 National Institute of Science Education conference and synthesized several theories of change that appear to be prevalent in efforts to reform undergraduate STEM instruction. Although such overviews and syntheses are valuable, they are not systematic reviews of the literature.

In addition to literature reviews within the three target disciplines, there has been significant literature review and syntheses of change strategies focused on K-12 education (Ellsworth, 2000; Fullan, 2001; Hutchinson & Huberman, 1994; Loucks-Horsley, Hewson, Love, & Stiles, 1998). These works tend to take a dissemination perspective in which new instructional strategies or materials are pushed from the developers to the teachers, who are expected to use them with some degree of fidelity. Although there are many aspects of this literature that are relevant to higher education, it is important to be cautions when applying results since the higher education environment has many contextual features that are quite different from K-12 education. These features, such as disciplinary affiliation (Biglan, 1973;

Braxton, 1995; Braxton, Eimers, & Bayer, 1996), loose coupling (e.g., Birnbaum, 1991), and reward structures (e.g., Fairweather, 1996) are thought to strongly impact the effectiveness of change strategies used.

There are, of course, other influential reviews of change strategies. Rogers' (2003) diffusion of innovations framework was built from the literature in a wide variety of fields and has significantly influenced thinking in FDR and SER fields. Van de Ven and Poole's (1995) framework of four core change processes was built from an interdisciplinary literature review and has influenced thinking in the HER field (e.g., Kezar, 2001). Other literature syntheses, such as Fixsen, Naoom, Blase, Friedman, and Wallace (2005) synthesis of literature on implementation of new activities or programs, appears highly relevant to higher education, but does not (yet) appear to have had much influence on change efforts in higher education.

Methods

Our methods for this literature synthesis followed careful sequential steps to ensure validity of the results. We describe them in detail here to help the reader understand the measures we took to achieve the results we present.

Identifying Articles

The literature search identified journal articles that describe efforts by change agents to promote instructional change in undergraduate STEM education. We selected journal articles as the focus of this review for several reasons. First, journal articles are often the sources of ideas for future studies and thus may greatly influence researchers and stakeholders in STEM education. Second, we wanted to look at the highest quality studies available and saw the peer-review process employed by most journals as the best mechanism to identify these studies. Finally, on a practical note, journal articles are the type of literature most easily identified through systematic searching procedures.

Articles were selected through a succession of inclusion and exclusion criteria. First, the most productive indexes were identified for initial searching: Web of Science, Psych Info, and ERIC. We set the date range of inclusion from 1995 to 2008 (our article database was largely finalized in early 2009). Multiple scans of these indexes were undertaken with different combinations of the terms ''change,'' ''development,'' ''teaching,'' ''instruction,'' ''instructional,'' ''improvement,'' ''higher education,'' ''undergraduate,'' ''college,'' and ''university.'' These combinations of search terms yielded hundreds of articles each. Abstracts were used to exclude articles that clearly did not meet the content criteria. For example, articles were eliminated if they were clearly situated in the K-12 environment or the focus of the article was on measuring student learning in response to a new pedagogical approach. Because the goal of the search was to identify all relevant articles, when there was a question about appropriateness, the default was to keep the article in the search list and have the research team review it further. Additionally, we specifically examined the journals that appeared to produce the greatest number of articles that fit our inclusion criteria to ensure that we had gleaned all the relevant literature they offered. We also examined the reference lists of selected articles for relevant articles not already in our database.

The criteria for selecting articles was that they describe or advocate for one or more change strategies in the context of higher education. Articles could (i) explicitly focus on impacting undergraduate STEM instruction, (ii) focus on improving undergraduate instruction in more than one discipline (of which STEM is or might be included), or (iii) focus on improving some aspect of a college or university (of which undergraduate STEM instruction might be included). Although these last two categories may seem odd to the reader given that

our goal was to identify strategies for changing undergraduate STEM instruction. We did not want to miss change strategies that were aimed, for example, at all undergraduate subjects (type ii) or change strategies that were targeted more broadly, for example new management practices that were aimed at academic as well as nonacademic areas of the university (type iii). We did exclude from the review articles related to descriptions of new teaching ideas developed by instructors with no emphasis on the dissemination of these ideas. There has been much work published in this area and descriptions of "best practices" are widely available in print (e.g., AAAS, 2004; National Research Council, 1999) and online (e.g., www.NSDL.org). We did not wish to catalog new instructional ideas, but rather to understand how these ideas can be sustained and spread. We admit that there is a fine line between articles describing teaching innovations and articles describing change strategies. Thus, we included articles that discussed change strategies that encourage curriculum development.

The initial database contained 295 articles that were used for the preliminary analysis (Phases I and II, as described below). During the preliminary analysis, some articles were removed as not meeting the inclusion criteria. Other articles were added based on references in the database articles, additional targeted searching, and recommendations from readers of the preliminary analyses. This led to a revised database of 300 articles that were used in the final analysis. This final set of articles, therefore, represents exhaustive searches in multiple databases, reviews of reference lists of chosen articles, and feedback from expert readers outside the research team. Even given the thorough nature of the search for relevant articles, we are aware that we have missed some articles that would appropriately be included in the data set. The difficulty in identifying relevant articles (due to a wide variety of key words used in relevant articles and a wide variety of journals in which these articles are published) emphasizes the need for work to synthesize this literature.

Analysis Procedures

The analysis of articles proceeded in three phases. The goals of Phases I and II were to use a subset of the articles to develop a categorization scheme and coding criteria that would capture important variations in the articles. These efforts and preliminary results are reported in detail elsewhere (Henderson, Beach, & Finkelstein, in press; Henderson, Beach, Finkelstein, & Larson, 2008a, b; Henderson, Finkelstein, & Beach, 2010) and are briefly summarized here. The goal of Phase III was to apply the categorization scheme and coding criteria that emerged from the first two phases to the entire set of articles and draw conclusions from the results.

Phase I involved an examination of 130 randomly chosen journal articles from the set of 295 that were identified as relevant. The process began by each member of the research team (initially consisting of the authors and R. Sam Larson) reading a common subset of the 130 articles and proposing potentially useful coding criteria. Based on multiple iterations over several weeks, a final coding sheet was agreed upon. This coding sheet can be found elsewhere (Henderson et al., 2008b). The remaining articles from the set of 130 were coded by at least two randomly assigned members of the research team. Minor disagreements were resolved by the pair, and major disagreements were discussed with the entire research team. Once the initial set of 130 articles was coded, the research team then met several times in person and by phone conference to discuss what we had learned from our initial reading of the chosen articles, ways to define the patterns and themes we saw emerging from the articles, and how we saw those patterns coming together to form potentially useful categories. From these initial coding approaches emerged two guiding questions that, when combined, form four categories of change strategies. These questions and categories, now in somewhat more

refined form, are presented in the results section. We then re-reviewed the 130 articles and placed them within the four categories developed.

In Phase II of the analysis, we selected 10 articles from each of the four categories identified in Phase I for more detailed qualitative examination. In this analysis round, we looked for the core change idea(s) either explicitly or implicitly evident in the change process described in the articles, the degree to which the authors grounded their work in established change literature, the evidence authors presented to support their change strategies, barriers to change reported by authors, and recommendations made by authors. Data were collected in the form of quotes, bullet lists, and synopses and arranged into a narrative summary of each article that could be analyzed for patterns. One member of the research team completed the first analysis of the articles within a particular category. This work was then examined and modified by another member of the research team, often with significant discussion. As a result of this closer analysis, several articles were moved from their initial categories. A total of 43 articles were ultimately analyzed in this second round.

At the conclusion of Phase II, we had a set of what we called preliminary results. These consisted of the four categories of change strategies and findings related to the articles within and across categories. Several publications (Henderson et al., in press, 2008a,b, 2010) and many presentations and workshops related to these preliminary results helped to ensure the validity and utility of the four categories of change strategies, to identify and correct weaknesses in the preliminary coding criteria, and to develop a final coding scheme. We arrived at a set of coding categories: general citation information about the article, information about the disciplinary backgrounds of each of the authors (FDR, HER, SER, other) as determined from the author information provided in the article or a web search, the degree to which the article meets the selection criteria, a primary categorization of the article into one of the four categories, a secondary categorization (if applicable) of the article into one of the remaining three categories, a subcategory for each of the four main categories, a rating of the level of connection that the article makes with other change literature, a rating of whether the article claims success of the change strategy studied, and a rating of the degree of evidence presented to support the claims of success (or lack of success). In addition, several open-ended boxes allowed researchers to comment on many of the coding areas and provide an article summary.

The goal of Phase III was to use the complete set of 300 articles and our improved understanding gained from the preliminary analysis to answer the five research questions posed in this article. Coding progressed in several rounds. The main coding was conducted by authors Beach and Henderson with the assistance of a post doctoral researcher, Yuhfin Lin. The first round consisted of each of the three researchers coding the same 25 randomly selected articles from the database. Rounds 2 and 3 each involved an additional 15 randomly selected articles. In-between each round of coding, the three researchers compared codes, discussed differences in coding and, when necessary, made small changes to the coding criteria. The degree of agreement between pairs of researchers on each of the five main coding areas (described above) was calculated using Cohen's Kappa and Percentage Agreement. Results for each round of analysis are shown in Table S1. The level of agreement reached at the end of Round 3 was deemed acceptable and the remaining articles were coded by only one researcher. To help ensure the quality of the coding, however, author Finkelstein independently coded a random selection of approximately 10% of the coded articles. He discussed and resolved discrepancies with the main coder. Larger coding problems were discussed by the entire four-member research team. The result of Phase III was a database containing the 300 coded articles.

Of the 300 articles in the database, we determined that 109 (36.3%) did not meet the selection criteria and should be removed. For these articles, failure to meet the selection criteria was not clear from the abstract and title. Common reasons for removal included a focus on K-12 education or a focus on subjects that could not include undergraduate STEM (e.g., instruction in business or nursing, graduate instruction). A breakdown of which of the selection criteria were met by the remaining 191 (63.7%) articles is shown in Table 1. These articles are listed in Spreadsheet_S1.xls that accompanies this manuscript.

Results

This section first presents our categorization of change strategies and then uses this categorization and analysis of the final coding of articles to answer the five research questions.

Categorization of Change Strategies

The four categories of change strategies that we developed from this review of the literature are based on the answers to two fundamental questions that emerged from Phase I. For the first question, ''What is the primary aspect of the system that the change strategy seeks to directly impact?'' we identified two discreet answers—*individuals* and *environments and structures*. Table 2 presents the definitions we developed for each response, and the underlying assumption about change that we saw driving each definition. For Individuals, the change seeks to address such factors as the beliefs and behaviors of instructors, assuming that they act of their own volition. Moving beyond individuals to environments and structures, the change seeks to impact the environments that are assumed to influence the actions of individuals.

For the second fundamental guiding question, ''To what extent is the intended outcome for the individual or environment known in advance?'' we identified two responses—*prescribed* and *emergent* (see Table 3). For prescribed outcomes, the change agent knows, upon initiating a change process, what kind of behavior or mental states in individuals or groups are expected and sought, driven by the assumption that the change agent(s) have the key knowledge needed to define the outcomes. For emergent outcomes, the end state, in terms of behaviors or beliefs, is determined as part of the change process, with the assumption that those involved in the change have important information needed to define the outcomes. Table 3 presents the definitions we developed for each response, and the underlying assumptions about change that we saw driving each definition.

Based on the possible combinations of responses to these two guiding questions, we developed a four-square typology of change strategies. Each category carries an underlying

Table 1
Breakdown of articles by selection criteria

Selection Criteria: Article Focuses ...	Number of Articles	Percentage of Articles
(i) Explicitly on impacting undergraduate STEM instruction	71	37.2
(ii) On improving undergraduate instruction in more than one discipline (of which STEM is or might be included)	103	53.9
(iii) On improving some aspect of a college or university (of which undergraduate STEM instruction might be included)	17	8.9
Total	191	100

Table 2
Overview of individuals versus environments and structures categorization criteria

What Is the Primary Aspect of the System That the Change Approach Seeks to Directly Impact?	Definition	Implicit Assumption
Individuals	The change intends to directly impact personal characteristics of single individuals, such as beliefs, knowledge, behaviors, etc.	Individuals' actions are primarily influenced by their own volition
Environments and structures	The change intends to directly impact characteristics of the system that are external to single individuals, such as rules, physical characteristics of the environment (e.g., room layout, technology), norms, etc.	Individuals' actions are primarily influenced by external environments

definition and assumption. The categories are explained in the text below and presented in Figure 1.

I. Individual/Prescribed—Focus on Disseminating Curriculum and Pedagogy. Within this category, defined by its intersection of individual focus and prescribed outcome, we identify the change strategy as disseminating curriculum and/or pedagogy. The emphasis of this type of intervention is on communicating the change agent's vision of good teaching to individual instructors. Change agents have a particular instructional strategy or conception about teaching and learning that they hope individual instructors will adopt. The role of the change agent is to use specialized knowledge to teach or tell others without that specialized knowledge about new ways to organize (curriculum) or teach (pedagogy) a subject as well as, perhaps, new conceptions that support the new curriculum or pedagogy. Change agents typically inform instructors about the target curriculum/pedagogy/conception and motivate them to adopt

Table 3
Overview of prescribed versus emergent categorization criteria

To What Extent Is the Intended Outcome for the Individual or Environment Known in Advance?	Definition	Implicit Assumption
Prescribed	The desired final state for the individual or environment is known at the beginning of the change process	Important knowledge relevant to change outcome is known to a few people (e.g., experts). Therefore a small group should determine the intended outcome
Emergent	The desired final state for the individual or environment is developed as part of the change process	Important knowledge relevant to change outcome exists in individuals throughout the system. Therefore a variety of stakeholders should be involved in determining the intended outcome

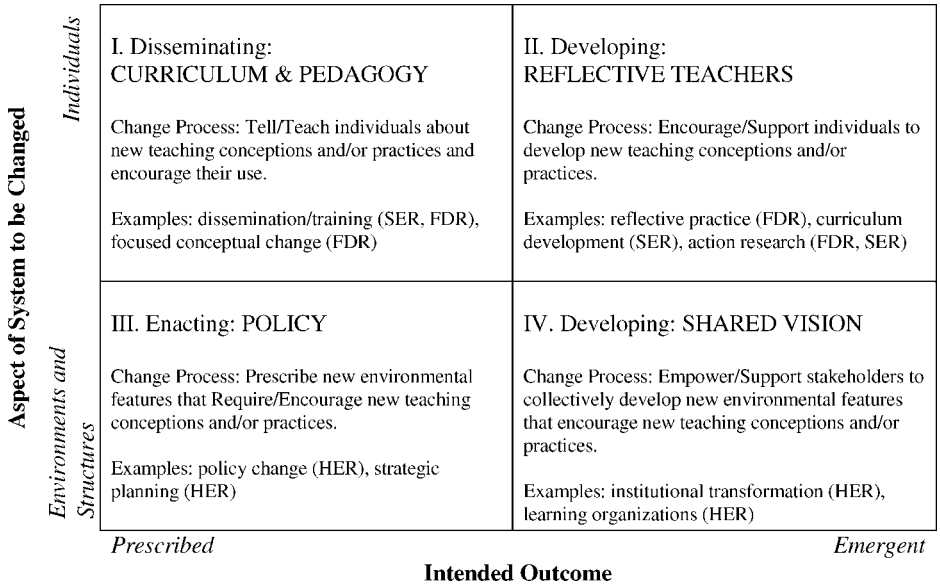

Figure 1. Overview and conceptual underpinnings of the four categories of change strategies.

it. Strategies in this category commonly employ workshops and talks as important change practices. Varying levels of support are offered to assist in adoption.

An example of an article that uses this change strategy is Silverthorn, Thorn, and Svinicki (2006). Through a grant from the National Science Foundation, the authors developed a set of ready-to-use activity modules focused on helping instructors actively engage students in introductory physiology topics. These modules were designed to be easily integrated into existing courses. After presentations and workshops at a summer professional conference, 56 instructors volunteered to field test the inquiry-oriented course modules. A subset of these ($N = 36$) was selected for this study and received instructor and student versions of the modules, assessment tools, suggestions for engaging students, and enrollment in a project listserv where they could receive support from the project leaders and other instructors. Almost half of them dropped out of the project before the Fall semester started, and none ended up using the modules. The research team turned their focus to study the reason behind the lack of use. They concluded that dissemination failed due to several types of obstacles and offered 8 lessons learned from this experience (e.g., Lesson 2: lack of instructor time was a formidable obstacle to translating interest to action).

II. Individual/Emergent—Focus on Developing Reflective Teachers. Within the category defined by its focus on individuals and emergent outcomes, we identify the change strategy as developing reflective teachers. The focus of this type of intervention is on encouraging teachers to use their own knowledge/experience/skill to improve their instructional practices. Information about various instructional strategies and materials may be provided, but the specific resources presented are not critical to the intervention. The role of the change agent is to encourage and support reflective practices by individual instructors that lead to instructor-identified and defined change outcomes. The change agent typically has a particular activity (e.g., action research, learning communities) in which instructors or groups of instructors

will engage to develop new (at least new to them) instructional strategies or conceptions. Varying levels of change agent support and control of the process are provided. The literature base that appeared most influential for articles in this category centers on reflective practice and action research and the primary role of the change agent is that of encouragement and facilitation. The articles in one of the subgroups (where change agents present a variety of options to faculty) is closest to the prescriptive focus of the curriculum/pedagogy category. The difference here is that faculty are seen as being in a strong position to choose appropriately.

An example of an article that uses this change strategy is Lynd-Balta, Erklenz-Watts, Freeman, and Westbay (2006). The article describes a faculty learning community (e.g., Cox, 2004), which was a group of 8 faculty (7 from STEM, 1 from education) who met for 3 hours one evening per month. The goal of the learning community was for faculty to explore incorporating more critical-thinking skills into their instruction. This was accomplished through learning about pedagogical theory, group discussions and self-reflection related to how the theory might be enacted, and action research in their classrooms (e.g., Kember, 2000). The group was co-facilitated by a biology professor and an education professor. Participants felt that the reflection and peer support provided by the learning community had helped them make improvements in their teaching.

III. Environments/Prescribed—Focus on Enacting Policy. Within the category defined by the intersection of focus on organizational structures or environments and prescribed outcomes, we identify the change strategy as enacting policy. The emphasis of this type of intervention is on developing appropriate environments (e.g., rules, reward systems, reporting requirements, investments in support structures) to facilitate instructors engaging in specific or desired activities. The change agent role is to use specialized knowledge to develop new environmental features that require or encourage new behaviors or attitudes that will lead to changes in instruction. A change agent has a particular vision toward which they wish instructors to work. Typically this means that an instructor adopts a prescribed activity, strategy, conception, or outcome. Whereas dissemination interventions that focus on individuals use internal motivation as the primary mechanism for change, here significant external incentives or requirements are used. The literature base most often invoked in these articles includes management theory, leadership writing, and some inclusion of organizational culture. The primary change agent role identified within these articles is directing, leading, or managing.

An example of an article that uses this change strategy is Patriquin et al. (2003). They studied the impact of post-tenure review through interviews with 36 faculty members at 2 public institutions that had adopted post-tenure review practices. Post-tenure review at these institutions focused on all aspects of faculty work, but one important goal was the improvement of undergraduate education. At both institutions, reviews of individual faculty were held every 5 years. If a faculty member was found deficient in some area a professional development plan would be developed to remedy the problem. The study authors conclude that the post-tenure review process has not been effective in changing faculty practices. Faculty report the post tenure review process as ''busy work'' and complained that the administration did not take any action (either positive or negative) based on the results of these reviews. The authors suggest that it may be possible for post tenure review to be productive if it were made to be more like a professional (peer) review and less like a legal (bureaucratic) review.

IV. Environments/Emergent—Focus on Developing Shared Vision. Within the category defined by its focus on environments and emergent outcomes, we identify the change strategy as developing shared vision. The emphasis of this type of intervention is on developing a new

collective vision for the department, institutional unit, or institution (and, on occasion, even supra-institutional entities) that will support new modes of instruction. The change agent role is to catalyze or empower individuals to come together and work toward collectively envisioned change. The change agent uses instructors and other stakeholders to develop a shared vision and to design new environments that are consistent with this vision. This approach commonly draws on the organizational change literature and that of community building.

An example of an article that uses this change strategy is Kressel, Bailey, and Forman (1999). The article describes a change strategy that uses an outside consultant working with individual departments to identify needs and then to develop and implement strategies that address these needs. The study was conducted with four academic departments at a major state university. Staff from the campus teaching excellence center first conducted confidential interviews with each faculty member from a particular department. They then prepared a report to the department, summarizing areas of faculty satisfaction as well as concern. The department met as a group, facilitated by teaching center staff, to decide which issues to address and how to address them. Based on surveys with faculty in the target departments and university administrators, the authors conclude that the process was successful in three of the four departments. In many cases the authors noted that the consultation process had not only led to productive instructional changes within the department but had also reduced conflict within the department. They noted that the type of collaborative process they promoted is not typically a part of the culture of higher education, which places a premium on the autonomy of individual faculty.

Which Change Strategies Are Most Common? Table 4 shows the primary change category of the 191 articles that met the selection criteria. There was an approximately even split between the number that fit into the Curriculum & Pedagogy, Reflective Teachers, and Policy categories, with far fewer that fit into the Shared Vision category. Only 28 articles (14.7%) were judged to not fit cleanly into a single category. These were also coded with a secondary category and will be discussed below.

Validity of the Four Categories Construct

Denzin and Lincoln (2005) describe the concept of validity of qualitative research results as the answer to the question, ''are these findings sufficiently authentic that I many trust myself in acting on their implications?'' (p. 205). Thus, before discussing these categories further it is important to address their validity. Lincoln and Guba (1985) identify four aspects of validity (what they refer to as trustworthiness): credibility, transferability, dependability, and confirmability.

Credibility refers to the extent to which the findings are believable. In this study, there is both internal and external evidence for the credibility of the four categories. First, within each of the categories of change strategies, the processes and assumptions are similar. Most

Table 4
Breakdown of articles by primary change category

Category	Number of Articles	Percentage of Articles
Curriculum and pedagogy	58	30.4
Teachers	64	33.5
Policy	53	27.7
Shared vision	16	8.4
Total	191	100

articles could be categorized into a single category. This suggests that there is something that the strategies within each category have in common that is different from strategies in the other categories. In addition, one of the categorical dimensions (individual vs. extra-individual distinction) has previously been proposed (Bronfenbrenner, 1979; Cole, 1996; Finkelstein, 2005). We also found, as discussed below, that the category of change strategy addressed in an article is correlated with the research community of the author. One would expect different research communities with different perspectives to employ different change strategies. This finding of differences between research communities was confirmed by a separate Delphi study (Beach & Henderson, 2010). Finally, we have presented the preliminary categories at conferences and in manuscripts. Feedback from both expert and non-expert audience members was generally that the categories seemed valid and were useful.

Transferability refers to the extent to which the findings can apply to other situations. Because the literature was drawn from a large number of journals in many different disciplines it is likely that the results are relevant throughout the area of change strategies in undergraduate STEM. Further, we suspect that the four proposed categories may be useful in understanding other aspects of educational change, such as K-12 education, or non-STEM areas of higher education. As mentioned earlier, feedback from experts and non-experts who have been introduced to these categories suggest that the categories are useful in diverse contexts.

Dependability essentially refers to the rigor and thoroughness of the research process. In this study there are several aspects that support the dependability of the results. First is that the research team represented researchers from a variety of backgrounds. A diverse research team was particularly important for this project since we sought to review literatures from a diverse set of research communities. In addition, the research team did not rush to quick conclusions. We explored, critiqued, and refined many different possible categorization schemes over an extended period of time. The studies of inter-rater reliability also confirm, at least internally, the dependability of these results. Finally, as discussed earlier, we sought external feedback at several points in the research process and found agreement with our approach and findings.

Confirmability refers to the ability of others to check the results presented. The rigor and record keeping involved in this project speak to its confirmability. We have attempted to describe the research processes in detail in this article and have previously presented detailed descriptions of our preliminary analysis procedures. In addition, we have made available the full set of articles used in this analysis as well as their categorizations. We primarily see this as a resource for other researchers to build on, but it also serves as a way for interested readers to assess our results.

Correlation of Change Strategies With Research Community of Authors

As described above, each article was coded for the research communities represented by the authors. Up to four authors for each article were coded as either FDR, HER, SER, or OTHER. Authors categorized as OTHER include researchers from fields such as psychology, sociology, business, medicine, and history. In a small number of cases, we were not able to find any information about the author affiliations from the article or on the web. These authors were also coded as OTHER. If an article had more than four authors (18 articles, or 9% of the database did) then the four author codes were used to exemplify the range of different communities represented. In all cases, the affiliation listed for the first author in the database corresponded to the research community of the first author of the article. The 191 articles

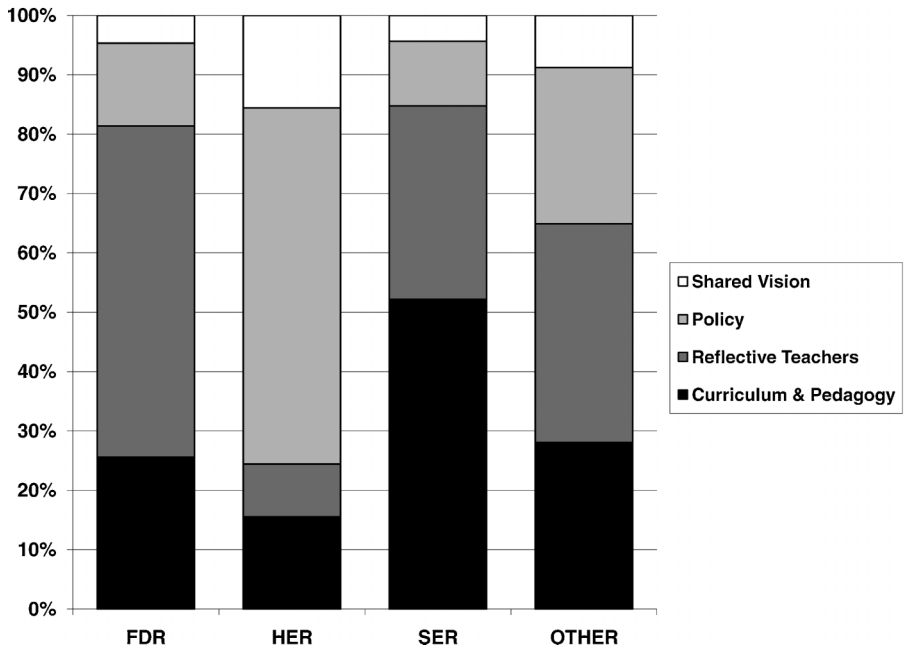

Figure 2. Primary change category of articles based on research community of first author.

were approximately evenly divided by the research community of the first author (FDR = 22.5%, HER = 23.5%, SER = 24.1%, OTHER = 29.8%).

Figure 2 shows a breakdown of the primary change category based on the research community of the first author. Each research community has a particular change strategy that they write about most frequently. FDR authors most frequently write about strategies in the Reflective Teachers category (24/43 = 56%). HER authors most frequently write about strategies in the Policy category (27/45 = 60%). SER authors most frequently write about strategies in the Curriculum and Pedagogy category (24/46 = 52%). Articles with authors categorized as OTHER are distributed similarly to the total set of articles.

Instead of looking at the first author, a similar analysis can be conducted using articles where all of the authors are from a particular disciplinary community. The results are quite similar for the HER and SER authors in that 72% and 85% respectively of the first authored articles have only authors from that discipline. Of the 43 articles with FDR first authors, only 25 (58%) have only FDR authors. Of these, 17 (68%) are in the developing reflective teachers category.

It is clear from the literature review that authors from different disciplinary communities write about different types of change strategies. It is not clear, however, to what extent the authors may have streamlined their descriptions of change strategies due to constraints of the publication process. For example, a HER author may pay great attention to shaping the beliefs of individual faculty during the implementation of a change strategy, but when reporting on this strategy in a HER journal emphasize the role of policy in shaping faculty behavior. To further examine the relationship between the research community of the authors and the change strategy used we conducted a Delphi study using expert informants within each community (Beach & Henderson, 2010). Preliminary results suggest that that the published literature does present a reasonable representation of the change strategies used by the three

disciplinary groups. The Delphi study began with three separate groups—one for each of the three disciplinary communities. In the final round, each group was presented with a summary of the responses from all three groups. FDR and SER experts most commonly described change strategies that focused on individual instructors. They did differ in that the FDR group named approaches that engaged instructors in reflection and exploration and SER experts named approaches that emphasized dissemination of information and demonstration of efficacy of approaches. The HER experts focused almost exclusively on organizational forces that could influence groups of instructors. When presented with these results, the experts saw them as a valid representation of their own experiences and logical extensions of the work each community does. As one HER expert noted, ''in my experience, these differences make sense. The HER group will take a broader perspective ... that includes environmental (institutional, unit-level, and program) factors that influence faculty work; they think about how organizational change happens at a more macro level. The FDR folks, by virtue of their professional positions, tend to focus on effective interventions with individuals in their own classrooms. The SER group—and this is a generalization—tends toward micro-level work on classrooms so the focus on individual and communities of folks working on improving teaching in their own courses (and perhaps their programs) also make sense to me.''

Connections to Other Literature

One key feature of scholarship within the scientific tradition is building on relevant theoretical or empirical ideas that have been developed previously (Burkhardt & Schoenfeld, 2003; Shavelson & Towne, 2002). As such we sought to code articles according to their connection to other change literature. By ''connections'' we mean that either the article bases the change strategy on other strategies or theories from the literature or that the results are connected to other relevant literature. We purposely left the definition of ''change literature'' rather broad, to include any literature-based connection with the change strategy studied or advocated. We identified four categories of connection. Level 1 meant that the article was strongly connected to the change literature. For example, Gibbs and Coffey (2004) based their study of the effectiveness of university-based faculty training programs on ideas from the professional development, reflective practice, and teacher conceptions literature. Level 2 meant that the article was weakly connected to the change literature. For example, Gaia, Corts, Tatum, and Allen (2003) begin their article by using literature to describe the training needs of graduate students and some general trends (e.g., the use of mentoring in other programs), but it was not clear how the details of their mentoring-based program built on ideas generated in these other programs or contributed to addressing issues that were previously unresolved in the literature. Level 3 meant that the article cited some change literature, but did make any connections between the literature cited and the change strategy studied or advocated. For example, Barlow and Antoniou (2007) begin their article on the experiences of new faculty (with a goal of identifying ways to better support new faculty) by citing other ''similar research.'' However, there is no indication that this literature (or any other literature) informed the study and the resulting recommendations (change strategies) that the authors believe will improve the situation. Finally, Level 4 articles do not cite any change literature. For example, Kerr and Runquist (2005) describe their experiences developing relationships with industry to work toward the improvement of undergraduate Chemistry in their department. Although they make several recommendations for change based on their experiences, they do not cite any other related literature.

In recognition of the rather subjective boundaries between the four levels, articles were assigned to the lowest reasonable code (we err on the side of over-emphasizing the

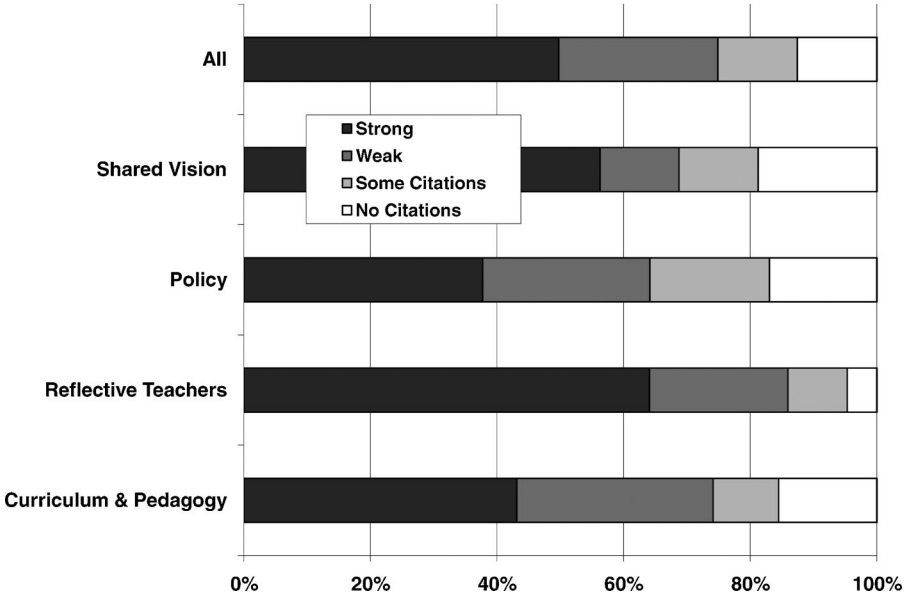

Figure 3. Connection to change literature based on primary categorization of articles.

connection of articles to prior literature). Nonetheless, the relatively low level of inter-rater reliability shown in Table S1 suggests that drawing distinctions between levels was quite difficult. On the other hand, it was highly unusual to have rater differences of more than one level. Thus, we believe that the literature connection results presented here do have meaning and, especially so, in aggregate.

Figure 3 shows that approximately half (49.7%) of the articles were categorized as strongly connected to change literature (Level 1), 25.1% were categorized as weakly connected to change literature (Level 2), 12.6% as making some connections (Level 3), and 12.6% as citing no change literature (Level 4). There did not seem to be any strong trends in terms of differences in degree of literature connections between articles from the four categories of change strategies. These numbers are low from our perspective given the low bar we set for being strongly connected. Ideally all scholarly articles would make strong connections to previously published work.

In addition, we undertook an analysis of the types of change literature used in the Level 1 articles. This involved one of the authors analyzing the arguments made in each of the Level 1 articles and then identifying what types of literature the arguments were connected to. These interpretations were then challenged by another member of our research team. By ''type of literature'' we are not necessarily referring to specific works, but rather to a set of works that have a common theme. This analysis was not conducted for the Developing Shared Vision category, due to the small number of articles it contained.

Of the articles in the Disseminating Curriculum and Pedagogy category 25 were found to have strong connections to existing change literature. Four types of literature were identified as common to at least 25% of the articles. The most common type of literature (used by 16 articles) was that related to teacher conceptions (e.g., Trigwell & Prosser, 1996b). In these articles, changes in teacher conceptions were seen as a necessary precursor to changes in their teaching practices. Thus, the articles sought to identify teacher conceptions and/or to modify

these conceptions to match a particular instructional style. For example, Dahlgren, Castensson, and Dahlgren (1998) found that teacher beliefs about the role of the instructor could hinder the implementation of Problem Based Learning. The next most common type of literature, used by 10 articles, applies general learning theories (e.g., Bransford, Brown, & Cocking, 1999) to the problem of helping faculty learn how to use new curriculum. For example, Hardré (2005) makes significant use of learning theory to advocate for an approach to providing professional development to graduate teaching assistants. Two additional types of literature were each used by 7 of the articles. One is the literature related to theories or models of change (e.g., Fullan, 2001; Rogers, 1995). For example, Pundak and Rozner (2008) use diffusion of innovations (Rogers, 1995) as an organizing perspective for promoting and studying change in faculty teaching practices. The final type of literature recognizes the importance of attending to institutional and structural barriers and rewards. For example, Skelton (2004) draws from literature on the low value of teaching in higher education (e.g., Hutchings & Shulman, 1999) in an evaluation of a particular national-level teaching award program designed to increase teaching innovation by recognizing a small number of good teachers and supporting them to disseminate their good practices.

Of the articles in the Developing Reflective Teachers category, 41 were identified as having a strong literature connection. Two types of literature were identified that were common to at least 40% of the articles. The most common type of literature, used by 20 articles, was that of reflective practice (e.g., Schön, 1983). In these articles, reflective practice by teachers was seen as an important component of instructional improvement. Many articles also identified specific forms of reflective practice, such as action research (Kember, 2000) or the scholarship of teaching and learning (SoTL) (e.g., Boyer, 1990). For example, Hubball, Collins, and Pratt (2005) describe a year-long voluntary Faculty Certificate Program at the University of British Columbia. The main goal of the program is to help the participants become reflective teachers. Eighteen different types of reflection activities were used in the program to scaffold faculty in meeting this goal. The second most common type of literature, used by 18 articles, was that of faculty beliefs or conceptions (e.g., Prosser & Trigwell, 1999; Samuelowicz & Bain, 1992). These articles all argued that an effective change strategy must help faculty change their conceptions from teacher-centered to student-centered. For example, Kember and Kwan (2000) used interviews with faculty to identify a strong correlation between conceptions of teaching and approaches to teaching. They conclude that changes in teaching require changes in conceptions of teaching and learning. In contrast to the articles in the Disseminating Curriculum and Pedagogy category, which also drew on the literature on teacher conceptions, these Developing Reflective Teachers articles did not measure conceptions against a particular instructional style. The third most common type of literature, used by only 7 articles, is that of disciplinary differences (e.g., Biglan, 1973). This literature was used to argue that different change strategies should be used for different disciplines. For example, Lueddeke (2003) used survey research to identify disciplinary differences relevant to the use of the scholarship of teaching.

Of the articles in the Enacting Policy category, 20 were identified as having a strong literature connection. Three types of literature were identified that were common to at least 40% of the articles. The most common type of literature, used by 10 articles, was that of faculty norms and academic culture. This type of literature suggests that there are a common set of beliefs, values, and behaviors that are shared by academics in general (e.g., Dill, 1982), academics at a particular type of institution (e.g., Berquist, 1992), or academics within a particular discipline (e.g., Becher, 1989). These articles typically see norms or culture as being important in improving undergraduate education since change strategies must be in

alignment with the norms or culture of an institution or department (e.g., Eimers, Braxton, & Bayer, 2001) or the change strategy must seek to change some aspects the existing norms or culture of an institution or department (e.g., Kezar & Eckel, 2002). Two additional types of literature were each used by 8 articles. One of these was literature related to the difficulty creating change using top-down approaches. This literature (e.g., Cuban, 1999; McLaughlin, 1991) was typically used to argue for the importance of developing more sophisticated and multi-faceted change strategies (e.g., Hearn, 1996). The second type was literature on SoTL (e.g., Boyer, 1990; Shulman, 1993). Articles using this literature often began with the assumption that the faculty use of SoTL would lead to improvements of undergraduate education. These articles would then identify various ways that SoTL could be incorporated, for example, through the use of teaching portfolios (Cox, 1995) or the appropriate use of post-tenure review (Patriquin et al., 2003).

This analysis of change literature shows that these three categories of change strategies use largely different types of change literature. There are, however, a few exceptions. The most significant overlap between categories is the idea that reflective practice or scholarship of teaching is an effective change strategy. This was found in 20/41 of the Reflective Teachers articles and 8/20 of the Policy articles. The Reflective Teachers articles were focused on teaching faculty how to be reflective and providing support for reflection (e.g., via consultations, faculty learning communities) while the Policy articles were focused on developing a system that would support or require faculty to be reflective. The other type of literature with a significant overlap is that of the importance of faculty conceptions and beliefs in the change process. This was found in 16/25 Curriculum and Pedagogy articles and 18/41 Reflective Teachers articles. It was used in more or less the same way in both cases—to support the importance of understanding and changing faculty beliefs or conceptions in order to change their teaching practices. The main differences were in the degree of prescription of the desired changes. Articles in the Curriculum and Pedagogy category had specific target instructional practices. These were supported by change strategies built on dissemination models or learning theories. Articles in the Reflective Teachers category did not have specific target instructional practices except for, in some cases, the general desirability of student-centered instruction. These were supported by change strategies built on reflective practice. Finally, there was some overlap in the use of literature on the importance of matching change strategies to the cultures and norms of each discipline. This literature was used by 10/20 of the Policy articles and 7/41 of the Reflective Teachers articles.

The finding of limited connection to prior literature is also supported by a separate citation analysis. The citation analysis found that there are few citations between articles in the database and also that there are few articles or books outside of the database that were cited by multiple articles within the database. For example, there are only 13 articles or books cited by more than 5 of the 191 articles in the database. Of these, the most frequently cited is Boyer (1990) which, as discussed above, is often referenced for its arguments in favor of the SoTL. The most frequently cited article within the database, Handelsman et al. (2004), was cited by only 6 of the other 190 articles in the database. This article was typically cited to justify the need to reform undergraduate STEM instruction. Thus, as suggested by the more qualitative content analysis described above, this citation analysis supports the finding that the articles in our database do not draw on any common set of literature.

Use of Evidence to Support Claims

Another cornerstone of scholarship, especially within modern calls of "scientific" educational research, is that of supporting claims with evidence (Burkhardt & Schoenfeld, 2003;

Shavelson & Towne, 2002). During coding, articles were separated into those that studied the implementation of a specific intervention versus those that did not study a specific intervention. Articles in the latter group ($N = 86$, 45.0%) included a wide variety of types, from opinion pieces to empirical studies with implications for change. For example, Trigwell and Prosser (1996b) conducted an empirical study to examine the relationship between faculty conceptions of teaching and their teaching strategies. They found that the faculty participants' teaching approaches were strongly affected by their stated intentions toward teaching. They conclude that professional development should focus on helping faculty use self-reflection to examine and change their conceptions of teaching. Such changes would lead to changes in practice. This approach is a change strategy that falls within the reflective teachers category. However, the article did not study the implementation of this change strategy.

The 105 articles (55.0%) that did describe the implementation of a change strategy were coded both in terms of the claims of success made about the intervention as well as the evidence presented to support these claims. Figure 4 shows that most of these articles (66.7%) claimed that the strategy studied was successful. Much smaller percentages of articles claimed mixed success (16.2%) or lack of success (12.4%). In a small percentage of articles (4.8%) we were unable to identify a claim of success or failure of the strategy studied even though such a claim would be expected.

In addition to making a clear claim about the success or failure of the intervention studied, we expected these 105 articles to present evidence from their study to support their claims of success or failure. In such empirical studies, the use of evidence to support claims is a (if not the) hallmark of this scientific and scholarly approach to research (Burkhardt &

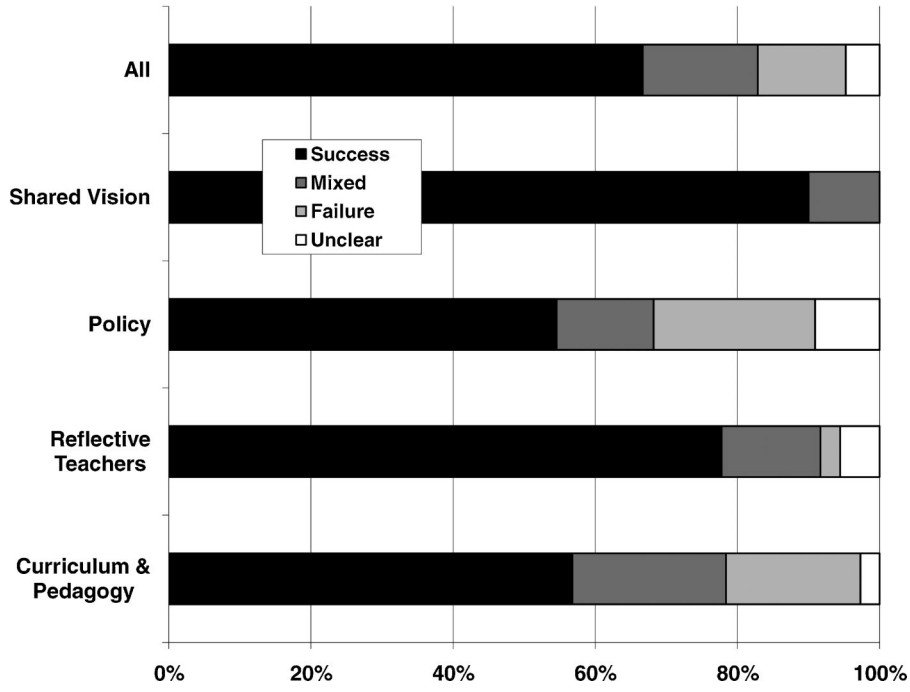

Figure 4. Claims of success of change strategy implemented based on primary categorization of articles.

Schoenfeld, 2003; Shavelson & Towne, 2002) and rooted in the history of science itself. We rated the level of evidence in one of four categories: Strong, Adequate, Poor, or None. A strong article had well-explained methods and made clear connections between the claims and supporting evidence. For example, Hubball et al. (2005) used the teaching perspectives inventory (TPI) to evaluate the effectiveness of an 8-month intervention called the Faculty Certificate Program. Participants ($N = 44$) engaged in weekly individual and group reflection on their teaching practices and beliefs. Pre/post-TPI testing showed significant positive changes in TPI scores. Quantitative results are augmented with qualitative interview data. An adequate article presented evidence to support claims of success, but the evidence or methods are not fully explained or fully convincing. For example, Schneider and Pickett (2006) studied a collaborative effort between an engineering instructor and a science education instructor to redesign an upper-level engineering course. The engineer felt there was substantial improvement while the educator felt the course did not achieve its intended student-centered goal. The article reports the collection of data during all of the interactions between the two participants and makes conclusions about the interaction between the participants, but details about the analysis procedures were minimal. A poor article presented anecdotal, vague, and/or undefined evidence to support claims made. For example, Angelique, Kyle, and Taylor (2002) describe how a group of faculty developed a peer mentoring network for new faculty. Anecdotes are presented to support claims of success. Finally, some articles presented no evidence to support their claims of success or failure. For example, Dirckinck-Holmfeld and Lorentsen (2003) describe a change strategy used to strengthen use of information and communication technology (ICT). Their sophisticated, multi-faceted, change strategy was based and justified very strongly in appropriate change literature. However, there was no description of the evidence used to support claims of success.

Figure 5 shows that only 21.0% of the articles that studied an intervention were categorized as presenting strong evidence to support claims of success or failure of the change strategy studied. Additionally, 27.6% were categorized as presenting adequate evidence, 39.0% as presenting poor evidence, and 12.4% as presenting no evidence. These numbers are low from our perspective, given that articles were already filtered to those that studied specific interventions and made claims about them. Ideally all scholarly articles that make claims about specific interventions would strongly ground their claims in evidence.

Best Practices for Creating Change

In this section we summarize the lessons learned from empirical studies within each of the categories. That is, we use the articles within each category to address, to the extent possible, the question of what works to promote change. We also address the question of what does not work. Because we wish to make these claims based on empirical evidence presented in the articles we will confine our analysis to articles that did describe the implementation of a change strategy and that we judged as presenting strong or adequate evidence to support their claims of the success or failure of the strategy. The 50 articles (26.2% of all articles) that met this criteria were divided between disseminating curriculum and pedagogy ($N = 22$), developing reflective teachers ($N = 16$), and enacting policy ($N = 12$). There were no articles in the developing shared vision category that met the criteria.

Disseminating Curriculum and Pedagogy. One thing that is clear from the articles within this category that presented at least adequate evidence to support their claims is that the common change strategy of developing and testing ''best practice'' curricular materials and then making these materials available to other faculty does not work (Clark, Froyd, Merton,

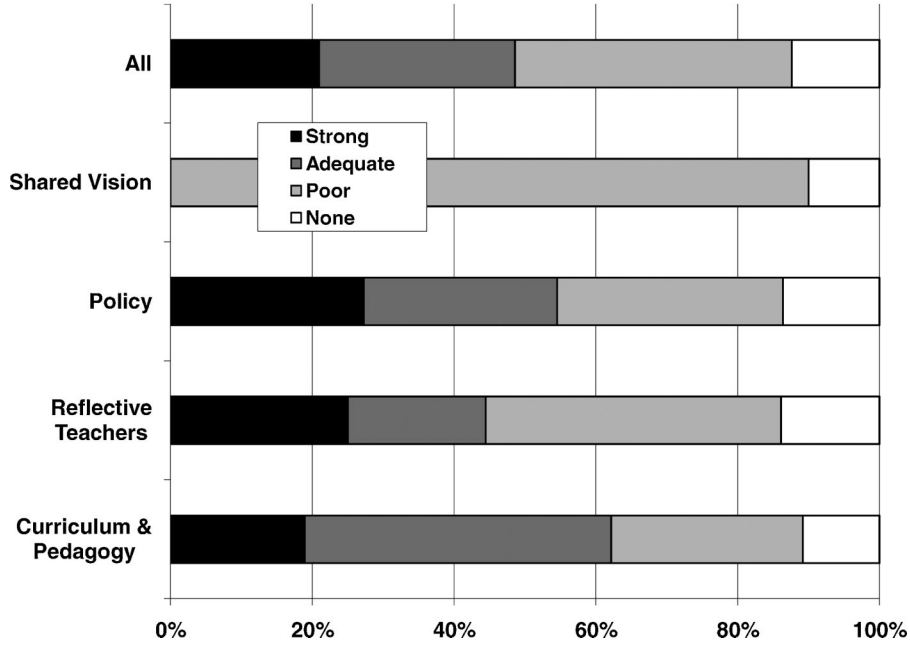

Figure 5. Evidence used to support claims of success or failure of change strategy implemented based on primary categorization of artifacts.

& Richardson, 2004; Penberthy & Millar, 2002; Sharp & McLaughlin, 1997; Silverthorn et al., 2006; Skelton, 2004). This is the strategy that Seymour (2001) describes as being based on the unproven theory that ''good ideas, supported by convincing evidence of efficacy, will spread 'naturally'—that, on learning about the success of particular initiatives, others will become convinced enough to try them'' (Seymour, 2001, p. 92). It is also a strategy that has been shown to be ineffective in change efforts within K-12 educational settings (Hutchinson & Huberman, 1994). Unfortunately, as Seymour (2001) points out, one reason that this change strategy may persist is that it is often reflected in proposal requirements of funding agencies.

Successful strategies focused on disseminating curriculum and pedagogy typically involve more than one of the following components: coordinated and focused efforts lasting over an extended period of time (Gallos, van den Berg, & Treagust, 2005; Gibbs & Coffey, 2004; Ho, Watkins, & Kelly, 2001; McShannon et al., 2006; Pundak & Rozner, 2008; Volkmann & Zgagacz, 2004), use of performance evaluation and feedback (Gallos et al., 2005; Hampton & Reiser, 2004; McShannon et al., 2006; Pundak & Rozner, 2008), and deliberate focus on changing faculty conceptions (Gibbs & Coffey, 2004; Ho et al., 2001). Each component will be briefly discussed.

Successful strategies include coordinated and focused efforts lasting over an extended period of time. The strategies that were shown to be successful lasted 4 weeks (Ho et al., 2001), one semester (McShannon et al., 2006; Volkmann & Zgagacz, 2004), or longer (Gallos et al., 2005; Gibbs & Coffey, 2004; Pundak & Rozner, 2008). The corollary of this is that one-time workshops (Davidovitch & Soen, 2006) or strategies that include a collection of

minimally related workshops (Walczyk & Ramsey, 2003) are not highly effective. This result is consistent with a previous review from the FDR community that suggests that successful faculty development programs benefit from interventions that last a full semester, year, or more (Emerson & Mosteller, 2000). The exception to this finding is change strategies that aim to bring about very localized changes in instruction, such as the incorporation of new technology. In such cases, it appears that one-time workshops can be successful (Campbell et al., 2007; Kahn & Pred, 2002).

Successful strategies provide performance evaluation and feedback. For example, Gallos, van den Berg, and Treagust (2005) describe the implementation of a revised first-year general chemistry course. The new course was pilot tested by two instructors in the department and revised. In the first stage of the change process, during the pilot testing, 14 additional instructors in the department were exposed to the new instructional approach through observations of the new course in action. These instructors also participated in weekly training sessions. In the second stage of the change process, these instructors implemented the new course in parallel sections. In small groups, the instructors met weekly with the course designer to discuss implementation problems. The entire group also met monthly to discuss the course. The course developer and two other colleagues regularly observed lessons and provided feedback and assistance to the instructors. Based on a variety of data collected, including classroom observations and videotapes, the authors conclude that 9 out of the 13 instructors were able to make significant changes in their instructional styles. The importance of consultation and feedback in dissemination efforts is a strong finding of a review of the implementation research (Fixsen et al., 2005). This review found that information dissemination and training by themselves, no matter how well done, are not highly effective, but that dissemination can be effective when it is more comprehensive, often involving such components as performance evaluation and coaching (Fixsen et al., 2005).

Successful strategies deliberately focus on changing faculty conceptions. Although changes in conceptions can occur through self-initiated reflection on the part of participants in a change initiative (Volkmann & Zgagacz, 2004), lack of such reflection and conceptual change can limit the success of change strategies (Weiss, Feldman, Pedevillano, & Copobianco, 2003). Strategies that deliberately focus on conceptual change appear to have high levels of success in creating meaningful conceptual change in faculty that result in changes in practice (Gibbs & Coffey, 2004; Ho et al., 2001). For example, Ho et al. (2001) describe an intervention that overlaps significantly with strategies designed to promote conceptual change in students (Posner, Strike, Hewson, & Gertzog, 1982). Their process involves several steps: self-reflection (to elicit current conceptions); exposure to alternative conceptions, reflection and analysis on current conceptions (to develop dissatisfaction with current conceptions), examining examples of teaching which demonstrate more elaborated conceptions (to show that new conceptions are fruitful); and putting new conceptions into practice (Ho et al., 2001, p. 147). Although lasting only 4 weeks, this intervention was found to lead to conceptual change and changes in teaching practice in two-thirds of the participants. The importance of conceptions is consistent with Fullan's (2001) conclusion, drawn from the K-12 literature, that meaningful educational change requires changes in beliefs. Fullan also cautions that changes in beliefs are difficult because new beliefs often challenge core values and many beliefs are held implicitly (Fullan, 2001, p. 44). The difficulty of changing beliefs was exemplified by one of the reviewed articles that described limited success of an ongoing change effort due to the lack of the change strategy to create dissatisfaction in the beliefs of the participant studied (Weiss et al., 2003).

Developing Reflective Teachers. All articles in this category focus on strategies designed to empower individual instructors to enact change. By focusing on the development of individuals, these approaches seek to allow those who are instructing to make choices for their classrooms—to enact new curricula or pedagogical approaches. From the studies of successful and unsuccessful change efforts in the developing reflective teachers category, we find common best practices. Successful strategies in this category tend to focus on individual or group consultation (Hativa, 1995; Henderson, 2008; Hubball et al., 2005; McShannon & Hynes, 2005; Romano, Hoesing, O'Donovan, & Weinsheimer, 2004) or collective communities that share approaches (Gess-Newsome, Southerland, Johnston, & Woodbury, 2003). Two key themes spanned the successful individual and collective change strategies: focus on feedback to instructors (McShannon & Hynes, 2005; Penny & Coe, 2004) and encouragement of reflection by the change agents (Hubball et al., 2005; Piccinin, Cristi, & McCoy, 1999). Meanwhile, strategic barriers, common across many change efforts, included both the challenges of institutional structures that were not aligned with the change effort (Luft, Kurdziel, Roehrig, & Turner, 2004; McShannon & Hynes, 2005), and the existing beliefs of faculty involved in the change effort (Gess-Newsome et al., 2003; Luft et al., 2004; Romano et al., 2004).

The majority of change efforts focused on developing reflective teachers employed strategies that utilize an external facilitator or support structure. These external agents varied from individual consultants to centers for teaching excellence or national workshops. A common strategy focuses on pairing an individual consultant with an individual educator or group to help shift classroom practices, faculty beliefs, and ultimately educational outcomes. For example, Hativa (1995) describes the success (and associated challenges) of bringing a consultant in to work with individual faculty in a physics department. As a result of class observations and student feedback, faculty began focusing on ''essential teaching abilities'' and student ratings of faculty improved. Other efforts focus on a group level—facilitated workshops and collective work. For example, Romano et al. (2004) found that cohorts of mid-career faculty at the University of Minnesota participating in the Center for Teaching and Learning Services program, report these externally mediated efforts improved their classroom practices, confidence, and approaches in classroom. Less common, but still prevalent was the creation of working groups, or collective efforts in which faculty were more self-directed. Gess-Newsome et al. (2003) reported a co-teaching effort in which three faculty collectively engaged in designing and teaching a new course, each bringing different resources to transform a science course. Differing levels of change were observed among the participant faculty, owing to variation in instructors' conceptions of teaching.

Whether externally facilitated or not, individual or collective, successful change efforts tended to focus on making feedback and reflection key components of the faculty development efforts. Providing feedback for teachers on their own practices is a key role of external facilitators, who can serve to reflect on practices and make suggestions about possible ways to address challenges in transforming educational environments. For example, in the GRASP program described by McShannon and Hynes (2005), observations of faculty practice focused on teaching strategies rather than content. This feedback from observers helped faculty adapt their strategies and implement new strategies that led to improved student performance. Another common form of feedback and reflection is an action research approach, whereby faculty participate in the study of their own classes and interventions. Kember and McKay (1996) advocate for action research as a key tool of faculty development. They apply this reflective approach to a case study of faculty change that not only successfully supported

faculty development and course transformation, but also to the spread of educational transformation across the department.

Efforts to support the development of reflective teachers for educational transformation are bound to find barriers. The two most common barriers cited are institutional structures and the existing beliefs of faculty participants. Romano et al.'s (2004) study of mid-career faculty capture the institutional challenges. ''The intrinsic motivation to change, while important, may not be sufficient to sustain a teaching enrichment program for mid-career faculty members. Especially in research universities, where there is an emphasis on scholarly research production'' (Romano et al., 2004, p. 44). Luft et al. (2004) continue, ''First, it is evident that rewards and incentives need to be instituted so that faculty and staff involved in undergraduate education are held accountable'' (Luft et al., 2004, p. 230). Of course, institutional support structures may be necessary but are not sufficient. Because individuals are the actors within this strategic framework, individual faculty must be willing to engage in development. Gess-Newsome et al. (2003) find that faculty who are most able to recognize a ''mismatch between personal beliefs and practices'' were more likely to find ''pedagogical dissatisfaction'' (p. 762). And, ''an individual's dissatisfaction can lead to the development of new knowledge and belief'' (Gess-Newsome et al., 2003, p. 762), which itself can influence classroom practices.

Enacting Policy. The articles within the category of Enacting Policy were, for the most part, grounded in organizational theory and emphasized the need to see change as a complex process requiring action that supports individualized solutions (rather than uniform solutions applied without regard to context). Successful strategies within this category promote a move away from strict top-down mandates and from assumptions that single policy solutions exist for multiple departments or disciplines within an institution or among institutions. In addition, policies that take into account and work within local (department, institution) culture are more likely to be successful.

One important conclusion from articles in this category is that power alone, in the form of ''top-down'' policy implementation meant to influence instructional practices, is insufficient to leverage change (Browne, 2005; Colbeck, 2002a,b; O'Meara, 2003; Patriquin et al., 2003; Ursin, Huusko, Aittola, Kiviniemi, & Muhonen, 2008). For example, Colbeck (2002b) investigated the impact of policies regarding teaching created at the state level on faculty within multiple disciplines in multiple institutions. She found that policy set from such an organizational distance from faculty does not have the impact intended by the policymakers. Departments and institutions often intervene to soften or re-interpret policy, or respond to it without involving faculty. These results were consistent with those found in Great Britain (Hannan, 2005), where national initiatives met with resistance at the institutional level.

Successful change initiatives must align with cultural and operational norms at lower levels of the system (Aune, 1995; Kezar & Eckel, 2002). Kezar and Eckel (2002) present three case studies involving institutional change to support their claim that change strategies must be aligned with institutional culture in order to be successful. They found that problems arose when desired changes violated institutional culture. Policies that seek uniformity and deter individualized solutions are not likely to be effective in promoting change (Colbeck, 2002a,b; Hannan, 2005; Major & Palmer, 2006; Ursin et al., 2008).

Summary. Several overlapping themes emerge from the three categories examined. First, individual faculty conceptions of teaching and learning are both supports and barriers to instructional change, whether prescribed or emergent. Individual instructors must understand

their own practice and the conceptions of teaching that influence it in order to fully embrace change. Within the Policy category, this self-examination is reflected in the need to align policymaking efforts with institutional culture. At all levels, the message is ''know thyself'' as a baseline for change. The challenge is to find effective ways to help faculty ''problematize'' their own thinking, and for institutional change leaders to ''get on the balcony'' (Heifetz, 1997 as quoted in Kezar & Eckel, 2002) and see their institutional cultures clearly. Second, support for change must be on-going. Studies in both the Curriculum and Pedagogy and Reflective Teachers categories emphasized the need for sustained engagement on the part of instructors and their facilitators. Within the Policy category, the articles carry an underlying assumption of sustained support over an extended time period. Third, the articles identified many barriers to change. The barriers to individual change noted by authors within the Curriculum and Pedagogy and Reflective Teachers categories include lack of recognition and rewards for improved instruction, lack of time, and lack of support. The Policy articles mention the same barriers, but generally in a more proactive way, describing them as levers that should be used to promoted change. Yet, none of the studies we found presented results of experiments with significantly new recognition and reward systems. Thus, although it appears to be possible to enact change within this set of barriers, there also appear to be opportunities for more widespread change through the development of strategies that can remove these barriers.

Articles Spanning Categories

Most of the articles describing both successful and unsuccessful strategies within each category describe the importance of thinking about educational environments as complex systems. Thus, we suggest that creating changes to these systems will likely require analysis and action at multiple points in the system. Since each of our categories of change strategies is designed to impact the system in a different way, we hypothesize that it may be possible to develop change strategies based on coordinated actions that span more than one category. This hypothesis is explored below using the small number of articles in our literature review that actually did span more than one category.

Of the 191 articles in the review, 28 (14.7%) could not be placed cleanly in a single category. In these cases, the articles were given a primary and a secondary categorization. Most of these articles ($N = 19$) involve a combination of disseminating curriculum and pedagogy and either developing reflective teachers ($N = 11$) or enacting policy ($N = 8$). Overall, the articles that spanned multiple categories were somewhat different from the single-category articles with respect to the data reported earlier. They were more likely to make strong connections to other literature than the single-category articles (57.1% vs. 49.7%), they were less likely to claim success of the change strategy studied (55.0% vs. 69.0%), and they were more likely to present strong evidence in support of their claims of success or failure (27.8% vs. 21.0%). Thus, on all three of these measures the two-category articles are in the direction that is representative of higher quality scientific research. Although Chi-square tests found that none of these individual differences between the ratings of two-category articles versus one-category articles was statistically significant, the fact that all three differences were in the direction of higher quality suggests that there may be some improvement in quality associated with two-category articles.

Articles that combine curriculum and pedagogy with developing reflective teachers often advocate for or study change strategies that combine dissemination to faculty along with significant faculty autonomy. In these approaches, the disseminated curricula or conceptions are typically thought of as suggestions and faculty are expected to use their expertise to

decide which suggestions to implement. Also, significant alteration or personalization of the suggested methods are expected and often encouraged. For example, Calkins and Light (2007), describe a faculty development program for untenured faculty at Northwestern University. The 8-month program has participants focus on an instructional project. The program also seeks to change instructor conceptions by introducing them to several specific active-learning based approaches. By examining participant written project descriptions, they report that most of the program participants changed their conceptions of teaching to be more student-centered and that these changes in conceptions were reflected in their teaching plans.

Articles that combine curriculum and pedagogy with developing policy often advocate for or study change strategies that combine dissemination along with removal of environmental barriers and/or addition of environmental supports. For example, Brent and Felder (2003) describe the SUCCEED faculty development program. This program sought to scale-up faculty use of a set of innovations in engineering courses that had been previously developed. The program had two main foci. One was the dissemination of these ideas through a series of workshops. The other was to make targeted changes in the institutional environment that would favor the use of these new ideas. Two important environmental changes were to have a faculty member within engineering designated to coordinate faculty development activities and to provide institutional incentives for teaching improvement. A web survey administered to program participants at eight participating institutions found that most faculty had participated in the faculty development program and that the program had led to increases in the use of targeted instructional activities as well as perceptions of improved student learning.

Based on the discussion above, there is reason to be optimistic about strategies that span more than one category. These strategies are theoretically stronger since they can focus on multiple aspects of the system at once. There are also examples of strategies of this type that have been very successful. However, significant additional exploration of these multi-category strategies is warranted before any strong conclusions can be reached.

Conclusions

Although there is some cause for optimism, based on this review of the literature, we conclude that the state of change strategies applied to undergraduate STEM instruction is not strong. On the positive side, we did find that the literature contains many studies related to change strategies in STEM education and that there are instances of significant scholarship in this area. However, we also found substantial weaknesses in this body of literature: most change agents belong to one of three isolated research communities, each using a subset of available strategies; new work often does not build on prior empirical or theoretical work; and most published results claim success of the change strategy studied, but the evidence presented is often not strong.

A likely reason that each group has its own preferred (or default) type of change strategy is that disciplinary groups have developed separately and tend to focus on areas within their interest and influence. For example, SER typically have strong STEM backgrounds and, thus, commonly focus on new ways to teach the specific STEM content, while FDR typically seek to create changes in disciplines where they may have only limited knowledge, and, thus, focus on helping faculty be more reflective about their teaching. Similarly, SER are typically active teachers within STEM departments and, thus, often develop curriculum materials for the classes they are teaching, while FDR typically find themselves charged with helping faculty outside of their discipline and, thus, focus on ways of supporting faculty in developing or adapting teaching strategies for their subject. It seems highly productive for each group to

learn about the strategies of the other groups and incorporate aspects of these strategies that may solve current problems.

As Shavelson and Towne (2002) discuss, two important features of scientific research in education are (i) the use of evidence to test theories, hypotheses, or conjectures and (ii) the explicit connection of research to relevant theory. Many of the 191 articles reviewed fell short on one or both of these criteria.

There are, however, some bright spots. Some authors do appear to be treating this problem scientifically. A good example is the ongoing research program of Michael Prosser, Keith Trigwell, and their coworkers. Their articles are the most interconnected in our database. Their work also show a clear progression of ideas over time, from initial work on categorizing faculty conceptions (Prosser, Trigwell, & Taylor, 1994), then relating conceptions to approaches to teaching (Trigwell & Prosser, 1996a,b), and finally to more recent work that explores change-related issues beyond individual instructors, such as departmental-level leadership (Martin, Trigwell, Prosser, & Ramsden, 2003; Ramsden, Prosser, Trigwell, & Martin, 2007).

Finally, the results of this literature review allow us to make some claims about what works and does not work for creating change. Although much of the literature is not methodologically strong enough to warrant such claims, 26.2% of the articles reviewed did present results of an empirical study that we judged to be at least adequately supported by the data presented. One thing that we can conclude from these is that two commonly used change strategies do not work: developing and testing ''best practice'' curricular materials and then making these materials available to other faculty (Curriculum and Pedagogy category) and ''top-down'' policy-making meant to influence instructional practices (Policy category). On the other hand, several claims can be made about what makes change strategies successful. First, effective change strategies must be aligned with or seek to change the beliefs of the individuals involved. Second, change strategies need to involve long-term interventions, lasting a semester, a year, and longer. Third, colleges and universities are complex systems. Developing a successful change strategy means first understanding the system and then designing a strategy that is compatible with this system.

Implications

Clearly, the improvement of change strategies in undergraduate STEM instruction is an important area that is in need of significant research attention. We hope that the analysis presented here can serve as a springboard for such attention. The four categories of change strategies and the categorization criteria can be used to promote discussion within and between research communities. In addition, the collection of 191 articles reviewed is available for download (see Supplementary Materials, Spreadsheet_S1.xls). It contains the article categorization, short summary, and bibliographic information. We hope that other researchers can make use of this set of articles in their work in this area.

Researchers on change and change agents alike can improve the quality of scientific research related to change strategies in higher education. In particular, researchers need to be careful to connect evidence to conclusions as well as to make connections between their study and prior empirical and theoretical work. It is also important for authors to identify how their studies can lead to improvements in theory and practice. Change strategies that are broader, and span traditional boundaries, appear to be fruitful avenues for future research efforts.

Although there is still much for researchers to learn, this review points to some consistent results within many of the categories of change strategies that are also consistent with literature outside of higher education. We consider these results robust enough to inform practice.

Change strategies that span multiple categories appear to be fruitful and change agents would be wise to learn about strategies outside of their typical practice and to work with other change agents across disciplinary boundaries. The collection of 191 articles reviewed makes an excellent starting point for change agents interested in exploring change strategies outside of their typical practice. Of course, there are many reasons for traditional boundaries and change agents seeking to span these boundaries must be aware of the costs of communication and coordination that arise with such efforts.

This material is based upon work supported, in part, by the National Science Foundation under Grants DRL-0723699 and SES-0623009. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors wish to recognize the important work of the following researchers who contributed to aspects of this project: Brian Cole, R. Sam Larson, Yuhfin Lin, and Jinhai Zhang.

References

AAAS. (2004). Invention and impact: Building excellence in undergraduate science, technology, engineering and mathematics (STEM) education. Washington, DC: American Association for the Advancement of Science.

Amundsen, C., Abrami, P., McAlpine, L., Weston, C., Krbavac, M., Mundy, A., & Wilson, M. (2005). The what and why of faculty development in higher education: A synthesis of the literature. Paper presented at the American Educational Research Association 2005 (April) Annual Meeting, Faculty Teaching, Development and Evaluation SIG.

Angelique, H., Kyle, K., & Taylor, E. (2002). Mentors and muses: New strategies for academic success. Innovative Higher Education, 26(3), 195–209.

Aune, B. P. (1995). The human dimension of organizational change. Review of Higher Education, 18(2), 149–173.

Barlow, J., & Antoniou, M. (2007). Room for improvement: The experiences of new lecturers in higher education. [Article]. Innovations in Education and Teaching International, 44(1), 67–77.

Barr, R. B., & Tagg, J. (1995). From teaching to learning—A new paradigm for undergraduate education. Change, 27(6), 13–25.

Beach, A., & Henderson, C. (2010). Practical knowledge versus literature-based knowledge of STEM Education Change Agents: Results of a National Delphi Process. Paper presented the American Educational Research Association Annual Conference.

Becher, T. (1989). Academic tribes and territories. Buckingham: Open University Press.

Berquist, W. H. (1992). The four cultures of the academy: Insights and strategies for improving leadership in collegiate organizations. San Francisco: Jossey-Bass.

Biglan, A. (1973). Relationships between subject matter characteristics and the structure and output of university departments. Journal of Applied Psychology, 57(3), 204–213.

Birnbaum, R. (1991). How colleges work. San Francisco, CA: Jossey-Bass.

Bok, D. (2006). Our underachieving colleges: A candid look at how much students learn and why they should be learning more. Princeton, NJ: Princeton University Press.

Boyer Commission on Undergraduates in the Research Universities. (1998). Reinventing undergraduate education: A blueprint for America's research universities.

Boyer, E. L. (1990). Scholarship reconsidered: Priorities of the professorate. San Francisco: Jossey-Bass.

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

Braxton, J. M. (1995). Disciplines with an affinity for the improvement of undergraduate education. In N. Hativa & M. Marincovich (Eds.), Disciplinary differences in teaching and learning: Implications for practice (pp. 59–64). San Francisco: Jossey-Bass.

Braxton, J. M., Eimers, M. T., & Bayer, A. E. (1996). The implications for teaching norms for the improvement of undergraduate education. Journal of Higher Education, 67, 603–625.

Brent, R., & Felder, R. M. (2003). A model for engineering faculty development. International Journal of Engineering Education, 19(2), 234–240.

Bronfenbrenner, U. (1979). The ecology of human development. Cambridge, MA: Harvard University Press.

Browne, E. (2005). Structural and pedagogic change in further and higher education: A case study approach. Journal of Further and Higher Education, 29(1), 11.

Burkhardt, H., & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. Educational Researcher, 32(9), 3–14.

Calkins, S., & Light, G. (2007). Promoting learning focused teaching through a project based faculty development program. To Improve the Academy, 26, 217–229.

Campbell, A. M., Ledbetter, M. L. S., Hoopes, L. L. M., Eckdahl, T. T., Rosenwald, A., Tonidandel, S., Bucholtz, B., & Gottfried, G. (2007). Genome consortium for active teaching: Meeting the goals of BIO2010. Life Science Education, 6, 109–118.

Center for Science Mathematics and Engineering Education. (1999). Transforming undergraduate education in science, mathematics, engineering, and technology. Washington, DC: National Academy Press.

Clark, C. M., Froyd, J., Merton, P., & Richardson, J. (2004). The evolution of curricular change models within the foundation coalition. Journal of Engineering Education, 93(1), 37–47.

Colbeck, C. L. (2002a). Assessing institutionalization of curricular and pedagogical reforms. Research in Higher Education, 43(4), 397–421.

Colbeck, C. L. (2002b). State policies to improve undergraduate teaching: Administrator and faculty responses. Journal of Higher Education. 73(1), 3–25.

Cole, M. (1996). Cultural psychology: A once and future discipline. Cambridge, MA: Harvard University Press.

Cox, M. D. (1995). A department-based approach to developing teaching portfolios: Perspectives for faculty and department chairs. Journal on Excellence in College Teaching, 6(1), 117–143.

Cox, M. D. (2004). Introduction to faculty learning communities. In M. D. Cox & L. Richlin (Eds.), Building faculty learning communities: New directions for teaching and learning, no. 97 (Vol. 2004, pp. 5–23). San Francisco: Jossey-Bass.

Cuban, L. (1999). How scholars trumped teachers: Change without reform in university curriculum, teaching, and research 1890–1990. New York: Teachers College Press.

Dahlgren, M. A., Castensson, R., & Dahlgren, L. O. (1998). PBL from the teachers' perspective. Higher Education, 36(4), 437–447.

Davidovitch, N., & Soen, D. (2006). Using students' assessments to improve instructors' quality of teaching. Journal of Further and Higher Education, 30(4), 351–376.

Denzin, N. K., & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), Qualitative research (3rd ed., pp. 191–215). Thousand Oaks, CA: Sage Publications.

Dill, D. D. (1982). The management of academic culture: Notes on the management of meaning and social integration. Higher Education, 11, 303–320.

Dill, D. D., & Friedman, C. P. (1979). An analysis of frameworks for research on innovation and change in higher education. Review of Educational Research, 49(3), 411–435.

Dirckinck-Holmfeld, L., & Lorentsen, A. (2003). Transforming University Practice through ICT–integrated perspectives on organizational, technological, and pedagogical change. Interactive Learning Environments, 11(2), 91–110.

Eimers, M. T., Braxton, J. M., & Bayer, A. E. (2001). Normative support for improving undergraduate education in teaching-oriented colleges. [Article]. Research in Higher Education, 42(5), 569–592.

Ellsworth, J. B. (2000). Surviving change: A survey of educational change models (report no. IR-109). Washington, DC: Office of Educational Research and Improvement.

Emerson, J. D., & Mosteller, F. (2000). Development programs for college faculty: Preparing for the twenty-first century. In R. M. Branch & M. A. Fitzgerald (Eds.), Educational media and technology yearbook 2000 (Vol. 25, pp. 26–42). Englewood, CO: Libraries Unlimited, Inc.

Fairweather, J. S. (1996). Faculty work and public trust: Restoring the value of teaching and public service in American academic life. Boston: Allyn and Bacon.

Finkelstein, N. (2005). Learning physics in context: A study of student learning about electricity and magnetism. International Journal of Science Education, 27(10), 1187–1209.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). Implementation research: A synthesis of the literature, Louis de la Parte Florida Mental Health Institute Publication #231. Tampa, Florida: University of South Florida, National Implementation Research Network.

Fullan, M. (2001). The new meaning of educational change (3rd ed.). New York: Teachers College Press.

Gaia, C. A., Corts, D. P., Tatum, H. E., & Allen, J. (2003). The GTA mentoring program: An interdisciplinary approach to developing future faculty as teacher-scholars. College Teaching, 51, 61–65.

Gallos, M. R., van den Berg, E., & Treagust, D. F. (2005). The effect of integrated course and faculty development: Experiences of a university chemistry department in the Philippines. [Article]. International Journal of Science Education, 27(8), 985–1006.

Gess-Newsome, J., Southerland, S. A., Johnston, A., & Woodbury, S. (2003). Educational reform, personal practical theories, and dissatisfaction: The anatomy of change in college science teaching. American Educational Research Journal, 40(3), 731–767.

Gibbs, G., & Coffey, M. (2004). The impact of training of university teachers on their teaching skills, their approach to teaching and the approach to learning of their students. Active Learning in Higher Education the Journal of the Institute for Learning and Teaching, 5(1), 87–100.

Hampton, S. E., & Reiser, R. A. (2004). Effects of a theory-based feedback and consultation process on instruction and learning in college classrooms. [Article]. Research in Higher Education, 45(5), 497–527.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R, Gentile, J., Lauffer, S., Stewart, J., Tilghman, S., & Wood, W. (2004). EDUCATION: Scientific teaching. Science, 304(5670), 521–522.

Hannan, A. (2005). Innovating in higher education: Contexts for change in learning technology. British Journal of Educational Technology, 36(6), 975–985.

Hardré, P. (2005). Instructional design as a professional development tool-of-choice for graduate teaching assistants. Innovative Higher Education, 30(3), 163–175.

Hativa, N. (1995). The department-wide approach to improving faculty instruction in higher education: A qualitative evaluation. [Article]. Research in Higher Education, 36(4), 377–413.

Hearn, J. C. (1996). Transforming U.S. higher education: An organizational perspective. Innovative Higher Education, 21(2), 141–154.

Henderson, C. (2008). Promoting instructional change in new faculty: An evaluation of the physics and astronomy new faculty workshop. American Journal of Physics, 76(2), 179–187.

Henderson, C., Beach, A., & Finkelstein, N. (in press). Four categories of change strategies for transforming undergraduate instruction. In P. Tynjälä, M. L. Stenström, & M. Saarnivaara (Eds.), Transitions, transformations and transgressions in learning and education. Springer.

Henderson, C., Beach, A., Finkelstein, N., & Larson, R. S. (2008a). Facilitating change in undergraduate STEM: Initial results from an interdisciplinary literature review. In C. Henderson, M. Sabella, & L. Hsu (Eds.), Proceedings (peer reviewed) of the 2008 AAPT Physics Education Research Conference (Vol. 1064, pp. 331–334). Melville, NY: American Institute of Physics.

Henderson, C., Beach, A., Finkelstein, N., & Larson, R. S. (2008b). Preliminary categorization of literature on promoting change in undergraduate STEM. Paper presented the Facilitating Change in Undergraduate STEM Symposium. Retrieved from http://www.wmich.edu/science/facilitating-change/PreliminaryCategorization.pdf.

Henderson, C., Finkelstein, N., & 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.

Ho, A., Watkins, D., & Kelly, M. (2001). The conceptual change approach to improving teaching and learning: An evaluation of a Hong Kong staff development programme. Higher Education, 42, 143–169.

Hubball, H., Collins, J., & Pratt, D. (2005). Enhancing reflective teaching practices: Implications for faculty development programs. Canadian Journal of Higher Education, 35(3), 57–81.

Hutchings, P., & Shulman, L. S. (1999). The scholarship of teaching: New elaborations, new developments. Change, 31(5), 10–15.

Hutchinson, J. R., & Huberman, M. (1994). Knowledge dissemination and use in science and mathematics education: A literature review. Journal of Science Education and Technology, 3(1), 27–47.

Kahn, J., & Pred, R. (2002). Evaluation of a faculty development model for technology use in higher education for late adopters. Computers in the Schools, 18(4), 127–150.

Kember, D. (2000). Active learning and action research: Improving the quality of teaching and learning. London: Kogan Page Limited.

Kember, D., & Kwan, K. P. (2000). Lecturers' approaches to teaching and their relationship to conceptions of good teaching. [Article]. Instructional Science, 28(5–6), 469–490.

Kember, D., & McKay, J. (1996). Action research into the quality of student learning—A paradigm for faculty development. [Article]. Journal of Higher Education, 67(5), 528–554.

Kerr, S., & Runquist, O. (2005). Are we serious about preparing chemists for the 21st century workplace or are we just teaching chemistry? Journal of Chemical Education, 82(2), 231.

Kezar, A. J. (2001). Understanding and facilitating organizational change in the 21st century: Recent research and conceptualizations. ASHE-ERIC Higher Education Report, 28(4), 1–162.

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

Kressel, K., Bailey, J. R., & Forman, S. G. (1999). Psychological consultation in higher education: Lessons from a university faculty development center. [Article]. Journal of Educational and Psychological Consultation, 10(1), 51–82.

Labov, J. B., Singer, S. R., George, M. D., Schweingruber, H. A., & Hilton, M. L. (2009). Effective practices in undergraduate STEM education. Part 1: Examining the evidence. CBE Life Science Education, 8(3), 157–161.

Larson, R. S. (2006). STEM Innovation s and dissemination: Improving teaching and learning in science, technology, engineering and mathematics (introduction to special issue devoted to STEM innovation and dissemination). Metropolitan Universities Journal, 17(4), 5–8.

Levinson-Rose, J., & Menges, R. J. (1981). Improving college teaching: A critical review of research. Review of Educational Research, 51(3), 403–434.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage.

Loucks-Horsley, S., Hewson, P., Love, N., & Stiles, K. (1998). Designing professional development for teachers of science and mathematics. Thousand Oaks, CA: Corwin Press.

Lueddeke, G. R. (2003). Professionalising teaching practice in higher education: A study of disciplinary variation and 'teaching-scholarship'. [Article]. Studies in Higher Education, 28(2), 213–228.

Luft, J. A., Kurdziel, J. P., Roehrig, G. H., & Turner, J. (2004). Growing a garden without water: Graduate teaching assistants in introductory science laboratories at a doctoral/research university. [Article]. Journal of Research in Science Teaching, 41(3), 211–233.

Lynd-Balta, E., Erklenz-Watts, M., Freeman, C., & Westbay, T. D. (2006). Professional development using an interdisciplinary learning circle: Linking pedagogical theory to practice. Journal of College Science Teaching, 35(4), 18–24.

Major, C. H., & Palmer, B. (2006). Reshaping teaching and learning: The transformation of faculty pedagogical content knowledge. [Article]. Higher Education, 51(4), 619–647.

Martin, E., Trigwell, K., Prosser, M., & Ramsden, P. (2003). Variation in the experience of leadership of teaching in higher education. [Article]. Studies in Higher Education, 28(3), 247–259.

McLaughlin, M. W. (1991). The RAND change agent study: Ten years later. In A. R. Odden (Ed.), Education policy implementation. Albany, NY: State University of New York Press.

McShannon, J., & Hynes, P. (2005). Student achievement and retention: Can professional development programs help faculty GRASP it? Journal of Faculty Development, 20(2), 87–94.

McShannon, J., Hynes, P., Nirmalakhandan, N., Venkataramana, G., Ricketts, C., Ulery, A, & Steiner, R. (2006). Gaining retention and achievement for students program: A faculty development program. [Article]. Journal of Professional Issues in Engineering Education and Practice, 132(3), 204–208.

Millar, S. B. (Ed.). (1999). Indicators of success in postsecondary SMET education: Shapes of the future. Synthesis and Proceedings of the Third Annual NISE Forum (1998). Madison, WI: Wisconsin Center for Education Research.

National Research Council. (1999). How people learn: Brain, mind, experience, and school. Washington, DC: National Academy Press.

O'Meara, K. (2003). Believing is seeing: The influence of beliefs and expectations on posttenure review in one state system. Review of Higher Education, 27(1), 17–43.

Patriquin, L., Bensimon, E. M., Polkinghorne, D. E., Bauman, G., Bleza, M. G., Oliverez, P. M., & Soto, M. (2003). Posttenure review: The disparity between intent and implementation. Review of Higher Education, 26(3), 275–297.

Penberthy, D. L., & Millar, S. B. (2002). The ''hand-off'' as a flawed approach to disseminating innovation: Lessons from chemistry. Innovative Higher Education, 26(4), 251–270.

Penny, A. R., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. [Article]. Review of Educational Research, 74(2), 215–253.

Piccinin, S., Cristi, C., & McCoy, M. (1999). The impact of individual consultation on student ratings of teaching. International Journal for Academic Development, 42(2), 75–88.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66(2), 211–227.

Prosser, M., & Trigwell, K. (1999). Understanding learning and teaching: The experience in higher education. Great Britain: St. Edmundsbury Press.

Prosser, M., Trigwell, K., & Taylor, P. (1994). A phenomenographic study of academics' conceptions of science learning and teaching. Learning and Instruction, 4, 217–231.

Pundak, D., & Rozner, S. (2008). Empowering engineering college staff to adopt active learning methods. Journal of Science Education and Technology, 17(2), 152–163.

Ramsden, P., Prosser, M., Trigwell, K., & Martin, E. (2007). University teachers' experiences of academic leadership and their approaches to teaching. [Article]. Learning and Instruction, 17(2), 140–155.

Redish, E. F. (2003). Teaching physics with the physics suite. Hoboken, NJ: John Wiley & Sons.

Rogers, E. M. (1995). Diffusion of innovations (4th ed.). New York: Free Press.

Rogers, E. M. (2003). Diffusion of innovations (5th ed.). New York: Free Press.

Romano, J. L., Hoesing, R., O'Donovan, K., & Weinsheimer, J. (2004). Faculty at mid-career: A program to enhance teaching and learning. Innovative Higher Education, 29(1), 21–48.

Samuelowicz, K., & Bain, J. D. (1992). Conceptions of teaching held by academic teachers. Higher Education, 24, 93–111.

Schneider, R., & Pickett, M. (2006). Bridging engineering and science teaching: A collaborative effort to design instruction for college students. School Science and Mathematics, 106(6), 259.

Schön, D. A. (1983). The reflective practitioner. New York: Basic Books.

Seymour, E. (2001). Tracking the process of change in US undergraduate education in science, mathematics, engineering, and technology. Science Education, 86, 79–105.

Sharp, S., & McLaughlin, P. (1997). Disseminating development initiatives in British higher education: A case study. [Article]. Higher Education, 33(3), 309–329.

Shavelson, R. J., & Towne, L. (Eds.). (2002). Scientific research in education. Washington, DC: National Academy Press.

Shulman, L. (1993). Teaching as community property: Putting an end to pedagogical solitude. Change, 25(6), 6–7.

Silverthorn, D. U., Thorn, P. M., & Svinicki, M. D. (2006). It's difficult to change the way we teach: Lessons from the Integrative Themes in Physiology curriculum module project. [Article]. Advances in Physiology Education, 30(4), 204–214.

Skelton, A. (2004). Understanding ''teaching excellence'' in higher education: A critical evaluation of the national teaching fellowships scheme. Studies in Higher Education, 29(4), 451–468.

Smith, K. A., Linse, A., Turns, J., & Atman, C. J. (2004). Engineering change. Proceedings of the 2004 American Society for Engineering Education Annual Conference & Exposition: American Society for Engineering Education.

Spalter-Roth, R. M., Fortenberry, N. L., & Lovitts, B. (2007). The acceptance and diffusion of innovation: A cross-curricular perspective on instructional and curricular change in engineering. Washington, DC: American Sociological Association.

The White House. (2009). Press release: President Obama launches 'educate to innovate' campaign for excellence in science, technology, engineering & mathematics (STEM) education. Retrieved May 28, 2010, from http://www.whitehouse.gov/the-press-office/president-obama-launches-educate-innovate-campaign-excellence-science-technology-en.

Trigwell, K., & Prosser, M. (1996a). Changing approaches to teaching: A relational perspective. [Article]. Studies in Higher Education, 21(3), 275–284.

Trigwell, K., & Prosser, M. (1996b). Congruence between intention and strategy in university science teachers' approaches to teaching. Higher Education, 32(1), 77–87.

Ursin, J., Huusko, M., Aittola, H., Kiviniemi, U., & Muhonen, R. (2008). Evaluation and quality assurance in Finnish and Italian universities in the Bologna process. Quality in Higher Education, 14(2), 109–120.

Van de Ven, A. H., & Poole, M. S. (1995). Explaining development and change in organizations. The Academy of Management Review, 20(3), 510–540.

Volkmann, M. J., & Zgagacz, M. (2004). Learning to teach physics through inquiry: The lived experience of a graduate teaching assistant. Journal of Research in Science Teaching, 41(6), 584–602.

Walczyk, J. J., & Ramsey, L. L. (2003). Use of learner-centered instruction in college science and mathematics classrooms. Journal of Research in Science Teaching, 40(6), 566–584.

Weimer, M., & Lenze, L. F. (1991). Instructional interventions: A review of the literature on efforts to improve instruction. In J. C. Smart (Ed.), Higher education: Handbook of theory and research (Vol. VII, pp. 294–333). New York: Agathon Press.

Weiss, T. H., Feldman, A., Pedevillano, D. E., & Copobianco, B. (2003). The implications of culture and identity: A professor's engagement with a reform collaborative. International Journal of Science and Mathematics Education, 1, 333–356.

Wood, W. B. (2009). Innovations in teaching undergraduate biology and why we need them. Annual Review of Cell and Developmental Biology, 25(1), 93–112.