

PROMISING PRACTICES IN  
UNDERGRADUATE SCIENCE,  
TECHNOLOGY, ENGINEERING,  
AND MATHEMATICS EDUCATION  
SUMMARY OF TWO WORKSHOPS

Natalie Nielsen, *Rapporteur*

Planning Committee on Evidence on Selected Innovations in  
Undergraduate STEM Education

Board on Science Education  
Division of Behavioral and Social Sciences and Education

NATIONAL RESEARCH COUNCIL  
*OF THE NATIONAL ACADEMIES*

THE NATIONAL ACADEMIES PRESS  
Washington, D.C.  
**[www.nap.edu](http://www.nap.edu)**

THE NATIONAL ACADEMIES PRESS 500 Fifth Street, N.W. Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by Grant No. DUE-0745112 between the National Academy of Sciences and the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number-13: 978-0-309-18723-7

International Standard Book Number-10: 0-309-18723-0

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, <http://www.nap.edu>.

Copyright 2011 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

Suggested citation: National Research Council. (2011). *Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics Education: Summary of Two Workshops*. Natalie Nielsen, Rapporteur. Planning Committee on Evidence on Selected Innovations in Undergraduate STEM Education. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

## THE NATIONAL ACADEMIES

### *Advisers to the Nation on Science, Engineering, and Medicine*

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

**[www.national-academies.org](http://www.national-academies.org)**



**PLANNING COMMITTEE ON EVIDENCE ON SELECTED  
INNOVATIONS IN UNDERGRADUATE STEM EDUCATION**

**SUSAN SINGER** (*Chair*), Department of Biology, Carleton College

**MELVIN GEORGE**, President Emeritus and Professor of Mathematics  
Emeritus, University of Missouri

**KENNETH HELLER**, School of Physics and Astronomy, University of  
Minnesota

**DAVID MOGK**, Department of Earth Sciences, Montana State University

**WILLIAM B. WOOD**, Department of Molecular, Cellular, and  
Developmental Biology, University of Colorado, Boulder

**HEIDI SCHWEINGRUBER**, *Study Director*

**JAY LABOV**, *Senior Advisor for Education and Communications*

**MARGARET HILTON**, *Senior Program Officer*

**NATALIE NIELSEN**, *Rapporteur*

**REBECCA KRONE**, *Program Associate*

## BOARD ON SCIENCE EDUCATION

**HELEN R. QUINN** (*Chair*), Stanford Linear Accelerator Center, Stanford University

**PHILIP BELL**, Learning Sciences, University of Washington, Seattle

**GEORGE BOGGS**, American Association of Community Colleges (retired), Washington, DC

**WILLIAM B. BONVILLIAN**, Washington, DC, Office, Massachusetts Institute of Technology

**JOSEPH S. FRANCISCO**, Department of Chemistry, Purdue University

**ADAM GAMORAN**, Center for Education Research, University of Wisconsin, Madison

**JERRY P. GOLLUB**, Natural Sciences and Physics Departments, Haverford College

**MARGARET A. HONEY**, New York Hall of Science, New York

**JANET HUSTLER**, Partnership for Student Success in Science (PS<sup>3</sup>), Synopsys, Inc., Mountain View, California

**SUSAN KIEFFER**, Department of Geology, University of Illinois, Urbana

**BRETT D. MOULDING**, Utah Partnership for Effective Science Teaching and Learning, Ogden

**CARLO PARRAVANO**, Merck Institute for Science Education, Rahway, New Jersey

**SUSAN R. SINGER**, Department of Biology, Carleton College

**WILLIAM B. WOOD**, Department of Cellular and Developmental Biology, University of Colorado, Boulder

**MARTIN STORKSDIECK**, *Director*

**HEIDI A. SCHWEINGRUBER**, *Deputy Director*

**MARGRET HILTON**, *Senior Program Officer*

**THOMAS E. KELLER**, *Senior Program Officer*

**NATALIE NIELSEN**, *Senior Program Officer*

**SHERRIE FORREST**, *Research Associate*

**REBECCA KRONE**, *Program Associate*

## Acknowledgments

This workshop summary is based on discussions at two workshops convened by the Board on Science Education of the National Research Council (NRC) on June 30 and October 13-14, 2008. We thank our colleagues who served on the planning committee, each of whom brought deep and varied expertise to the process of planning the workshop. The planning committee members identified presenters, organized the agenda, selected paper authors, and facilitated the discussion, although they did not participate in the writing of this report. This summary reflects their diligent efforts, the excellent papers and presentations by other experts at the workshop, and the insightful comments of the many workshop participants. The workshop would not have become a reality without the generous support of the National Science Foundation.

This summary has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the NRC's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published summary as sound as possible and to ensure that the summary meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the process. We wish to thank the following individuals for their review of this summary: George M. Bodner, Arthur Kelly Professor of Chemistry, Engineering and Education, Department of Chemistry, Purdue University; Paula Heron, Department of Physics, University of Washington; Julie Libarkin, Department of Geological Sciences, Division of Science and Mathemat-

ics Education, Cognitive Science Program, Center for Research on College Science Teaching and Learning, Michigan State University; and William B. Wood, Distinguished Professor, Emeritus, Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder.

Although the reviewers listed above provided many constructive comments and suggestions, they were not asked to endorse the content of the report, nor did they see the final draft of the report before its release. The review of this report was overseen by Kendall Starkweather, executive director of the International Technology Education Association. Appointed by the NRC, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the author and the institution.

We are grateful for the leadership and support of Robert Hauser, executive director of the NRC Division of Behavioral and Social Sciences and Education, and Martin Storksdieck, director of the Board on Science Education. We also thank Margaret Hilton, senior program officer, for her valuable contributions to the design and implementation of the workshop agenda, her collaborations with the commissioned paper authors, and her considerable work on this summary; Natalie Nielsen for serving as rapporteur for the two workshops; and Rebecca Krone for her flawless logistical support throughout the project.

Susan Singer, *Chair*  
Heidi Schweingruber, *Study Director*  
Planning Committee on Evidence on  
Selected Innovations in  
Undergraduate STEM Education



# Contents

1	INTRODUCTION	1
	Project Origin, 2	
	Report Overview, 3	
2	LINKING LEARNING GOALS AND EVIDENCE	5
	Examples from the Disciplines, 5	
	The State of Evidence in Discipline-Based Education Research, 15	
3	SURVEYING PROMISING PRACTICES	21
	Promising Practices for Faculty and Institutions and Predicting Success in College Science, 21	
	Small-Group Discussions and Final Thoughts, 24	
4	SCENARIO-, PROBLEM-, AND CASE-BASED TEACHING AND LEARNING	26
	Problem-Based Learning, 27	
	Case-Based Teaching, 29	
	Use of Complex Problems in Teaching Physics, 31	
	Discussion, 33	
5	ASSESSMENT TO GUIDE TEACHING AND LEARNING	35
	Concept Inventories in the Sciences: Examples from the Geosciences Concept Inventory, 35	
	Concept Inventories in Engineering, 37	

	Identifying and Addressing Student Difficulties in Physics, 38 Discussion, 40	
6	STRUCTURING THE LEARNING ENVIRONMENT Studio Courses, 42 Redesigning Large Classes for Learning, 44 Doing Science: Providing Research Experiences, 50 Small-Group Discussions, 52	42
7	FACULTY PROFESSIONAL DEVELOPMENT Professional Development of Future Faculty, 53 Workshops by a Professional Society for New Physics Faculty, 55 Changing Instruction, 57	53
8	SYSTEMIC CHANGE: BARRIERS AND OPPORTUNITIES Diffusion of Promising Practices, 60 Reflections on Linking Evidence and Promising Practices in STEM, 64 Future Directions and Next Steps, 66	60
	REFERENCES	69
	APPENDIXES	
A	June Workshop Agenda and Participants List	75
B	October Workshop Agenda and Participants List	80

## 1

## Introduction

Today, a quiet revolution is under way in the teaching of undergraduate science, mathematics, engineering, and technology. Courses that have resembled nothing so much as their 19th century precursors are beginning to change, as students and instructors realize that employment and citizenship in the 21st century will require radically different kinds of skills and knowledge. A new generation of faculty is questioning the contemporary constraints of academic life and looking at new ways to balance the teaching of students with other priorities. Departments and institutions are acknowledging that their responsibilities extend beyond producing the next generation of scientists, engineers, mathematicians, and technicians; they are recognizing that the challenge also is to equip students with the scientific and technical literacy and numeracy required to play meaningful roles in society. (National Research Council, 1996, p. 1)

In the mid-to-late 1990s, the National Research Council (NRC) and the National Science Foundation (NSF) wrote reports on the state of undergraduate education in science, technology, engineering, and mathematics—the disciplines collectively referred to as STEM (see National Research Council, 1996, 1999; National Science Foundation, 1996). As the quoted passage above suggests, these reports reflected past innovations and encouraged future innovations in STEM education at 2-year and 4-year postsecondary institutions. In the decade after their release, NSF, other government agencies, and several private foundations dedicated hundreds of millions of dollars to improve the quality of STEM undergraduate education.

Since then, numerous teaching, learning, assessment, and institutional innovations in undergraduate STEM education have emerged. Because virtually all of these innovations have been developed independently of one another, their goals and purposes vary widely. Some focus on making science accessible and meaningful to the vast majority of students who will not pursue STEM majors or careers; others aim to increase the diversity of students who enroll and succeed in STEM courses and programs; still other efforts focus on reforming the overall curriculum in specific disciplines. In addition to this variation in focus, these innovations have been implemented at scales that range from individual classrooms to entire departments or institutions.

### PROJECT ORIGIN

By 2008, partly because of this wide variability, it was apparent that little was known about the feasibility of replicating individual innovations or about their potential for broader impact beyond the specific contexts in which they were created. The research base on innovations in undergraduate STEM education was expanding rapidly, but the process of synthesizing that knowledge base had not yet begun. If future investments were to be informed by the past, then the field clearly needed a retrospective look at the ways in which earlier innovations had influenced undergraduate STEM education.

To address this need, NSF asked the NRC to convene an ad hoc steering committee to plan and implement a series of two public workshops focused on a thoughtful examination of the state of evidence of impact and effectiveness of selected STEM undergraduate education innovations. The steering committee was appointed and charged with identifying selection criteria and selecting STEM innovation “candidates” from reform efforts in teaching, curriculum, assessment, and faculty development. Of particular interest were STEM innovations in which the evidence of impact is strong and rich enough to analyze its effect on the “uptake” and sustainability of an innovation over time. The committee adopted the term “promising practices” to refer to innovations in STEM learning, teaching, and assessment.

The first workshop took place in June 2008 and focused on the challenge of aligning the learning goals of—and evidence of effectiveness for—promising practices within and across the science disciplines. In the second workshop, held in October 2008, participants delved more deeply into a select group of the promising practices in undergraduate STEM education that came to light at the June meeting. In planning both workshops, the committee focused in particular on innovations associated with the first two years of undergraduate STEM education. The innovations discussed in October represent a small proportion of the many promising practices

in undergraduate STEM education—time constraints during the workshop, the availability of promising practices with known evidence of effectiveness, and the availability of speakers influenced the innovations that were discussed at the October meeting.

In addition to planning a broad exploration of the evidence, the committee sought to connect education researchers from different disciplinary fields and to provide foundational information for a parallel NSF-funded initiative by the Wisconsin Center for Education Research. That initiative, *Engaging Critical Advisors to Formulate a New Framework for Change: Expansion of “Toward a National Endeavor to Marshal Postsecondary STEM Education Resources to Meet Global Challenges,”* focused on future directions for STEM and aimed to identify new strategies for organizing and implementing STEM undergraduate education practices. It underscored the need for the STEM community to take stock of what has been learned and to attend to the evidence base for drawing conclusions.

## REPORT OVERVIEW

This volume summarizes the two NRC workshops on promising practices in undergraduate STEM education. Chapters 2 and 3 summarize the first workshop: Chapter 2 focuses on the link between learning goals and evidence, and Chapter 3 presents a range of promising practices at the individual, faculty, and institutional levels. Subsequent chapters address the topics that were taken up in the second workshop, which involved deeper explorations of selected promising practices in STEM undergraduate education. Chapters 4-6 address a range of classroom-based promising practices: scenario-, problem-, and case-based teaching and learning (Chapter 4); assessments (Chapter 5), and improving student learning environments (Chapter 6). Chapter 7 focuses on professional development for future faculty, new faculty, and veteran faculty. The volume concludes with a broader examination of the barriers and opportunities associated with systemic change (Chapter 8).

It is important to be specific about the nature of this report, which documents the information presented in the workshop presentations and discussions. Its purpose is to lay out the key ideas that emerged from the two workshops and that should be viewed as an initial step in examining the research. The report is confined to the material presented by the workshop speakers and participants. Neither the workshop nor this summary is intended as a comprehensive review of what is known about the topic, although it is a general reflection of the field. The presentations and discussions were limited by the time available.

This report was prepared by a rapporteur and does not represent findings or recommendations that can be attributed to the steering committee.

Indeed, the report summarizes views expressed by workshop participants, and the committee is responsible only for its overall quality and accuracy as a record of what transpired at the workshops. Also, the workshops were not designed to generate consensus conclusions or recommendations but focused instead on the identification of ideas, themes, and considerations that contribute to understanding.

## 2

## Linking Learning Goals and Evidence

This chapter and the next summarize the June workshop, which focused on different learning goals for undergraduate students in science, technology, engineering, and mathematics (STEM) and different types of evidence related to those goals.

### EXAMPLES FROM THE DISCIPLINES

In the first session related to this topic, moderator Adam Gamoran (University of Wisconsin, Madison) introduced three panelists who used examples from chemistry, evolutionary ecology, and physics to address the following questions:

1. What are and what should be some of the most important learning goals for science students in lower division courses?
2. What types of evidence would be needed to conclude that a specific goal had been achieved?
3. Are there some types of evidence that carry more weight? If so, what makes that evidence particularly compelling?

### Chemistry

Cathy Middlecamp (University of Wisconsin, Madison) explained that the American Chemical Society sponsors *Chemistry in Context*, a long-term curriculum development project. The curriculum breaks the mold of traditional general chemistry courses by integrating key chemistry concepts

within a coherent framework focused on real-world issues. The placement of chemical principles and concepts is driven by what students need to know in order to understand the science related to each real-world issue (Middlecamp, 2008).

The curriculum targets two types of learning goals: (1) goals for student attitudes and motivation and (2) goals for student knowledge. The motivation goals are to give students a positive learning experience in chemistry and to motivate them to learn chemistry. The specific goals for student knowledge are to promote broader chemical literacy; to help students better meet the challenges of today's world; and to help students make choices, informed by their knowledge of chemistry, to use natural resources in wise and sustainable ways.

Middlecamp then turned to the evidence. She noted that there has been no formal evaluation of *Chemistry in Context*, and there is no ongoing assessment of student learning. In addition, no evidence has been collected on the number of faculty members using the curriculum or about why they select it. Most of the available evidence related to the motivation goals and student knowledge goals is gathered locally by instructors for the purpose of improving instruction and is not disseminated beyond the department or campus. Evidence of progress toward motivation goals includes student attitude surveys, evaluations of the instructor, and student behaviors after taking the course (such as taking further chemistry courses or participating in discussions of chemistry in informal settings). As an example, Middlecamp presented survey data from more than 2,000 students she taught using *Chemistry in Context*.

Evidence of student knowledge goals includes direct measures of student performance in class (tests, demonstrated skills), student surveys, and course-level data (e.g., class completion rate). To illustrate, Middlecamp presented a breakdown of responses from 1,172 students who had taken the course. When asked about the extent of their learning gains in “connecting chemistry to your life,” more than 450 students (38 percent) responded that they had gained “a lot” and another 400 (34 percent) reported “a great deal.” In response to the statement, “the lecturer makes the course interesting,” 74 percent strongly agreed, and 16 percent agreed. Reflecting on the quality of this evidence, Middlecamp noted that, while compelling to individual instructors, it is local, anecdotal, and nonsystematic.

Middlecamp argued that, despite the weakness of the evidence collected to date, *Chemistry in Context* is successful in terms of two larger goals of the project—to be adopted and adapted widely and to catalyze development of STEM curricula that take a similar approach. Success in achieving these goals is measured by different types of evidence, including the number of textbooks sold, the continued attendance at faculty workshops, and the translation of the book into other languages. For example, data indicat-



ing that sales have risen from about 6,000 for the first edition, published in 1994, to an estimated 23,000 for the sixth edition, published in 2008, show that adoption of the curriculum is growing. Translations into other languages and other regional and cultural contexts are evidence that the curriculum is adaptable.

Middlecamp suggested that two factors—the role of professional societies and the sustainability challenge—have helped advance the goals of wide adoption and catalyzing development of similar curricula. The American Chemical Society’s sponsorship of *Chemistry in Context*, including its active role in dissemination, played a role in the early success of the project, she said. In addition, an initiative on liberal education by the American Association of Colleges and Universities calls for undergraduates to develop science knowledge through engagement with “big questions, both contemporary and enduring” (American Association of Colleges and Universities, 2008). By recommending this learning outcome, the professional society supports the adoption of *Chemistry in Context* and also encourages development of other science curricula that take a similar, real-world approach.

At the same time, the global challenge of sustainability drives a need for scientifically and technologically informed citizens and encourages higher education institutions and professional societies to focus STEM curricula on this real-world challenge. For example, the Curriculum for the Bio-region Initiative of the Washington Center for Improving the Quality of Undergraduate Education has engaged STEM faculty to define sustainability learning outcomes (see <http://www.evergreen.edu/washcenter/project.asp?pid=62>). The American Association for the Advancement of Science focused its 2009 annual meeting on sustainability with the theme Our Planet and Its Life: Origins and Futures.

Middlecamp closed by proposing that STEM higher education faculty target curriculum and instruction to the areas of intersection among their own vision of teaching and learning, what students care about, the challenges facing the planet.

### Evolutionary Ecology

Bruce Grant (Widener University) began his presentation by emphasizing the importance of addressing students’ alternative conceptions of evolution. He noted that the United States ranked near the bottom in a recent comparative international study on the proportion of the public that accepts the theory of evolution (Miller, Scott, and Okamoto, 2006). Grant suggested that this lack of acceptance of a well supported theory reflects a larger ideological struggle in American society over the basic concept that evidence matters. He explained that he was motivated to change his teach-

ing approach because of these concerns and because a large proportion of students fail introductory biology classes or drop biology as a major field of study.

Grant then described his practitioner research, arguing that it has improved his freshmen students' conceptual acceptance of evolution by natural selection. He has conducted research on student learning among eight cohorts of freshmen enrolled in an evolutionary ecology course each year from 2000 to 2007, revising the course based on his research. He observed that, because practitioner research incorporates many aspects of traditional scientific epistemology but excludes other aspects, it constitutes a unique and complementary "way of knowing" that can improve science teaching and student learning.

Grant said he administered a standardized final examination at the end of the course each year to assess student learning and their response to his course revisions. The examination includes the prompt, "Please offer a brief and concise definition of evolution." Since 2005, he has also used this prompt as a pretest. In addition, he has administered a standardized assessment item designed to measure students' conceptions about evolution (Ebert-May, 2000).

Beginning in fall 2005, Grant conducted frequent short-answer surveys of students' preconceptions about key topics before they were discussed in class, but the assessment results showed only slight improvement in the learning of basic concepts. Beginning in fall 2006, Grant directly confronted his students with their alternative conceptions, as indicated by their responses to the short-answer surveys and the pretests. He presented students with histograms of their responses and, at the same time, revised the course syllabus to address the alternative conceptions. In addition, he asked them in guided discussions to reflect on the kinds of evidence and arguments he should present that would help them understand the key topics. Finally, he substantially reduced the content and shifted class time toward increased writing and classroom discourse.

These changes yielded significant gains in student learning in the more recent classes, in comparison with earlier classes. The fraction of correct responses to the prompt, "Please offer a brief and concise definition of evolution" rose from about 50 percent in the period 2000 to 2005 to 90 percent in December 2006 and 80 percent in December 2007. Students' mean scores on the standardized final exam went from 6.44 in December 2002 to 9.51 in December 2006 and 8.79 in December 2007. Grant also found large gains in student scores on the standardized question on evolution. From 2000 through 2005, only about 3 percent of students scored 8, 9, or 10 on this 10-point question, but in 2006 and 2007, about 54 percent achieved a score of 8, 9, or 10. The mean scores on this item also improved significantly, from 4.38 to 7.36.

Grant concluded that the revisions he instituted in fall 2006 significantly decreased students' misconceptions and improved their learning about the concept of evolution and the process of evolution by natural selection. In addition, he learned new approaches to teaching that rely on the evidence generated by his practitioner research. He promised to continue to redesign and improve the course and described plans to increase his use of published concept inventories and to engage students in research on their own learning. He encouraged other STEM faculty to engage in practitioner research.

### Physics

Jose Mestre (University of Illinois, Urbana-Champaign) presented his perspective on learning goals and evidence. He explained that his view of important learning goals reflects the current problem that, because of the explosion of scientific and technological knowledge, students in introductory courses are asked to learn an increasing body of knowledge, only to forget it weeks after the course is over. He suggested three learning goals:

1. Structure instruction to help students learn a few things well and in depth.
2. Structure instruction to help students retain what they learn over the long term.
3. Help students build a mental framework that serves as a foundation for future learning.

Mestre proposed that that evidence of achievement of the first goal would include understanding of concepts underlying problem solutions (depth) and the ability to apply concepts within and across domains (breadth). Measures of students' ability to understand and apply concepts obtained months after the course was over would provide evidence of achieving the second goal (retention). Finally, students' ability to learn new material more efficiently would constitute evidence of achievement of the third goal.

Mestre views these types of evidence as most compelling, and he argued against using evidence of student gains in factual or procedural knowledge to demonstrate that an instructional practice is effective. He noted that the latter type of gains do not indicate that students have developed a conceptual organizing framework, nor do they reflect flexible, durable learning. However, current assessment practices emphasize short-term recall of facts and procedures. Few studies have been conducted on transfer or retention of STEM knowledge months after a course is over. As a result, there are gaps in the available evidence related to the three student learning goals he listed.

Mestre said that the quality of evidence related to learning goals has an important effect on the adoption of promising practices. In physics, the development of the Force Concept Inventory (Hestenes, Wells, and Swackhamer, 1992), which provides high-quality evidence of student misconceptions, led to dramatic increases in the use of new teaching and learning approaches designed to engage students and eliminate misconceptions (Mestre, 2005). Mestre described as good news the development of similar tests of misconceptions in other disciplines (see Chapter 6).

### Plenary Discussion

In the discussion following the presentations, Kimberly Kastens (Columbia University) asked whether the goals of the approaches described by the speakers included changing student behavior related to societal issues, such as global warming. Middlecamp responded that *Chemistry in Context* aims to influence students' behavior in making choices; specifically, the goal is to help them make informed choices about issues that affect themselves and others. Mestre said that, although his physics classes do not focus on societal issues, he does seek to change students' behavior in constructing scientific arguments and responding to other students' arguments. He noted that it is difficult to change students' behavior in this area, as they want him to simply present the scientific reasoning that leads to the correct answer.

Edward (Joe) Redish (University of Maryland) asked Grant whether he had evidence to support his claim that student scores improved because he had acknowledged and validated their struggles with learning the concepts and had made learning more personal, relevant, and accessible to them. Grant acknowledged that he lacked evidence for the claims and called for research on how students develop a learning community and become motivated to learn science.

Gamoran noted that Grant used an interrupted time-series research design. He pointed out that although this design is useful to demonstrate that a change occurred, it cannot determine whether the "interruption" (i.e., the change in instruction) caused the outcome. Other factors that may have caused test scores to increase cannot be ruled out. In addition, because Grant introduced a package of changes, including eliminating some of the content, instituting short-answer surveys at the beginning of class, and confronting students with their misconceptions, it is difficult to untangle the specific changes that may have caused the gains. Nevertheless, Gamoran described Grant's research as a valuable "existence proof," demonstrating that it is possible to reduce the level of students' alternative conceptions.

Committee chair Susan Singer (Carleton College) asked the speakers about their use of cognitive research and theory, such as research on devel-

opment of expertise. Mestre replied that some of his early research focused on the differences between novices and experts in physics thinking and problem solving and indicated that he frequently draws on the cognitive research when investigating and revising his teaching practices. Middlecamp said she had learned more from her own experiences of people being intimidated by chemistry or wondering about its relevance than from the research. These experiences, she said, increased her awareness of the problems of traditional chemistry teaching and motivated her to develop a different approach. Grant said that he has made concerted efforts to learn from the cognitive research, despite the difficulty of deciphering the jargon in this rapidly developing field.

Robin Wright (University of Minnesota) said she has been surprised and frustrated by the hesitation of faculty to accept research evidence supporting new teaching methods. She asked how to improve transfer of this research evidence to STEM faculty. Mestre suggested inviting a skeptical faculty member to test his or her students' understanding of basic concepts in the discipline. He predicted that this approach would demonstrate that the students in lecture courses do not understand these concepts as well as ones who have been taught using active learning methods. Middlecamp responded to Wright by proposing a different question, "What . . . might tip change more quickly than the evidence?" Mestre suggested that it is important to ask not only how to transfer research findings to faculty, but also how to transfer the findings from faculty. He noted that faculty members require training in how to collect valid evidence to measure the effects of instruction, but administrators place a low priority on this type of research on instruction.

In response to a question about his goal of having students develop a mental framework to serve as a foundation for future learning, Mestre described a study showing that children who knew a lot about spiders could more easily recall new information about spiders than other children with less background knowledge. He speculated that helping students develop mental frameworks in physics would make it easier to teach them new material in physics. He noted that it would be difficult to assess whether learning really becomes more efficient if students develop such mental frameworks.

Carol Snyder (American Association of Colleges and Universities) observed that change in undergraduate STEM might be supported more effectively by collaborative work in departments than by the efforts of individual faculty members. Heidi Schweingruber (National Research Council) asked what the speakers are doing to measure the goals of developing positive attitudes toward science, noting that increasing students' science knowledge will not necessarily lead to changes in their behavior or degree of motivation to learn. For example, some physicians smoke, despite their

knowledge of the overwhelming evidence of the health dangers of smoking. Middlecamp responded that she believes that the *Chemistry in Context* goal to develop students' motivation for lifelong science learning is more important than the goal to help them learn specific chemistry content. This is why the *Chemistry in Context* team selects chemistry content that matters to people for inclusion in the curriculum, she said.

Linda Slakey (National Science Foundation) asked Middlecamp about directly introducing chemistry concepts related to real-life issues, without first introducing students to the basic topics of general chemistry, which she considers to be the scaffolding on which to build new understanding. Celeste Carter (National Science Foundation), who teaches a 9-month course in biotechnology at a community college, said that many of her students, including some with advanced degrees, do not understand basic concepts in biology. She said she tries to build conceptual understanding through discussion of scientific methods and through laboratory activities. James Stith (American Institute of Physics) asked how departments can be held accountable for ensuring that prerequisite courses provide the basic understanding students need to benefit from more advanced courses.

Middlecamp responded to Slakey that a scaffold in *Chemistry in Context* would be the real-life issue, such as “the air we breathe” or “the water we drink.” Grant said that people have very different definitions of the word “scaffold”; he thinks of it as an awareness of one's own learning process and how one builds understanding in response to instruction. Mestre responded that he uses class time to scaffold student learning, drawing on his own expertise, and asks students to read more basic content material outside class. Responding to the question about holding departments accountable, Mestre observed that new doctoral graduates lack knowledge of active learning strategies and proposed that departments should be held accountable for bringing their newly hired faculty up to speed on the findings of cognitive research and their implications for instruction.

Moderator Adam Gamoran offered three concluding remarks. First, he observed that the evidence underlying promising practices in STEM is thin, as each speaker had described local, anecdotal evidence. Second, he suggested that cognitive scientists, educational testing experts, and disciplinary experts collaborate to develop new forms of assessment to guide STEM teaching and learning. Third, he called for increasing the scope and scale of research to support development of approaches that are useful across different faculty members, departments, and institutions.

Workshop participants then formed small groups for further discussion of learning goals and evidence.

### Small-Group Discussions

In small groups, workshop participants discussed the learning goals in the STEM disciplines, their views about the most important of these goals, and the types of evidence needed to establish effectiveness in terms of the most important goals. They also considered whether the desired learning goals and associated types of evidence differ across the STEM disciplines. Following the discussions, session moderator Susan Singer invited a reporter from each group to briefly describe that group's response to these questions.

James Stith reported that his group explored the following issues:

- Should there be different goals for students majoring in a STEM discipline and for other students, who require only a general knowledge of the subject matter?
- Although professional societies have promulgated science learning goals, faculty members may not understand or even be aware of these goals.
- Expert faculty members find it challenging to represent the material in their discipline to the novice learner and help him or her make the connection between representations and the real world.
- It is important to help students understand that a STEM field has an underlying structure and is not simply a collection of facts.
- What are the best ways to teach students about the nature of science, including the role of experimental methods and the relationships between facts and theory?

Robin Wright explained that her group focused on three types of learning goals for students: (1) core concepts and ways of knowing in the particular STEM discipline; (2) skills in communication, critical thinking, and asking good questions; and (3) positive attitudes toward STEM. She reported several group observations related to these goals:

- What counts as evidence of learning outcomes differs across STEM disciplines.
- Test questions should be aligned with specific desired learning outcomes.
- Surveys can be helpful to assess the development of positive attitudes toward STEM.
- Assessment should take place not only within a single course, but also across courses, levels of education, and even lifetimes.

Brock Spencer (Beloit College) shared the following points from his group's discussion:

- The goals for general education students may include more emphasis on societal issues than the goals for STEM majors.
- Important goals related to student understanding of the nature of science include knowledge of experimental methods, the ability to make judgments and deal with uncertainty, the capacity to build a scientific argument based on physical evidence, and understanding the explanatory power of scientific models.
- The Force Concept Inventory may be more effective in changing faculty behavior than in creating evidence of student learning. It is easier to administer than other, more labor-intensive assessments, but it also provides a less detailed view of students' thinking and learning.
- Current efforts to develop new assessments of students' skills and attitudes will provide new types of evidence in the future.
- Scientists are sometimes skeptical of evidence obtained using qualitative or ethnographic methods.
- It is valuable to identify common, cross-disciplinary goals and also to identify important learning goals in each discipline.

Dexter Perkins (University of North Dakota) reported that members of his group discussed the following ideas:

- Cross-disciplinary goals, such as problem solving, communication, and critical thinking, are important, in addition to more specific goals for what students should know and be able to do after completing a particular class.
- What are the best ways to build instruction to achieve these cross-disciplinary goals?
- Assessing student progress toward cross-disciplinary goals is difficult.
- There are many different kinds of evidence related to learning goals and no single best way to collect these kinds of evidence.
- Pre- and posttests are valuable to measure change in specific abilities or attitudes, and grading rubrics are very helpful to ensure that pre- and posttests are graded consistently.
- It would be valuable to obtain evidence of students' later learning and performance, after they leave a particular STEM class.
- One way to demonstrate the effectiveness of instructional changes is to obtain multiple measures on a cohort of students as they progress through the STEM curriculum.
- Much more evidence is needed, but it is difficult to obtain.



Perkins concluded that, despite the skepticism of their STEM colleagues about new types of teaching, the group members are motivated by the fun and satisfaction of researching student learning and revising instruction to improve learning.

### THE STATE OF EVIDENCE IN DISCIPLINE-BASED EDUCATION RESEARCH

Opening a second session on learning goals and evidence, session moderator Kenneth Heller (University of Minnesota) introduced three panelists who had been invited to summarize the major findings from discipline-based education research in their respective disciplines and to identify the most promising directions for future research.

#### Physics Education Research

Edward (Joe) Redish opened his remarks by using examples from the established field of physics education research to disagree with Adam Gamoran's earlier observation that the evidence underlying promising practices in STEM is primarily local and anecdotal. Redish said research in physics education has been under way for 30 years and that the physics education research community includes a literature base and regular conferences. He pointed out that the online peer-reviewed journal *Physical Review Special Topics-Physics Education Research* has been available since 2005. In addition, a 1999 bibliography cites more than 200 papers in physics education research conducted at the university level (McDermott and Redish, 1999).

Redish said that physics education researchers frequently rely on interviews as a source of evidence of effectiveness, asking students to explain the process they used to solve a physics problem. Researchers also use pre- and posttests, and for the past 10 years they have collected ethnographic data, including videotapes of students at work in the physics classroom.

Turning to his summary of findings from physics education research, Redish said that the findings support constructivist theories of education, indicating that students assemble their responses to instruction from what they already know. In the process, students sometimes develop incorrect, but robust, alternative conceptions. A relatively small number of alternative conceptions dominate students' responses to instruction. These alternative conceptions may exist even among students who are successful in using algorithms to solve problems.

The research shows that physics learning is highly dependent on context. A student may develop alternative conceptions on the fly in response to new information. The existing knowledge he or she draws on when developing either a correct conception or an alternative conception can

be dramatically affected by how he or she perceives contextual factors. A student may even hold contradictory ideas about a phenomenon without noticing the contradiction.

Other findings illuminate how students compile their understanding of physics. When they have learned a concept well, they develop automatic thought patterns and may no longer be aware of the components in these patterns. Similarly, students may hold intuitions that they find hard to explain. Instructors, who have also developed automatic thought patterns, may find it difficult to understand why students do not just see it, as they do. To support students in this situation, instructors must reverse-engineer their own knowledge to identify its components and the relationships among them.

Finally, the research findings address instructional reform. First, the research has demonstrated that it is possible to create instructional environments that substantially improve student performances on tests of conceptual understanding. Second, research has shown that these instructional environments can be transferred to other institutions and implemented successfully. Third, the evidence suggests that a critical element in successfully implementing these instructional environments appears to be getting students mentally engaged.

Cautioning that much more research is needed to understand the specific factors involved in student learning of physics, Redish (2008) identified the following four promising areas for future research:

1. Investigate what prior knowledge, expectations, and attitudes students bring to physics class and when and how they apply prior knowledge, expectations, and attitudes in response to instruction.
2. Deconstruct students' alternative conceptions and identify underlying components that may be easier to realign than to replace.
3. Study how students come to understand their own construction of their mental structures of interrelated concepts and principles—which are fundamental to learning physics—and learn when to apply knowledge they already possess.
4. Conduct interdisciplinary research that carefully links physics education research with cognitive and neuroscience research.

### Life Sciences Research

William Wood (University of Colorado, Boulder) opened his remarks by describing the context for life sciences education research—the discipline of biology—as fragmented into subfields. Many of the professional societies associated with these subfields have begun to conduct research on teaching and learning and establish education research journals; however, most

faculty members read only the education journal that is specific to their own professional society.

Wood said that, with growing awareness and interest, life sciences education research is where physics education research was in the late 1980s. Active learning strategies for teaching large biology classes, based partly on this research, are being actively disseminated. For example, the National Academies Summer Institutes on Undergraduate Education in Biology drew more than 200 participants in 2004-2008, including faculty representing 65 institutions in 36 states. Wood said that these participants, in turn, have applied their learning, impacting an estimated 80,000 students.

New approaches to biology instruction are informed by several types of evidence. First, life scientists depend heavily on evidence that has emerged from physics education research. Second, they often conduct “design research,” testing the effectiveness of their own changes in instruction over time, but often without a control group for comparison. There have been only a handful of quasi-experimental studies in life sciences education research and no controlled experimental studies that randomly assign students to different types of instruction.

Wood presented findings from a quasi-experimental study he conducted with a colleague, focusing on upper level undergraduates enrolled in a required course in developmental biology (Knight and Wood, 2005). Over the course of two successive semesters, the authors presented the same course syllabus using two different teaching styles: in fall 2003, the traditional lecture format; and in spring 2004, decreased lecturing and increased student participation and cooperative problem solving during class time, including frequent in-class assessment of understanding. They found significantly higher learning gains and better conceptual understanding in the more interactive course; when they repeated the interactive course in spring 2005, they found similar results.

Wood raised several important questions for the future of life sciences education research. First, the field lacks a strong theoretical framework that integrates and interprets the research to date, similar to the volume in physics education research (Redish, 2004). This leads to two questions:

1. Does all physics research apply to learning life sciences, or does the higher requirement for factual knowledge in the life sciences require new research models?
2. Under what circumstances does student-centered instruction result in more learning than traditional lecture classes, and under what circumstances does it not?

Second, Wood highlighted an important question about the practical impact of life sciences education research: What kinds of evidence/interventions/

interactions result in meaningful change in the way postsecondary institutions and their faculties view student learning and design their instructional practices?

Addressing this final question, Wood suggested that discipline-based education researchers in all STEM disciplines could be instructed by the study of Henderson and colleagues (2008) related to change in STEM higher education. That study identified four integral elements of undergraduate STEM education: (1) teachers, (2) culture, (3) curriculum/pedagogy, and (4) policy. The authors propose that an effective change strategy would address all four elements, but they found that most change strategies emerging from discipline-based STEM education research address only the element of curriculum/pedagogy (Henderson et al., 2008).

### Geosciences Education Research

Helen King (Helen King Consultancy) opened her remarks with a description of the current context supporting education research in the geosciences. Although geosciences education research is a relatively young subdiscipline, it includes a strong and growing community of researchers and practitioners at all levels of education. Knowledge is shared in the *Journal of Geoscience Education* and within and across national and international professional associations. The community is beginning to establish research methodologies, and the field is gaining legitimacy, as evidenced by the rapidly growing number of tenure-track education positions in geosciences departments.

In this context, King said, faculty members are developing new teaching practices based partly on general cognitive research and partly on findings from research in other STEM disciplines, as well as on findings emerging from geosciences education research. These new teaching practices include promoting active learning, deploying an array of assessment strategies, engaging students in problem solving while in the field, using visualizations and other applications of computer technology, and creating relevant case studies.

King cited a study of teaching practices employed by geology faculty in the United States which stated, “there is no question that research on learning and resulting recommendations for best classroom practice . . . have had an impact on geosciences classes” (Macdonald et al., 2005, p. 237). She then identified the major themes of geosciences education research, including how students learn important concepts and skills, the nature of discovery in geosciences, and students’ alternative conceptions of the discipline and of particular topics.

The research on geosciences education has identified several sticking points in student learning, including the development of systems thinking,

understanding complexity and uncertainty, and transfer of knowledge from mathematics and physics to solve problems in the geosciences. Research on the development of expertise in the geosciences, including the difficult process of developing spatial thinking and the ability to think about geological time, has potential to help novices advance toward such expertise. Finally, researchers are beginning to gain understanding of how different learning environments and contexts, including the classroom, laboratory, the field, and the workplace, affect students' learning. This has included investigations of how contexts influence students' values, beliefs, and feelings, and how these influences may, in turn, affect learning.

King concluded with an outline of progress in geosciences education research. This progress includes professional development for faculty, with training in important findings from geosciences education research; research funding and collaboration across institutions, disciplines, and nations; and dissemination of research findings to raise the profile of the research and encourage application of its findings.

### Discussion

In the discussion following the presentations, Adam Fagen (National Research Council) asked Heller and Redish about the applicability of what has been learned in physics education research to the other science disciplines. Redish and Heller agreed that there are not only some real differences, but also similarities across the disciplines. Heller said he reminds his physics colleagues that learning is a biological process, and that content, skills, and attitudes are inseparable from a biological perspective.

Ginger Holmes Rowell (National Science Foundation) asked what can be learned from recent learners (i.e., students who have taken a course in the previous semester) and how they might help to design more effective learning environments. Addressing the second question, Redish cited the University of Colorado's Learning Assistance Program as an interesting and exciting use of recent learners by giving them instruction in pedagogy and folding them back into the classroom. He explained that many students from that program are recruited to become K-12 science teachers.

In response to a question from Robin Wright, Wood said that the field needs better assessments to measure higher order thinking skills, such as problem-solving. Redish agreed and described his own efforts to create assessments that address the higher levels of Bloom's taxonomy, such as by including essay questions and multiple-choice problems that are difficult to answer without a solid conceptual understanding of a physical system.

Wood and King discussed the importance of being transparent with students about learning goals and methods as a way to promote learning. Wood noted that Dee Silverthorn (2006) has written beautifully on the need

to let students know why inquiry and problem solving are important for their futures before suddenly requiring them to do things in biology courses that they have never before been asked to do. King referred to the phenomenon, discussed earlier in the day, of sharing exam questions with students, which she said can be a strong motivator to learn the required content.

Heidi Schweingruber asked about the relative emphasis in the disciplines on deep conceptual knowledge versus thinking about how students understand inquiry and the nature of science. Redish responded that, although there is strong agreement about the importance of conceptual knowledge, it is integrated differently into the different epistemologies of the disciplines. Wood added that teaching conceptual knowledge is relatively similar across the disciplines, but inquiry within each discipline is more specialized.

Responding to another question, Wood said that inquiry is probably not as much a tradition in the lower level courses as it should be. He explained that biologists teach more about facts because they think students have to know the facts before they can start thinking about inquiry. King added that, when she was pursuing a degree in geology, no one explicitly told her about the nature of knowledge and inquiry in geology. She suggested that it is important to help students better understand the nature of the discipline they are studying and the role of inquiry.

## 3

# Surveying Promising Practices

### PROMISING PRACTICES FOR FACULTY AND INSTITUTIONS AND PREDICTING SUCCESS IN COLLEGE SCIENCE

Moderator Melvin George (University of Missouri) introduced three panelists to discuss a range of promising practices. Each panelist was asked to address the following questions:

1. How would you categorize the range of promising practices that have emerged over the past 20 years? Consider practices that are discipline-specific as well as those that are interdisciplinary.
2. What types of categories do you find are most useful in sorting out the range of efforts that have emerged? Why did you choose to aggregate certain practices within a category?
3. As you chose exemplars for your categories, what criteria did you use to identify something as a promising practice?

Jeffrey Froyd (Texas A&M University) began by describing a framework that he developed to categorize promising undergraduate teaching practices in science, technology, engineering, and mathematics (STEM).<sup>1</sup> The framework begins with a set of decisions that faculty members must make in designing a course:

---

<sup>1</sup>For more detail about this framework, see the workshop paper by Froyd (see [http://www.nationalacademies.org/bose/Froyd\\_Promising\\_Practices\\_CommissionedPaper.pdf](http://www.nationalacademies.org/bose/Froyd_Promising_Practices_CommissionedPaper.pdf)).

- Expectations decision: How will I articulate and communicate my expectations for student learning?
- Student organization decision: How will students be organized as they participate in learning activities?
- Content organization decision: How will I organize the content for my course? What overarching ideas will I use?
- Feedback decision: How will I provide feedback to my students on their performance and growth?
- Gathering evidence for grading decision: How will I collect evidence on which I will base the grades I assign?
- In-classroom learning activities decision: In what learning activities will students engage during class?
- Out-of-classroom learning activities decision: In what learning activities will students engage outside class?
- Student-faculty interaction decision: How will I promote student-faculty interaction?

The next component of Froyd's framework relates to two types of standards against which faculty members are likely to evaluate a promising practice: (1) implementation standards and (2) impact standards. Implementation standards include the relevance of the promising practice to the course, resource constraints, faculty comfort level, and the theoretical foundation for the promising practice. Student performance standards relate to the available evidence on the effectiveness of the promising practice, which may include comparison studies or implementation studies.

Froyd then identified eight promising practices related to teaching in the STEM disciplines and analyzed each in terms of his implementation and student performance standards (see Table 3-1).

Jeanne Narum (Project Kaleidoscope) identified three characteristics of institutional-level promising practices in STEM, noting that they (1) connect to larger goals for what students should know and be able to do upon graduation, (2) focus on the entire learning experience of the student, and (3) are kaleidoscopic (Narum, 2008). She explained that promising practices can focus on student learning goals at the institutional level, the level of the science discipline, and the societal level. To illustrate these points, Narum described examples of institutional transformation at the University of Maryland's Baltimore Campus, Drury University, and the University of Arizona. As she explained, each institution set specific learning goals, designed learning experiences based on the goals, and assessed the effectiveness of the learning experiences. Narum also provided examples of other institutions engaged in promising practices related to assessment and pedagogies of engagement. In closing, Narum said that the best institutional practices arise when administrators and faculty share a common



TABLE 3-1 Summary of Promising Practices

Promising Practices	Rating with Respect to Implementation Standards	Rating with Respect to Student Performance Standards
1: Prepare a set of learning outcomes	Strong	Good
2: Organize students in small groups	Strong	Strong
3: Organize students in learning communities	Fair	Fair to good
4: Scenario-based content organization	Good to strong	Good
5: Providing students feedback through systematic formative assessment	Strong	Good
6: Designing in-class activities to actively engage students	Strong	Strong
7: Undergraduate research	Strong or fair	Fair
8: Faculty-initiated approaches to student-faculty interactions	Strong	Fair

NOTE: Strong = easy and appropriate to implement, good = slightly less so, and fair = even less so.

SOURCE: Froyd (2008). Reprinted with permission.

vision of how the pieces of the undergraduate learning environment in STEM fit together and a commitment to work together as an institution to realize that vision.

Philip Sadler (Harvard University) focused on lessons from precollege science education. He described a large-scale survey that he and his colleagues conducted of students in introductory biology, chemistry, and physics courses at 57 randomly chosen postsecondary institutions. The focus of the study was on certain aspects of high school STEM education (e.g., advanced placement courses, the sequencing of high school science courses) that predict students' success or failure in their college science courses. Sadler reported that 10 percent of students in introductory science courses had previously taken an advanced placement (AP) course in the same subject in high school, and those students performed only slightly better in their introductory college courses than non-AP students. Moreover, AP students who took introductory (101-level) courses did better in 102-level courses than AP students who began with 102-level courses. These findings led Sadler to recommend against AP courses for most high school students.

Next, Sadler discussed the effect of high school science-course taking on students' performance in introductory college science courses. Overall, students who took more mathematics in high school performed better in all of their science courses than students who took fewer mathematics courses. Moreover, students who took multiple high school courses in a given science discipline performed better in college science courses in that

discipline. However, Sadler and his colleagues found no cross-disciplinary effects, meaning that students who took multiple chemistry courses did not perform significantly better in college biology; students who took multiple high school physics courses did not perform better in college chemistry; and so on. Sadler also reported that the use of technology in high school science classes did not predict success in college science; however, experience in solving quantitative problems, analyzing data, and making graphs in high school did seem to predict success in college science courses.

### SMALL-GROUP DISCUSSIONS AND FINAL THOUGHTS

In small groups, participants identified what they considered to be the most important promising practices in undergraduate STEM education. The following list emerged from the small-group reports:

1. Teaching epistemology explicitly and coherently.
2. Using formative assessment techniques and feedback loops to change practice.
3. Providing professional development in pedagogy, particularly for graduate students.
4. Allowing students to “do” science, such as learning in labs and problem solving.
5. Providing structured group learning experiences.
6. Ensuring that institutions are focused on learning outcomes.
7. Mapping course sequences to create a coherent learning experience for students.
8. Promoting active, engaged learning.
9. Developing learning objectives and aligning assessments with those objectives.
10. Encouraging metacognition.
11. Providing undergraduate research experiences.

To close the workshop, steering committee members reflected on the main themes that were covered throughout the day. Susan Singer focused on the question of evidence and observed that the workshop addressed multiple levels of evidence. Explaining that assessment and evidence are not synonymous, she pointed out that classroom assessment to inform teaching generates one type of evidence that workshop participants discussed. Another type of evidence is affective change, and she observed that some people gather evidence to convince their colleagues to change their practice. Singer said the workshop clearly showed that scholars in some disciplines have given careful thought to the meaning of evidence and have begun to gather it to build a general knowledge base.

Melvin George began his reflections by asking, “Why do we need any evidence at all?” He noted that one reason for gathering evidence is to discover what works in science education, but he said that evidence alone does not cause faculty members to change their behavior. Suggesting that the problem might lie with ineffectual theories of change rather than a lack of evidence, George proposed that it might be more productive to direct more attention and resources to making change happen.

David Mogk (University of Montana) observed that the participants discussed a continuum of promising practices ranging from individual classroom activities to courses to curricula to departments to institutional transformation. Discussing the day’s themes, Mogk described a desire to identify promising practices that promote mastery of content and skills while addressing barriers to learning, and he recalled discussions about the difficulty of articulating and assessing some of those skills. He identified the use of technology as a promising practice that cuts across disciplines and suggested a need to examine the cognitive underpinnings of how people learn in each domain. Mogk called for better alignment of learning goals, teaching and learning activities, and assessment tools.

William Wood reflected on the issue of domain-specific versus generic best practices. He noted that many of the practices discussed during the workshop seem universally applicable across disciplines and even across different levels, such as the classroom, department, and institution as a whole. He also suggested that university faculty might apply some of these principles when encouraging their colleagues to transform their teaching practice. Rather than transmitting the evidence in a didactic manner and expecting colleagues to change, Wood proposed taking a more constructivist approach to build their understanding of promising practices.

Kenneth Heller remarked on the different grain sizes of the promising practices that the participants discussed. He noted that the different goals and different kinds of evidence associated with each grain size present a challenge to generating useful evidence about promising practices. He agreed with previous speakers that evidence is important but not sufficient to drive change. Heller concluded by using a quote from the poet Voltaire as a cautionary message about gathering more evidence instead of putting existing research into practice: “The best is the enemy of the good.”

## 4

## Scenario-, Problem-, and Case-Based Teaching and Learning

The primary purpose of the October workshop was to thoughtfully examine the evidence behind a select set of promising practices that came to light during the June workshop. Susan Singer opened the October workshop by linking its agenda to key themes of the June workshop (see Chapter 3). Although these practices are not perfect and do not represent the universe of evidence-based innovations, she said, they are recognized by experts as promising, and each is supported by some evidence.

The promising practices discussed include scenario-, problem-, and case-based teaching and learning (this chapter); assessments to guide teaching and learning (Chapter 5); efforts to restructure the learning environment (Chapter 6); and faculty professional development (Chapter 7). Singer explained that the presentations were based on papers prepared following a template the steering committee developed after the June workshop.<sup>1</sup> The authors were asked to describe the context in which the promising practice was implemented, identify examples of how the practice was used, and provide evidence to support the claim that the practice was promising, including evidence of its impact or efficacy.

---

<sup>1</sup>The template for the papers is available at [http://www.nationalacademies.org/bose/Commissioned\\_Paper\\_Template.pdf](http://www.nationalacademies.org/bose/Commissioned_Paper_Template.pdf) and the papers are available at [http://www.nationalacademies.org/bose/PP\\_Commissioned\\_Papers.html](http://www.nationalacademies.org/bose/PP_Commissioned_Papers.html).

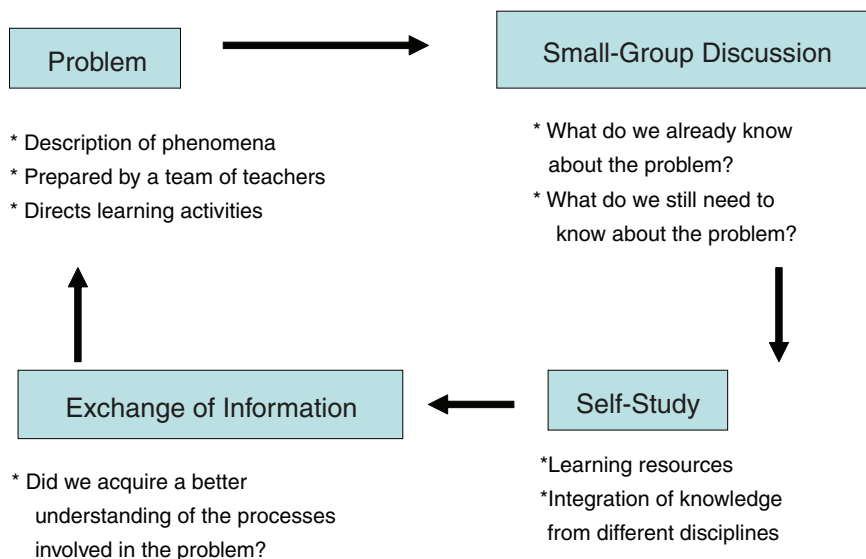


FIGURE 4-1 Problem-based learning.  
SOURCE: Gijbels (2008). Reprinted with permission.

## PROBLEM-BASED LEARNING

David Gijbels (University of Antwerp) described the cycle of problem-based learning (see Figure 4-1). After the instructors present a problem to the class, students meet in small groups to discuss what they know about it and what they need to learn. During a short period of independent self-study, students gather the needed resources to solve the problem. They then reconvene their small groups to re-assess their collective understanding of the problem. When they solve the problem, the instructor provides a different problem and the cycle begins anew.

Noting that problem-based learning has many possible definitions and permutations, Gijbels nonetheless stressed the importance of identifying a core set of principles that characterize this type of learning. Having a core definition enables researchers to compare problem-based learning with other types of learning environments. In his research, Gijbels uses a model developed by Howard Barrows (1996) that identifies six characteristics of problem-based learning:

1. Student-centered learning.
2. Small groups.
3. Tutor as a facilitator or guide.

4. Problems first.
5. The problem is the tool to achieve knowledge and problem-solving skills.
6. Self-directed learning.

Gijbels then described a meta-analysis conducted to examine the effects of problem-based learning on students' knowledge and their application of knowledge, and to identify factors that mediated those effects (Dochy et al., 2003). The meta-analysis focused on empirical studies that compared problem-based learning with lecture-based education in postsecondary classrooms in Europe, and almost all of the studies that met the criteria focused on medical education.<sup>2</sup> Through the analysis, Gijbels and his colleagues found the following:

- Students' content knowledge was slightly lower in problem-based learning courses than in traditional lecture courses.
- Although students in problem-based learning environments demonstrated less knowledge in the short term, they retained more knowledge over the long term.
- Students in problem-based learning settings were better able to apply their knowledge than students in traditional courses.

These findings prompted Gijbels and his colleagues to undertake a deeper analysis of the assessment of problem-based learning (Gijbels et al., 2005). That analysis focused on three levels of knowledge that were assessed in the selected studies: (1) knowledge of concepts, (2) understanding of principles that link concepts, and (3) the application of knowledge. Gijbels noted that of the 56 studies in the analysis, 31 focused on concepts, 17 focused on principles, and 8 focused on the application of knowledge. The analysis revealed the following:

- Students in problem-based learning environments and traditional lecture-based learning environments exhibited no differences in the understanding of concepts.
- Students in problem-based learning environments had a deeper understanding of principles that link concepts together.
- Students in problem-based learning environments demonstrated a slightly better ability to apply their knowledge than students in lecture-based classes.

---

<sup>2</sup>The study is described in the workshop paper by Gijbels (see [http://www.nationalacademies.org/bose/Gijbels\\_CommissionedPaper.pdf](http://www.nationalacademies.org/bose/Gijbels_CommissionedPaper.pdf)).

Gijbels concluded by stating that problem-based learning has not completely fulfilled its potential. He suggested that students might become better problem solvers if faculty members assessed them more on problem solving. Noting that students often do not develop a sense of shared cognition when working in teams in problem-based learning environments, he also stressed the importance of attending to group developmental processes when implementing problem-based learning.

### CASE-BASED TEACHING

Mary Lundeberg (Michigan State University) defined some key elements of case-based teaching. In the paper she wrote for the workshop (Lundeberg, 2008, p. 1), she said:

Cases involve an authentic portrayal of a person(s) in a complex situation(s) constructed for particular pedagogical purposes. Two features are essential: interactions involving explanations, and challenges to student thinking. Interactions involving explanations could occur among student teams, the instructor and a class; among distant colleagues; or students constructing interpretations in a multimedia environment. Cases may challenge students' thinking in many ways, e.g., applying concepts to a real life situation; connecting concepts [and/or] interdisciplinary ideas; examining a situation from multiple perspectives; reflecting on how one approaches or solves a problem; making decisions; designing projects; considering ethical dimensions of situations. Brief vignettes, quick examples, or unedited documents are not cases.

She presented four examples to illustrate the wide range of cases that might be used in undergraduate science, technology, engineering, and mathematics (STEM) education:<sup>3</sup>

1. The Deforestation of the Amazon: A Case Study in Understanding Ecosystems and Their Value, a problem-based case used in a biology seminar for nonmajors.
2. Cross-Dressing or Crossing-Over: Sex Testing of Women Athletes, a historical case used in large lecture courses with clicker technology (handheld wireless devices through which students register their responses to multiple-choice questions that are projected on a screen).
3. Case It!, in-depth problem-based multimedia cases used in biology labs.
4. Project-based scenarios used in engineering.

---

<sup>3</sup>For more detail on these cases, see the workshop paper by Lundeberg (see [http://www.nationalacademies.org/bose/Lundeberg\\_CommissionedPaper.pdf](http://www.nationalacademies.org/bose/Lundeberg_CommissionedPaper.pdf)).

Citing the National Research Council (2002), Lundeberg identified three types of research questions often investigated in studies of educational activities—those that focus on description, cause, and process. She explained that there is much more descriptive research (i.e., faculty and student perceptions of what is happening) than research showing causal effects or describing the process of learning.

Lundeberg described the research that she and her colleagues have conducted on case-based learning. The descriptive aspects of their research involved surveys of 101 faculty members in 23 states and Canada who were using cases from the National Center on Case Study Teaching and Science (see <http://library.buffalo.edu/libraries/projects/cases/case.html>). On the surveys, faculty members reported that cases make students more engaged and active learners and help them to develop multiple perspectives, gain deeper conceptual understanding, engage in critical thinking, enhance their communications skills, and develop positive peer relationships (Lundeberg, 2008). Lundeberg also reported that faculty members cited the increased time needed to prepare lessons and assess students as obstacles to implementing case-based learning.

To identify the systematic effects of case-based learning, Lundeberg and her colleagues conducted a year-long study of the use of cases in large undergraduate biology classes equipped with clickers. The study combined a design involving random assignment to experimental and control groups with an A-B-A-B design in which 12 participating faculty members alternated the use of cases and lectures systematically across two semesters. They found that “students ( $n = 4,366$ ) who responded to cases using ‘clicker’ technology performed significantly better than their peers on five of the eight biology topics (cells, Mendelian genetics, cellular division, scientific method, and cancer), and in five of the eight areas in which they were asked to transfer information (cells, cellular division, scientific method, microevolution and DNA)” (Lundeberg, 2008, p. 8).

Students in the clicker classes also performed significantly better on tests of data interpretation than students in lecture classes. However, students who used cases with clicker technology showed no difference or lower effects on standardized tests measuring accumulated medical knowledge, on one topic in biology (characteristics of life), and on standardized tests of critical thinking.

Lundeberg argued that cases are effective for several reasons. First, stories are a powerful mechanism for organizing and storing information. In addition, the real-life context engages students. Cases also challenge students’ thinking and require them to integrate knowledge, reflect on their ideas, and articulate them. Lundeberg noted that role-playing during case-based education engages students and enables them to consider multiple perspectives.



In closing, Lundeberg reiterated that cases have an impact on understanding, scientific thinking, and engagement. She cited the need for more multiyear, mixed-methods studies on the effectiveness of case-based teaching, particularly classroom experiments that do not confound instructor or student effects. She also identified several gaps in the knowledge base at the undergraduate level: Which students benefit from cases? What content is most suitable? What benefits do different types of cases afford? What kinds of interaction between students and faculty matter? Do cases promote scientific literacy?

### USE OF COMPLEX PROBLEMS IN TEACHING PHYSICS

Tom Foster (Southern Illinois University) discussed the use of complex problems in teaching physics. He explained that complex problems are rooted in cooperative group problem solving, which is characterized by the following traits (Foster, 2008):

- positive interdependence among group members;
- individual accountability;
- monitoring of interpersonal skills;
- frequent processing of group interactions and functioning; and
- aspects of the task or learning activity that require ongoing conversation, dialogue, exchange, and support.

Foster emphasized the importance of designing the appropriate task in using this teaching method. He noted that if the problems are simple enough to be solved moderately well alone, students will not abandon their independence to work in a group. Students also will not abandon their independence if the problems are too complex for the group to initially succeed in solving them.

Context-rich problems are one example of an appropriate task for group problem solving. Foster creates such problems by converting traditional end-of-chapter problems into complex problems that students solve cooperatively, placing students in the problem by using the word “you.” Foster and his colleagues prefer not to include pictures in the problem, as a way of encouraging the group to decide whether and how to illustrate it. According to Foster, context-rich problems also provide many other decision points to foster ongoing interaction among group members. For example, problems might include extra information, omit information, or leave variables unnamed. These problems also “hide the physics” by avoiding technical terms and focusing on real-world settings. By hiding the physics, the problems demonstrate that the world is rich in physics and require students to determine which fundamental physical principles to apply (Foster, 2008).

In physics, context-rich problems are closed-ended, meaning that there is essentially one correct answer that is dictated by the rules of mathematics and physics. Even though they are closed-ended, the problems still require creativity to define and apply the correct principles and equations. Citing Schwartz, Bransford, and Sears (2005), Foster said that this balance between effectiveness and innovation is vital to the transfer of knowledge from one situation to another.

Foster noted that context-rich problems are an excellent way to challenge students' misconceptions about problem solving. For example, students often believe that the aim of solving a physics problem is to reduce it to a mathematical exercise, and that it is always necessary to use all the information in a problem. Faculty members can address these misconceptions by structuring the problems differently, as described in previous paragraphs.

In Foster's experience, it is easy to make context-rich problems too difficult. He and his colleagues have developed a set of 21 "difficulty traits" that fall into the broad categories of approach, analysis of the problem, and mathematical solution. Faculty members can use the traits as a checklist to design context-rich problems and to assess and adjust their level of difficulty.

Turning to the evidence, Foster explained that he uses traditional instruments, such as the Force Concept Inventory and conceptual surveys on electricity and magnetism, to measure students' concept development. He has found that students who solve context-rich problems in cooperative group settings score as well on these measures as their peers who are taught using other interactive methods. To assess problem solving, Foster uses a rubric developed at the University of Minnesota that includes five dimensions: (1) description of the problem, (2) physics approach (i.e., whether students used the correct physics), (3) specific application of the physics, (4) mathematical procedures, and (5) logical progression. Foster reported that students' problem-solving abilities improve through the use of context-rich problems, but he cautioned that the method does not result in quantum leaps in problem-solving abilities. Foster called his evidence on students' attitudes and behaviors about context-rich problems anecdotal but positive.

He closed by identifying future directions for this method of physics instruction. Citing the need to create more context-rich problems in physics, he mentioned problems that begin with an answer and require the formulation of a question (such as on the television show "Jeopardy!") as well as problems in which students identify and correct errors. He also stressed the importance of developing context-rich problems outside physics to assess the transfer of knowledge from one domain to another.

## DISCUSSION

Remarking on the differences in terminology across disciplines, Karen Cummings (Southern Connecticut State University) observed that these differences pose a challenge for researchers. She asked Gijbels how he distinguished between knowledge of concepts and application of knowledge in his study. Gijbels agreed and explained that for his review of the literature he examined actual assessment questions to determine what type of knowledge they were assessing. Lundeberg added that it was a challenge for the faculty members in her study to develop assessments that measure higher order thinking, because it is easier for them to write questions that focus on definitions and conceptual knowledge.

Martha Narro (University of Arizona) asked Gijbels to clarify some of the findings that he discussed in his presentation. He explained that, across studies that assessed student learning of concepts, there was no significant difference between students in problem-based and traditional settings. Across studies that assessed student learning of principles and application of conceptual knowledge, however, students in problem-based environments performed better. He also pointed out that the findings varied depending on the context (specifically, whether the students were in their first or last year of medical school) and the curriculum, and that he was reporting on the overall trends in the data.

Responding to another question, Lundeberg and Foster discussed the issue of relevance when constructing scenarios, problems, and cases. They agreed that there is very little research on what it means to be relevant. Lundeberg related several examples of cases that faculty members designed to be relevant but that did not resonate with students. In her experience, allowing students to design their own cases is a powerful way to make the cases relevant. Foster added that many college students are still developing their identities, which makes the notion of relevance more challenging. An audience member, referring to a paper by Mayberry (1998) about pedagogies that encourage students to develop their own sense of science, cautioned faculty members to be careful about coming across as knowing more than students about what is relevant.

Following another question, the speakers engaged in a discussion about the importance of longitudinal research to understand the longer term impact of these pedagogical strategies. Lundeberg mentioned some examples of longitudinal studies of innovative instructional strategies that show mixed results. Foster added that it is difficult to measure long-term knowledge or to trace it back to its origins. As an example, he said that although students might not demonstrate understanding of a concept after a certain

course, the exposure they gained to that concept might facilitate later learning. In that situation, the initial course had an effect that is impossible to measure. The panelists noted that longitudinal research is important, difficult to conduct, difficult to fund, and relatively rare.

## 5

## Assessment to Guide Teaching and Learning

This chapter summarizes presentations and discussions related to promising practices in assessment, including the use of concept inventories and an example of how research and assessment can inform instructional improvements.

### CONCEPT INVENTORIES IN THE SCIENCES: EXAMPLES FROM THE GEOSCIENCES CONCEPT INVENTORY

Julie Libarkin (Michigan State University) discussed concept inventories in the sciences. She explained that concept inventories are multiple-choice assessments that are designed to diagnose areas of conceptual difficulty prior to instruction and evaluate changes in conceptual understanding related to a specific intervention (Libarkin, 2008). Incorrect response options for each question often are written to reflect students' misconceptions.

Libarkin said she views concept inventories as a valuable and necessary first step to investigate science learning across institutions. She remarked on their proliferation, noting that she found 23 inventories in various science domains as she was preparing for the workshop.

Using the geosciences concept inventory (GCI) as an example, Libarkin described the development cycle for concept inventories. She and her colleagues began the development process by reviewing textbooks to identify the most important geosciences concepts to cover. Although most inventories target a specific concept in the sciences (e.g., force or natural selection), the GCI covers the geosciences as a whole; it is a bank of 69 questions that are related through a psychometric technique called item-response theory.

Libarkin explained that it is possible to create subinstruments from the CGI to focus on specific topics, but it is unique among concept inventories because each subinstrument is statistically related to the others and to the whole.

The next step was to collect data on students' alternate conceptions through interviews and open-ended surveys. After that, an external team of science educators, psychometricians, and geologists reviewed the instrument. Using information from students and the external reviewers, the developers created and field-tested a pilot concept inventory.

Faculty members whose students were involved in the pilot test also reviewed the instrument. Libarkin described a situation in which this review resulted in changes to the inventory. One question asked about the coexistence of humans and dinosaurs. The 30th person to review the instrument, a biology professor, pointed out that birds are dinosaurs. Because students who know that birds are classified as dinosaurs might respond that humans and dinosaurs coexisted, the GCI development team reworded that question.

After pilot-testing the inventory, the development team performed statistical analyses on the items, conducted interviews with students to better understand their responses to the questions, and revised the instrument. In all, Libarkin said the development of the GCI took two and one-half years.

Cautioning that data are only as good as the tools used to gather them, Libarkin identified some of the considerations that are involved in developing concept inventories. First, she reviewed the terminology related to multiple-choice questions. The question itself is called the stem, and incorrect response options are called distractors.

Libarkin then provided a checklist for developing multiple-choice assessment questions. The checklist began with guiding questions, such as "Is the topic covered by this question important for geosciences understanding?" "From the perspective of an expert geoscientist, does the question actually measure some aspect of geosciences understanding?" "Would a test-taker interpret this question, including both the stem and the response options, in the same way as intended by the test developer?"

The checklist also included several rules for creating sound multiple-choice questions. Using those rules as a guideline, Libarkin analyzed a question from the first version of the GCI. She noted that the question violated several of the rules and explained how the development team revised it to be more consistent with the rules.

Observing that concept inventories serve several purposes, Libarkin explained that the importance of the question quality varies with the purpose. For example, if the purpose is to document alternative conceptions to "wake up" faculty, the style of the questions might not matter. The question

format matters more if the purpose is to evaluate learning for instruction, and it is very important if the purpose of the concept inventory is to assess learning for research.

## CONCEPT INVENTORIES IN ENGINEERING

Teri Reed-Rhoads (Purdue University) observed that although engineering lags behind science in terms of developing concept inventories, the few engineering concept inventories available are increasingly being used for such purposes as accreditation, grant proposals, and grant project accountability. In addition, she explained that engineering faculty members are beginning to use concept inventories to facilitate changes in pedagogy aimed at increasing student learning.

Reed-Rhoads defined engineering concept inventories as those that are developed by engineers, either on their own or in collaboration with others. Using this definition, Reed-Rhoads identified 21 engineering concept inventories, 6 of which she labeled as science, technology, engineering, and mathematics (STEM) concept inventories, which were developed by or in conjunction with engineers and focused on nonengineering-related subjects.<sup>1</sup>

Discussing the relative maturity of engineering concept inventories, Reed-Rhoads pointed out that many more examinees have taken the statics concept inventory than the other engineering-related concept inventories, and that its growth has been exponential. For example, between year 2 and year 3 of its existence, the cumulative number of examinees for the statics inventory jumped from about 300 to about 1,700, further increasing to 2,700 in year 4 (Reed-Rhoads and Imbrie, 2008). In contrast, the cumulative number of examinees for the systems and signals inventory steadily grew from about 300 in year 1 to about 500 in year 2 to slightly more than 800 in year 3. She also explained that, because concept inventories take years to develop (as noted by Libarkin), there is often a significant lag time between their development and a discernible effect on instructional practices.

In engineering, concept inventory developers initially were slow to analyze the psychometric properties of engineering concept inventories, said Reed-Rhoads. She observed, however, that developers are increasingly collaborating with psychometricians to analyze and validate their instruments. She also noted that the research base on students' engineering misconceptions is lagging behind those in some of the other sciences. This lag complicates the development of the concept inventories; in other disciplines

---

<sup>1</sup>The specific concept inventories are listed in the workshop paper by Reed-Rhoads (see [http://www.nationalacademies.org/bose/Reed\\_Rhoads\\_CommissionedPaper.pdf](http://www.nationalacademies.org/bose/Reed_Rhoads_CommissionedPaper.pdf)).

the inventory developers draw on existing research about misconceptions, whereas in engineering, the concept inventories drive the definitions of the misconceptions (Reed-Rhoads and Imbrie, 2008).

Reed-Rhoads identified gaps in the research related to engineering concept inventories. First, she explained that concept inventories have been used only in the basic engineering courses so far, which means that upper division courses and subject areas are sparsely represented. In addition, although some research indicates that examinees' attitudes and beliefs about a field of study might influence assessment results in that field (Gal and Ginsburg, 1994), few of the engineering concept inventories have related instruments that measure the affective and cognitive domains.

Another gap in the research is that engineering concept inventories have not been extensively studied for the various types of bias that might be included in the questions (Reed-Rhoads and Imbrie, 2008). These biases include how gender, race/ethnicity, native language, and culture might affect student scores on the inventories. The understanding of bias in engineering concept inventories is limited because not enough students from different subpopulations have used the instrument; with such low sample numbers, the statistics for each subgroup are not reliable. However, Reed-Rhoads noted that although women are the most underrepresented population in engineering, enough women have used the concept inventories to allow for some statistical testing related to gender bias.

Reed-Rhoads also observed that the relationships among concept inventories is important but not well understood. She emphasized the need to track students' conceptual development, which requires greater knowledge of how the concept inventories fit together. She argued that this need is becoming increasingly important as concept inventories proliferate.

The final gap relates to helping faculty members use concept inventories to change their practices. To this end, Reed-Rhoads and her colleagues created a community of inventory developers, faculty members, and students called ciHUB (short for concept inventory hub) to provide access to resources that can facilitate collaboration and the use of research-based tools to improve instruction.

### IDENTIFYING AND ADDRESSING STUDENT DIFFICULTIES IN PHYSICS

Karen Cummings delivered a presentation by Paula Heron (University of Washington) on work by Heron and her colleagues in the University of Washington's Physics Education Group.<sup>2</sup> This group conducts a coordi-

---

<sup>2</sup>See the workshop paper by Heron (see [http://www.nationalacademies.org/bose/Heron\\_CommissionedPaper.pdf](http://www.nationalacademies.org/bose/Heron_CommissionedPaper.pdf)).



nated program in which research, curriculum development, and instruction are tightly linked in an iterative cycle. One of the group's major curriculum development projects, *Tutorials in Introductory Physics* (McDermott, Shaffer, and the Physics Education Group at the University of Washington, 2002), was the focus of the presentation.

Cummings explained that the Physics Education Group developed the tutorials to supplement instruction in an introductory, calculus-based physics course at the University of Washington that is required for all physics majors. Approximately 1,000 students are enrolled in the course at any time. The course meets for three 50-minute classes and one 3-hour laboratory each week. Each course also has a 50-minute tutorial each week, and students have weekly online homework that is linked to the lecture material. They also are assessed through three mid-term exams and a final exam that contain material from the lectures, labs, and tutorials. Because the course is similar in structure and content to many others in colleges and universities throughout the United States, the setting is well suited for the development and assessment of instructional materials that can be adopted at other institutions.

In the weekly tutorials, students work through carefully structured worksheets in small groups, and instructors question them in a semi-Socratic manner. Designed to fit within the constraints imposed by large lecture-based courses, the research-based tutorials foster the development of reasoning skills and conceptual understanding.

Tutorial development depends on systematic investigations of student learning at the beginning of, during, and after instruction, including ongoing individual student interviews to probe their understanding in depth (Heron, Shaffer, and McDermott, 2008). Based on those interviews, the researchers write open-ended questions to ascertain the prevalence of specific difficulties. They also conduct descriptive studies in the classroom to further inform the development of their curriculum materials.

These tutorials have been assessed extensively at the University of Washington and at many of the dozens of institutions that have adopted them. At the University of Washington, students who completed the tutorial were given a posttest with questions that could not be answered by memorization. Eighty percent of the students gave a correct or nearly correct answer (compared with 20 percent without the tutorial) (Heron, Shaffer, and McDermott, 2008). Results from other institutions that have used the University of Washington tutorials include the following:

- Learning gains in introductory physics courses that used tutorials at the University of Colorado were much higher than is typical in introductory courses (Finkelstein and Pollack, 2005).
- At Montana State University, a longitudinal study showed that nonmajors retained gains they made in understanding force—as

measured by the Force Concept Inventory (FCI)—up to 3 years after completing an introductory physics that used the tutorials (Francis, Adams, and Noonan, 1998).

- In Harvard University physics classes that used a variety of interactive strategies—including the University of Washington tutorials—the gender gap between the FCI scores of male and female students disappeared (Lorenzo, Crouch, and Mazur, 2006).
- After a large introductory physics course at the University of Colorado that used tutorials, Finkelstein and Pollack (2005) did not observe the shift toward unfavorable attitudes about physics that typically occurs in those courses.

Based on these results, Heron, Shaffer, and McDermott (2008) posited that additional assessments would be valuable in the areas of student reasoning skills, student ability to transfer conceptual knowledge to quantitative problems, and student ability to apply concepts and principles in subsequent courses.

At the workshop, Cummings characterized *Tutorials in Introductory Physics* as an example of how research can guide the improvement of instruction within the practical constraints of courses with large enrollments. She explained that the tutorials and other research-based instructional materials are most successful when the developers invest sustained effort in their continuous improvement and in supporting adopters. She ended by noting that the growth in STEM departments of groups and individuals who devote their scholarly effort to conducting research on teaching and learning in the science disciplines is the truly promising practice in STEM education (Heron, Shaffer, and McDermott, 2008).

## DISCUSSION

Before taking questions from the audience, the panelists reflected on each others' presentations. Cummings remarked about the dearth of published concept inventories in chemistry and noted that researchers in all disciplines would benefit from the information Libarkin and Reed-Rhoads presented about the process of developing concept inventories. Reed-Rhoads agreed that disseminating information about the development and appropriate use of concept inventories is important. She stressed the need for a “*Good Housekeeping* seal of approval” for concept inventories. She and Libarkin also discussed the need to warehouse and analyze the data collected from concept inventories. Libarkin added that she would like to see the disciplinary communities be trained to use and improve the tools.

David Mogk and William Wood expressed concerns about the inappropriate dissemination and use of concept inventories. In response, Libarkin

explained her view that concept inventories are useful as a snapshot of students' understanding of one or more targeted concepts, and that other assessment methods provide a deeper look at students' mental models. She agreed that it is important for concept inventories to be aligned with the assessment purpose. Reed-Rhoads expressed the view that widespread dissemination is beneficial as long as the authors of the concept inventories have access to the resulting data so they can improve the instrument.

Kenneth Heller pointed out that the FCI is not about forces and is not a concept inventory. Rather, it is an instrument about misconceptions that is based on the misconception research. Although the instrument is reliable, Heller stressed that it is not a predictor of students' success in introductory physics. He asked the presenters whether they are trying to replicate the success of the FCI or develop concept inventories that may or may not have the same properties as the FCI. Libarkin and Reed-Rhoads said their respective communities (geosciences and engineering) are trying to do both. Cummings agreed with Heller's assessment of the FCI and emphasized the importance of being clear about what these instruments measure.

Heidi Schweingruber asked the concept inventory developers to elaborate on the link between concept inventories and instructional change. Cummings responded that the University of Washington Physics Education Research Group gets feedback on strategies that work and do not work to foster conceptual understanding and uses that feedback to develop curriculum materials. The Physics Education Research Group works with professors who adopt the materials to ensure that they have the support they need to implement the materials effectively.

Responding to a question from Jay Labov (National Research Council), Libarkin and Cummings said that concept inventories do not measure whether students will have enough knowledge of science to make informed decisions later in their lives. Cummings added that this gap suggests a need for additional research and instrumentation.

## 6

## Structuring the Learning Environment

This session of the workshop focused on the role of learning environments in supporting science, technology, engineering, and mathematics (STEM) learning. Speakers presented different approaches to addressing the challenges that large introductory courses can pose to students' academic success. These approaches include a variety of strategies to make large classes more interactive, as well as programs to engage undergraduate students in research experiences.

### STUDIO COURSES

Karen Cummings discussed a studio physics course at Rensselaer Polytechnic Institute (RPI). Like studio art, studio physics involves learning by doing. Consequently, studio instruction is a whole-course modification that involves collaborative, hands-on learning in specially designed classrooms. The focus on hands-on activities requires longer class periods than is typical in introductory physics courses; at RPI, the studio courses meet twice a week for 2 hours per session. Instructors in studio physics courses use technology in various ways to maximize instructional time and to improve learning outcomes.

Studio physics has its origins in the work of the University of Washington's Physics Education Research Group (see Chapter 5), which gave rise to the development of *Workshop Physics* at Dickenson College, a calculus-based physics course with a published curriculum (Jackson, Laws, and Franklin, 2003) that is taught without lectures. According to Cummings, studio physics is a more efficient model of *Workshop Physics*, and it differs from *Workshop*

*Physics* because it is not a curriculum. Instead, it is a pedagogical approach and a classroom structure.

The first studio physics course was established at RPI in 1993. By 2008, all introductory physics courses at RPI were studio courses. Cummings said that there are 15 to 20 sections of studio physics at RPI every semester, and each section contains approximately 50 students.

In one evaluation, Cummings compared a traditional lecture course with two forms of the studio course, one of which incorporated interactive lecture demonstrations and cooperative problem solving that was shown to be effective in previous research. Studying 10 sections of approximately 50 students each, she used student surveys, students' formal course evaluations, and validated instruments to measure conceptual learning outcomes and attitudinal outcomes. The students were divided into two groups: standard studio and "studio plus" (the studio that incorporated the lecture demonstrations and cooperative problem solving). Both groups did the same homework, saw the same lectures, took the same exams, and had the same classrooms. The only difference was that studio plus incorporated research-based curricular materials.

The standard studio course was more efficient than the traditional lecture because lecture and laboratory time was combined, but no more effective in terms of learning outcomes (Cummings, 2008). When instructors incorporated research-based curricular materials, however, students at all levels made significant improvements on the force concept inventory and its associated attitudinal survey. In Cummings's view, these data suggest that the studio format alone is not sufficient to improve students' conceptual understanding.

Cummings also described an introductory biology course at RPI that blends a studio-style course with a web-based learning activity that students can pursue outside the time and space constraints of the classroom (asynchronous learning). To evaluate this course, McDaniel and colleagues (2007) administered a survey that assessed knowledge of biological concepts to students in a standard lecture course and the studio course with the asynchronous component. They measured normalized gains, or the ratio of how much students learned compared with how much room they had to learn based on their pretest scores.<sup>1</sup> Students in the studio course performed significantly better in ecology and evolution than students in the traditional biology lecture course.

Studio courses are expensive to implement. As a result, instructors at many institutions are implementing less expensive hybrid models. With these

---

<sup>1</sup>As defined by Hake (1998), normalized gain = (posttest – pretest)/(100 – pretest). For example, students who score 80 on the pretest and 90 on the posttest gain only 10 percentage points, but those 10 percentage points represent half of what they did not know.

models, instructors increase interactivity during lectures and create tight links between lecture materials and laboratory activities without modifying the classroom space or schedule. Based on her research, Cummings said that if instructors use physics education research–based materials in these hybrid models, students’ conceptual understanding can improve significantly.

Although hybrid models can yield appreciable gains in conceptual understanding, the appeal of the studio model lies in its ability to promote other skills. With studio models, Cummings said, students are more responsible for their own learning and develop lifelong learning skills. For example, they are required to communicate about scientific content and their intentions in applying the scientific method. They also must work efficiently in groups that they did not select, which mirrors many work environments. For Cummings, these potential gains raise the question of “What can the studio environment be proven to do that less expensive models and implementation cannot?”

## REDESIGNING LARGE CLASSES FOR LEARNING

### Project SCALE-UP

Robert Beichner (North Carolina State University) discussed the SCALE-UP (Student-Centered Active Learning Environment for Undergraduate Programs) project, which aims to restructure classes with large enrollments following the studio model (see <http://www.ncsu.edu/per/scaleup.html>). More than 50 colleges and universities have adapted the SCALE-UP approach in physics, chemistry, mathematics, engineering, and literature courses. Although the implementation of SCALE-UP varies by institution, its central feature is a redesigned learning environment to facilitate collaborative, hands-on learning and interaction among students and instructors. SCALE-UP classrooms typically have round tables with an instructor station in the middle of the room, and some contain whiteboards, public thinking spaces, and storage facilities that make equipment accessible to students. Students in SCALE-UP courses are formally assigned to mixed-ability groups of students who sit at the round tables, and each table has several networked laptops.

Similar to the studio approach described in the previous section, a typical SCALE-UP class meets for five or six hours a week, combining lecture time and lab time. Classes often begin with a short lecture to set the stage for the day’s activities and relate them to the previous class. Students spend the remainder of the time in activities called “tangibles,” which are hands-on observations or measurements; “ponderables,” which are complex, real-world questions; and simulations. Classes typically end with a whole-group follow-up discussion and a brief summary lecture.

Data from several institutions show that across all performance levels (top, middle, and bottom of the class), students in SCALE-UP (studio) physics courses made greater normalized gains in their conceptual knowledge than students in lecture courses. Students in the top third of the class made greater normalized gains than students in the middle or bottom third. These gains were particularly pronounced at the Massachusetts Institute of Technology, where students learn the content by teaching each other. Similarly, in a study on physics problem solving at North Carolina State University, students in the SCALE-UP course outperformed their peers on eight of the nine exam questions; the SCALE-UP course had not yet covered the content of the ninth question (Beichner, 2008a).

Researchers at the institutions that are implementing SCALE-UP have also studied the program's effect on other outcomes. For example:

- Attendance at North Carolina State University is not required, yet attendance in SCALE-UP courses there averages 93 percent (Beichner et al., 2007).
- At Florida International University, the drop, failure, withdraw (DFW) rate for studio-based courses is one-fourth the rate for traditional courses. Enrollment requests for those courses exceed capacity by roughly four times, and faculty and student evaluations of the courses are overwhelmingly positive. After taking the course, 10-20 percent of the students pursue physics majors or minors (Kramer, Brewes, and O'Brien, 2008).
- Clemson uses a SCALE-UP model for all introductory math courses. DFW rates in those courses have dropped from 44 to 22 percent (Biggers et al., 2007).
- A 5-year study with 16,000 students at North Carolina State University showed that failure rates are significantly lower for students in SCALE-UP courses than for students in traditional courses, even though course requirements for SCALE-UP are more rigorous (Beichner, 2008a).
- Female students in SCALE-UP courses at Pennsylvania State University, Erie, had significantly lower pretest scores in a variety of mathematics and science areas. By the end of the semester, their grades were the same as males (Beichner, 2008b).
- At North Carolina State University, students with SAT mathematics scores of less than 500 fail an advanced engineering course 17 percent of the time if they take an introductory SCALE-UP physics course as the prerequisite. If their introductory course is lecture-based, they fail the later course 31 percent of the time (Beichner, 2008b).

Beichner summarized some of the objectives that he and his colleagues have measured for SCALE-UP physics and the methods they have used to assess those objectives (see Table 6-1).

Based on his research and experiences with SCALE-UP, Beichner identified three issues that warrant further study. First, he explained that in many team settings the input of underrepresented groups is devalued. This phenomenon does not occur in SCALE-UP courses, and it is important to understand why. Second, he called for research on the factors that influence the adoption of these reforms, similar to the work of Henderson and Dancy (2007, 2008a; see Chapter 8, this volume). Finally, he said that a large-scale, international study of SCALE-UP implementation would shed light on how it varies from generation to generation and the effect of different implementations on learning outcomes and affective outcomes.

### Online Problem-Based Learning Case Discussions

Marcy Osgood (University of New Mexico) presented the work that she and her colleagues have done to redesign large, introductory biochemistry courses. She explained that the University of New Mexico is a large university with a racially and ethnically diverse student body. Many students come from rural high schools, and many are older students returning to college after serving in the military.

Osgood teaches in the Department of Biochemistry in the School of Medicine. Because her department administers an undergraduate major in the School of Arts and Sciences, she and her colleagues frequently work with undergraduate students. Drawing on their experiences at the medical school, in the early 2000s they transformed a large undergraduate biochemistry class into a hybrid class with lectures and small-group, problem-, and

**TABLE 6-1** Summary of Objectives and Assessment Methods for SCALE-UP Physics

Objective	Assessment Method
Conceptual Understanding	Pre-posttests, interviews, portfolios
Problem Solving	Comparison tests, interviews, portfolios
Laboratory	Practical testing, portfolios
Technology	In-class observations, practical testing, portfolios
Communication	In-class observations, video recording, interviews
Attitudes	Maryland Physics Expectations Survey, interviews, in-class observations
Positive Learning Experience	Course evaluations, interviews, focus groups

SOURCE: Beichner (2008b). Reprinted with permission.



case-based tutorials. They adopted this approach because they knew that problem-based cases effectively engage students in the content, and they believed that an interactive approach would provide more opportunities for diverse students to excel.

After two years of implementing problem-based learning in large biochemistry classes, Osgood and her colleagues found that students in those classes performed better on content-based exams than students in traditional courses. However, the approach was time- and space-intensive, so they shifted to an online format.

In the online format, groups of approximately six to eight students use the scientific method to solve vague problems posed by the instructors. Through iterative postings on the course project's website over a period of weeks, the groups develop hypotheses about what the problem means, develop an approach to solving the problem, and design experiments to investigate the problem. Instructors provide students with data based on their experimental design. Students integrate their data analysis with the course content, reflect on what they have learned, and identify how they might further address the problem.

With the online approach, Osgood can proctor 10 small groups at once, as opposed to proctoring one face-to-face group at a time. She has developed a rubric to grade student postings for content and evaluate group dynamics and progress; the rubric allows her to grade 10 groups in approximately one hour per day. She converts the rubrics to bar graphs that illustrate the groups' progress (see Figure 6-1). The first group in Figure 6-1 was successful because the graph shows a steady increase in the content in the students' postings, and the group had nearly 80 postings in a two-week period. In the end, that group solved the case. The second group was slower to start and eventually failed the assignment.

In addition to evaluating the whole group, Osgood and her colleagues use the rubrics to analyze individual student contributions. On the basis of these analyses, they have developed a typology of students. As Osgood explained, one category is the "serial" or "shotgun investigator." These students conduct all possible tests without checking the results, considering cost-benefit analysis, or asking their colleagues what might be happening. "Summarizers" constitute a second category. As the name suggests, these students summarize the results of their colleagues' experiments and identify the next steps without conducting any experiments of their own. The third category is "the lonely scientist." Students in this category conduct all of the steps themselves and typically are the only ones posting to their groups. A final category is the "beginning expert," who understands the concepts, integrates the methods and content appropriately, and brings the rest of the class along with him or her in the understanding of the problem.

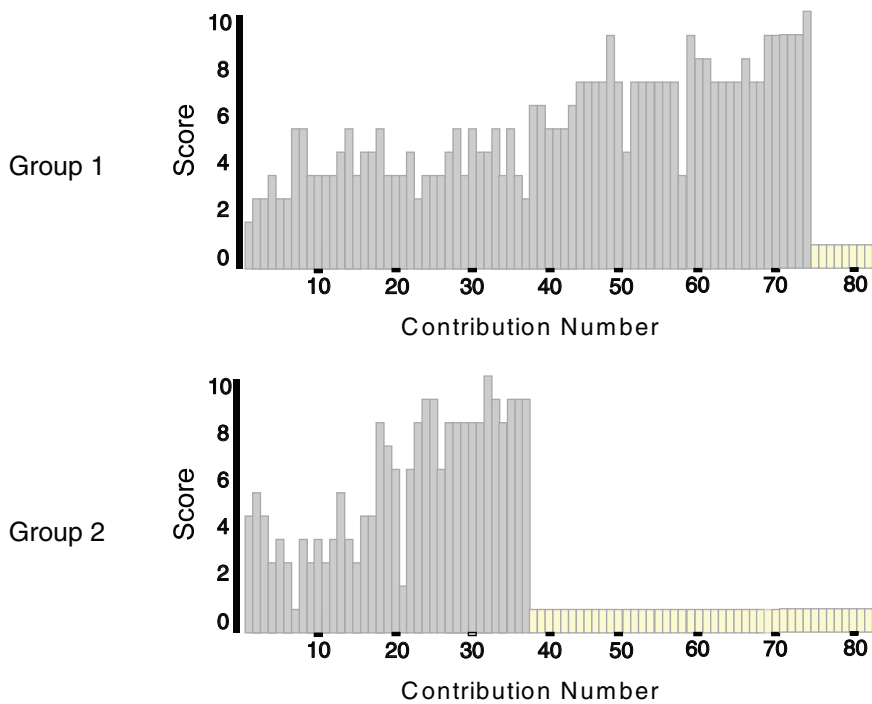


FIGURE 6-1 Graph for assessing the progress and dynamics of small, online groups in a large biochemistry course.

SOURCE: Osgood, Mitchell, and Anderson (2008). Reprinted with permission.

Osgood believes that understanding the relationship between students' practice in groups and their practice as individuals will help instructors to offer assistance that targets students' specific needs.

### Active Learning Strategies for Introductory Geology Courses

David McConnell (North Carolina State University) discussed his efforts to redesign introductory geology courses at the University of Akron. With colleagues in cognitive psychology and science education, McConnell sought to

- determine if students are prepared to use higher order thinking skills;
- teach an introductory course for nonmajors in which students improve their higher order thinking skills and conceptual understanding; and

- identify strategies that others can use to assess ongoing student learning.

To address the first goal, McConnell and his colleagues assessed 741 introductory geology students using the Group Assessment of Logical Thinking (GALT) (Roadranga, Yeany, and Padilla, 1982, 1983), a 12-item test in which students answer questions and explain why they answered the way they did. Success on the GALT requires competence in proportional reasoning, controlling variables, combinational reasoning, probabilistic reasoning, and correlational reasoning. On the basis of their results, students are placed on a continuum of ability to think abstractly. Concrete thinkers who prefer a fact-based approach and rely on memorization are at the low end of the continuum, and abstract thinkers who can understand previously unseen ideas are at the high end. Transitional thinkers, who prefer to apply ideas in a practical way, fall in the middle of the continuum.

A total of 43 percent of the University of Akron students were classified as capable of broad, abstract thought in physical science, based on their GALT scores (McConnell, 2008). The remaining 57 percent required support to grasp abstract concepts. As a result, McConnell and his colleagues sought to design learning environments that would foster students' ability to grasp abstract information, including the concrete and transitional thinkers who required additional support in this area. Drawing on similar work at other institutions, they divided lectures into small segments, assigned students to work together in groups, and used formative assessments during class to determine student understanding and progress. Because McConnell and his colleagues implemented these changes in the context of aging lecture halls in which the seats are bolted down and closely spaced, he believes that this implementation is one of the least expensive permutations of redesigning a large learning environment.

At several points during each lecture, instructors gave students a variety of opportunities for collaborative learning. According to McConnell, these exercises targeted different levels of Bloom's taxonomy. For example, the tasks required students to confront their preconceptions, allowed them to reflect on their understanding of key concepts, linked information to previous knowledge, and asked questions requiring the use of a range of thinking skills. Other course activities ranged from assigned reading and homework, to concept tests (asking and answering questions among peers), to graphical work products (concept maps, Venn diagrams) that demonstrated analysis and conceptual understanding.

Discussing the results of in-class assessments, McConnell explained that after three days of lecturing, fewer than half of students responded correctly to a question about the number of tectonic plates. After discussing

the topic in groups, 75 percent of students answered the question correctly on a retest. When instructors used rudimentary models rather than standard lecture to introduce plate tectonics, 56 percent of students answered the question correctly the first time, and 84 percent answered correctly after discussion in groups.

McConnell also shared results from a study he and others conducted in his classes about the use of models to explain the seasons (McConnell et al., 2005). Students in two control classes learned about the seasons through standard lecture with some demonstration. Students in six experimental classes received rudimentary models—a foam ball on a skewer with a small flashlight—and instructions about how to model different scenarios related to the seasons. Students in the experimental classes had favorable views about using the models and showed greater gains in their conceptual understanding of the seasons than students in the control classes. In addition, students in the experimental classes made greater gains in their logical thinking skills as measured by the GALT (McConnell et al., 2005).

### DOING SCIENCE: PROVIDING RESEARCH EXPERIENCES

Another way to address the challenges that large introductory classes can pose to academic success is to engage students in research. Research experiences allow students to work directly with, and learn from, individual science faculty. Noting that the best way to learn science is by doing science, committee member David Mogk introduced speakers to discuss two programs that provide research experiences for undergraduate students.

#### University of Michigan Undergraduate Research Opportunity Program

Sandra Gregerman (University of Michigan) discussed the Undergraduate Research Opportunity Program (UROP), which was launched in 1988 to increase the retention and academic success of underrepresented minority students at the University of Michigan. In this year-long program, first- and second-year students spend 6-12 hours per week conducting research on ongoing faculty projects in the sciences and other disciplines. The program contains academic and social support components, including peer advising, skill-building workshops, and research peer groups in which students discuss a variety of research-related issues. Each year, the program culminates in a symposium; in 2008, 750 students presented their research in poster form and 20 students delivered oral presentations on their research (Gregerman, 2008).

Gregerman and her colleagues have conducted many studies of the program over the years. Results of one longitudinal study with an experi-

mental design<sup>2</sup> show that participating in UROP increases retention rates for some students. For example, 75 percent of African American men who participate complete their degrees, compared with 56 percent who do not participate (Gregerman, 2008). To better understand these results, evaluators conducted interviews and focus groups with students in the experimental and control groups. In those interviews, UROP students were more likely than students in the control group to mention that faculty members and graduate students cared about their success and to discuss the possibility of graduate school. They also were more likely than students in the control group to report going to faculty members' office hours and seeking help from someone in their network instead of the library. A survey of alumni revealed that UROP participants also were significantly more likely to attend graduate or professional school (82 versus 56 percent of nonparticipants).

### Center for Authentic Science Practice in Education

Gabriela Weaver (Purdue University) and Donald Wink (University of Illinois, Chicago) discussed their work with the Center for Authentic Science Practice in Education (CASPiE), a multi-institutional partnership to increase student retention in the sciences through authentic research experiences. The partner institutions include a wide range of 2- and 4-year colleges and universities (see <http://www.purdue.edu/discoverypark/caspie/partners.html>). These partners have developed a model in which first- and second-year science students participate in faculty research projects as part of their regular coursework. Undergraduate research experiences through CASPiE include skill-building workshops, access to sophisticated research equipment, guidance and mentoring from faculty, and opportunities for peer networking and support.

Evaluation results indicate that CASPiE participants learn chemistry as well as nonparticipants and are more likely to perceive their labs as authentic and relevant to the future (Wink and Weaver, 2008). Evaluation data also suggest that CASPiE students increase their ability to communicate the meaning of their work, despite the absence of prescribed steps in their lab manuals.<sup>3</sup>

---

<sup>2</sup>In this study, researchers matched program applicants on the basis of demographic and academic characteristics and randomly accepted every other applicant. Students who were accepted to the program constituted the experimental group and those who were not selected represented the control group (Gregerman, 2008).

<sup>3</sup>For more detail about the evaluation methods and results, see the workshop paper by Wink and Weaver ([http://www7.nationalacademies.org/bose/Wink\\_Weaver\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Wink_Weaver_CommissionedPaper.pdf)).

### SMALL-GROUP DISCUSSIONS

In small groups, participants discussed the day's presentations. The following points emerged during the summaries of those discussions:

- Systemic reform is difficult and takes time. The research base is more developed than it was 10 years ago, but practice has not changed on a broad scale. Gaps in the evidence still exist, and evidence alone is not sufficient to drive change.
- The evidence suggests that teaching methods matter and that some instructional strategies are more effective than others. For example, active, cooperative learning seems to work in different contexts.
- The research does not fully illustrate why certain practices work, for which students, and in which contexts. Additional gaps include research on the affective domain, instructor effects (implementation of the promising practice, relationship with students, and belief in students' abilities), the effect of culture, students' social construction of knowledge, the expert-novice continuum, departmental and institutional change, and cost-benefit analyses.
- Dissemination of promising practices could be more effective. The disparate pieces have not been pulled together into a coherent whole.
- The learning goals of a particular promising practice should determine what evidence and methods are required to determine its effectiveness.
- Different stakeholders—students, faculty, administrators, industry—have different standards of evidence and different metrics for success.

## 7

## Faculty Professional Development

If the promising practices described in the previous chapters are to take hold, then faculty members require support and training to implement them. This chapter describes several efforts to provide faculty members with professional development that is targeted at reformed instruction. These efforts span the continuum from future faculty to new faculty to veteran faculty.

### PROFESSIONAL DEVELOPMENT OF FUTURE FACULTY

Donald Gillian-Daniel (University of Wisconsin, Madison) discussed the issue of professional development for future faculty members—that is, graduate students and postdoctoral students in science, technology, engineering, and mathematics (STEM). He described the Delta Program in Research, Teaching, and Learning at the University of Wisconsin, Madison, which is designed to help current and future faculty succeed in the changing landscape of science, engineering, and mathematics higher education (see <http://www.delta.wisc.edu/index.html>).

The Delta Program is a prototype of the Center for the Integration of Research, Teaching, and Learning, which seeks to develop and advance effective teaching practices (see <http://www.cirtl.net>). According to Gillian-Daniel, Delta and programs like it have three aims: (1) to improve undergraduates' learning by better preparing the faculty who will teach them, (2) to prepare future faculty for the demands of their jobs, and (3) to change the culture of graduate education.

Delta is based on three core ideas: (1) teaching as research, (2) learning communities, and (3) learning through diversity. Teaching as research is the

idea that graduate students can apply disciplinary research skills to address questions about teaching and student learning in their classroom. Learning communities bring individuals together across disciplinary and generational boundaries to create and share knowledge. Learning through diversity is grounded in the view that each individual's background enriches the learning environment. Gillian-Daniel hypothesized that the combination of these elements is crucial to the Delta program's effectiveness.

Gillian-Daniel presented two examples to illustrate the Delta Program's impact on teaching and learning.<sup>1</sup> The first example addressed the effect of improved teaching on student learning. In that study, a Delta Program alumnus and his colleagues examined whether the combination of a multimedia learning object, lectures, and laboratory improved student learning about fuel cells (Lux et al., 2007). The researchers assessed the effect of the learning object with pre- and post-quizzes and used a web-based questionnaire to elicit student opinions about the value of the different course components. Correct responses on the quizzes increased from 42 percent in the pretest to 80 percent after the instructors introduced the learning object. In addition, 100 percent of the students in the laboratory were able to create a functional fuel cell (Lux et al., 2007).

The second example focused on the development of skills and pedagogical techniques in faculty members. In this example, a Delta Program alumna examined whether students who were taught with active learning strategies changed their views about such strategies in their own teaching (McNeil and Ogle, 2008). The researchers developed a seminar course that required students to prepare a 45-minute lecture on a topic in their discipline that incorporated one or more active learning techniques. Pre-post course evaluations included questions such as "If you were preparing a lecture, list the steps that you would go through." After the course, students reported that they would take more steps to prepare for a lecture, including ones related to integrating active learning components (McNeil and Ogle, 2008).

Discussing gaps in the research, Gillian-Daniel cited the need for longitudinal studies to understand how professional development programs for future faculty affect their teaching practice throughout their careers. In a related vein, he called for longitudinal studies to examine how reformed teaching in introductory courses affects undergraduate students over the course of their college careers. He also stressed the importance of identifying the effective elements of existing programs, which would involve developing common metrics or benchmarks to measure program outcomes.

---

<sup>1</sup>For additional examples of the Delta program's effectiveness, see the workshop paper by Gillian-Daniel (see [http://www.nationalacademies.org/bose/Gillian\\_Daniel\\_Commissioned\\_Paper.pdf](http://www.nationalacademies.org/bose/Gillian_Daniel_Commissioned_Paper.pdf)).



Finally, he said it would be useful to create a repository of instruments and data on various promising practices for researchers to use.

### WORKSHOPS BY A PROFESSIONAL SOCIETY FOR NEW PHYSICS FACULTY

Ken Krane (Oregon State University) discussed the New Faculty Workshop in Physics and Astronomy, which he and his colleagues have been running since 1996. With financial support from the National Science Foundation, the workshop is sponsored by the American Association of Physics Teachers in partnership with the American Physical Society and the American Astronomical Society.

Krane and his colleagues developed the workshop to improve physics teaching at research universities, which they defined as any institution that awards an M.S. or a Ph.D. in physics. These institutions represent a high leverage point to affect teaching because they enroll the vast majority of students in introductory physics, produce the majority of physics majors, and hire the majority of physics faculty.

The New Faculty Workshop is an annual event. Over the course of 3 days, Krane explained, workshop developers seek to provide a coherent and interconnected set of paradigms for improving instruction. The workshops also promote research-based reforms that new faculty can adopt with minimal time commitment and minimal risk to their tenure status, according to Krane. Small-group and plenary sessions offer opportunities for new faculty to connect with innovators in physics education and physics education research and to form their own communities of practice as they implement effective teaching strategies.<sup>2</sup>

Krane and his colleagues measure the workshop's success in terms of the following three goals:

1. Involve a significant fraction of the newly hired faculty in physics and astronomy.
2. Familiarize participants with recent and successful pedagogic developments.
3. Effect an improvement in physics and astronomy teaching when new pedagogies are implemented at home institutions.

Addressing these goals, Krane reported results from an evaluation of the program by Charles Henderson (2008). Henderson found that the

---

<sup>2</sup>For more information, see the workshop paper by Krane (see [http://www.nationalacademies.org/bose/Krane\\_CommissionedPaper.pdf](http://www.nationalacademies.org/bose/Krane_CommissionedPaper.pdf)) and the New Faculty Workshop home page (see <http://www.aapt.org/Conferences/newfaculty/nfw.cfm>).

workshop involves 20 to 25 percent of all the new hires in physics and astronomy. In addition, a survey of participants revealed the following (Henderson, 2008):

- 94 percent of current participants reported the desire to incorporate new ideas from the workshop into their teaching.
- 70 percent of former participants rate their teaching as more innovative than their colleagues' teaching.
- 96 percent report changes in teaching methods since attending the workshop, and 40-60 percent of those indicate most or all of the changes are a direct result of workshop participation.

Krane (2008) also shared the following testimonial from a department chair at one of the institutions that sends a large number of participants to the New Faculty Workshop:

As a department chair, I believe that these workshops are more effective than I could ever be at convincing new professors that both the teaching and research they do will be recognized by their profession. . . . I believe the workshops have helped change the culture at [university] to place greater value on excellent physics teaching. Our younger faculty have come to believe this with an enthusiasm with which they are gradually infecting the entire faculty of my Department. I offer, as an indication of the progress which a dedicated cadre of faculty can achieve, the statistic that the number of physics majors graduated at [university] last spring was the largest in at least two decades. The improvement is not a statistical fluctuation, and represents a thorough reversal of the depressing decline in the number of majors at [university] through the 80s and 90s.

Three factors have contributed to the workshop's success in the physics community, according to Krane. First, introductory physics courses across the country are remarkably similar, with similar challenges and approaches to addressing those challenges. As a result, a well established set of best practices exists around active engagement in physics classrooms. Second, the small size of the physics community means that one workshop can reach a significant portion of new faculty each year. Finally, Krane credited much of the workshop's success to strong support from the physics professional societies. In particular, the backing of the research-based professional societies has enhanced the workshop's credibility at the research universities, making department heads more likely to support faculty participation.

## CHANGING INSTRUCTION

### Rethinking Professional Development in Undergraduate STEM Education

Diane Ebert-May (Michigan State University) discussed her evaluations of two established faculty professional development programs: the NSF-funded Faculty Institutes for Reforming Science Teaching (FIRST) project and the National Academies' Summer Institutes, funded by the Howard Hughes Foundation.<sup>3</sup> The evaluations are guided by three research questions.

1. Do faculties change in response to professional development?
2. Are those changes in teaching sustained over time?
3. What factors contribute to the change pedagogy?

Of the 134 workshop participants in the institutes, 75 were involved in the evaluation study. The numbers of tenured and nontenured faculty were roughly equal, and 56 percent of study participants were female. Although most study participants were teaching at R1 institutions (institutions that focus primarily on research), Ebert-May said the study also included faculty from a variety of 2- and 4-year colleges and universities.<sup>4</sup>

Evaluators used the Reformed Teaching Observation Protocol (RTOP) to rate participants' videotaped lessons shortly after the institutes and again up to 2 years later. Developed by Evaluation Facilitation Group of the Arizona Collaborative for Excellence in the Preparation of Teachers, the RTOP is designed to determine the extent to which instructors are using reformed teaching in undergraduate science and mathematics courses (Piburn et al., 2000).

Ebert-May discussed five categories of teaching addressed by the RTOP, which represent a continuum from teacher-centered to student-centered activities. As she explained, category I is pure lecture; category II is lecture with some demonstration and minor student participation; category III involves significant student engagement with some minds-on and hands-on involvement; category IV includes active student participation in the critique and in carrying out experiments; and category V constitutes active student involvement in open-ended inquiry resulting in alternative hypotheses, several explanations, and critical reflection.

In Ebert-May's evaluations, the majority of instructors fell into catego-

---

<sup>3</sup>For more detailed information about the FIRST workshops, see <https://www.msu.edu/~first4/Index.html>. For more information about the National Academies Summer Institutes, see <http://www.academiessummerinstitute.org/>.

<sup>4</sup>Research universities 1 (R1) offer a full range of baccalaureate programs and give high priority to research.

ries I and II. More than half of all study participants did not change their practice from the first videotaped lesson to the next; 25 percent of instructors in categories I and II moved toward more learner-centered strategies from the first lesson to the next; and 15 percent of instructors who started in the more learner-centered categories moved toward more instructor-centered practices over time.

Multivariate analyses of these data showed that years of teaching experience and class size influence RTOP scores. For example, instructors with more teaching experience were less likely to engage with students and have them work in cooperative groups, leading to lower RTOP scores. In addition, larger class sizes were associated with lower RTOP scores (i.e., scores that involve more lecture) (Ebert-May, 2000). However, these and other variables explained only 25 percent of the variation in RTOP scores, leaving 75 percent of the variation unexplained. In Ebert-May's view, additional research is required to better understand why teaching varies.

### Addressing Disciplinary and Institutional Culture

Cathy Manduca (Carleton College) spoke about her work with professional societies and at the departmental level to improve instruction in the geosciences. Data from the geosciences, she explained, indicate that faculty attend professional development workshops, learn new ideas there, and subsequently change their practice. Despite the success of professional development efforts, however, the geosciences community is frustrated that change is not happening quickly enough.

In Manduca's view, it is possible to understand the change process by examining the cultures in which faculty members operate. She posited that faculty live in two different cultures—a disciplinary community, which emphasizes scientific research, and a broader institutional community, which is focused on the education enterprise. These cultures exert a strong influence on the extent to which faculty members change their teaching practice.

Discussing her work with professional societies, Manduca explained that uninformed faculty are at one end of the spectrum and those who actively research the impact of specific curriculum changes are at the other end. Informed faculty who make use of the research and observe how their teaching affects student learning are in the middle. Manduca's efforts focus on disseminating information to increase the number of informed faculty. In contrast to other presenters at the workshop, she said that evidence alone is sufficient for geosciences faculty to change their practice.

Journal articles and meetings of professional societies, such as the American Geophysical Union, represent one vehicle for disseminating research and best practices to the geosciences community. On the Cutting

Edge, a project of the National Association of Geoscience Teachers, is another important mechanism to help faculty stay abreast of geosciences research and teaching methods. According to the website (see <http://serc.carleton.edu/NAGTWorkshops/about.html>):

The workshop series and website combine to provide professional development opportunities, resources, and opportunities for faculty to interact online and in person with colleagues around the world who are focused on improving their teaching. An integral aspect of the project is development of an expanding community of geoscience educators with a strong and diverse leadership.

In all, 20 percent of geosciences faculty in the United States have participated in On the Cutting Edge, and 46 percent know about the program (Manduca, 2008a). Faculty from a wide variety of institutions, including R1 institutions, participate. Manduca said the workshop has legitimized teaching as a topic of discussion, oriented disciplinary research networks toward education, and created a culture of sharing information and resources.

Given that geosciences faculty turn to their colleagues for information on teaching, Manduca explained that departments are the most proximal source of support or discouragement for changes in practice. Departments are also important leverage points because they sit at the intersection of the institutional and disciplinary cultures described above. Acknowledging the importance of departments, Manduca described the Building Strong Geoscience Departments Program, which is designed to strengthen discussions of departmental issues in the disciplinary communities.<sup>5</sup> According to Manduca (2008b), early data indicate that “this effort can claim to have developed a community within the discipline that is discussing departmental issues and sharing their collective wisdom internally. The results of this work have demonstrably raised the level of discussion of accreditation. It cannot yet claim to be reaching the majority of departments” (p. 11).

---

<sup>5</sup>For more detailed information, see the workshop paper by Manduca (see [http://www.nationalacademies.org/bose/Manduca\\_CommissionedPaper.pdf](http://www.nationalacademies.org/bose/Manduca_CommissionedPaper.pdf)).

## 8

## Systemic Change: Barriers and Opportunities

In the final sessions of the workshop, speakers offered systemic perspectives on the issue of changing undergraduate education in science, technology, engineering, and mathematics (STEM). Small groups and committee members also reflected on the workshop's proceedings to identify future directions and next steps.

### DIFFUSION OF PROMISING PRACTICES

Melissa Dancy (Johnson C. Smith University) and Charles Henderson (Western Michigan University) discussed their research on reform and science education at the undergraduate level. Dancy began by noting that the research clearly shows that the traditional lecture-based method is ineffective and that alternative methods yield better outcomes. Although there is still room for additional research and development, Dancy said the problems generally are well documented and solutions are available to address them. However, anecdotal evidence suggests that the impact of this research has been minimal in undergraduate science classrooms and that typical classroom practice remains largely lecture-based.

According to Dancy, change is not happening quickly because change strategies are based largely on a development and dissemination model. With this model, education researchers develop and test specific innovations and disseminate the results to instructors. Typically, this model involves telling instructors that the methods they currently use are ineffective and introducing the evidence for alternative practices in the hopes that instructors will adopt them in their classrooms. Dancy said this approach

fails to consider contextual factors that influence practice and the ability to change.

The development and dissemination model, in Dancy's view, also ignores instructors as an important part of the development process, creating fractious relationships between researchers and instructors. Change agents blame instructors for the lack of change. They assume instructors do not realize that their methods are ineffective, are unaware of alternative options, or do not value effective teaching. For their part, instructors blame the change agents. Interviews with five tenured physics faculty who are considered by their peers to be effective teachers revealed high levels of frustration with the research community (Henderson and Dancy, 2008a). Those faculty members reported that education research is dogmatic and sends the message that everything faculty members are doing is wrong and detrimental to student learning. They expressed a desire to be part of the solution, rather than mere targets of the research.

To improve these relationships and accelerate the change process, Dancy offered several ideas. First, she said curriculum developers can provide easily modifiable materials that instructors can adapt to their own situations as their professional judgment warrants. Second, dissemination can focus on the principles behind a curriculum, not just the curriculum itself. And finally, to acknowledge the constraints faculty face at different institutions, she is in favor of conducting explicit research on the conditions for transferring a reform to different environments.

Dancy presented a model to explain the discontinuity between beliefs and actions regarding implementing reformed instruction (see Figure 8-1). The model shows how individual beliefs interact with context to influence practice. When the two are aligned, belief and action are consistent; when they are not aligned, actions are less consistent with beliefs. For example, faculty members who have progressive beliefs about instruction might teach in environments that do not support innovation—the chairs are bolted down, large numbers of students have expectations for traditional instruction, or their colleagues do not use innovative instructional strategies. Because of contextual constraints, these instructors are likely to use more traditional methods than they otherwise might, according to Dancy. For this reason, she said, any change strategies need to consider the context.

In studying the implementation of promising practices, the research community has focused more on the individual than the environment. However, in Dancy's view, the individual might not represent the greatest point of leverage. Instead, she argued, it would be fruitful to direct more attention to structural changes that could remove barriers to progressive instruction. She also recommended that the research community intensify its efforts to develop models of change beyond the development and dissemination model.

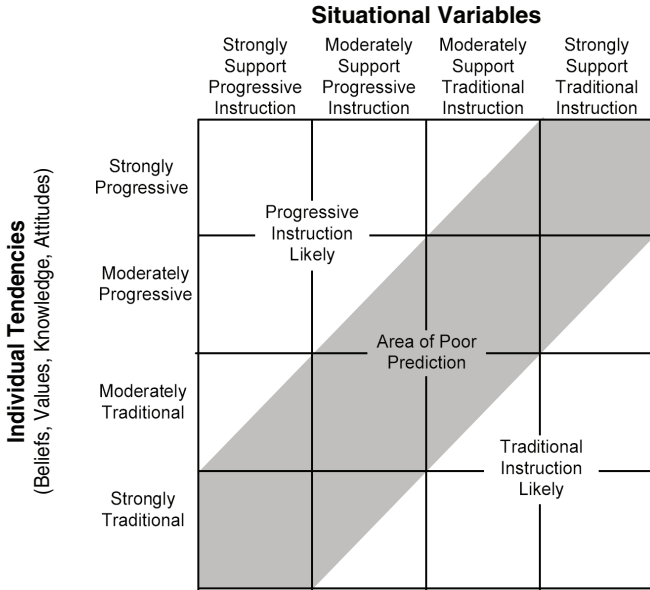


FIGURE 8-1 Belief/action discontinuity.

SOURCE: Reprinted figure with permission from Henderson, C., and Dancy, M. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics: Physics Education Research*, 3(2), 020102. Copyright 2007 by the American Physical Society.

Building on Dancy’s points, Henderson discussed the literature on undergraduate STEM reform. He began by identifying three stakeholder groups: disciplinary STEM education researchers (generally in STEM departments), faculty development researchers (generally in centers for teaching and learning), and higher education researchers (generally in schools of education). Each group has its own journals, conferences, and professional societies. According to Henderson, the literature from all three stakeholder groups is similar and reflects a shift toward a focus on student learning and away from instructors and instruction. However, these groups are conducting their research in isolation from each other, with no overlapping references.

Henderson and his colleagues conducted a systematic study of the literature of the three stakeholder groups and other relevant literature bases (Henderson, Finkelstein, and Beach, 2010). From this review, they developed four categories of change strategies along the dimensions of research



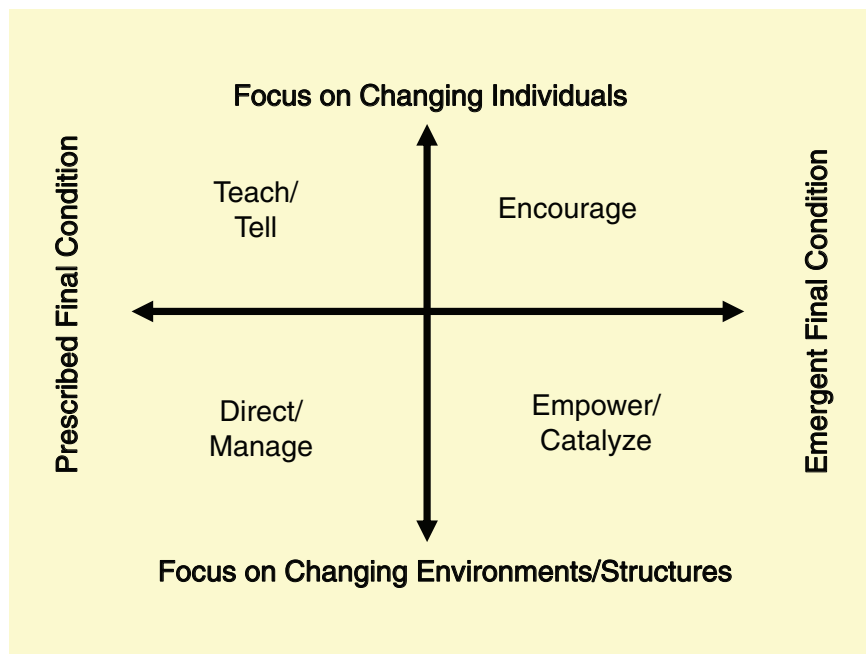


FIGURE 8-2 Change agent roles and strategies.

SOURCE: Henderson and Dancy (2008b). Reprinted with permission.

focus—individual change versus environmental or structural change—and the extent to which the measure of success is prescribed in advance—prescribed versus emergent outcome.<sup>1</sup>

As Figure 8-2 shows, each category has a different change strategy. For the first category—prescribed final condition and a focus on changing individuals—the change strategy is to teach or tell individuals about new teaching ideas or practices. This category represents the development and dissemination model that is common to the STEM education research community and to faculty development researchers. In the second category, the focus remains on changing individuals, but the final condition is emergent. The change strategy is to encourage or support individuals to develop new teaching practices; faculty developers are the primary community employing this strategy. Third, with a prescribed final condition and changing environments or structures, the strategy is to develop new environmental

<sup>1</sup>A prescribed final outcome means that the change agent defines what constitutes success before implementing the change (i.e., if this strategy is successful, student learning will increase).

features that require or encourage new teaching conceptions or practices (e.g., policy change, strategic planning). Higher education researchers are doing most of the work in this area. The fourth category combines a focus on changing the environment with an emerging final condition. Higher education researchers are the primary change agents in this category, and the strategy is to empower the collective development of environmental features that support new teaching ideas or practices (e.g., institutional transformation and learning organizations).

In closing, Henderson underscored Dancy's point that STEM change agents primarily use a development and dissemination model to effect change. They do not draw on approaches from other groups or other disciplines, and they rarely test the effectiveness of the development and dissemination approach. A more fruitful approach, he said, would be to use knowledge from both inside and outside the STEM community to develop better change models and collect empirical data on their effectiveness. In short, he said, such an approach would more closely follow the scientific method.

### REFLECTIONS ON LINKING EVIDENCE AND PROMISING PRACTICES IN STEM

James Fairweather (Michigan State University) observed that most efforts to reform undergraduate STEM education start from a presumptive model based on classroom innovation and the teaching and learning process. The premise, he explained, is that hundreds, if not thousands, of individual faculty improvements will lead to a substantial aggregate change. He pointed out, however, that the aggregate effect has not yet reached desired levels, which underscores the need to advance the conversation about reform.

Fairweather labeled the existing body of reforms as a collection of solutions in search of problems. He identified some common goals that are targeted by reforms:

- Increasing public awareness of STEM or generally improving STEM literacy.
- Stoking the STEM pipeline by attracting K-12 students into STEM, recruiting college students into STEM majors, and improving retention in the majors.
- Enhancing the preparation of STEM college students for their professions.
- Improving various student learning outcomes, including increased content knowledge, the longer term retention of knowledge, application, synthesis, and problem solving.
- Reforming the curriculum.

These goals are divergent and necessitate different approaches, said Fairweather. For example, an effort in one classroom may increase students' retention of content knowledge, but it might not improve their problem-solving skills or stimulate interest in the field or retention in the major. It would be useful, he said, to identify what each innovation is trying to achieve. Such an analysis would uncover redundancies and gaps and would make it easier to target additional reform efforts that address those gaps.

Fairweather observed that researchers make several assumptions about the nature of evidence in reforming STEM undergraduate education. First, they assume that STEM faculty administrators require empirical evidence to convince them of the success of education reforms. Second, they assume that the quality of empirical evidence will be judged according to scientific standards in STEM rather than in education. Third, they assume that the demonstration of evidence alone is sufficient to prompt change; in reality, Fairweather said, empirical evidence is necessary but not sufficient.

Fairweather went on to observe that evaluation practices themselves sometimes confound the ability to truly determine the effectiveness of innovative practices. For example, most evaluation in undergraduate STEM education focuses on in-class events, making it difficult to compare and characterize the entire body of knowledge. In addition, researchers rely more on self-report data than on the gold standard of pre-post comparisons. It is also relatively uncommon to link learning objectives, instructional approaches, and evaluation tools. Finally, said Fairweather, although longitudinal studies, in-depth studies, and studies of systemic reform would yield more nuanced understandings, they are the exceptions rather than the rule.

These observations prompted him to list some useful steps related to evaluating promising practices:

- Distinguish between what is required for any effective teaching or learning environment (e.g., having clear objectives) from what is required to implement innovative pedagogical innovations (e.g., group work).
- Recognize that initial results from studies of innovative practices might not be positive, especially if students are engaging in practices that are new to them.
- Describe the context, with case studies, in sufficient detail so readers can determine whether the results are applicable to them.
- Identify statistical measures (e.g., effect sizes, significance levels) that reflect reasonable and meaningful changes in outcomes.
- Distinguish between evaluations for different audiences and purposes, such as helping a faculty member implement an innovation, helping a faculty member document the effects of a classroom

innovation, or convincing other faculty members to try the new instructional approach.

- Recognize that curriculum reform involves political and cost-effectiveness concerns as well as evidence of impact.

Fairweather also identified some factors that influence the success of innovative strategies. First he noted that focusing on future versus current faculty seems to be an effective way to promote reform (see Chapter 7). It is also important, he said, to understand the implicit change model involved with any innovation. Specifically, it is important to recognize whether the change is expected to happen in a linear or nonlinear way; to identify structural impediments to reform; to understand the role of professional societies and accreditation; and to take into account the role of available institutional resources, including professional development. He concluded by emphasizing that “more effort needs to be expended on strategies to promote the adoption and implementation of STEM reforms rather than on assessing the outcomes of these reforms. Additional research can be useful, but the problem in STEM education lies less in not knowing what works and more in getting people to use proven techniques” (Fairweather, 2008, p. 28).

## FUTURE DIRECTIONS AND NEXT STEPS

After the presentations, participants broke into small groups to reflect on the two workshops and identify future directions for promoting innovations in undergraduate STEM education. Committee members offered some final thoughts.

### Reports from Small-Group Discussions

All of the small groups emphasized the importance of increasing collaboration among the various stakeholders in undergraduate STEM education. They cited the need to forge stronger connections among discipline-based instructors, discipline-based education researchers, education researchers, cognitive scientists, higher education policy researchers, and disciplinary societies. Strengthening these connections, they said, would further scholarship with respect to STEM education and provide opportunities for professional development targeted at implementing research-based practices. Some groups saw value in jointly identifying an umbrella set of challenges that faculty in the STEM disciplines could tackle as a united community.

All of the small groups mentioned the importance of research. Some favored drawing more heavily on existing research. Specifically, they mentioned the extensive literature from other disciplines on faculty develop-

ment and the idea of requiring National Science Foundation grantees to base curriculum proposals on existing research. Several groups identified the need for additional research, particularly on institutional change and its relation to STEM education. Ideas in this regard included a concerted research initiative around the broad question of what influences faculty members' teaching decisions; research that examines the drivers for change, the resistance for change, and strategies for overcoming that resistance; the role of influential leaders in promoting change; and a deeper analysis of change strategies that do not work.

Finally, the groups mentioned the importance of disseminating research in a way that makes it enticing and easy for “hungry adopters” to change their practice. The process would take into account the role of textbooks and textbook developers and would involve understanding why more faculty are not adopting innovations and identifying those who might be amenable to changing their practice. According to the small groups, dissemination efforts might include a design manual articulating research-based guidelines for structuring courses and mechanisms for sharing information about innovations within and across disciplines.

### Final Thoughts

Kenneth Heller observed that many of the teaching strategies discussed during the workshop (e.g., case-based learning, problem-based learning, using closed-ended problems or context-rich problems) involved a common set of elements. For example, they all include cooperative group learning, connection to a real problem, and coaching—and these methods seem to be effective.

David Mogk focused on next steps. He cited a need for resources and networks that will engage more faculty in the scholarship of learning and help them become agents of change in their classes, departments, and institutions. Drawing parallels between the scientific method and education research and assessment, he encouraged workshop participants to help their colleagues engage in assessment for the betterment of STEM education and for the health of science and society.

Melvin George remarked on the dearth of discussion about the purpose of improving STEM education, stressing the need to identify a compelling sense of purpose that will generate support for reforms. He also agreed with the need to create a design manual for “hungry adopters.” He concluded by underscoring the points made by Fairweather, Dancy, and Henderson about directing more resources to understanding the factors that influence change versus continuing to study which practices are effective.

Building on George's points, William Wood added that it is important to understand the role students play—positive and negative—in the change

process. He noted that students' facility with technology and access to information have required instructors to shift away from teaching facts (Prensky, 2001). However, in his experience, students pose barriers to reform because they often resist new pedagogies and are unfamiliar with how to learn. For this reason, in addition to educating instructors about better instruction, Wood stressed a need to educate students about how to learn.

Susan Singer commented on the fact that several people view further research on effective practices and further research on implementing change as mutually exclusive. She observed that, similar to scientific research, the process of change is iterative and requires both types of research. She also cited a need to develop a broader theoretical framework to guide STEM education research within and across disciplines, expressing the hope that this workshop series is the beginning of a conversation along those lines, rather than the end.

## References

- American Association of Colleges and Universities. (2008). *College learning for the new global century*. Washington, DC: Author.
- Barrows, H.S. (1996). Problem-based learning in medicine and beyond. In L. Wilkerson and W.H. Gijsselaers (Eds.), *Bringing problem-based learning to higher education: Theory and practice*. New directions for teaching and learning, No. 68. (pp. 3-13). San Francisco: Jossey-Bass.
- Beichner, R.J. (2008a, October). *The SCALE-UP project: A student-centered active learning environment for undergraduate programs*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Beichner\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Beichner_CommissionedPaper.pdf).
- Beichner, R.J. (2008b, October). *SCALE-UP: Student-centered active learning environment for undergraduate programs*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Beichner\\_Presentation\\_Workshop2.pdf](http://www7.nationalacademies.org/bose/Beichner_Presentation_Workshop2.pdf).
- Beichner, R.J., Saul, J.M., Abbott, D.S., Morse, J.J., Deardorff, D.L., Allain, R.J., Bonham, S., Dancy, M.H., and Risley, J.S. (2007). Student-centered activities for large enrollment undergraduate programs (SCALE-UP) project. In E.F. Redish and P.J. Cooney (Eds.), *Reviews in Physics Education Research (PER), volume 1: Research-based reform in university physics*. College Park, MD: American Association of Physics Teachers.
- Biggers, S., Benson, L., Moss, W., Ohland, M., Orr, M., and Schiff, S. (2007, June). *Adapting and implementing the SCALE-UP approach in statics, dynamics, and multivariable calculus*. Paper presented at ASEE Annual Conference and Exposition, Honolulu, Hawaii.
- Cummings, K. (2008, October). *The Rensselaer studio model for learning and teaching: What have we learned?* Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Cummings\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Cummings_CommissionedPaper.pdf).

- Dochy, F., Segers, M., Van den Bossche, P., and Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, 13, 533-568.
- Ebert-May, D. (2000, August). *Workshop #2: Assessment of student learning: Implementing the scholarship of teaching and research in ecology (biology) education*. Workshop presented at the Annual Meeting of the Ecological Society of America, Snowbird, UT.
- Fairweather, J. (2008). *Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, October, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Fairweather\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Fairweather_CommissionedPaper.pdf).
- Finkelstein, N.D., and Pollock, S.J. (2005). **Replicating and understanding successful innovations: Implementing tutorials in introductory physics**. *Physical Review Special Topics—Physics Education Research*, 1, 010101.
- Foster, T. (2008, October). *Use of complex problems in teaching physics*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Foster\\_Presentation\\_Workshop2.pdf](http://www7.nationalacademies.org/bose/Foster_Presentation_Workshop2.pdf).
- Francis, G.E., Adams, J.P., and Noonan, E.J. (1998). Do they stay fixed? *The Physics Teacher*, 36, 488-490.
- Froyd, J. (2008, June). *White paper on promising practices in undergraduate STEM education*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Froyd\\_Promising\\_Practices\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Froyd_Promising_Practices_CommissionedPaper.pdf).
- Gal, I., and Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education*, 2(2).
- Gijbels, D. (2008, October). *Effectiveness of problem-based learning: Scored goals and how to play a better second half*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Gijbels\\_Presentation\\_Workshop2.pdf](http://www7.nationalacademies.org/bose/Gijbels_Presentation_Workshop2.pdf).
- Gijbels, D., Dochy, F., Van den Bossche, P., and Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research*, 75(1), 27-61.
- Gregerman, S.R. (2008, October). *The role of undergraduate research in student retention, academic engagement, and the pursuit of graduate education*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Gregerman\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Gregerman_CommissionedPaper.pdf).
- Hake, R. (1998). Interactive engagement versus traditional methods: A six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64-74.
- 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., and Dancy, M. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics: Physics Education Research*, 3(2), 020102.
- Henderson, C., and Dancy, M. (2008a). Physics faculty and educational researchers: Divergent expectations as barriers to the diffusion of innovations. *American Journal of Physics (Physics Education Research Section)*, 76(1), 79-91.



- Henderson, C., and Dancy, M. (2008b, October). *Barriers and promises in STEM reform part 2: The study and improvement of STEM change strategies*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Henderson\\_Presentation\\_Workshop2.pdf](http://www7.nationalacademies.org/bose/Henderson_Presentation_Workshop2.pdf).
- Henderson, C., Beach, A., Finkelstein, N., and Larson, R.S. (2008, June). *Preliminary categorization of literature on promoting change in undergraduate STEM*. Paper presented at the Facilitating Change in Undergraduate STEM symposium, June, Augusta, MI. Available at <http://www.wmich.edu/science/facilitating-change/PreliminaryCategorization.pdf>
- Henderson, C., Finkelstein, N., and Beach, A. (2010). Beyond dissemination in college science teaching: An introduction to four core change strategies. *Journal of College Science Teaching*, 39(5), 18-25.
- Heron, P.R.L., Shaffer, P.S., and McDermott, L.C. (2008, October). *Identifying and addressing student conceptual difficulties: An example from introductory physics*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Heron\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Heron_CommissionedPaper.pdf).
- Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30, 159-166.
- Jackson, D.P., Laws, P.W., and Franklin, S.V. (2003). *Explorations in physics: An activity-based approach to understanding the world*. New York: John Wiley.
- Knight, J.K., and Wood, W.B. (2005). Teaching more by lecturing less. *Cell Biology Education*, 4(4), 298-310.
- Kramer, L., Brewe, E., and O'Brien, G. (2008). Improving physics education through a diverse research and learning community at Florida International University. *American Physical Society Forum on Education Newsletter* (Summer), 29-32.
- Krane, K. (2008, October). *The workshop for new faculty in physics and astronomy*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Krane\\_Presentation\\_Workshop2.pdf](http://www7.nationalacademies.org/bose/Krane_Presentation_Workshop2.pdf).
- Libarkin, J. (2008, October). *Concept inventories in higher education science*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Libarkin\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Libarkin_CommissionedPaper.pdf).
- Lorenzo, M., Crouch, C.H., and Mazur, E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics*, 74, 118-122.
- Lundeberg, M.A. (2008, October). *Case pedagogy in undergraduate STEM: Research we have; research we need*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Lundeberg\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Lundeberg_CommissionedPaper.pdf).
- Lux, K.W., D'Amato, M.J., Anderegg, B., Walz, K.A., and Kerby, H.W. (2007). Introducing new learning tools into a standard classroom: A multi-tool approach to integrating fuel-cell concepts into introductory college chemistry. *Journal of Chemical Education*, 84(2), 248.
- Macdonald, R.H., Manduca, C.A., Mogk, D.W., and Tewksbury, B.J. (2005). Teaching methods in undergraduate geoscience courses: Results of the 2004 On the Cutting Edge survey of U.S. faculty. *Journal of Geoscience Education*, 55(3), 237-252.

- Manduca, C. (2008a, October). *Changing undergraduate STEM education: Addressing disciplinary and institutional culture*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Manduca\\_Presentation\\_Workshop2.pdf](http://www7.nationalacademies.org/bose/Manduca_Presentation_Workshop2.pdf).
- Manduca, C.A. (2008b, October). *Working with the discipline: Developing a supportive environment for education*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Manduca\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Manduca_CommissionedPaper.pdf).
- Mayberry, M. (1998). Reproductive and resistant pedagogies: The comparative roles of collaborative learning and feminist pedagogy in science education reform. *Journal of Research on Science Teaching*, 35(4), 443-459.
- McConnell, D., (2008, October). *Less talk, more action: Active learning in introductory geoscience courses*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/McConnell\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/McConnell_CommissionedPaper.pdf).
- McConnell, D.A., Steer, D.N., Owens, K., and Knight, C. (2005). How students think: Implications for learning in introductory geoscience courses. *Journal of Geoscience Education*, 53(4), 462-470.
- McDaniel, C., Lister, B., Hanna, M., and Roy, H. (2007). Increased learning observed in redesigned introductory biology course that employed web-enhanced, interactive pedagogy. *CBE Life Science Education*, 6(3), 243-249.
- McDermott, L.C., and Redish, E.F. (1999). Resource letter on physics education research. *American Journal of Physics*, 67(9), 755.
- McDermott, L.C., Shaffer, P.S., and the Physics Education Group at the University of Washington. (2002). *Tutorials in introductory physics*. Upper Saddle River, NJ: Prentice-Hall.
- McNeil, E., and Ogle, B. (2008, October). *Active learning through teaching in an extracellular matrix engineering course for biomedical engineering graduate students*. Paper presented at the ASEE North Midwest Sectional Conference, Platteville, Wisconsin.
- Mestre, J.P. (2005). Facts and myths about pedagogies of engagement in science learning. *Peer Review*, 7(2), 24-27.
- Middlecamp, C. (2008, June). *Chemistry in Context: Goals, evidence, gaps*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/PP\\_Middlecamp\\_Presentation.pdf](http://www7.nationalacademies.org/bose/PP_Middlecamp_Presentation.pdf).
- Miller, J.D., Scott, E.C., and Okamoto, S. (2006). Science communication: Public acceptance of evolution. *Science*, 313, 765-766.
- Narum, J. (2008, June). *Promising practices in undergraduate STEM education*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/PP\\_Narum\\_Presentation.pdf](http://www7.nationalacademies.org/bose/PP_Narum_Presentation.pdf).
- National Research Council. (1996). *From analysis to action: Undergraduate education in science, mathematics, engineering, and technology*. Center for Science, Mathematics, and Engineering Education. Washington, DC: National Academy Press.
- National Research Council. (1999). *Transforming undergraduate education in science, mathematics, engineering, and technology*. Committee on Undergraduate Science Education, Center for Science, Mathematics, and Engineering Education. Washington, DC: National Academy Press.

- National Research Council. (2002). *Scientific research in education*. Committee on Scientific Principles for Education Research, R.J. Shavelson and L. Towne, Eds. Washington, DC: The National Academies Press.
- National Science Foundation. (1996). *Shaping the future: New expectations for undergraduate education in science, mathematics, and technology*. Advisory Committee to the National Science Foundation, Directorate of Education and Human Resources. Arlington, VA: Author.
- Osgood, M., Mitchell, S., and Anderson, A. (2008, October). *Tracking student problem-solving strategies in online PBL case discussions: A method to target interventions to individuals and groups most in need of help*. [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Osgood\\_Presentation\\_Workshop2.pdf](http://www7.nationalacademies.org/bose/Osgood_Presentation_Workshop2.pdf).
- Piburn, M., Sawada, D., Falconer, K., Turley, J., Benford, R., and Bloom, I. (2000). *Reformed teaching observation protocol (RTOPI)*. Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers.
- Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, 9(5).
- Redish, E.F. (2004). A theoretical framework for physics education research: Modeling student thinking. In E.F. Redish and M. Vicentini (Eds.), *Proceedings of the International School of Physics "Enrico Fermi" course CLVI*, (pp.1-64). Amsterdam: IOS Press. Available at <http://www.physics.umd.edu/perg/papers/redish/Redish%20VarennaPre.pdf>
- Redish, E.F. (2008, June). *What is the state of evidence in DBER?* [Presentation slides]. Presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/PP\\_Redish\\_Presentation.pdf](http://www7.nationalacademies.org/bose/PP_Redish_Presentation.pdf).
- Reed-Rhoads, T., and Imbrie, P.K. (2008, October). *Concept inventories in engineering education*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Reed\\_Rhoads\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Reed_Rhoads_CommissionedPaper.pdf).
- Roadrangka, V., Yeany, R.H., and Padilla, M.J. (1982). *GALT, Group Assessment of Logical Thinking*. Athens: University of Georgia.
- Roadrangka, V., Yeany, R.H., and Padilla, M.J. (1983, April). *The construction and validation of Group Assessment of Logical Thinking (GALT)*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Dallas.
- Schwartz, D.L., Bransford, J.D., and Sears, D. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1-51). Greenwich, CT: Information Age.
- Silverthorn, D.U. (2006). Teaching and learning in the interactive classroom. *Advances in Physiology Education*, 30, 135-140.
- Wink, D.J., and Weaver, G.C. (2008, October). *Evaluation of the Center for Authentic Science Practice in Education (CASPiE) model of undergraduate research*. Paper presented at the National Research Council's Workshop Linking Evidence to Promising Practices in STEM Undergraduate Education, Washington, DC. Available at [http://www7.nationalacademies.org/bose/Wink\\_Weaver\\_CommissionedPaper.pdf](http://www7.nationalacademies.org/bose/Wink_Weaver_CommissionedPaper.pdf).



# Appendix A

## June Workshop Agenda and Participants List

### AGENDA

#### Workshop on Linking Evidence and Promising Practices in STEM Undergraduate Education

*Monday June 30, 2008*

8:00 a.m. Introductions

8:30 a.m. Overview of the workshop goals  
*Susan Singer, Carleton College*

8:45 a.m. Panel: Linking Evidence and Learning Goals  
*Moderator: Adam Gamoran, University of Wisconsin,  
Madison*  
*Panelists: Cathy Middlecamp, University of Wisconsin,  
Madison*  
*Jose Mestre, University of Illinois,  
Urbana/Champaign*  
*Bruce Grant, Widener University*

Following the meeting, each panelist will write a brief paper based on his/her presentation and input from the discussion. Panelists were asked to address the following questions in their papers and will select specific areas to highlight in their presentations.

1. What are and what should be some of the most important learning goals for science students in lower division courses? We are interested in goals over a range of grain sizes from activities within an individual course to college-wide efforts.
2. In the context of the learning goals you identified, what types of evidence would be needed in order to conclude that a specific goal had been achieved?
3. With so many forms of evidence available to us in science education, are there some types of evidence that carry more weight in your experience? If so, what makes that evidence particularly compelling?
4. As you consider learning goals and evidence, where are the biggest gaps in evidence in science undergraduate education?
5. How important has the quality of evidence been in influencing or guiding the widespread uptake of a promising practice? Can you identify specific examples where the presence or absence of evidence of effectiveness has had a major impact on dissemination or use?

9:30 a.m. Audience discussion of panel

10:00 a.m. Break and transition to small groups

10:15 a.m. Small groups to discuss learning goals and evidence

Each group will hold a discussion, using the following questions as guidance. Please take notes for the report out following the discussion.

Questions to guide small-group discussion:

- What are the varied learning goals in your discipline? Of these, what do you consider to be the most important learning goals?
- What types of evidence are needed to establish effectiveness given the goals identified?
- Are there differences across disciplines in the desired learning goals? In what counts as evidence of effectiveness?

11:00 a.m. Report out by small groups

11:30 a.m. Panel: What Is the State of Evidence in Discipline-Based Education Research?

*Moderator: Kenneth Heller, University of Minnesota*

*Panelists: William Wood, University of Colorado, Boulder  
Edward Redish, University of Maryland  
Helen King, Consultant*

Each panelist was asked to respond to the following:

1. Summarize the major findings from discipline-based education research in your discipline.
2. Identify the most promising or important directions for future research.

12:15 p.m. Audience discussion of panel

12:45 p.m. Lunch and informal discussion of morning sessions

1:30 p.m. Panel: Surveying Promising Practices  
*Moderator: Melvin George, University of Missouri*  
*Panelists: Jeffrey Froyd, Texas A&M University*  
*Philip Sadler, Harvard University*  
*Jeanne Narum, Project Kaleidoscope*

Following the meeting, each panelist will write a brief paper based on his/her presentation and input from the discussion. Panelists were asked to address the following questions in their papers and will select specific areas to highlight in their presentations.

1. How would you categorize the range of promising practices that have emerged over the past 20 years? Consider practices that are discipline-specific as well as those that are interdisciplinary.
2. What types of categories do you find are most useful in sorting out the range of efforts that have emerged? Why did you choose to aggregate certain practices within a category?
3. As you chose exemplars for your categories, what criteria did you use to identify something as a promising practice?

2:30 p.m. Audience discussion of panel

3:00 p.m. Break and transition to small groups

3:15 p.m. Small-group discussion of promising practices

Start this session with a one-minute written response to the following question:

Reflecting on the panel discussion, from your experience what top three promising practices would you identify? Please list the promising practice, related outcomes, goals, audience, and context in which the practice is best suited.

In a round robin format, discuss why these were the top picks and what the state of the evidence is related to each practice.

4:15 p.m. Report out by small groups

4:45 p.m. Steering committee's and participants' final reflections

5:30 p.m. Adjourn

## PARTICIPANTS

### *Speakers*

Jeffrey Froyd, Texas A&M University

Bruce Grant, Widener University

Jose Mestre, University of Illinois, Urbana/Champaign

Cathy Middlecamp, University of Wisconsin, Madison

Helen King, Helen King Consultancy

Jeanne Narum, Project Kaleidoscope

Edward Redish, University of Maryland

Philip Sadler, Harvard University

William Wood, University of Colorado, Boulder

### *Invited Guests*

Susan Albertine, Association of American Colleges and Universities

Robert Beichner, North Carolina State University

Myles Boylan, National Science Foundation

Celeste Carter, National Science Foundation

Amber Coleman, Board on Chemical Sciences and Technology, National Research Council

Mark Connolly, University of Wisconsin, Madison

Malcolm Drewery, National Academy of Engineering

Adam Fagen, Board on Life Sciences, National Research Council

Adam Gamoran, University of Wisconsin, Madison

Pamela Hines, American Association for the Advancement of Science



Kimberly Kastens, Columbia University  
Mary M. Kirchhoff, American Chemical Society  
David Mandel, National Center on Education and the Economy  
Tina Masciangioli, Board on Chemical Sciences and Technology, National  
Research Council  
Lillian McDermott, University of Washington  
Susan Millar, University of Wisconsin, Madison  
Michael Moloney, Board on Physics and Astronomy, National Research  
Council  
Lina Patino, National Science Foundation  
Dexter Perkins, University of North Dakota  
Ginger Holmes Rowell, National Science Foundation  
Carol Schneider, Association of American Colleges & Universities  
Dee Silverthorn, University of Texas, Austin  
Linda Slakey, National Science Foundation  
Carol Snyder, American Association of Colleges and Universities  
Brock Spencer, Beloit College  
James Stith, American Institute of Physics  
Larry Suter, National Science Foundation  
Partibha Varma-Nelson, National Science Foundation  
Jodi Wesemann, American Chemical Society  
Karl Wirth, Macalester College  
Robin Wright, University of Minnesota  
Terry Woodin, National Science Foundation

## Appendix B

### October Workshop Agenda and Participants List

#### AGENDA

#### Workshop on Linking Evidence and Promising Practices in STEM Undergraduate Education

*Monday, October 13, 2008*

- |                 |  |
|-----------------|--|
| 8:00 a.m.       | Introductions  |
| 8:30 a.m.       | Framing the workshop<br><i>Susan Singer, Carleton College</i>  |
| 9:00-10:15 a.m. | Session 1: Scenario-, Problem-, and Case-Based<br>Teaching and Learning<br><i>Moderator: Kenneth Heller, University of<br/>Minnesota</i> |
| 9:00 a.m.       | Effectiveness of Problem-Based Learning<br><i>David Gijbels, University of Antwerp</i>   |
| 9:15 a.m.       | Evaluating Case-Based Teaching<br><i>Mary Lundeberg, Michigan State University</i>   |
| 9:30 a.m.       | Use of Complex Problems in Teaching Physics<br><i>Tom Foster, Southern Illinois University</i>   |
| 9:45-10:15      | Discussion of presentations  |

- 10:15-10:30 a.m. Break
- 10:30 a.m.-12:15 p.m. Session 2: Assessment to Guide Teaching and Learning  
*Moderator: Susan Singer, Carleton College*
- 10:30 a.m. Concept Inventories in the Sciences  
*Julie Libarkin, Michigan State University*
- 10:45 a.m. Concept Inventories in Engineering  
*Teri Reed-Rhoads, Purdue University*
- 11:00 a.m. Identifying and Addressing Student Difficulties in Physics  
*Paula Heron, University of Washington*
- 11:15 a.m.-12:00 p.m. Discussion of presentations
- 12:00-12:45 p.m. Working lunch: Discuss morning presentations
- 12:45-2:30 p.m. Session 3: Structuring the Learning Environment  
*Moderator: William B. Wood, University of Colorado, Boulder*
- 12:45 p.m. The Effectiveness of Studio Courses at RPI  
*Karen Cummings, Southern Connecticut State University*
- 1:00 p.m. Redesigning Large Classes for Learning (1): Project SCALE-UP  
*Robert Beichner, North Carolina State University*
- 1:15 p.m. Redesigning Large Classes for Learning (2): Developing and Assessing Problem-Solving Skills in Online Student Groups  
*Marcy Osgood, University of New Mexico*
- 1:30 p.m. Redesigning Large Classes for Learning (3): Active Learning Strategies for Introductory Geology Courses  
*David McConnell, North Carolina State University*
- 1:45-2:30 p.m. Discussion of presentations
- 2:30-2:45 p.m. Break and transition to small groups

2:45-3:30 p.m.                      Small-group discussion of Sessions 1-3

Discussion questions:

- Which practices have the strongest evidence?
  - Where are the gaps in the evidence?
- What kinds of outcomes are commonly assessed?
  - Are these sufficient for establishing effectiveness?
- What kinds of assessments were used to measure these outcomes?
  - How adequate are these assessments, and are new assessments needed to accurately measure all possible outcomes?
- Do you see ways that the evidence across the different practices converges?
- What are the implications for broad dissemination of the practices?
- What are the implications for future research on these practices?

3:30-4:00 p.m.                      Report out by small groups

4:00-5:15 p.m.                      Session 4: Doing Science—Providing Research Experiences  
*Moderator: David Mogk, Montana State University*

4:00 p.m.                      Evaluation of the University of Michigan UROP Program  
*Sandy Gregerman, University of Michigan*

4:15 p.m.                      Center for Authentic Science Practice in Education  
*Donald Wink, University of Illinois, Chicago*  
*Gabriela Weaver, Purdue University*

4:30-5:15 p.m.                      Discussion of presentations

5:15 p.m.                              Adjourn for the day

*Tuesday October 14, 2008*

8:00-8:30 a.m.                      Introductions

8:30-10:15 a.m.                      Session 5: Faculty Professional Development  
*Moderator: Kenneth Heller, University of Minnesota*

- 8:30-8:45 a.m. Professional Development of Graduate Students/  
Teaching Assistants  
*Donald Gillian-Daniel, University of Wisconsin,  
Madison*
- 8:45-9:00 a.m. Workshops by a Professional Society for New  
Physics Faculty  
*Ken Krane, Oregon State University*
- 9:00-9:15 a.m. Changing Undergraduate STEM Instruction  
*Cathy Manduca, Carleton College*
- 9:15-9:30 a.m. Effectiveness of Faculty Professional  
Development  
*Diane Ebert-May, Michigan State University*
- 9:30-10:15 a.m. Discussion of presentations
- 10:15-10:30 a.m. Break
- 10:30-11:15 a.m. Session 6: Systemic Change in Undergraduate  
STEM  
*Moderator: Melvin George, University of  
Missouri*
- 10:30-11:00 a.m. Diffusion of Promising Practices  
*Melissa Dancy, Johnson C. Smith University  
Charles Henderson, Western Michigan University*
- 11:00-11:30 a.m. Discussion of presentations
- 11:30 a.m.-12:00 p.m. Small-group discussions

Discussion questions:

- Discuss the evidence related to faculty professional development.
  - How strong is the evidence base? What does it tell us about how best to support faculty development?
  - What does it tell us about the role of faculty development in reform of undergraduate STEM education?
- Discuss the role of evidence in diffusion of promising practices and implications for future directions for both research and practice.

- 12:00-12:45 p.m.      Lunch and continue small-group discussions
- 12:45-1:15 p.m.      Report out by small groups
- 1:15-2:30 p.m.      Session 7: Future Directions  
*Moderator: Susan Singer, Carleton College*
- Reflections on Linking Evidence and Promising  
 Practices in STEM  
*James Fairweather, Michigan State University*  
 Responses and Next Steps  
*Workshop Steering Committee*  
 Final questions and answers
- 2:30 p.m.              Adjourn

## PARTICIPANTS

### *Speakers*

Robert Beichner, North Carolina State University  
 Karen Cummings, Southern Connecticut State University  
 Melissa Dancy, Johnson C. Smith University  
 Diane Ebert-May, Michigan State University  
 James Fairweather, Michigan State University  
 Tom Foster, Southern Illinois University  
 David Gijbels, University of Antwerp  
 Don Gillian-Daniel, University of Wisconsin, Madison  
 Sandra Gregerman, University of Michigan  
 Charles Henderson, Western Michigan University  
 Paula Heron, University of Washington  
 Ken Krane, Oregon State University  
 Julie Libarkin, Michigan State University  
 Mary Lundeberg, Michigan State University  
 Cathy Manduca, Carleton College  
 David McConnell, North Carolina State University  
 Marcy Osgood, University of New Mexico  
 Teri Reed-Rhoads, Purdue University  
 Gabriela Weaver, Purdue University  
 Donald Wink, University of Illinois, Chicago

*Invited Guests*

Susan Albertine, American Association of Colleges and Universities  
 Deborah Allen, National Science Foundation  
 Myles Boylan, National Science Foundation  
 David Burns, National Center for Science and Civic Engagement  
 Beth Cady, National Academy of Engineering  
 Heather Dobbins, University of Maryland  
 Catherine Frey, National Science Foundation  
 Jeffrey Froyd, Texas A&M University  
 Howard Gobstein, National Association of State Universities and Land-Grant Colleges  
 Elizabeth Godfrey, University of Auckland  
 Bruce Grant, Widener University  
 Jack Hehn, American Institute of Physics  
 Helen King, Helen King Consultancy  
 Mary Kirchhoff, American Chemical Society  
 James Lancaster, Board on Physics and Astronomy, National Research Council  
 David Mandel, Carnegie-IAS Commission  
 Cathy Middlecamp, Wisconsin Center for Education Research  
 Susan Millar, Wisconsin Center for Education Research  
 Martha Narro, iPlant Collaborative, University of Arizona  
 Jeanne Narum, Project Kaleidoscope  
 Karen Kashmanian Oates, National Science Foundation  
 Catherine O’Riordan, American Institute of Physics  
 Greg Pearson, National Academy of Engineering  
 Dexter Perkins, University of North Dakota  
 Muriel Poston, Skidmore College  
 Jennifer Presley, National Association of State Universities and Land-Grant Colleges  
 Christine Maidl Pribbenow, Wisconsin Program for Science Teaching  
 Linda Ravan, American Geological Institute  
 Merilie Reynolds, American Geological Institute  
 Terry Rhodes, American Association of Colleges and Universities  
 James Stith, American Institute of Physics  
 Jodi Wesemann, American Chemical Society  
 Suzanne Westbrook, iPlant Collaborative, University of Arizona

