



# **Doing the Scholarship of Teaching and Learning in Mathematics**

Cover Art: The source text for the wordle ([www.wordle.net](http://www.wordle.net)) consisted of Chapter 20 Synthesis of the Value and Benefits of SoTL Experienced by the Contributors with the names of individuals and the word “chapter” removed.

© 2015 by  
*The Mathematical Association of America (Incorporated)*

*Library of Congress Control Number 2014953115*

Print ISBN 978-0-88385-193-7

Electronic ISBN 978-1-61444-318-6

*Printed in the United States of America*

Current Printing (last digit):

10 9 8 7 6 5 4 3 2 1

# Doing the Scholarship of Teaching and Learning in Mathematics

Edited by

Jacqueline M. Dewar

and

Curtis D. Bennett

*Loyola Marymount University*



*Published and Distributed by*  
The Mathematical Association of America

The MAA Notes Series, started in 1982, addresses a broad range of topics and themes of interest to all who are involved with undergraduate mathematics. The volumes in this series are readable, informative, and useful, and help the mathematical community keep up with developments of importance to mathematics.

### **Council on Publications and Communications**

Frank Farris, *Chair*

### **Committee on Books**

Fernando Q. Gouvêa, *Chair*

### **Notes Editorial Board**

Michael K. May *Editor*

Michael C. Axtell	Anneke Bart
Christopher L. Frenzen	Louis M. Friedler
Hugh Howards	Theresa Jeevanjee
Elizabeth W. McMahan	Dan Sloughter
Stephen J. Willson	Joe Yanik

### **MAA Notes**

14. *Mathematical Writing*, by *Donald E. Knuth, Tracy Larrabee, and Paul M. Roberts*.
16. *Using Writing to Teach Mathematics*, *Andrew Sterrett*, Editor.
17. *Priming the Calculus Pump: Innovations and Resources*, Committee on Calculus Reform and the First Two Years, a subcommittee of the Committee on the Undergraduate Program in Mathematics, *Thomas W. Tucker*, Editor.
18. *Models for Undergraduate Research in Mathematics*, *Lester Senechal*, Editor.
19. *Visualization in Teaching and Learning Mathematics*, Committee on Computers in Mathematics Education, *Steve Cunningham and Walter S. Zimmermann*, Editors.
20. *The Laboratory Approach to Teaching Calculus*, *L. Carl Leinbach et al.*, Editors.
21. *Perspectives on Contemporary Statistics*, *David C. Hoaglin and David S. Moore*, Editors.
22. *Heeding the Call for Change: Suggestions for Curricular Action*, *Lynn A. Steen*, Editor.
24. *Symbolic Computation in Undergraduate Mathematics Education*, *Zaven A. Karian*, Editor.
25. *The Concept of Function: Aspects of Epistemology and Pedagogy*, *Guershon Harel and Ed Dubinsky*, Editors.
26. *Statistics for the Twenty-First Century*, *Florence and Sheldon Gordon*, Editors.
27. *Resources for Calculus Collection, Volume 1: Learning by Discovery: A Lab Manual for Calculus*, *Anita E. Solow*, Editor.
28. *Resources for Calculus Collection, Volume 2: Calculus Problems for a New Century*, *Robert Fraga*, Editor.
29. *Resources for Calculus Collection, Volume 3: Applications of Calculus*, *Philip Straffin*, Editor.
30. *Resources for Calculus Collection, Volume 4: Problems for Student Investigation*, *Michael B. Jackson and John R. Ramsay*, Editors.
31. *Resources for Calculus Collection, Volume 5: Readings for Calculus*, *Underwood Dudley*, Editor.
32. *Essays in Humanistic Mathematics*, *Alvin White*, Editor.
33. *Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results*, *James J. Kaput and Ed Dubinsky*, Editors.
34. *In Eves Circles*, *Joby Milo Anthony*, Editor.
35. *You're the Professor, What Next? Ideas and Resources for Preparing College Teachers*, The Committee on Preparation for College Teaching, *Bettye Anne Case*, Editor.
36. *Preparing for a New Calculus: Conference Proceedings*, *Anita E. Solow*, Editor.
37. *A Practical Guide to Cooperative Learning in Collegiate Mathematics*, *Nancy L. Hagelgans, Barbara E. Reynolds, SDS, Keith Schwingendorf, Draga Vidakovic, Ed Dubinsky, Mazen Shahin, G. Joseph Wimbish, Jr.*
38. *Models That Work: Case Studies in Effective Undergraduate Mathematics Programs*, *Alan C. Tucker*, Editor.
39. *Calculus: The Dynamics of Change*, CUPM Subcommittee on Calculus Reform and the First Two Years, *A. Wayne Roberts*, Editor.
40. *Vita Mathematica: Historical Research and Integration with Teaching*, *Ronald Calinger*, Editor.
41. *Geometry Turned On: Dynamic Software in Learning, Teaching, and Research*, *James R. King and Doris Schattschneider*, Editors.
42. *Resources for Teaching Linear Algebra*, *David Carlson, Charles R. Johnson, David C. Lay, A. Duane Porter, Ann E. Watkins, William Watkins*, Editors.
43. *Student Assessment in Calculus: A Report of the NSF Working Group on Assessment in Calculus*, *Alan Schoenfeld*, Editor.

44. Readings in Cooperative Learning for Undergraduate Mathematics, *Ed Dubinsky, David Mathews, and Barbara E. Reynolds*, Editors.
45. Confronting the Core Curriculum: Considering Change in the Undergraduate Mathematics Major, *John A. Dossey*, Editor.
46. Women in Mathematics: Scaling the Heights, *Deborah Nolan*, Editor.
47. Exemplary Programs in Introductory College Mathematics: Innovative Programs Using Technology, *Susan Lenker*, Editor.
48. Writing in the Teaching and Learning of Mathematics, *John Meier and Thomas Rishel*.
49. Assessment Practices in Undergraduate Mathematics, *Bonnie Gold*, Editor.
50. Revolutions in Differential Equations: Exploring ODEs with Modern Technology, *Michael J. Kallaher*, Editor.
51. Using History to Teach Mathematics: An International Perspective, *Victor J. Katz*, Editor.
52. Teaching Statistics: Resources for Undergraduate Instructors, *Thomas L. Moore*, Editor.
53. Geometry at Work: Papers in Applied Geometry, *Catherine A. Gorini*, Editor.
54. Teaching First: A Guide for New Mathematicians, *Thomas W. Rishel*.
55. Cooperative Learning in Undergraduate Mathematics: Issues That Matter and Strategies That Work, *Elizabeth C. Rogers, Barbara E. Reynolds, Neil A. Davidson, and Anthony D. Thomas*, Editors.
56. Changing Calculus: A Report on Evaluation Efforts and National Impact from 1988 to 1998, *Susan L. Ganter*.
57. Learning to Teach and Teaching to Learn Mathematics: Resources for Professional Development, *Matthew Delong and Dale Winter*.
58. Fractals, Graphics, and Mathematics Education, *Benoit Mandelbrot and Michael Frame*, Editors.
59. Linear Algebra Gems: Assets for Undergraduate Mathematics, *David Carlson, Charles R. Johnson, David C. Lay, and A. Duane Porter*, Editors.
60. Innovations in Teaching Abstract Algebra, *Allen C. Hibbard and Ellen J. Maycock*, Editors.
61. Changing Core Mathematics, *Chris Arney and Donald Small*, Editors.
62. Achieving Quantitative Literacy: An Urgent Challenge for Higher Education, *Lynn Arthur Steen*.
64. Leading the Mathematical Sciences Department: A Resource for Chairs, *Tina H. Straley, Marcia P. Sward, and Jon W. Scott*, Editors.
65. Innovations in Teaching Statistics, *Joan B. Garfield*, Editor.
66. Mathematics in Service to the Community: Concepts and models for service-learning in the mathematical sciences, *Charles R. Hadlock*, Editor.
67. Innovative Approaches to Undergraduate Mathematics Courses Beyond Calculus, *Richard J. Maher*, Editor.
68. From Calculus to Computers: Using the last 200 years of mathematics history in the classroom, *Amy Shell-Gellasch and Dick Jardine*, Editors.
69. A Fresh Start for Collegiate Mathematics: Rethinking the Courses below Calculus, *Nancy Baxter Hastings*, Editor.
70. Current Practices in Quantitative Literacy, *Rick Gillman*, Editor.
71. War Stories from Applied Math: Undergraduate Consultancy Projects, *Robert Fraga*, Editor.
72. Hands On History: A Resource for Teaching Mathematics, *Amy Shell-Gellasch*, Editor.
73. Making the Connection: Research and Teaching in Undergraduate Mathematics Education, *Marilyn P. Carlson and Chris Rasmussen*, Editors.
74. Resources for Teaching Discrete Mathematics: Classroom Projects, History Modules, and Articles, *Brian Hopkins*, Editor.
75. The Moore Method: A Pathway to Learner-Centered Instruction, *Charles A. Coppin, W. Ted Mahavier, E. Lee May, and G. Edgar Parker*.
76. The Beauty of Fractals: Six Different Views, *Denny Gulick and Jon Scott*, Editors.
77. Mathematical Time Capsules: Historical Modules for the Mathematics Classroom, *Dick Jardine and Amy Shell-Gellasch*, Editors.
78. Recent Developments on Introducing a Historical Dimension in Mathematics Education, *Victor J. Katz and Costas Tzanakis*, Editors.
79. Teaching Mathematics with Classroom Voting: With and Without Clickers, *Kelly Cline and Holly Zullo*, Editors.
80. Resources for Preparing Middle School Mathematics Teachers, *Cheryl Beaver, Laurie Burton, Maria Fung, and Klay Kruczek*, Editors.
81. Undergraduate Mathematics for the Life Sciences: Models, Processes, and Directions, *Glenn Ledder, Jenna P. Carpenter, and Timothy D. Comar*, Editors.
82. Applications of Mathematics in Economics, *Warren Page*, Editor.
83. Doing the Scholarship of Teaching and Learning in Mathematics, *Jacqueline M. Dewar and Curtis D. Bennett*, Editors.

MAA Service Center  
P.O. Box 91112  
Washington, DC 20090-1112  
1-800-331-1MAA      FAX: 1-301-206-9789



# Contents

## Foreword    **xiii**

*David M. Bressoud*

## Preface    **xv**

*Jacqueline M. Dewar and Curtis D. Bennett*

## Part I. Getting Started in SoTL as a Mathematician

### 1. Understanding SoTL and Its Potential Benefits    **3**

*Jacqueline M. Dewar and Curtis D. Bennett*

- The Origins of SoTL in Higher Education    3
- Distinguishing SoTL from “Good” and “Scholarly” Teaching    4
- Forging Disciplinary Connections for SoTL    4
- SoTL’s Emergence in Mathematics    5
- SoTL and RUME    6
- Evaluating and Valuing SoTL    8
- Special Considerations for Junior Faculty Members    8
- The Benefits of SoTL    9

### 2. Initiating a SoTL Investigation    **13**

*Jacqueline M. Dewar and Curtis D. Bennett*

- A Typical Starting Point    13
- A Taxonomy of SoTL Questions    14
- Using Disciplinary Knowledge to Narrow a Question    15
- Undertaking a Literature Search    16

### 3. Gathering and Analyzing Evidence    **19**

*Jacqueline M. Dewar*

- Triangulating Data    19
- Challenges that Educational Research Design Presents for SoTL    19
  - Dealing with the Lack of a Control Group    20
  - Aligning the Assessment Measure with What Is Being Studied    22
  - To Seek Statistical Significance or Not?    22
  - Which Measure of Change?    23
- Quantitative versus Qualitative Data    23
- Data Sources Familiar to Most Mathematicians    24
  - Survey Design Considerations    25
    - Essential Steps in Designing a Survey    25
- Data Sources Less Familiar to Mathematicians    27
  - Knowledge Surveys    27
  - Think-Alouds    31
    - Advice for Conducting a Think-Aloud    32



- Focus Groups 33
      - Model for Conducting a Focus Group 34
  - Analyzing Qualitative Data 35
    - Rubrics 36
      - How to Create a Rubric 36
    - Coding Data (Content Analysis) 39
      - Process for Coding Data 40
      - Assessing the Quality of Content Analysis 41
- 4. Resources for Pursuing SoTL and Going Public 45**

*Jacqueline M. Dewar*

  - Human Subjects Considerations 45
  - Finding Collaborators and Support 46
  - Going Public with Results 47
    - Conferences 47
    - Journals 47

## Part II. Illustrations of SoTL Work in Mathematics

### *Theme 1: Experiments with Approaches to Teaching*

#### **5. Assessing the Effectiveness of Classroom Visual Cues 51**

*Gretchen Rimmasch and Jim Brandt*

*Can instructor use of visual cues improve student performance in recognizing and correctly applying procedural rules? Specifically, do the cues help students to simplify exponential expressions and calculate derivatives?*

SoTL Process—Pilot study; Control group; IRB; Evidence: pre- and post-tests and surveys; use of a rubric.

#### **6. (Re)discovering SoTL Through a Fundamental Challenge: Helping Students Transition to College Calculus 59**

*Rann Bar-On, Jack Bookman, Benjamin Cooke, Donna Hall, and Sarah Schott*

*Can algebra review and study group sessions improve student retention and performance in beginning calculus?*

SoTL Process—Evolution of the intervention and the study; Collaboration between faculty members and academic support staff; IRB; Evidence: participation rates, course grades.

#### **7. A Quantitative and Qualitative Comparison of Homework Structures in a Multivariable Calculus Class 67**

*Lynn Gieger, John Nardo, Karen Schmeichel, and Leah Zinner*

*How does the use of online homework affect student attitudes, motivation, and performance? Will it lead students to study more frequently, be more engaged and responsible for their learning, and lead to improved homework grades, test grades, or course grades?*

SoTL Process—Collaboration across multiple disciplines; Evidence: pre- and post-surveys, focus group, homework grades; Coding qualitative data.

**8. Playing Games to Teach Mathematics 77***Edwin Herman*

*Does the use of games in class for practice or review result in better learning or improve student attitudes and satisfaction?*

SoTL Process—Evolution of the intervention and the study; Quasi-control group experiment; Evidence: student evaluations, chapter exam grades, final exam grades.

**9. Investigating How Students' Linking Historical Events to Developments in Mathematics Changed Engagement in a History of Mathematics Course 87***Pam Crawford*

*Does the guided discovery teaching technique of having students investigate links between historical events and developments in mathematics improve students' engagement with historical material when compared with my previous teaching methods?*

SoTL Process—IRB, including previously collected data; Evidence: pre- and post-course surveys, exam essay questions; SoTL compared to RUME.

**Theme 2: Crafting Learning Experiences around Real-World Data or Civic Engagement****10 Using SoTL to Assess the Outcomes of Teaching Statistics Through Civic Engagement 99***Cindy Kaus*

*Will incorporating civic engagement into a general education statistics course increase retention and student confidence in their ability to do and communicate statistics?*

SoTL Process—Control groups and ethical concerns; Evidence: pre- and post-course surveys, student evaluations, course grades; Survey response rates.

**11. A Pedagogical Odyssey 107***Michael C. Burke*

*A vision of the possible: How would a mathematics course that used data-based integrative writing assignments affect students?*

SoTL Process—Evolution of the intervention and the study; Evidence: systematic analysis of essays, student reflective writing, final exams.

**12. Presenting Evidence for the Field That Invented the Randomized Clinical Trial 117***John Holcomb*

*A vision of the possible: Can team writing projects using real data lead to individual acquisition of desired data analysis skills?*

SoTL Process—Control groups and ethical concerns; IRB; Evidence: take home data analysis components of midterm and final exams, surveys; Survey response rates; Survey design issues; Coding qualitative data.

### *Theme 3: Using Assigned Reading Questions to Explore Student Understanding*

#### **13. Conceptual or Computational? Making Sense of Reading Questions in an Inverted Statistics Course** 127

*Derek Bruff*

*In an introduction to statistics course for engineering majors, what are students able to learn by reading their textbooks before class? What kinds of pre-class reading assignments, including questions about the reading, might help students learn more from reading their textbooks before class?*

SoTL Process—Quasi-control group experiment; Shifting from a *What works?* to a *What is?* investigation and back again; IRB, including previously collected data; Evidence: pre- and in-class quizzes, survey, whole class interview; Coding qualitative data; Bloom's taxonomy.

#### **14. An Investigation Into the Effectiveness of Pre-Class Reading Questions** 137

*Mike Axtell and William Turner*

*What types of pre-class reading questions best facilitate independent learning, and the retention of the material learned, in our students?*

SoTL Process—Narrowing a research question; Literature search; Evidence: pre-class reading questions, in-class reading quizzes.

### *Theme 4: Exploring Student Understanding of the Nature of Mathematics*

#### **15. Liberal Arts Mathematics Students' Beliefs About the Nature of Mathematics: A Case Study in Survey Research** 145

*Stephen D. Szydluk*

*What do students enrolled in a problem-based general education mathematics course believe about the nature of mathematics? Do these beliefs change from the beginning to the end of the course?*

SoTL Process—IRB process and Informed Consent Form; Evidence: pre- and post-survey; Survey development, validity, and reliability; Coding qualitative data; Choosing a publication venue.

#### **16. The Mathematics of Symmetry and Attitudes towards Mathematics** 157

*Blake Mellor*

*How does a course on the mathematics of symmetry change the perspectives on and attitudes towards mathematics in a class for students majoring in the liberal arts?*

SoTL Process—Evidence: pre- and post-surveys, student teaching evaluations; Survey response rate and administration issues; Coding qualitative data.

#### **17. Mathematics Research Experiences for Preservice Teachers: Investigating the Impact on Their Beliefs** 171

*Wendy A. O'Hanlon, David D. Barker, Cynthia W. Langrall, John A. Dossey, Sharon M. McCrone, and Saad I. El-Zanati*

*Prior to a Research Experience for Undergraduates (REU) what are the beliefs of the preservice teacher participants about the teaching and learning of mathematics and how do their beliefs change after participating in the REU experience?*

SoTL Process—Evidence: pre- and post-surveys; Survey development and reliability.

**Theme 5: Tackling Large Questions****18. The Question of Transfer: Investigating How Mathematics Contributes to a Liberal Education 183***Curtis D. Bennett and Jacqueline M. Dewar**How does mathematics contribute to liberal learning?*

SoTL Process—Narrowing a research question; Evidence: surveys, interviews with students and faculty members, focus group; Coding qualitative data.

**19. Using SoTL Practices to Drive Curriculum Development 191***Rikki Wagstrom**How might integrating civic issues into a college algebra prerequisite course improve quantitative reasoning, algebraic skill development, and student confidence and interest in studying mathematics?*

SoTL Process—Refining a question; Control groups and ethical concerns; Pilot study; Literature search; IRB; Evidence: pre- and post-test questions, surveys; Choosing a publication venue.

**Epilogue****20. Synthesis of the Value and Benefits of SoTL Experienced by the Contributors 203***Curtis D. Bennett and Jacqueline M. Dewar***Index 207**



# Foreword

David M. Bressoud  
*Macalester College*

In our progression as teachers of mathematics, first comes good teaching. This includes learning how to build rapport with one's students and to motivate them by finding what is exciting and sharing it. It entails seeking ways to challenge, encourage, and support them. That is a powerful start, but, eventually, a good teacher is not content with attentive faces and good teaching evaluations. Eventually, a really good teacher realizes that many of his or her students are having difficulty with fundamental concepts that had been so clearly explained and seemingly firmly fixed in their minds but then somehow were lost.

At the next stage, we become reflective teachers. We try to understand what went wrong, how we might reach more of our students more effectively. We seek advice from more experienced colleagues, learn about other approaches to the teaching of a given topic, and experiment within our own classrooms. Reflective teachers are constantly learning about teaching, adjusting their courses, seeking to do even better at challenging, encouraging, and supporting their students.

There is a third stage: scholarly teaching. A reflective teacher might discover an approach to a lesson or a course that really works. They want to share it. A scholarly teacher seeks to back this up with evidence that it is working and has the potential to work for others. More than that, a scholarly teacher wants to understand *why* it works. No one else is going to pick up your insights and implement them precisely the way you did. A scholarly teacher tries to identify the critical core of what makes this approach work, so that those who would adopt it know what is essential and what can be adapted.

There is another motivation for scholarly teaching. It comes when nothing seems to work. Scholarly teaching can arise from the decision to confront pedagogical or curricular difficulties as an intellectual problem worthy of full attack. This involves studying what actually happens in the classroom and exploring why students are having trouble. Framing the appropriate question can involve serious scholarship. Often this requires work aimed at obtaining a deeper understanding of student difficulties. It includes a thorough search of the relevant literature to refine the question and to place it in the context of what is known. It requires careful design of the investigation with attention to how the data, either quantitative or qualitative, will be collected and analyzed.

When made public, the fruits of scholarly teaching constitute the Scholarship of Teaching and Learning (SoTL). This volume is about how to do it, but it also describes insights gained across a broad range of topics from the use of Reading Questions to the incorporation of Civic Engagement to grappling with the issue of exactly how Mathematics contributes to a Liberal Education. It has been edited by two of the experts in the field and includes fifteen accounts by practitioners who explain both how they have engaged in this scholarship and what they have learned. The result is a description of the process of scholarly teaching in multiple voices and from multiple perspectives. In their concluding Epilogue, the editors draw from all of their contributors to explain how this scholarship can help us avoid the common pathologies of teaching: *amnesia*—the tendency to forget what has been successful and what has not, *fantasia*—the lack of understanding of what actually transpires in our classrooms, and *inertia*—the reluctance to change that which is not working.

This book is filled with great ideas and insights. I borrowed the idea of a progression from good teaching to reflective teaching to scholarly teaching from the writers of Chapter 6. For those who might be uncertain about how SoTL differs

from Research in Undergraduate Mathematics Education (RUME), there are excellent discussions of the distinctions by the editors in Chapter 1 and by Pam Crawford in Chapter 9.

It is impossible to discuss the development of SoTL without acknowledging the role of Brian Winkel. I want to close this foreword with a personal note of gratitude. He was the founding editor of *PRIMUS*, one of the foremost venues for the publication of the results of scholarly teaching in mathematics. He established it in 1991, as SoTL was just beginning to emerge within the mathematical community, and guided it for twenty years, for most of that time without any editorial assistance. To qualify as scholarship, SoTL requires peer review. While there are opportunities for such peer review, including MAA's own journals, they are limited, and would be much more so without *PRIMUS*. Brian has been tireless in his encouragement of scholarly teaching and insistent on maintaining the standards of scholarship. All of us who aspire to scholarly teaching owe him a tremendous debt.

# Preface

Jacqueline M. Dewar and Curtis D. Bennett  
*Loyola Marymount University*

The scholarship of teaching and learning is a growing field of inquiry in which faculty members bring disciplinary knowledge to investigate questions of teaching and learning and systematically gather evidence to support their conclusions. Submitting the results to peer review and making them public for others to build on have generally become expected components of the scholarship of teaching and learning, or SoTL. Individual faculty members, their students, departments, and institutions can all benefit from this work. As one of the contributors to this volume (Edwin Herman, Chapter 8) observes: “The process of doing SoTL research can be even more important than the results obtained. Framing and researching a question and designing a project encourages the researcher to experiment within the classroom, much as a painter experiments with styles on a canvas. As the project progresses, the question (or questions) become more refined, more interesting, and the answers can both inform and improve the way you teach” (p. 83).

This Notes volume is written for collegiate mathematics instructors who want to know more about conducting scholarly investigations into their teaching and their students’ learning. Faculty members in related disciplines, such as engineering, computer science, or the sciences should also find the book of interest, as should high school mathematics teachers. Conceived and edited by two mathematics faculty members, the volume serves as a how-to guide for doing SoTL in mathematics. It contains information and resources for undertaking scholarly investigations into teaching and learning, and includes many examples.

SoTL is a topic of increasing interest in the mathematics community. Well-attended MAA minicourses on how to get started in SoTL have been presented at the 2006, 2007, and 2008 Joint Mathematics Meetings and at the 2009 MathFest. Successful SoTL paper sessions have been offered at the joint mathematics meetings annually since 2007. Project NExT fellows have shown interest in the topic by inviting speakers to address SoTL on panels they organized for the 2009, 2011, 2012, and 2013 Joint Mathematics Meetings.

The book is divided into two parts followed by an epilogue. The four chapters in Part I provide background on this form of scholarship and specific instructions, a how-to guide, for undertaking a SoTL investigation in mathematics. The authors of Part I (the editors of this volume) target their advice to mathematicians. As two mathematicians who learned to do SoTL in the Carnegie Scholars program of the Carnegie Academy for the Scholarship of Teaching and Learning, they are able to write from experience about undertaking SoTL work without a formal background in educational research.

Chapter 1 presents the history of the SoTL movement and differentiates SoTL from related ideas (good teaching and scholarly teaching) and work (research in mathematics education). It describes the benefits of SoTL for faculty members, their departments, and their institutions. The rewards include improving the teaching and learning of individuals, providing evidence of effective teaching or scholarly publications for tenure and promotion portfolios, and making contributions to departmental and institutional initiatives.

Chapter 2 takes the reader through the initial steps of a typical SoTL investigation, beginning with a question or problem about teaching and learning and describing how to reframe it into a researchable question. It shows faculty members how to draw upon their disciplinary knowledge and teaching experience to advance their investigations and how to use the taxonomy of SoTL questions (*What is? What works? What could be?*) to monitor the development of their research questions. It offers suggestions for doing searches into education literature and explains how they can aid in launching a project.

Chapter 3 is about the design of SoTL studies. A significant portion of this chapter addresses various types of evidence used in SoTL work, including methods that may not be familiar to mathematicians, such as focus groups



and think-alouds. It discusses ways to analyze the qualitative data that result from these methods. It also contains guidelines for designing surveys and creating rubrics for use in SoTL studies.

Chapter 4 provides information and resources for pursuing SoTL and going public with findings. It addresses a practical issue unfamiliar to many teaching mathematicians, namely needing to obtain human subjects clearance. It offers suggestions for finding collaborators and advice on locating venues to present or publish SoTL work.

Part II contains fifteen examples of SoTL projects in mathematics from fourteen institutions, both public and private, spanning the spectrum of higher educational institutions from community colleges to research universities. Among the twenty-five contributing authors are four Carnegie scholars, three former participants in the MAA minicourses on Scholarship of Teaching and Learning, two faculty members with training in mathematics education, and several collaborating non-mathematicians. Their projects are rooted in a variety of mathematical topic areas: remedial mathematics, quantitative literacy, mathematics for liberal arts, pre-service teacher preparation, a freshman course for mathematics majors focused on mathematical problem solving and communication, precalculus, first-semester calculus, multivariable calculus, and statistics. The projects cover the gamut of methodologies; that is, there are examples of SoTL work that use both quantitative and qualitative methods, as well as SoTL work that involves the three different types of SoTL questions.

The purpose of the chapters in Part II is to reveal the process of doing SoTL. They serve as models for carrying out a SoTL investigation, the components of which are outlined in Part I. The authors describe how their studies began, and the design decisions they made. The authors are candid about the difficulties they encountered and the limitations of their work. They discuss lessons learned about doing SoTL and offer recommendations. An Editors' Commentary prefaces each chapter to highlight certain aspects of the process of doing SoTL revealed by its authors. Each chapter reports the benefits that accrued to the authors and their careers from engaging in SoTL.

The fifteen illustrations of SoTL are grouped into five themes:

- Experiments with Approaches to Teaching
- Crafting Learning Experiences around Real-World Data or Civic Engagement
- Using Assigned Reading Questions to Explore Student Understanding
- Exploring Student Understanding of the Nature of Mathematics
- Tackling Large Questions

This volume is intentionally interconnected. The chapters in Part II illustrate many of the concepts, issues, methods, and procedures discussed in Part I. For example, when the taxonomy of SoTL questions (Chapter 2), the construction and use of a rubric (Chapter 3), or the institutional review process for human subjects research (Chapter 4) is discussed, the reader is referred to chapters in Part II for additional descriptions derived from the authors' experiences. The Table of Contents previews each chapter by listing the questions being investigated and the SoTL process topics addressed.

In the Epilogue (Chapter 20) we present a synthesis of the authors' perceptions of the value of SoTL. While our primary goal is to assist mathematics instructors interested in undertaking a scholarly study of their teaching practice, a secondary goal is to promote a greater understanding of SoTL work and its value to the mathematics community. This final chapter allows us to reflect on the outcomes and benefits that accrued to the 25 authors as a result of their scholarly inquiries into teaching and learning.

### Acknowledgements

We would like to express our gratitude to the journal *PRIMUS: Problems, Resources, Issues in Mathematics Undergraduate Studies* for allowing us to draw from and expand our 2012 article, "An Overview of the Scholarship of Teaching and Learning in Mathematics" (*PRIMUS*, 22(6), 458–473). We want to thank our contributing authors for providing the wonderful illustrations of SoTL work in Part II. We are extremely grateful to Dr. Stephanie August, Associate Professor of Computer Science at Loyola Marymount University, for reading and commenting on the initial draft of Part I, and to Dr. Stephen Maurer, Notes Series Editor during the development of this book, who by patiently and promptly answering all our questions assisted us all along the way.



# **Getting Started in SoTL as a Mathematician**



# 1

## Understanding SoTL and Its Potential Benefits

Jacqueline M. Dewar and Curtis D. Bennett  
*Loyola Marymount University*

### Introduction

Scholarship of teaching and learning (SoTL) is a scholarly activity whose history is generally not well known to teaching or research mathematicians. Many activities are labeled SoTL, some appropriately and others not. In light of this, Chapter 1 has several goals. It aims to inform the reader about the origins of the scholarship of teaching and learning, the efforts to forge connections between SoTL and academic disciplines, and the emergence of SoTL within mathematics. It attempts to set SoTL apart from good teaching, scholarly teaching, and, to the extent possible, research in undergraduate mathematics education (RUME). The chapter addresses the issue of evaluating and valuing this work for tenure and promotion, a matter of great concern for junior faculty members. It closes with a brief discussion of the benefits of SoTL, a topic that we revisit in the Epilogue, Chapter 20, where we present a synthesis of the benefits our authors experienced from participating in SoTL.

### The Origins of SoTL in Higher Education

In 1990, Ernest Boyer, President of the Carnegie Foundation for the Advancement of Teaching, introduced the expression “scholarship of teaching” into the vocabulary of higher education. His book, *Scholarship Reconsidered* (Boyer, 1990), called for colleges and universities to embrace a broader vision of scholarship in order to tap the full range of faculty talents across their entire careers and to foster vital connections between academic institutions and their surrounding communities. Boyer argued for the recognition of four types of scholarship: discovery, application, integration, and teaching. The scholarship of discovery refers to what is traditionally called research in most disciplines. The scholarship of application, now frequently called scholarship of engagement, refers to applying knowledge to consequential problems, often conducted with and for community partners. The scholarship of integration makes connections between disciplines. More and more, interdisciplinary work is being recognized as essential to solving complex real world problems. These scholarships may have different interpretations depending on the discipline and type of institution. For example, for engineers, consulting work is often considered scholarship of application. Boyer’s description of the fourth form, the scholarship of teaching, contained many of the characteristics of what is now called the “scholarship of teaching and learning,” or “SoTL,” but it failed to include peer review and making results public. Later on, for many in the SoTL movement, a fully developed definition of SoTL included these elements

(Hutchings & Shulman, 1999; Richlin, 2001, 2003; Smith, 2001). So both the name and the concept of this form of scholarship have evolved.

As President of the Carnegie Foundation for the Advancement of Teaching, Boyer was able to bring national and international attention to SoTL, but others had discussed similar concepts prior to the publication of his book. For example, Cross (1986) had argued that faculty should undertake research on teaching and learning in their classrooms in order to discover more effective teaching methods and to establish a body of knowledge about college teaching that would maximize learning. More than twenty years later, calls to do this continue (Schmidt, 2008).

After Boyer, the next Carnegie President, Lee Shulman, along with Carnegie Vice President Pat Hutchings stated that the scholarship of teaching is integrating the experience of teaching with the scholarship of research (Hutchings & Shulman, 1999). More recently, Hutchings and Carnegie Senior Scholar Mary Huber (Huber & Hutchings, 2005) have “come to embrace a capacious view of the topic, wanting to draw this movement in the broadest possible terms” (p. 4). In their view, SoTL could range from modest investigations that document the teaching and learning in a single classroom to broad studies with elaborate research designs.

We define SoTL as follows:

SoTL is the intellectual work that faculty members do when they use their disciplinary knowledge (in our case, mathematics) to investigate a question about their students’ learning (and their teaching), submit their findings to peer review, and make them public for others to build upon.

This definition emphasizes the investigation into student learning over teaching practice, as that shift was essential to the transformation of Boyer’s scholarship of teaching into Carnegie’s scholarship of teaching and learning. As Hutchings and Shulman (1999) wrote: “Indeed, our guidelines for the Carnegie Scholars program call for projects that investigate not only teacher practice but the character and depth of student learning that results (or does not) from that practice” (p. 13).

## Distinguishing SoTL from “Good” and “Scholarly” Teaching

One of the sources of confusion about SoTL and its value to higher education is the lack of clear distinctions among three overlapping concepts: “good teaching,” “scholarly teaching,” and the “scholarship of teaching and learning.” These terms have many interpretations, even in this volume. We rely on Smith (2001), who wrote that good (or better) teaching is defined and measured by the quality of student learning, while scholarly teaching means something else. Scholarly teachers and their teaching must be informed not only by the latest developments in their fields but also by research about instructional design and how students learn. Based on this research, scholarly teachers make choices about instruction and assessment. When they do more than simply draw on this research to become scholarly teachers, but also seek to contribute to the knowledge base by carrying out research on teaching and learning, then they are engaging with the scholarship of teaching and learning.

The preceding discussion represents but one way to parse these terms and does not even address another closely related term, “reflective teaching.” Regardless of how the terms are defined, in the end SoTL research usually involves aspects of discovery, application, and integration (Boyer, 1990) and is undertaken to improve practice, both within and beyond the investigators’ classrooms. To be SoTL, the inquiry must also satisfy the three additional features of being public, open to critique and evaluation by peers, and in a form that others can build on.

What Laurie Richlin wrote in 2001 remains true today: “Although a decade has passed since the idea of a scholarship of teaching entered the lexicon of American higher education, the concept remains intertwined with the activities of scholarly teaching. Only by separating the different activities and focusing on the scholarly process can we give each the honor and rewards it deserves” (p. 87).

## Forging Disciplinary Connections for SoTL

In 1998, Carnegie initiated new programs to promote SoTL under the name CASTL (Carnegie Academy for the Scholarship of Teaching and Learning). The first was the CASTL Scholars program, which selected 158 post-secondary faculty members, many distinguished researchers in their disciplines, to populate six cohorts of CASTL scholars over

nine years. They worked on individual scholarship of teaching and learning projects and many went on to become leaders in the SoTL movement. Carnegie also initiated the Scholarly and Professional Societies Program to encourage recognition of SoTL by the disciplines.

During the same period several Carnegie publications and speeches approached SoTL via the disciplines (Huber & Morreale, 2002; Shulman, 2005). In the summer of 2000, Carnegie invited representatives of disciplinary societies to a meeting and offered grant money for disciplinary action in SoTL. Two dozen scholarly societies participated, including the Mathematical Association of America (MAA). The MAA sent two representatives (Executive Director Tina Straley and President Tom Banchoff, himself a 1999 CASTL scholar), and the MAA applied for and received a grant.

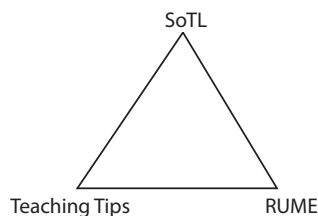
Halfway through the decade-long CASTL scholars program, Carnegie began two new programs aimed at building support for SoTL at the institutional level, the Campus Cluster program (2003–06) and the Leadership program (2006–09). Over 250 colleges and universities participated. *Scholarship of Teaching and Learning Reconsidered* (Hutchings, Huber, & Ciccone, 2011) described in more detail the succession of programs that ran under Carnegie's long-term initiative to promote Boyer's vision of teaching as scholarly work. By the time the CASTL initiative came to a formal close in 2009, SoTL had become an increasingly international movement. The International Society for the Scholarship of Teaching and Learning ([www.issotl.org](http://www.issotl.org)), founded in 2004, has arguably become the leading international society promoting and supporting the scholarship of teaching and learning. Meanwhile, interest in improving undergraduate education, especially in STEM (Science, Technology, Engineering, and Mathematics) fields led national groups such as the National Academies (Board on Science Education, n.d.), the American Association of Colleges and Universities (Project Kaleidoscope, n.d.), and the National Center for Science and Civic Engagement (n.d.) to shared agendas that draw on or contribute to the advancement of the scholarship of teaching and learning. (We use the American Psychological Association's APA citation style throughout this publication as that is common practice for SoTL and RUME publications. The notation "n.d." appearing in the previous sentence indicates a publication with "no date" in APA style.)

## SoTL's Emergence in Mathematics

SoTL has a growing presence in the mathematical community. Of the six cohorts of CASTL scholars selected by Carnegie, a total of 12 mathematicians were present in the second (1999), third (2000), fifth (2003), and sixth (2005) cohorts. Among the eight mathematicians in the 1999 and 2000 cohorts were the then president of the MAA, three members of major committees of the American Mathematical Society (AMS) and the MAA, one recipient of the MAA's distinguished teaching award, and a winner of the CASE-Carnegie national teacher of the year award. Four were from doctoral granting institutions. None of the four mathematicians in the final two CASTL cohorts were from doctoral granting institutions. This supports the theory that CASTL initially focused on recruiting individuals well-connected within their disciplines and preparing them to do SoTL work, so as to position highly regarded SoTL ambassadors in the disciplines. For more on this topic, see Dewar and Bennett (2010).

The last eight years have witnessed a number of SoTL-related activities in mathematics aimed at increasing awareness and understanding of SoTL, attracting new practitioners, and providing venues for dissemination. An MAA minicourse (four hours spread over two days) entitled "Beginner's Guide to the Scholarship of Teaching and Learning in Mathematics" has been offered at a national meeting of the Mathematical Association of America in 2006, 2007, 2008, and 2009. More than 100 mathematics faculty members from all types of institutions and a few graduate students have attended.

To provide a venue for presenting SoTL work at the joint mathematics meetings, two of the minicourse organizers (the editors of this volume) decided to propose a SoTL paper session. In 2006 in collaboration with several others, they began submitting proposals for SoTL paper sessions. Despite competition for slots on the program, a SoTL paper session has run six times (2007, 2010, 2011, 2012, 2013, and 2014) as a contributed paper session sponsored by the MAA and open to submissions from anyone. In addition, a SoTL paper session was offered twice as a special session jointly sponsored by the AMS and the MAA (2008 and 2009), where most, but not all, of the slots were filled by invitation, and once as an invited paper session co-sponsored by the MAA and the AMS (2010), where all the speakers were invited. Some topics presented in the early SoTL paper sessions, such as getting students to read the



**Figure 1.1.** The teaching commons

text, anticipated growing interest on the part of the mathematical community and later appeared as the specific focus of contributed paper sessions run by others.

Project NExT, an organization for mentoring junior mathematics faculty members, has shown interest in SoTL. At the 2009 joint mathematics meetings, NExT sponsored a panel titled “Engaging in and Publishing the Scholarship of Teaching and Learning” with speakers from the SoTL and the RUME communities. Between 2011 and 2013, NExT sponsored three panels on or including the topic of SoTL: “Getting involved in the Scholarship of Teaching and Learning,” “Turning Teaching into Scholarship,” and “Scholarship of Teaching and Learning: Ideas to Publication.” Interest in SoTL on the part of junior faculty members in mathematics community bodes well for SoTL’s future in mathematics.

## SoTL and RUME

As the SoTL movement unfolded and called for faculty to treat teaching and learning in a scholarly fashion, mathematics already had an existing community of scholars doing educational research into college level mathematics teaching. As previously noted, this community is called RUME, for Research in Undergraduate Mathematics Education (see [sigmaa.maa.org/rume](http://sigmaa.maa.org/rume)). Currently there are 75 doctoral programs in mathematics education in the United States, 70 of which grant a Ph.D. (as opposed to an Ed.D.) degree (Hsu, 2011). In 2000, RUME was the first Special Interest Group to be officially recognized by the MAA (see [sigmaa.maa.org/sigmaahistory.html](http://sigmaa.maa.org/sigmaahistory.html)). RUME is set off from disciplinary mathematical research by its very different questions of interest, methodologies, and epistemologies (Schoenfeld, 2000), but how it differs from SoTL carried out in mathematics may not be so clear.

Mathematicians and Carnegie scholars Tom Banchoff and Anita Salem (2002) saw SoTL as potentially bridging the gap between RUME and teaching mathematicians, but we prefer a different metaphor: SoTL as a broadened landscape of scholarly work related to teaching. Our model, shown in Figure 1.1, places the labels “Teaching Tips,” “SoTL,” and “RUME” at the three vertices of an equilateral triangle. The result is a space that can be populated by all sorts of work, including that of the “Math Tech Ed” community that is concerned with technology in teaching and learning mathematics. (Within the MAA, a contact for this community is the Special Interest Group SIGMAA WEB, which is concerned with the use of the world wide web and other technologies for teaching mathematics; see [sigmaa.maa.org/web](http://sigmaa.maa.org/web). Also, the MAA’s Mathematical Sciences Digital Laboratory at [mathdl.maa.org/mathDL](http://mathdl.maa.org/mathDL) provides resources for teaching with technology.) Our description of this shared space aligns with Huber and Hutchings’ (2005) view of SoTL as increasing the teaching commons, “a conceptual space in which communities of educators committed to inquiry and innovation come together to exchange ideas about teaching and learning and use them to address the challenges of educating students for personal, professional, and civic life” (p. x).

In our model, teaching tips refers to a description of a teaching method or innovation that an instructor reports having tried “successfully” and that the students “liked.” If the instructor begins to systematically gather evidence from students about what, if any, cognitive or affective effect the method had on their learning, she is moving toward scholarship of teaching and learning. When this evidence is sufficient to draw conclusions, and those conclusions are situated in the literature, peer reviewed, and made public, the instructor has produced a piece of SoTL work. However, the work may differ in form and scope from that found at the third vertex of the triangle in Figure 1.1.

Mathematics education research or RUME is more in line with Boyer’s “scholarship of discovery” wherein research methodologies, theoretical frameworks, empirical studies, and reproducible results would command greater importance. This naturally influences the questions asked or considered worth asking, the methods used to investigate them,

and what the community accepts as valid. The work of Schwab (1964) on the structure of the disciplines provides four ways of identifying and describing the differences between SoTL and RUME:

1. *Content boundaries:* RUME is focused on mathematical knowledge and mathematical learning, and interdisciplinary in that it draws on studies in education, psychology, sociology, anthropology, and linguistics as well as its own body of literature. SoTL, while being interdisciplinary and having a growing body of literature, is a multidisciplinary community with practitioners in disciplines ranging from the arts and humanities, to business, to the sciences and engineering. Mathematicians in the SoTL community encounter SoTL investigations in these other disciplines, encouraging their interest in questions that cross boundaries, for example, questions about student voice (that is, student perspectives and insights on learning, teaching, and curriculum and considerations of how to incorporate them into decision-making processes) or the impact of community-based learning experiences.
2. *Skills and habits employed by practitioners:* SoTL questions virtually always come from practice. SoTL researchers encounter a teaching or learning problem they want to fix, a question they want to answer, or a phenomenon they desire to understand or describe in depth. Sometimes SoTL practitioners simply aim to show what is possible in a certain situation. (See, for example, Chapter 12 by John Holcomb.) RUME researchers may happen upon a question through their teaching practice, but more frequently their questions arise from theory or methodological concerns. Doctoral training in mathematics education research gives them a formal background in theoretical frameworks for mathematical knowledge and learning to draw upon. In contrast, SoTL researchers may utilize the taxonomy of SoTL questions—*What works? What is? What could be?*—to refine their questions (Hutchings, 2000). (The process of developing and refining SoTL research questions by using this taxonomy and disciplinary knowledge is the focus of Chapter 2.) They may call on the theoretical frameworks developed by the RUME community and must, of course, appropriately place their work in the context of previous studies in order to publish it.
3. *Modes of inquiry:* Compared to the typical SoTL practitioner, the typical RUME researcher has far more expertise and knowledge about research methodologies and ways of communicating methodological concerns. Still, the SoTL researcher seeks to triangulate her evidence, that is, to use data from several sources, in support of a conclusion, and may draw on quantitative and qualitative data. (Chapter 3 is intended to inform teaching mathematicians about methods for gathering and analyzing evidence that could prove useful in a SoTL study.)
4. *Purposes or outcomes for the disciplinary work:* The SoTL practitioner seeks to answer her question about teaching and learning that most likely arose in her practice, and by making the answer public she seeks to improve practice and student learning elsewhere. As a research discipline of its own, RUME seeks to increase what is known about mathematics learning at its deepest level and to develop and extend theories about that learning. RUME also seeks to improve teaching practice and student learning.

The intended audience is another way to differentiate SoTL from mathematics education research. Since its inception SoTL has been aimed at instructors in all fields of higher education, with the intent to improve teaching and learning and to provide a basis for others to build upon. As a research field of its own, other RUME researchers naturally constitute the primary audience for original RUME work. But, the line of demarcation can be blurry. Two members of the RUME community, Marilyn Carlson and Chris Rasmussen (2008), edited a volume in the MAA Notes series, *Making the Connection: Research and Teaching in Undergraduate Mathematics Education*, which is “intended for mathematicians and mathematics instructors who want to enhance the learning and achievement of students in their undergraduate mathematics courses” (p. vii). Of course, almost no form of scholarship fits neatly into any one camp. SoTL work has been presented at RUME conferences and RUME researchers have presented work in SoTL paper sessions at the national mathematics meetings. Thus the current situation in the mathematical community seems to support Huber and Hutchings (2005) in their claim that the teaching commons can be a big tent whose purpose is to improve teaching and learning as a whole.

One of our contributors, Pam Crawford, holds a doctorate in mathematics with a concentration in teaching collegiate mathematics, which gave her training and experience in RUME. She also participated in her university’s SoTL scholars program. In Chapter 9, she draws on her dual background to comment on the ways her SoTL project differed from her RUME dissertation.



## Evaluating and Valuing SoTL

In mathematics departments, SoTL-type work is not traditionally viewed as disciplinary research (unless one is working in mathematics education), so the question naturally arises in mathematics just as it does in other disciplines: By what standards should SoTL work be evaluated?

The question is not new. Soon after the publication of *Scholarship Reconsidered* (Boyer, 1990), it became clear that an essential piece of promoting a broader definition of scholarship was missing. The effort to broaden the meaning of scholarship simply could not succeed until institutions had clear standards for evaluating this wider range of scholarly work. Faculty and administrators accord full academic value only to work they can confidently judge, through a process usually involving peer review. That prompted senior staff members at the Carnegie Foundation to undertake a new project. They collected and examined hundreds of documents (tenure and promotions guidelines, federal grant and private foundation funding criteria, academic journal submission guidelines, etc.) in an attempt to distill a common set of standards. The result, *Scholarship Assessed* (Glassick, Huber, & Maeroff, 1997), became another foundational SoTL publication. According to Glassick et al. (1997), the most remarkable feature of this collection of guidelines was how much they shared common elements. Their synthesis of these materials provided a clear response to the problem of how to judge SoTL: SoTL, or any of the other nontraditional forms of scholarship, is to be judged by the same criteria as the traditional scholarship of discovery. Their book delineated six standards for assessing any scholarly work, and provided questions to flesh out the meaning of each standard:

1. Clear goals—Does the scholar state the basic purposes of his or her work clearly? Does the scholar define objectives that are realistic and achievable? Does the scholar identify important questions in the field?
2. Adequate preparation—Does the scholar show an understanding of existing scholarship in the field? Does the scholar bring the necessary skills to his or her work? Does the scholar bring together the resources necessary to move the project forward?
3. Appropriate methods—Does the scholar use methods appropriate to the goals? Does the scholar apply effectively the methods selected? Does the scholar modify procedures in response to changing circumstances?
4. Significant results—Does the scholar achieve the goals? Does the scholar's work add consequentially to the field? Does the scholar's work open additional areas for further exploration?
5. Effective presentation—Does the scholar use a suitable style and effective organization to present his or her work? Does the scholar use appropriate forums for communicating work to its intended audiences? Does the scholar present his or her message with clarity and integrity?
6. Reflective critique—Does the scholar critically evaluate his or her own work? Does the scholar bring an appropriate breadth of evidence to his or her critique? Does the scholar use evaluation to improve the quality of future work? (Glassick et al., 1997, p. 36)

This widely cited work has, in theory, answered the question of how SoTL work can be evaluated.

## Special Considerations for Junior Faculty Members

Still, how much individual colleagues, departments, and institutions count SoTL as research in tenure and promotion decisions varies widely. At some institutions SoTL will be relegated to the teaching section of a tenure dossier where it can perform a useful function as evidence of teaching effectiveness and of going beyond scholarly teaching. Others may consider a SoTL article as little more than an addendum to a required level of traditional disciplinary research, especially when published in an interdisciplinary SoTL journal rather than in a mathematics or mathematics education publication. Because of the lack of SoTL publishing opportunities in mathematics journals (Dewar & Bennett, 2010), the downgrading based on publication venue is a problem, particularly for pre-tenure faculty members. On the other hand, some institutions fully embrace SoTL's inclusion as research (see, for example, the comments of Lynn Gieger and her colleagues in Chapter 7 and Cindy Kaus in Chapter 10). Since the conversation initiated by Boyer in 1990 to move beyond a narrow definition of research has yet to come to a full and satisfying conclusion (O'Meara & Rice, 2005), pretenure faculty members should seek guidance and clarity from their departments and institutions on how SoTL work will be valued.

## The Benefits of SoTL

The landscape of higher education has changed tremendously in the last several decades. Our student bodies have become more diverse, with ever-larger percentages of high school students entering college (KewalRamani, Gilbertson, Fox, & Provasnik, 2007). Technology has offered many new options for instruction. Neuroscience has made discoveries about the biological basis of learning (Bransford, Brown, & Pellegrino, 2000; Doyle & Zakrajsek, 2013; Leamson, 1999; Zull, 2002). Calls have been made for the use of “high impact practices” (e.g., community-based learning, writing intensive courses, undergraduate research) to promote more engaged learning (Brownell & Swaner, 2010; Kuh, 2008). Each of these developments has implications for teaching that open avenues for SoTL investigations and for improving practice.

SoTL has something to offer faculty and their institutions as they wrestle with the developments just mentioned. Participation in SoTL promotes more reflective teaching and improved teaching effectiveness. It is common for the details of what was good and bad in student work from previous semesters to fade from an instructor’s memory. But, reflecting on student work from a SoTL perspective helps instructors capture the details. By asking and answering SoTL questions, faculty members can find out how well they are teaching and their students are learning and gain insights for making improvements in their classrooms. As a consequence, SoTL offers a means other than student or peer evaluations to document teaching effectiveness and student learning in applications for merit, tenure, or promotion. On a more personal level, SoTL investigations can be energizing and deeply rewarding to faculty members at all stages of their careers, in part because of the collegial connections SoTL fosters across disciplines and institutions.

Hutchings, Huber, and Ciccone (2011) presented evidence for SoTL’s impact on classroom teaching, professional development, institutional assessment, and the recognition and reward of pedagogical work. SoTL can lead to institutional involvement in national or even international higher education initiatives as described in Valleau (2010). Publicly embracing SoTL is one way in which an institution can demonstrate its student-centeredness. In addition, because faculty members who ask and attempt to answer SoTL questions have to gather and analyze evidence that goes beyond grades on assignments and tests, they can help drive institutional assessment efforts to be a more meaningful process aimed at curriculum development and pedagogical improvement. In particular, when discussing curricular issues in a department, SoTL work can be a worthy ally. All in all, it seems that SoTL provides a method for addressing a number of the challenges facing mathematics departments and the higher education community in general.

We have just offered general examples of the benefits of SoTL. The fifteen chapters in Part II describe SoTL projects, and the authors explain how participating in their projects informed or benefitted their teaching, their careers, their departments, or other institutional efforts. In the Epilogue (Chapter 20) we present a synthesis and summary of their views. We encourage you to read and take into account their perspectives on the rewards of participating in SoTL.

## References

- Association of American Colleges and Universities. (n.d.) *Project Kaleidoscope (PKAL)—Advancing what works in STEM education*. Retrieved from [www.aacu.org/pkal/index.cfm](http://www.aacu.org/pkal/index.cfm)
- Banchoff, T., & Salem, A. (2002). Bridging the divide: Research versus practice in current mathematics teaching and learning. In M. T. Huber & S. P. Morreale (Eds.), *Disciplinary styles in the scholarship of teaching and learning: Exploring common ground* (pp. 181–196). Washington, DC: American Association for Higher Education.
- Board on Science Education. (n.d.). *Status, Contributions, and Future Direction of Discipline-Based Education Research (DBER)*. Washington, DC: The National Academies. Retrieved from [sites.nationalacademies.org/DBASSE/BOSE/DBASSE.072106](http://sites.nationalacademies.org/DBASSE/BOSE/DBASSE.072106)
- Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. San Francisco, CA: Jossey-Bass.
- Bransford, J., Brown, A., & Pellegrino, J. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Brownell, J., & Swaner, L. (2010). *Five high-impact practices: Research on learning outcomes, completion, and quality*. Washington, DC: Association of America Colleges and Universities.

- Carlson, M., & Rasmussen, C. (Eds.). (2008). *Making the connection: Research and teaching in undergraduate mathematics education* (MAA Notes, #73). Washington, DC: Mathematical Association of America.
- Cross, K. (1986). A proposal to improve teaching or what “taking teaching seriously” should mean. *AAHE Bulletin*, 39(1), 9–14.
- Dewar, J., & Bennett, C. (2010). Situating SoTL within the disciplines: Mathematics in the United States as a case study. *International Journal of the Scholarship of Teaching and Learning*, 4(1).
- Doyle, T., & Zakrajsek, T. (2013). *The new science of learning: How to learn in harmony with your brain*. Sterling, VA: Stylus.
- Glassick, C., Huber, M., & Maeroff, G. (1997). *Scholarship assessed: Evaluation of the professoriate*. San Francisco, CA: Jossey-Bass.
- Hsu, E. (2011). *Spreadsheet of North American doctoral programs in math education*. Retrieved from [sigma.maa.org/rume/phd.html](http://sigma.maa.org/rume/phd.html)
- Huber, M. T., & Hutchings, P. (2005). *The advancement of learning: Building the teaching commons*. San Francisco, CA: Jossey-Bass.
- Huber, M. T., & Morreale, S. P. (Eds.). (2002). *Disciplinary styles in the scholarship of teaching and learning: Exploring common ground*. Washington, DC: American Association for Higher Education.
- Hutchings, P. (Ed.). (2000). *Opening lines: Approaches to the scholarship of teaching and learning*. Palo Alto, CA: The Carnegie Foundation for the Advancement of Teaching.
- Hutchings, P., Huber, M., & Ciccone, A. (2011). *Scholarship of teaching and learning reconsidered*. San Francisco, CA: Jossey-Bass.
- Hutchings, P., & Shulman, L. (1999). The scholarship of teaching: New elaborations, new developments. *Change*, 31(5), 10–15.
- KewalRamani, A., Gilbertson, L., Fox, M., & Provasnik, S. (2007). *Status and trends in the education of racial and ethnic minorities, NCES 2007-039*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Kuh, G. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter*. Washington, DC: Association of American Colleges and Universities.
- Leamson, R. (1999). *Thinking about teaching and learning: Developing habits of learning with first year college and university students*. Sterling, VA: Stylus.
- National Center for Science and Civic Engagement. (n.d.). *Assessment initiatives: The scholarship of teaching and learning*. Retrieved from [www.sencercer.net/Assessment/assessmentinitiatives.cfm](http://www.sencercer.net/Assessment/assessmentinitiatives.cfm)
- O’Meara, K., & Rice, E. (Eds.). (2005). *Faculty priorities reconsidered: Rewarding multiple forms of scholarship*. San Francisco, CA: Jossey-Bass.
- Project Kaleidoscope—PKAL. (n.d.). *Advancing what works in STEM education*. Washington DC: Association of American Colleges and Universities. Retrieved from [www.aacu.org/pkal/index.cfm](http://www.aacu.org/pkal/index.cfm)
- Richlin, L. (2001). Scholarly teaching and the scholarship of teaching. In C. Kreber (Ed.), *Scholarship revisited: Perspectives on scholarship of teaching and learning*, New Directions for Teaching and Learning: No. 86 (pp. 57–68). San Francisco, CA: Jossey-Bass.
- Richlin, L. (2003). Understanding, promoting and operationalizing the scholarship of teaching and learning: A message from the editor. *Journal on Excellence in College Teaching*, 14(2/3), 1–4.
- Schmidt, P. (2008, October 13). Harvard’s Derek Bok: Professors, study thine own teaching. *Chronicle of Higher Education*. Retrieved from [chronicle.com/article/Harvards-Derek-Bok-/1239/](http://chronicle.com/article/Harvards-Derek-Bok-/1239/)

- Schoenfeld, A. H. (2000). Purposes and methods of research in mathematics education. *Notices of the AMS*, 47(6), 641–649.
- Shulman, L. (2005, February). *The signature pedagogies of the professions of law, medicine, engineering and the clergy: Potential lessons for the education of teachers*. Presentation at the Math Science Partnerships (MSP) Workshop: Teacher Education for Effective Teaching and Learning, hosted by the National Research Council's Center for Education, Irvine, California. Retrieved from [hub.mspnet.org/index.cfm/111172](http://hub.mspnet.org/index.cfm/111172)
- Smith, R. (2001). Formative evaluation and the scholarship of teaching and learning. In C. Knapper & P. Cranton (Eds.), *Fresh approaches to the evaluation of teaching*, New directions for teaching and learning: No. 88 (pp. 51–62). San Francisco, CA: Jossey-Bass.
- Valleau, A. (2010). *Transformative dialogues: Teaching and learning eJournal*, 4(1). Retrieved from [www.kpu.ca/td/past-issues/4-1](http://www.kpu.ca/td/past-issues/4-1)
- Zull, J. E. (2002). *The art of changing the brain: Enriching the practice of teaching by exploring the biology of learning*. Sterling, VA: Stylus Publishing.



# 2

## Initiating a SoTL Investigation

Jacqueline M. Dewar and Curtis D. Bennett  
*Loyola Marymount University*

### Introduction

In Chapter 1 we defined SoTL as

the intellectual work that faculty members do when they use their disciplinary knowledge (in our case, mathematics) to investigate a question about their students' learning (and their teaching), submit their findings to peer review, and make them public for others to build upon.

This chapter considers questions and situations that might prompt a SoTL study. It presents a taxonomy of SoTL questions derived from the work of Carnegie scholars that can be useful in guiding the development of a project. We discuss how disciplinary knowledge can be brought to bear on framing SoTL research questions. We describe how literature searches can inform SoTL studies and give suggestions for conducting a search. The chapter includes illustrative examples and points to additional examples in Part II.

### A Typical Starting Point

In one of the formative articles of the scholarship of teaching and learning movement in the United States, Randy Bass (1999) discussed the different reactions that “teaching problems” and “research problems” typically garner from faculty members, with the latter engendering far more positive interest and reaction. He posited that one of the tenets of SoTL is that a teaching problem should be viewed as an invitation to a scholarly investigation, similar to how most faculty members view a research problem.

Here is a problem that any teacher can relate to:

*My students aren't as prepared for class as I would like them to be.*

An attempt to fix the problem might lead to the following sequence of thoughts or actions.

- *What if I give my students reading assignments in the textbook?* (Few will actually read and those who do won't read carefully enough.)
- *What if I give them reading questions for the assigned reading?* (They may read just enough to answer the questions.)

- *What if I also ask them to generate their own questions after reading?* (I probably won't be satisfied with the questions they ask.)
- *How can I get them to ask better questions?* (What do I mean by better questions?)
- *What kind of questions are they asking now?*

Many SoTL investigations develop along these lines. A teaching problem (about students not being prepared for class) leads to a sequence of questions about the effect of interventions (involving students reading before class) to address the problem. Then a new problem arises (one about getting students to ask better questions about the readings). Finally, a question emerges that seeks to deepen the faculty member's understanding of something the students are doing (asking questions after reading).

## A Taxonomy of SoTL Questions

Seeing this sequence occur many times in scholars' work in the Carnegie Academy for the Scholarship of Teaching and Learning enabled Carnegie Vice President Pat Hutchings (2000) to identify the following taxonomy of SoTL questions:

*What works?*—This question asks whether a particular teaching method, assignment, or approach is effective.

*What is?*—This question seeks detailed information about what students are really doing in a particular situation.

*What could be?*—This question examines a case where something really interesting happened or a situation involving something the instructor is passionate about or has an opportunity to do. This question type aims to show *what is possible*, so it is sometimes referred to as *a vision of the possible*.

Hutchings (2000) also acknowledged *theory-building* as a fourth but less common type of SoTL work. As we noted in Chapter 1, theory-building is far more common in RUME work than in SoTL.

With the taxonomy in hand, we can identify the following components of a typical SoTL investigation emerging in our example:

The original teaching problem—

*My students aren't as prepared for class as I would like them to be.*

Three *What works?* questions—

- *Will giving reading assignments be effective in preparing my students for class?*
- *Will having students answer questions on the reading prepare them for class?*
- *Will asking them to generate their own questions after reading be effective?*

A new teaching problem—

*I am not satisfied with the questions my students are generating, how can I get them to ask better questions?*

This teaching problem could lead to another set of ideas, interventions, and questions about what might work to get them to ask better questions. At some point it will be necessary to describe how and why some questions are better than others. This naturally leads to a *What is?* question:

- *What kind of questions are they asking now?*

We turn to a different situation, one that resulted in a published SoTL study, to provide an example of a *What could be?* question. For Dewar and two of her colleagues, the opportunity afforded by a grant to add group projects on community issues to a quantitative literacy course (Dewar, Larson, & Zachariah, 2011) led to a *What could be?* question:

- *What happens if group projects on community issues are added to a quantitative literacy course?*

Consideration of this question prompted more questions:

- *Will students learn as much or more?*
- *Will their attitudes toward mathematics change?*
- *What else might they gain?*

Several of these subordinate questions were then re-framed as *What works?* questions:

- *Will students who do group projects on community issues in a quantitative literacy course learn more mathematics than students taught in the regular lecture-discussion format?*
- *Will they leave the course with more positive attitudes toward mathematics and its usefulness?*

These examples demonstrate that SoTL projects can involve several different types of questions simultaneously. The types of questions asked will have an influence on the design of the study and the kind of data, quantitative, qualitative, or both, that are collected. We treat the connection between the taxonomy questions and the research design and data in greater detail in Chapter 3.

In Part II, twenty-five mathematicians from fourteen institutions, including a community college, primarily undergraduate institutions, and research-intensive universities, describe fifteen SoTL projects. This collection provides more examples of each question type. Michael Burke (Chapter 11) and John Holcomb (Chapter 12) provide compelling examples of *What could be?* questions. The study by Curtis Bennett and Jacqueline Dewar in Chapter 18 focuses entirely on a *What is?* question, while Derek Bruff's project (Chapter 13) illustrates an interplay between *What is?* and *What works?* questions. In Chapter 6 a professor's desire to address an instructional challenge leads to collaborative work by group of mathematics instructors and academic support professionals that took the form of a *What works?* investigation. All the chapter contributors in Part II try to show the process of doing SoTL. Many of these SoTL scholars use the taxonomy in formulating their questions and developing their plans to gather and analyze evidence. Another example of using the SoTL taxonomy to frame two research questions (prompted by a second semester calculus project that exposed student misconceptions about the role of error calculations in applications of Simpson's rule) can be found in [Bennett and Dewar \(2012\)](#).

## Using Disciplinary Knowledge to Narrow a Question

The knowledge that mathematics faculty members have about the discipline can prove especially helpful in developing SoTL questions. To illustrate this, we return to our first example:

The original teaching problem—

*My students aren't as prepared for class as I would like them to be.*

Let us suppose that after trying the intervention of assigning reading from the textbook before class, I find that students don't learn much from reading a mathematics text. This leads to the following:

A new teaching question—

*How do I help students to read a textbook more effectively?*

Calling upon disciplinary knowledge about reading mathematics texts that most mathematics faculty members possess through their own experience reveals underlying disciplinary issues, such as:

- Reading mathematics is different from reading a novel or reading a history text.
- Interacting with a mathematics text involves methods that differ from reading novels or history texts.
- Most students don't know how to build their own understanding when reading definitions, theorems, or examples.

Reflecting on these observations can produce a series of questions:

- *What do students do when they read the text?*
- *Does modeling how to read and understand a definition help them?*



- *What kinds of directed ways of interacting with a definition would help them learn to read and understand definitions?*
- *Does instruction on reading mathematics texts facilitate independent learning?*

The first of these is a *What is?* question and the second and fourth are *What works?* questions. The third question is not a research question; rather it prompts the instructor to design tasks for a *What works?* investigation. It might be a worksheet that directs students to develop examples that satisfy a definition and counterexamples that do not. The effect of using the worksheet on developing student skills in understanding definitions could then be studied as a *What works?* question.

This application of disciplinary knowledge has led to the beginnings of a different SoTL investigation than the one described earlier. On a global scale we moved away from the issue of preparing for class to developing reading skills for mathematics and then to independent learning of mathematics. In the process we asked both *What is?* and *What works?* questions. Along the way we limited the research question at one point by focusing in on how to read and understand definitions. We could have asked similar questions about theorems and examples.

Before going further with this investigation about reading mathematical definitions for understanding, it would be a very good idea to consult the literature to see what has already been studied about this. We offer some suggestions for searching educational literature in the next section.

## Undertaking a Literature Search

If publishing SoTL results in a peer-reviewed journal is a desired outcome, then at some point a literature search will be necessary in order to situate the work within the body of knowledge on the topic and write the literature review. For most mathematics faculty members, undertaking a search of educational literature will be an unfamiliar and possibly daunting task, one that may be tempting to postpone. There is no subject classification for SoTL work, like the Mathematics Subject Classification scheme used by the American Mathematical Society for papers on mathematical subjects, and the relevant literature for a SoTL study might appear in a wide variety of places. These range from pedagogically oriented publications within STEM disciplines, the social sciences, or education to interdisciplinary SoTL journals and higher education publications. To see the extent of this literature consult the journal list maintained by the Center for Excellence in Teaching and Learning at Kennesaw State University ([cetl.kennesaw.edu/teaching-journals-directory](http://cetl.kennesaw.edu/teaching-journals-directory)).

There are many reasons why it is valuable to seek out the relevant literature early in the project. Knowledge of what others have done can help formulate the project question, improve the design of the project, provide ideas or instruments for measuring outcomes, and save repeating the work of others. In Chapter 14 Mike Axtell and William Turner attest to this when they write: “We initially viewed the literature review as an onerous task to be completed for form’s sake, but we soon realized that the literature review was playing a central role in shaping our study question and our thoughts for how to design the study” (p. 139). Rann Bar-On and his colleagues describe in Chapter 6 how reading the literature advanced their understanding of the mathematical development of precalculus and calculus students. They saw this as a critical turning point in their movement from good teaching to reflective teaching to scholarly teaching, and eventually to a scholarship of teaching and learning project. Rikki Wagstrom (Chapter 19) tells how her literature search caused her to change the site of the investigation from a college algebra course to the prerequisite for that course.

Some suggestions for doing literature searches follow. Begin by defining the topic of interest, using synonyms to express it in two or three different ways. Those words or phrases become the search terms. Initial searches can be done with internet search engines such as Google or Google Scholar to get an idea of how effective the search terms are. If the result is too many or too few hits, refine the inputs. For too few hits, use a wildcard, usually an asterisk (\*), at the end of words that can have multiple endings to search for all the words simultaneously. For example, using “learn\*” will capture the terms “learn,” “learners,” and “learning.” The logical connector “OR” can produce more results since the search “gender” OR “equity” will find documents containing either word. For too many hits, do an “AND” and a “NOT” search, or limit the search to certain fields, such as the abstract or title. Next turn to the Education Resources Information Center or ERIC ([www.eric.gov](http://www.eric.gov)), the world’s largest digital library of education literature, and PsycINFO ([www.apa.org/pubs/databases/psycinfo/index.aspx](http://www.apa.org/pubs/databases/psycinfo/index.aspx)), a database of the American Psychology Association.

For doing SoTL in mathematics, the Math Ed Literature Database ([bfc.sfsu.edu/cgi-bin/rume.pl](http://bfc.sfsu.edu/cgi-bin/rume.pl)) is an important resource. This searchable, annotated collection of mathematics education literature contains about 400 references gathered by the Special Interest Group of the MAA on Research on Undergraduate Mathematics Education and more than 1000 entries on both statistics and mathematics education compiled by Professor Joan Garfield at the University of Minnesota. Also do a subject and keyword search of campus library holdings and enlist the help of the librarians in the reference department. Although ERIC dominated the educational research scene up until the early 2000s, now there are more database options available to researchers from various publishers (EBSCO, Gale, etc.). By consulting the local librarians, faculty members can determine if their campus subscribes to these or other databases, such as Education Full Text.

Scan the abstracts of the results from the searches to see if they will be helpful. For the ones that seem relevant take note of the descriptors, keywords, and authors' names. Try new searches based on these. If databases do not provide the full text of the resource, seek out items of interest from other sources such as the campus library, interlibrary loan, or the publisher. Be organized and keep good records of the search, what was found, and why it was of interest. Save the search statements for possible reuse or later modification. Bibliographic management tools such as EndNote or Zotero (for MsWord) or BibTeX (for LaTeX) make it possible to capture the citations in an appropriate format for the bibliography. Be aware that these tools are not 100% accurate.

Another tactic to expand the search is to use a snowball strategy in one of three ways. When reading the sources related to the study, check their bibliographies for older work on the subject. It may be possible to find newer literature that cites the sources already identified by using resources such as *Web of Science* ([thomsonreuters.com/thomson-reuters-web-of-science](http://thomsonreuters.com/thomson-reuters-web-of-science)). Check with campus librarians to see if there is access to this or a similar resource. Finally, from the literature, identify scholars who have done recent work on the topic and contact them to describe the study and inquire if they know of newer unpublished work.

As an example of some of these strategies put into in action we recommend Chapter 19, where Rikki Wagstrom provides details about her literature search. She notes that she found it more fruitful to begin her database search using the "All Text field" rather than the "KeyWord" or "Descriptor" fields. She writes that searching the ERIC database for articles pertaining to the integration of civic issues into algebra courses using "civic issues" AND "algebra" in the KW Identifier or DE Descriptor fields yielded no results. However, using the TX All Text field pulled up three records, including one that provided additional descriptors, among which were "numeracy," "algebra," "citizenship education," "comparative analysis," and "social responsibility." From this one search, she acquired many leads to expand her search.

Once a question is framed and positioned in the literature, the next step is to design the study. This involves planning what evidence will be gathered, from whom, and how it will be analyzed. Social scientists refer to this as the "methodology" of the study. The next chapter addresses these issues.

## References

- Bass, R. (1999). The scholarship of teaching and learning: What's the problem? *Inventio*, 1(1), 1–10.
- Bennett, C., & Dewar, J. (2012). An overview of the scholarship of teaching and learning in mathematics. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 22(6), 458–473. doi:10.1080/10511970.2011.585389
- Dewar, J., Larson, S., & Zachariah, T. (2011). Group projects and civic engagement in a quantitative literacy course. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 21(7), 606–637. doi:10.1080/10511970903579048
- Hutchings, P. (Ed.). (2000). *Opening lines: Approaches to the scholarship of teaching and learning*. Menlo Park, CA: The Carnegie Foundation for the Advancement of Teaching.



# 3

## Gathering and Analyzing Evidence

Jacqueline M. Dewar  
*Loyola Marymount University*

### Introduction

SoTL involves the systematic investigation of a question we have about student learning and we look for answers in evidence generated by students. After framing a researchable question, we have to gather and analyze evidence. So this chapter examines some basic considerations of research design, such as whether, and how, to gather quantitative data, qualitative data, or both. It is likely that one or more of the types of evidence discussed in this chapter will be unfamiliar to mathematicians. Many of them were new to Curtis Bennett and me as well, when we began doing SoTL. In this chapter I write about these methods from our experience in learning to use them.

### Triangulating Data

A SoTL researcher should develop a plan for systematically collecting multiple types of evidence. A diversity of evidence can help the researcher to form a convincing picture of student learning (Wiggins, 1998). This approach is called triangulation of the data. According to Webb, Campbell, Schwartz, and Sechrest (as cited in Shavelson & Towne, 2002, p. 64): “When a hypothesis can survive the confrontation of a series of complementary methods of testing, it contains a degree of validity unattainable by one tested within the more constricted framework of a single method.” In other words, claims or explanations supported by several types of evidence—for example, student work samples, interviews, and retention rates—are considered to be more accurate. This will be an asset if the work is submitted for publication in a peer-reviewed journal. We turn now to a discussion of the difficulties in approaching SoTL as standard educational research.

### Challenges that Educational Research Design Presents for SoTL

Designing an educational research study can pose a number of challenges to faculty members interested in SoTL. Many of the same challenges confront mathematics education research at all levels. In response to a query from the NSF, and with its funding, the American Statistical Association held a series of workshops for statisticians and mathematics education researchers to discuss whether the statistics community could “offer any contributions to improving the quality of mathematics education research” (Working Group on Statistics in Mathematics Education Research, 2007, p. 1). The resulting report, *Using Statistics Effectively in Mathematics Education Research*, describes some fundamental statistical issues that can arise in mathematics education research and offers guidelines for undertaking, reporting, and

evaluating that research. We take up some of those concerns now as they related to the scholarship of teaching and learning.

When attempting to show a causal relationship, the gold standard for research is the random assignment of subjects to an experimental group and a control group. How appropriate random assignment experiments are in educational research is a matter of considerable discussion by both educational researchers (Berliner, 2002; Cook & Payne, 2002) and SoTL practitioners (Grauerholz & Main, 2013; Poole, 2013). Poole discussed how disciplinary views differ on what is knowable and whether behavioral indicators alone suffice to measure educational outcomes. Grauerholz and Main described the near impossibility of controlling the factors that can influence the learning environment in a classroom. Indeed, many of the contributors to Part II discuss the problems or concerns they had with the methodology of random assignment to treatment and control groups. We begin with an example of a SoTL investigation and describe how it deviated from the gold standard.

In the study of an existing multi-section Quantitative Literacy (QL) course, five out of 20 sections of the course were revised to include semester-long group projects involving local community issues that students could investigate using the mathematical topics of the course. Students enrolled without knowing that some sections would have projects while others would not, so the assignment was at least blind, if not random. The instructors saw no evidence of students changing sections after seeing the syllabus. The course fulfilled a graduation requirement for all students whose majors did not require a mathematics course. The enrollment in the project-based sections reflected the typical mix of majors and years in school. Student background characteristics (gender, year in school or age, majors, entering SAT scores and the like) could have been more formally compared, but were not. The project-based experimental group included 110 students and the comparison group had more than twice that. Evaluation of the efficacy of the projects approach included pre- and post-comparisons between the two groups on a test and on an attitudinal survey. In addition, a focus group and a knowledge survey were conducted with the experimental group. We return to this example several times later in this chapter when we discuss knowledge surveys, focus groups, and the issue of alignment between assessment tools and the outcomes being measured. Readers who are interested in a full account of the study should consult Dewar, Larson, and Zachariah (2011).

## Dealing with the Lack of a Control Group

In a typical SoTL investigation, finding a comparison group is difficult, if not impossible for a number of reasons. The most basic is that we may only have the students in our classroom as our subjects and no access to a control group. Yet, as noted in Adhikari and Nolan (2002), investigating a *What works?* question requires some sort of comparison group or, at least, baseline data. We describe several ways to clear this hurdle, each with its own set of advantages and disadvantages.

It may be possible to teach two sections of the same course in the same semester. If so, one section could be taught using the traditional approach and one using the experimental approach. Of course, as with many forms of classroom research, a question naturally arises concerning the instructor's impartiality. The instructor is trying a new approach that she thinks will be better in some way, and has formulated a research question about it. Will the instructor's beliefs or interests give the experimental section an advantage? It is natural to consider recruiting an impartial and capable colleague to teach the experimental section. However, if the instructor of the experimental group is not as familiar or comfortable with the teaching method or intervention, the results may not be the same as they would be for an instructor who is. Rikki Wagstrom (Chapter 19) decided to teach all the experimental sections herself, so that she did not have to subject her colleagues to the stress of learning a new curriculum and having their students assessed. In Chapter 14, Mike Axtel and William Turner express additional concerns about obtaining a control group. For them, a control group "made sense only if the two sections were taught in a very similar manner," and they observe that professors at small colleges like theirs "very rarely teach two sections of the same course in a semester" (p. 138). The lack of strict controls when comparing student performance across groups is one factor contributing to the perception that SoTL studies are less generalizable than formal educational research studies.

Sometimes, the intervention can be alternated between two sections of a course taught by the same instructor. This is called a "quasi-experimental" or "switching replications" design. For example, the instructor can teach material covered in the first and third quizzes in a course with the experimental approach in one section and with the standard approach in another, and then switch the approaches for the second and fourth quizzes. Both mathematics education

and SoTL researchers employ this tactic for some studies. It can greatly lessen the variability resulting from different levels of ability in students and/or sections, though it cannot be used to assess the cumulative benefit of a treatment over an entire semester. Edwin Herman (Chapter 8) describes how he alternated his intervention, using a game to review concepts, between two sections of a multivariate calculus courses across four exams. Derek Bruff also adopted a quasi-experimental design in his investigation of the types of information students can learn best by reading the text on their own. In Chapter 13 he tells of rotating the focus of his reading quizzes from notational to computational to conceptual questions across three groups. He reports that the design met with approval by his campus human subjects committee, since it minimized potential risk to his students' grades and learning. (See the section of Chapter 4 titled Human Subjects Considerations for further discussion of this topic.) When choosing to alternate an intervention between two sections, instructors may need to consider whether students could possibly attend the other section or study with students who do. This method can be more vulnerable to such cross-over students, since it measures shorter-term effects rather than those developed over an entire course.

A different approach to the lack of a control group when trying a new method is to rely on comparison data obtained from prior cohorts (Hutchings, 2002). We are able to offer two examples of this from work carried out at our institution. Blake Mellor (Chapter 5) wanted to examine whether a course in the mathematics of symmetry as a general education course would give students a better understanding of what it means to "do mathematics" and a greater appreciation for mathematics than the standard QL course does. He taught the standard QL course first, and gathered quantitative and qualitative data through surveys and reflective writing. This gave him baseline data from a comparison group. He could then teach his mathematics of symmetry course and use the same instruments.

The next example describes a different situation, one involving the development of a totally new course. Even so, the course developers found a way to make some comparisons to prior cohorts. In 1992, to try to improve retention in the major, the mathematics department at Loyola Marymount University introduced a workshop course for freshmen mathematics majors. It aimed to enhance their education in a number of ways, including developing skills in problem solving and mathematical writing, thereby increasing the likelihood of success in the major. To increase student motivation to persist in the major, the course also discussed mathematical people and careers to show the human side of the discipline and inform students of the variety of rewarding career paths open to mathematics majors. Several assessments of the course's effectiveness (Dewar, 2006) utilized comparisons with prior cohorts of mathematics majors. At the end of the 1992 introductory year, a survey about mathematical careers and mathematicians showed the first group of freshmen taking the course to be more knowledgeable than more advanced students who had not taken the course. Over the next several years, data gathered allowed comparison of the retention rates of the first few cohorts with rates prior to the introduction of the course. Other assessments involved pre- and post-tests of problem solving, a portfolio assignment that demonstrated improvement in writing from the beginning to the end of the course, and self-reported improvements in confidence at the end of the course, none of which made comparisons to prior cohorts. The dropout rate was nearly halved, and overall, the results were so convincing that the course became a requirement in the first year of the major. Students and department faculty members alike credit it with playing a critical role in creating a sense of community in each new freshmen class of majors.

Comparisons between an experimental and a control group can raise ethical questions. We hear of medical trials being interrupted because the results for the treatment group are so positive that it is deemed unethical to withhold the treatment from the placebo group. A similar sort of question, though not of the same life-threatening nature, can arise in SoTL studies. Is it fair to provide only some students with special learning experiences? Or, as Lynn Gieger and her colleagues at a small liberal art college write in Chapter 7, "Since the research was being conducted with students who expect personalized, high-quality learning experiences, we did not consider it ethical or reasonable to perform studies with a classical control and treatment strategy" (p. 69).

A different sort of ethical question can also arise in doing SoTL work. Is it ethical to ask students to perform tasks they were not prepared for? For example, is it fair to ask students taught in a traditional elementary statistics course to analyze real data and provide a written report if these tasks were not part of their course, but were the focus of the instruction in the section being studied in the SoTL investigation? John Holcomb discusses this dilemma and how it influenced the direction of his investigation in Chapter 12.

Another practical and ethical issue that can arise in SoTL studies is how much extra work the study or its assessments will require of the students. This concern influenced decisions Rikki Wagstrom (Chapter 19) made in her study. The SoTL researcher should aim to place little or no extra burden on students. By giving careful attention to design, many

assessment tasks intended to capture SoTL evidence can be embedded within course assignments. Ideally, much of the evidence gathered will emerge organically from learning or evaluation processes built into the course.

An alternate way of showing the efficacy of a teaching method or intervention is to measure achievement of stated learning goals. This is known as a “criterion-referenced” evaluation and it avoids the ethical dilemma inherent in making comparisons by assessing two groups of students using tasks for which only one group received instruction. A criterion-referenced approach requires specifying in advance the learning goals and designing assessments that measure them. Typically, this involves pre- and post-comparisons of the students in the course. This approach comes with its own set of concerns, including the importance of aligning the assessment instrument with the outcomes that are to be measured, as the following discussion illustrates.

### **Aligning the Assessment Measure with What Is Being Studied**

Recall the earlier example (p. 20) of the QL course that my colleagues and I revised to include a civic engagement component where students applied the elementary mathematics they were learning to semester-long projects involving campus or community issues. When the pre- and post-test results in the projects-based sections were compared to those of students taking sections of the course without the projects, there was improvement on only one question—the one that involved interpreting the meaning of a margin of error. Achieving improvement on only one question was a disappointing result that prompted serious reflection. Our initial explanation was that the regular approach was also doing a good job of teaching content. The projects approach was showing improvement not so much in skills but in attitudes and in understanding how mathematics connects to and is useful in daily life.

Then a famous physics education researcher (and Nobel prize winner), Carl Wieman, paid a visit to campus and I discussed the SoTL project with him. He pointed out that the other pre- and post-test questions might not be designed to find any distinctions. He was right. All the other questions were straightforward computational problems such as “What is the monthly payment on a car loan if . . . ?” The questions all tested factual knowledge and the ability to apply formulas and perform computational tasks. We would not expect the addition of research projects on community issues to improve those skills.

The question that showed a significant difference in performance favoring the projects-based approach was not testing facts or procedures. It required an evaluation of which of four given explanations for the meaning of a particular confidence interval was the best. Evaluation is a higher order task in Bloom’s taxonomy (Anderson & Krathwohl, 2001). All QL students, not just those in the projects-based sections, computed margins of error on homework and studied the procedure in preparation for quizzes and tests. The meaning of the margin of error was discussed in all classes, appeared in the text, and may have been in a few homework problems. However, in the projects-based sections many of the groups chose projects that involved gathering survey data. They had to compute the margin of error for their data and they had to explain what it meant in their own words in their written reports and in-class presentations. Of all the pre- and post-test questions, only the conceptual question on interpreting the margin of error could possibly capture the difference between the two approaches to teaching QL.

For another discussion of aligning assessments with the object of the study, see Chapter 5 where Gretchen Rimmasch and Jim Brandt describe how they had to alter their plan for analyzing the data they had collected in order to assess the real focus of their experiment with visual cues in a multivariate calculus class.

### **To Seek Statistical Significance or Not**

SoTL researchers often wrestle with the issue of whether or not to seek to show statistically significant differences. Several of the contributors in Part II did seek and, in some instances, found statistical significance (e.g., Herman, Chapter 8; Rimmasch & Brandt, Chapter 5). However, much SoTL is carried out in classrooms where the number of students involved is too small to seek statistical significance using common comparison methods such as *t*-tests. Some researchers will teach the experimental version several semesters in a row to aggregate enough data to meet sample size requirements. In Chapter 16, Mellor reports that he considered this idea and rejected it because during the several semesters time it would take him to gather sufficient data, he would be making adjustments to the course.

Mathematics instructors who want to embark on a SoTL project but lack a background in statistics need not be too concerned. Many worthwhile projects, especially those belonging to the *What is?* category of the SoTL taxonomy, require no inferential statistics. (For examples, see Chapters 13 and 18.) Mathematicians whose projects call for

making statistical comparisons can consult introductory level textbooks on applied statistics. Colleagues who teach elementary statistics courses in mathematics and psychology departments recommended these texts: [Brase and Brase \(2013\)](#), [Ramsey and Schafer \(2013\)](#), and [Gravetter and Wallnau \(2011\)](#). The internet provides many resources as well, such as [Trochim \(2006a\)](#). Because Excel is capable of performing most statistical tests, specialized software such as SPSS, SAS, or R is not required. Finally, as discussed in Chapter 4, a SoTL researcher can seek out a collaborator in the social sciences.

## Which Measure of Change?

When using pre- and post-comparisons, some consideration needs to be given to which measure of change to use. Possibilities include actual change, percent change, or normalized gain. The last of these was popularized by [Hake \(1998\)](#) in his large-scale study of Interactive-Engagement (active learning) teaching methods in introductory physics. “Normalized gain,” the ratio of actual gain to maximum potential gain, is defined as

$$g = \frac{\text{post} - \text{pre}}{100 - \text{pre}},$$

where the pre-test and post-test each have a maximum score of 100.

This measure of change was developed to allow comparisons of student populations with different backgrounds. Students who have significant knowledge about a topic, and therefore score well on a pre-test, are unable to show much actual or percent change on a post-test. This measurement difficulty is resolved by using the normalized gain, since it represents the fraction of the available improvement that is realized. The normalized gain has become so widely used in physics education that its properties have been extensively studied. For example, [Stewart and Stewart \(2010\)](#) have shown that it is insensitive to guessing.

Whenever designing an investigation to seek statistical significance for a *What works?* question becomes too difficult, an alternative worth considering is to treat the SoTL inquiry from the perspective of anthropology rather than educational research. That is, can the investigation be approached as a *What is?* rather than a *What works?* question? Seeking to understand the causes of the problem or attempting to describe accurately the phenomenon or student behavior or (mis)understanding that drew attention to the question can provide an alternate way to engage with the issue. A *What is?* approach can produce results that are useful in later attempts to formulate and assess a method to fix the problem. This approach proved fruitful for Derek Bruff (Chapter 13) in his study of reading questions.

We next shift our attention from the design difficulties encountered conducting SoTL research as educational research to considerations of the two types of data we might collect in a SoTL inquiry.

## Quantitative versus Qualitative Data

It seems easy enough to distinguish between the two types of data. Quantitative data are numerical data, such as test scores, grade point averages, time-on-task, or percent completion rates. Qualitative data are anything else, that is, non-numerical data. This includes solutions to problems, proofs, and reflective writing by students, text from interviews, video recordings of students solving problems in groups, or characteristics such as gender. Usually, mathematics instructors are more familiar with quantitative data. They tend to believe that quantitative data are more scientific, rigorous, or reliable, a view held by many, but not all, social science researchers as well. However, faculty members in disciplines such as anthropology or educational psychology may have a different view. From their perspective, qualitative data have advantages because they provide more in-depth or nuanced understandings of the subject under investigation. It allows the investigator to examine “Why?” or “How?” questions. In contrast, quantitative data are best suited to address “What?” questions. Recalling Hutchings’ (2000) taxonomy of SoTL questions from Chapter 2, quantitative data may be a better fit for *What works?* questions while qualitative data most naturally yield insights into *What is?* or *What could be?* questions.

However, [Trochim \(2006b\)](#) cautions that identifying data as quantitative or qualitative creates a false dichotomy. All quantitative data at some point almost certainly involved qualitative judgments. For instance, think about a math anxiety scale. Developing it required numerous decisions: how to define math anxiety, how to phrase potential scale items, how to ensure the items are clearly stated for the intended subjects, what cultural and language constraints might occur, and so on. The investigator who decides to use the scale has to make another set of judgments, such as whether



---

Scores
pre- and post-tests
course work and homework assignments
quizzes, mid-terms, or final exams
projects
papers
standardized scales and tests
Surveys
attitudes, beliefs, or satisfaction, often using a Likert scale from Strongly Disagree to Strongly Agree
student ratings of teaching
Frequency counts or percentages
multiple-choice test item responses
course completion rates
participation in class, on discussion boards, etc.
online homework system usage
office visits
Measures of time use
time spent online accessing homework systems or other resources
Institutional research data
grade point averages
grades
admission or placement test scores
retention data (e.g., in course, program, major, or institution)
enrollment in follow-up courses
student demographics

---

**Table 3.1.** Examples of familiar quantitative (numerical) data

the scale relates to his question and if it is appropriate for his subjects. When the subjects fill out the scale, they will make many more judgments: what the various terms and phrases mean, why the researcher is giving this scale to them, and how much energy and effort they want to expend to complete it.

If the discussion of the math anxiety scale, adapted from [Trochim \(2006b\)](#), seems too obscure, consider a typical calculus exam, which yields a percentage score for each student. The instructor decided how many and which problems to include, how they would be worded, and how partial credit would be given. The students made their own judgments about what the questions were asking, how much work to show, and how much time and effort to expend. In both examples, what may look like a simple quantitative measure is based on many qualitative judgments made by lots of different individuals.

Conversely, virtually all qualitative data can be expressed and analyzed numerically. Much like quantitative data, qualitative data can be categorized, counted, or turned into percentages, and used to compare populations. We will describe how to do this with qualitative data later in this chapter.

Rather than reinforce the quantitative-qualitative division by first discussing one and then the other, we consider data sources likely to be more familiar to mathematicians first and then those less familiar. Both categories contain quantitative and qualitative data types.

## Data Sources Familiar to Most Mathematicians

Mathematics instructors not trained in mathematics education are typically most familiar and comfortable with gathering and interpreting test scores, grade point averages, time-on-task, or percent completion rates. [Table 3.1](#) displays many kinds of familiar numerical data that could be gathered for a SoTL study.

While most of these data items seem straightforward to collect and interpret, there can be surprises lurking even in these very familiar forms of numerical data. As [Light, Singer, and Willett \(1990\)](#) put it: “You can’t fix by analysis

what you bungle by design” (p. viii). For example, in collecting time-on-task data from online homework systems, how long a student user has a webpage open may be misleading as an indicator of time-on-task. The earlier discussion of the project-based QL course highlighted another pitfall, namely, collecting and trying to interpret data that is not well aligned with the factors or attributes under investigation.

## Survey Design Considerations

Perhaps because most survey data are numerical and easy to obtain, many instructors interested in studying student learning or motivation decide to survey their students, often using a Likert-type scale with responses ranging from Strongly Disagree to Strongly Agree. The responses are converted to a numerical scale typically ranging from 1 to 5. With sufficiently many student participants, the pre- and post-results can be tested for statistical significance. The majority of the contributors to Part II employed surveys in their SoTL studies. One used a professionally developed survey (Holcomb, Chapter 12) and another (Kaus, Chapter 10) revised one to suit her purposes. The remainder chose to design their own, and some noted concerns and pitfalls they encountered either in the design or the administration of their surveys (for example, see Gieger et al., Chapter 7; Mellor, Chapter 16). Survey completion rates were discussed by several authors (Kaus, Chapter 10; Holcomb, Chapter 12; Mellor, Chapter 16), who offered suggestions for addressing this common concern. Stephen Szydlik (Chapter 15) provided a lengthy discussion of the steps he took to design, pilot, and validate his survey instrument. Unexpected survey results can suggest ways to collect new or more nuanced data as Wendy O’Hanlon and her colleagues reported in Chapter 17.

### Essential Steps in Designing a Survey

Because the use of self-constructed surveys is so common in SoTL studies, we present a summary of critical steps in designing surveys. They were drawn from materials provided by the Office of Assessment at Loyola Marymount University, which cited [Suskie \(1996\)](#).

Step 1. Identify the objectives and the target population for the survey.

- Before designing the survey clearly understand its purpose. Exactly what information is the survey intended to capture?
- Think carefully about the target population, that is, who will take the survey? The characteristics of the population (e.g., students enrolled in certain courses, students using online homework systems, faculty members, future teachers, etc.) may limit the type of information that can be collected from them. Consider whether there are factors that may influence their responses to the questions.
- Think how the information will be used for the study. To make comparisons between different groups of people, include questions that will enable appropriate grouping of respondents.

Step 2. Construct well-written questions.

The quality of the data gathered will depend on the quality of the questions. Good questions have the following characteristics:

- Each question is aligned with one or more of your objectives.
- The question is stated in the simplest possible language. Any complex terms used in the survey are defined.
- The wording is specific and precise so as to avoid confusion or misunderstanding.
- The question is concise and to the point. Long questions take time to read and are difficult to understand.
- The questions contain few, if any, negative words and they avoid double negatives. Negative statements tend to confuse survey takers.
- The questions are not “double-barreled.” A double-barreled question inquires about two or more parameters but allows only a single response. Asking students in a single question whether the online homework system provided *timely and useful feedback* is an example of a double-barreled question.
- The questions are not biased or leading. This might influence the response.

Step 3. Choose the format of the response options.

There is a wide variety of response options to choose from for survey questions. Here are the most commonly used response options along with special considerations for each.

- Yes/No options. Be careful using this simple response option as it allows for no grey area in between. The yes/no option is useful for contingency questions, which prompt the respondent with a yes/no question to determine if subsequent questions apply.
- Multiple-choice options. These provide a fixed set of answers to choose from and can be designed to allow the selection of one or multiple responses. Response options should be mutually exclusive. If the options are not distinct, respondents will not know which to select. Capture all possible responses while using the smallest number of categories possible. Including an “Other” category followed by a request to please “Explain” or “Specify” will capture overlooked options.
- Likert scales. These ask respondents to rate their preferences, attitudes, or subjective feelings along a scale, such as from Strongly Disagree to Strongly Agree. Each option in the scale should have a label and a decision must be made whether to include a Neutral option in the center of the scale. For each topic or dimension being surveyed, framing some items positively and some negatively makes it possible to check that respondents are reading carefully and to confirm expected correlations.
- Alternative responses. Another judgment to be made is whether to offer a response such as Don’t Know, Decline to State, Not Applicable, or Other as a way to allow survey-takers to avoid giving a direct answer.
- Open-ended responses. Open-ended responses allow survey-takers to answer in their own words. This type of response provides a rich source of information, but requires more time to analyze. The textual data will have to be coded for common themes, a process described later in this chapter.

For any format, order all response options logically and be consistent with the ordering for similar questions in the survey.

Step 4. Organize the survey with care.

A well-organized survey can improve the response rate. Here are some suggestions:

- Begin with an introduction that explains the purpose of the survey and how the person’s participation will help. Address privacy concerns by stating that participation is voluntary, and responses will be held confidential. If your survey is also anonymous, inform respondents of that. State who is conducting the survey and whom to contact with questions or concerns.
- Group questions by topic and put groupings in a logical order.
- Put demographic questions at the end of the survey and give respondents the option of “decline to state.”
- Make it as short as possible.
- Close with a thank you and repeat the contact information for any questions or concerns.

Step 5. Pilot test the survey and make any necessary changes.

Pilot testing is important because it is all too easy to write a confusing question unintentionally or to have an error in the logic of an online survey.

- Recruit a small group of people to take the survey and give you feedback on the clarity of the directions and the questions and how long the survey took.
- Have the test group include some individuals who are similar to your actual survey population.
- Review the completed responses for unexpected answers or inconsistencies.
- If you make a large number of changes based on the pilot results, consider conducting another pilot test.

[Adhikari and Nolan \(2002\)](#), two statisticians, offered additional pointers on the construction, administration, and analysis of surveys based on their experience evaluating summer programs for undergraduates at Mills College and

---

Solutions to
exam or quiz questions
homework problems
Writing assignments
projects or papers
portfolios
reflective writing or journals
Open-ended responses on surveys

---

**Table 3.2.** Examples of familiar qualitative (non-numerical) data

UC Berkeley. As a final bit of advice on surveys, inquire if there is any office or person on campus that supports the administration of surveys by faculty members doing assessment or research studies. They may be able to provide additional information or access to survey software.

Up to now, we have focused on familiar sources of quantitative data. Mathematics instructors are familiar with certain types of non-numerical data, especially if they are teaching proof-based courses or require students to write research papers. While most instructors have experience assigning, tallying, or grading some of the types of data listed in Table 3.2, they are probably far less familiar with methods for analyzing these data to answer a SoTL question.

Instructors often determine a numerical grade for solutions, papers, or projects without applying a rubric. However, when these qualitative data are used as evidence of student learning in a SoTL study, a rubric is often helpful. It can yield numerical data for multiple dimensions or characteristics of the work, resulting in a more nuanced assessment of learning than a single grade on a problem or a paper.

Another common approach to analyzing written work obtained from portfolios, journals, or surveys as part of a SoTL project is to code or categorize the text for commonalities. These might be themes found in students' reflective writing or common mistakes, false starts, or misunderstandings in proofs. These two methods of analyzing qualitative data, rubrics and coding data, are described in more detail later in this chapter.

## Data Sources Less Familiar to Mathematicians

Methods for investigating student learning that will be unfamiliar to many teaching mathematicians include knowledge surveys, think-alouds, and focus groups. We knew little, if anything, of these methods when we began doing SoTL work. Only one of these methods generates quantitative data, the knowledge survey, so we will begin with it.

### Knowledge Surveys

A knowledge survey is a tool designed to assess changes in students' confidence about their knowledge of disciplinary content. [Nuhfer and Knipp \(2003\)](#) first described this tool in *To Improve the Academy*, an annual publication for faculty developers, the people who direct campus teaching and learning centers. Knowledge surveys are useful for course development, revision, or improvement efforts. They have also been employed to document success or, more importantly, failure in student learning in courses or programs ([Dewar, 2010](#)). However, their use as a valid measure of student learning is not without controversy, as we shall describe. Nevertheless, we believe they have their place in SoTL work, provided that the researcher uses multiple measures to triangulate the data gathered from a knowledge survey. Their use offers additional benefits to the instructor, such as being a way to examine student background knowledge and to study course content sequencing or timing. They provide students with a preview of course content, a way of monitoring their progress throughout the semester, and a study guide for a final exam.

How do knowledge surveys work? Knowledge surveys are given to students at the beginning and the end of a course or a unit. They contain questions about every content item or learning objective, anything that students are supposed to know or be able to do at the end of the course or unit. This might be stating definitions, solving equations, finding a derivative or an area under a curve, explaining what it means for a series to converge, or interpreting a margin of error. These questions are phrased like quiz or test questions and are presented in the same chronological order as the content appears in the course (or unit or program). The interesting part is that students do not actually answer the

Question	Rating
1. Determine the number of square inches in 12 square yards.	1 2 3
2. The speed limit posted as you leave Tecate, Mexico is 50 km/hr. Find the corresponding limit in miles per hour.	1 2 3
3. A department store advertised an 80% off sale on fall apparel. The ad also contained a coupon for an extra 15% off to be applied to the reduced price of any sale or clearance purchase. Find the final price of a \$150 suit (ignore tax).	1 2 3
4. Is it possible to go on a diet and decrease your calorie intake by 125%? Explain why or why not.	1 2 3
5. Describe any likely sources of random or systematic errors in measuring the numbers of popped kernels in “large” boxes of popcorn at a movie theater.	1 2 3
6. Show that you understand the difference between absolute error and relative error by giving two examples: one where the absolute error is large but the relative error small, and the other where the absolute error is small but the relative error is large.	1 2 3
7. Use the appropriate rounding rules to answer the following with the correct precision or number of significant digits: Find the total weight of a 50 kg bag of sand and a 1.25 kg box of nails.	1 2 3
8. On what basis, if any, would you question the following statistic: The population of the United States in 1860 was 31,443,321?	1 2 3
9. Suppose you win a \$100,000 raffle. You wisely invest half of it in a savings account that pays interest with an APR of 5% that will be compounded quarterly. Find how much you will have in the account in 10 years.	1 2 3
10. What does the Annual Percentage Yield (APY) of an investment measure?	1 2 3
11. Suppose you want to start a savings program for a down payment on a house. In 10 years you would like to have \$125,000. Your financial advisor can find an account with an APR of 7% that will be compounded monthly. Find how much you will have to deposit into the account per month in order to have the \$125,000 in 10 years.	1 2 3
12. You wish to buy a new car and can afford to pay at most \$400 per month in car payments. If you can obtain a 4-year loan with an APR of 3.2% compounded monthly, what is the largest loan principal you can afford to take out? Round your answer to the nearest dollar.	1 2 3
13. You just received a \$1000 credit card bill, and your card has an annual interest rate of 18%. Your credit card company uses the unpaid balance method (i.e., charges interest on the unpaid balance) in order to calculate the interest you owe. Suppose you make a \$200 payment now, and make no new charges to your credit card in the next month. Find the balance on your next credit card bill a month from now.	1 2 3
14. The United States has a progressive income tax. Explain what that means.	1 2 3
15. Which is more valuable to a taxpayer, a tax deduction or a tax credit? Explain why.	1 2 3
16. Describe at least three misleading perceptual distortions that arise in graphics.	1 2 3
17. A company has 10 employees, making the following annual salaries: 3 make \$20,000, 2 make \$30,000, 4 make \$50,000, and 1 makes \$1,200,000 per year. Explain whether the median or the mean would be a better representation of the “average” salary at the company.	1 2 3
18. Two grocery stores have the same mean time waiting in line, but different standard deviations. In which store would you expect the customer to complain more about the waiting time? Explain.	1 2 3
19. What are quartiles of a distribution and how do we find them?	1 2 3
20. Give a five-number summary and depict it with a boxplot for the following set of data: {2, 5, 3, 4, 4, 6, 7, 5, 2, 10, 8, 4, 15}.	1 2 3
21. The body weights for 6-month old baby boys are normally distributed with a mean of 17.25 pounds and standard deviation of 2 pounds. Your 6-month old son Jeremiah weighs 21.25 pounds. Jeremiah weighs more than what percentage of other 6-month old baby boys?	1 2 3

**Figure 3.1.** A knowledge survey for a quantitative literacy course

22.	In order to determine how many students at LMU have ever used a fake ID to buy liquor, we survey the students in this class and find 40% of them have done so. We conclude 40% of LMU students have used a fake ID to buy liquor. Discuss possible sources of bias in the sample and comment if the conclusion is justified.	1	2	3
23.	A survey of 1,001 randomly selected Americans, age 18 and older, was conducted April 27–30, 2000, by Jobs for the Future, a Boston-based research firm. They found that 94% of Americans agree that “people who work full-time should be able to earn enough to keep their families out of poverty.” Explain what is meant by saying the margin of error for this poll at the 95% confidence interval is 3%.	1	2	3
24.	Formulate the null and alternative hypotheses for a hypothesis test of the following case: A consumer group claims that the amount of preservative added to Krunch-Chip brand of potato chips exceeds the 0.015 mg amount listed on the packages.	1	2	3
25.	Describe the two possible outcomes for the hypothesis test in #24.	1	2	3
26.	A random sample of Krunch-Chip potato chip bags is found to have a mean of 0.017 mg of preservative per bag. Suppose 0.03 is the probability of obtaining this sample mean when the actual mean is 0.015 mg preservative per bag as the company claims in #24. Does this sample provide evidence for rejecting the null hypothesis? Explain.	1	2	3

**Figure 3.1.** (Continued)

questions or work the problems. Instead they rate their confidence to answer each question correctly on a scale of 1 to 3, where 3 means they are confident they can answer correctly right now; 2 means they could get it 50% correct—or if given a chance to look something up, and they know just what to look up, they are sure they would then get the answer 100% correct; and 1 means they have no confidence that they could answer the question, even if allowed to look something up.

My department successfully employed knowledge surveys in an assessment of the engineering calculus sequence (Dewar, 2010). Knowledge surveys were one of several measures of student learning in the QL course mentioned earlier (p. 20) that incorporated a civic engagement component (Dewar, Larson, & Zachariah, 2011). Figure 3.1 contains the knowledge survey for that QL course and Figure 3.2 provides the double bar graph for the pre- and post-survey results for each question.

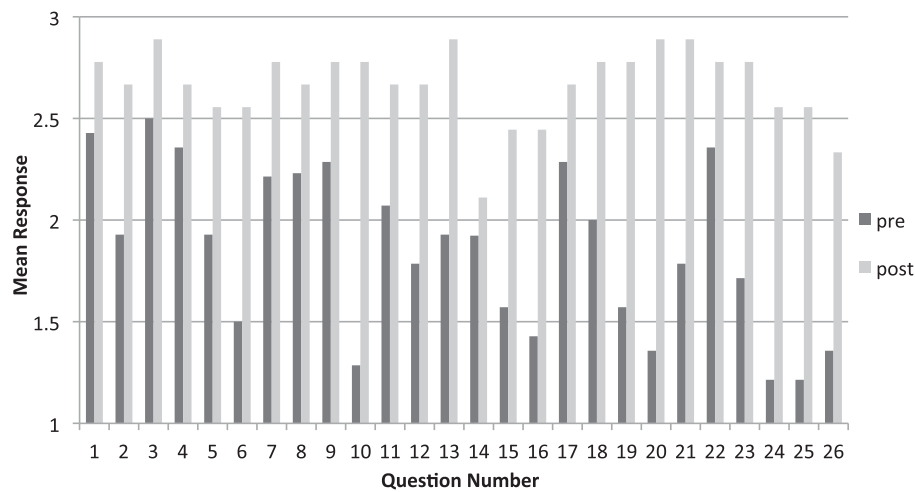
Several observations emerge almost immediately from the graph in Figure 3.2:

- The responses to question #14 indicate a problem with student learning.
- Confidence falls off for the topics at the end of the semester.
- Student confidence was fairly high for some of the items at the beginning of the semester.

We will elaborate on each of these observations after discussing the advantages knowledge surveys hold over pre- and post-tests. Unlike typical pre- and post-tests, knowledge surveys can be administered as take-home work rather than using class time for them. Since rating confidence about doing a problem takes a lot less time than actually carrying out the solution or writing an explanation, this instrument allows all course topics to be covered in detail. Many knowledge surveys contain 100 or more questions. The example in Figure 3.1 is brief as knowledge surveys go, because a significant amount of the coursework involved students working in groups to apply mathematics to a campus or community issue, and that work was assessed in a different way. Because making confidence judgments takes little of the students' time, the survey can include complex higher order questions (apply, explain, describe, justify, design) as well as basic skills questions (find, calculate). If students are given access to the survey during the semester, they can use it as a study guide, allowing them to map their progress in the course.

Preparing a knowledge survey at the beginning of a course promotes thoughtful and informed course planning and organization. The survey can indicate which topics students may already know a lot about, and so less time can be spent on them. It may suggest that more time needs to be spent on certain topics. In multi-section courses taught by multiple instructors, a knowledge survey can promote more consistent coverage of content, especially if the instructors collaborate on constructing it.

## Knowledge Survey Results - Spring 2006



**Figure 3.2.** Pre- and post-knowledge survey results

We return now to the earlier observations about the graph in Figure 3.2. The small improvement in confidence on item #14—a question about the meaning of a progressive income tax—disappointed the instructor (the author). This prompted me to reflect on how I had presented the topic in the course: by lecturing, with little student involvement, and no assigned homework or quiz questions on it. The knowledge survey enabled me to identify a failure of instruction, as well as a failure of learning. I concluded that if being able to explain what a progressive income tax means is an important learning outcome for the QL course, I had to revise my instruction to engage the students with the topic. And, the knowledge survey could be used to see if there was any improvement in students' confidence ratings after a change in the instruction.

The additional observations from the graph (Figure 3.2) that student confidence was fairly high for some of the items at the beginning of the semester, but fell off for topics at the end of the semester, also bear further discussion. Apparently, both are common occurrences. As [Nuhfer and Knipp \(2003\)](#) observed, students tend to be overconfident on the pre-survey, and typically the post-survey shows a drop in confidence for end-of-semester material, which may be due to a rush to cover topics before the end of the semester or students not yet having time to study the material. Student overconfidence at the beginning suggests that additional background knowledge checks should be undertaken before skipping or lightening coverage of critical topics.

How hard are knowledge surveys to make up? Anyone who has previously taught the course will probably find most of the needed material in old quizzes and tests or textbook exercises. It is important to check that all learning objectives are addressed, and that only material to be taught is included. Writing the survey for the QL course in Figure 3.1 took no more than a couple hours.

The instructor should avoid asking tricky questions on a knowledge survey. For example, in a knowledge survey for a first semester calculus course, including an integral problem that appears to be solvable by substitution but actually requires integration by parts, a technique not yet encountered, will mislead students. They are likely to rate their confidence high, when they probably cannot do the problem. This brings us to the accuracy of knowledge surveys as a measure of student learning.

[Nuhfer and Knipp \(2003\)](#) claimed that knowledge surveys provide a good representation of changes in student learning, and showed a graph that suggested final exam scores in an astronomy class were similar to post-survey results. They observed that instructors can usually offer explanations for gaps in learning revealed by knowledge surveys, as I was able to do for question #14 in the QL survey. Some ([Wirth & Perkins, 2005](#)) embraced knowledge surveys as a course design and assessment tool, while others ([Bowers, Brandon, & Hill, 2005](#)) presented evidence that they did not reliably predict final grades or final exam performance. This disagreement brought up issues related to reliability of exams ([Nuhfer & Knipp, 2005](#)) and students' ability to assess their own learning accurately ([Ehrlinger, Johnson, Banner, Kruger, & Dunning, 2008](#); [Kruger & Dunning, 1999](#)). Recent research ([Bell & Volckmann, 2011](#))

has suggested that confidence ratings on knowledge surveys are valid reflections of students' actual knowledge. However, students scoring high on the exams estimated their knowledge more accurately than the lower-scoring students, who were overconfident.

High ratings on knowledge surveys may not be foolproof evidence of student learning, but low ratings are strongly indicative of gaps in learning that could be investigated. Given their low burden in terms of time to construct and administer, their value in course planning, and their usefulness to students as a guide, knowledge surveys are worth serious consideration as one tool for triangulating evidence about changes in student learning.

We now consider two methods of gathering qualitative evidence that that we would expect few mathematics instructors to be familiar with: think-alouds and focus groups.

## Think-Alouds

A “think-aloud” is a method developed by social science researchers to study mental processes. Subjects are given a task and instructed to talk out loud as they perform the task. They are audio or video recorded. After completing the task, they can be asked to reflect back on and report all they remember about their thinking. The validity of these verbal reports as evidence of the subjects' thinking rests on their ability to verbalize the information that they are paying attention to as they perform the task. In other words, at the core of this method is the hypothesis that “the information that is heeded during the performance of a task is the information that is reportable; and the information that is reported is information that is heeded” (Ericsson & Simon, 1993, p. 167). There are several keys to eliciting accurate reporting that we discuss after we describe our experience using the method.

The think-aloud method proved effective when Curtis Bennett and I investigated the evolution of students' understanding of proof as they progressed through the Loyola Marymount University mathematics curriculum (Bennett & Dewar, 2007, 2013). As 2003–2004 Carnegie scholars and experienced college mathematics instructors, we wanted to investigate the level of understanding of proof our students had at critical points in our mathematics curriculum and then try to identify the learning experiences that promoted growth in their understanding. Toward these goals, we conducted think-alouds with 12 students in which they first investigated and then attempted to prove a simple conjecture from number theory. Our Carnegie mentors encouraged us to use this methodology and suggested several resources for doing so that we will pass along.

Before selecting the student subjects for the think-alouds we identified four courses in the curriculum that involved proof. We included at least two students who were at points in our curriculum just before and after these courses. A faculty member was recruited to serve as an expert subject for the think-aloud. The subjects were given pencil and paper and instructions to think-aloud as they investigated the following mathematical situation:

Please examine the statements:

For any two consecutive positive integers, the difference of their squares

(a) is an odd number, and

(b) equals the sum of the two consecutive positive integers.

What can you tell me about these statements?

Even if they made a general argument as part of their investigation, subjects were asked to try to write down proofs for their conjectures. Finally, students were given tasks to evaluate several sample proofs for correctness, to check whether they would apply the proven result, and to see if they thought a counterexample was possible after having a valid proof. Throughout, students were reminded to keep talking aloud.

The think-aloud transcripts provided us with more data than we could have derived from work written by students while performing the task. For example, they revealed affective influences on student performances. We found that lack of confidence could inhibit students from obtaining solutions and that a lack of motivation could prevent students from producing a proof that met standards for good mathematical writing. We obtained evidence that students could employ sophisticated mathematical thinking even though they lacked the tools or ability to communicate that thinking in a standard proof format. We saw how having additional knowledge might result in poorer performance on the task. We found that different types of knowledge (factual, procedural, strategic, epistemic, social, etc.) were required to produce an “expert” response. To describe accurately the spectrum of results revealed by the think-alouds, we developed a



typology of mathematical knowledge that, in part, drew on a typology of scientific knowledge (Shavelson & Huang, 2003). Our typology included six cognitive and two affective components. Looking at the student responses in detail, we were able to identify markers of student progress in several of the components. This allowed us to expand our typology into a two-dimensional description of mathematical knowledge and expertise by drawing on Alexander's 2003 model of domain learning. The end result was the development of a taxonomy of mathematical knowledge and expertise that took us far beyond answering the question about the evolution of our students' understanding of proof. This particular SoTL project exemplifies the interdisciplinary nature of SoTL work, because the findings resulted from employing disciplinary knowledge in combination with social science methodology, a typology from science education, and a model from K–12 education adapted to collegiate level mathematics (Bennett & Dewar, 2013). It also provides an example of theory-building occurring as an unintended outcome of a SoTL project.

### Advice for Conducting a Think-Aloud

Whether conducting a think-aloud for a single subject or for multiple subjects, a script, called a protocol, should be prepared for use. It should begin with specific instructions for thinking aloud and supply exercises so that the subject can experience thinking aloud. Subjects should practice reporting their thinking both concurrently and retrospectively (Ericsson & Simon, 1993). The concurrent report consists of what they say as they think-aloud while they perform the task. The retrospective report is obtained after they have completed the task by asking them to report anything else they can remember about their thinking while they did the task. We provide a sample warm-up script, derived from the protocol found in Ericsson and Simon (1993, pp. 377–378) and similar to the one used in our study. It contains practice exercises and requests both concurrent and retrospective reports. The words the researcher speaks to the subject appear in italics.

#### Sample Warm-up Script for Introducing a Subject to the Think-aloud Process

*Thanks for agreeing to be interviewed. The purpose of this interview is to take a close look at mathematics students' thinking processes. Your responses are completely voluntary and do not affect any coursework or grade. They will be kept confidential. If any portion of your response is made public, I will use a pseudonym so that you cannot be identified. You will be asked to do a "think aloud" while you are working on a problem or question. That is, I want you to tell me everything you are thinking from the time you first see the question until you give an answer. I would like you to talk aloud constantly from the time I present each problem until you have given your final answer. I don't want you to try to plan out what you say or try to explain to me what you are saying. It is not necessary to use complete sentences. Just act as if you are alone in the room speaking to yourself. Please face the desk. I'll sit behind you so that it seems more like you are alone and thinking aloud. It is most important that you keep talking. I will occasionally remind you to keep talking if you are silent for too long. Do you understand what I want you to do?*

*Good, first we will practice having you think aloud.*

*Here is the first practice question: How many windows are in the house or apartment where you live?*

After the subject completes the practice think-aloud task, Ericsson and Simon (1993) suggest that the researcher request a retrospective report in the following way.

*Good, now I want to see how much you can remember about what you were thinking from the time you read the question until you gave the answer. I am interested in what you actually can remember, rather than what you think you might have thought. If possible, I would like to tell me about your memories in the sequence in which they occurred while working on the question. Please tell me if you are uncertain about any of your memories. I don't want you to work on solving the problem again, just report all you can remember thinking about when answering the question.*

*Now tell me all that you can remember about your thinking.*

*Good, now I will give you two more practice problems before we proceed. I want you to do the same thing for each of these. I want you to think aloud as before as you think about the question. After you have answered it I will ask you to report all that you can remember about your thinking. Any questions?*

Additional practice exercises suggested by Ericsson and Simon include

- Name 20 animals (Subjects were told the researcher would keep track of the number.)
- Solve the anagram ALIOSC.

Because our task was mathematical, we wrote some mathematically oriented practice problems:

- If you have 87 candy bars and you need 106 candy bars for a Halloween party, how many more will you have to buy?
- If the sum of two whole numbers is odd, what can you conclude about their product, and why?
- If a right triangle has hypotenuse 25 and leg of 24, find the other leg.

The subject was told that he could use pencil and paper on the mathematical tasks. We alternated between the two types until the subject indicated he had enough practice.

As the sample protocol script indicates, the subject should sit facing away from the researcher. Sitting face-to-face makes it seem more like an interview to the subject. Because the subject is likely to be working on a mathematical task, we recommend having him sit at a desk and, if appropriate, have pencil and paper available to use. If the subject remains silent for too long, more than a few seconds, gently remind him to keep talking out loud by saying something neutral like, “Remember to keep talking out loud” or “Please tell me what you are thinking.”

The protocol script should cover the entire think-aloud, not just the initial practice task. A well-prepared script has the tasks aligned with the questions being investigated. In the case of multiple subjects, it helps the researcher ask each subject the same questions, in the same order.

Ericsson and Simon (1993) noted that sometime subjects give descriptions of the problem solving activity rather than thinking aloud. One indicator that subjects are describing rather than reporting is using the progressive tense, saying things like, “Now I am thinking I shall get one up there” (p. 244). This is not a report of the information the subject is paying attention to during the task; it is a response to the internal query, “What am I thinking about?” Ericsson and Simon considered these sorts of reports as less reliable. A researcher can gain more confidence in conducting think-alouds by experiencing one, that is, talking out loud while doing the practice tasks and the task the subjects will be asked to do. Conducting a trial think-aloud will also help.

Although the sample script said that the researcher would keep the identity of the subject confidential, and reporting one’s thinking while performing a mathematical task is not likely to reveal anything personal, the study needs to be cleared in advance as it represents research on human subjects. Chapter 4 provides more information on this.

Here is some additional practical advice for conducting think-alouds. Be sure that the audio or video recording devices are working and have sufficient battery power for the entire session. Once the sessions are recorded, the next step is to download the recordings as digital files to a computer. The files then require transcription, which takes either time or money. Surprisingly, it can take three or four hours to transcribe a single hour of taping. Transcription services can be hired to do this task for a fee (on the order of \$1.50 per minute of recording, depending on the quality of the audio). Choose a recording device that is reliable and compatible with the intended method of storage and transcription. Study the transcripts for insights or examples, using disciplinary knowledge or applicable theoretical frameworks. Coding transcript data for themes is a standard approach to analyzing this type of data. This process is described in the section titled Coding Data.

Next we discuss still another method of gathering evidence drawn from the social sciences, the focus group.

## Focus Groups

A focus group is a structured group process for obtaining the spread of opinions on an issue from a group of individuals. Unlike a one-on-one interview, a focus group allows participants to hear and interact with one another. It seeks a range of opinions, views, or perceptions, not a consensus. The number of participants can vary from as few as six to a dozen or more. One focus group takes far less time (usually 60 to 90 minutes) than conducting one-on-one interviews with each of the subjects. However, a personal interview would be the better choice if the topic to be explored is sensitive or if the researcher desires to explore the topic deeply with each individual subject.

John Holcomb (Chapter 12) stated that a focus group might have been a better method for him to explore his students' opinions about working in groups instead of the open-ended question he asked on a survey. Millis (2004) described a very efficient process for conducting a focus group and analyzing and reporting the findings. I found Millis's model easy to implement when conducting a focus group in the QL study with my colleagues (Dewar, Larson, & Zachariah, 2011).

### Model for Conducting a Focus Group

The following outline derived from Millis (2004) provides details for conducting a focus group.

#### Step 1. Preparation

- Find an impartial facilitator, someone not involved with the study, to conduct the focus group. The person should have some experience with conducting interviews, if not focus groups. Be sure that the person understands the procedures to be followed.
- Optional: Get an assistant facilitator to take notes or arrange to have the focus group recorded.
- Develop the questions to be asked based on careful consideration of the information desired.
- Recruit participants, typically students, whose perspectives are sought. Students can be offered pizza or snacks (before or after, not during) or extra credit homework points to participate. Let participants know that the focus group will start exactly on time, and there will be no late admissions.
- Optional: Prepare a short survey on paper for students who arrive early.
- Prepare a script for the facilitator, the presentation slides (containing the directions for the procedure outlined in Step 2), the roundtable sheet (one for each smaller team as noted in Step 2), the initial survey (if one is to be given to the early arrivals), and an index card for each participant.
- Have the seating arranged in a circle or U-shape.
- Start on time, close the door, and put a notice on it that late arrivals cannot enter. Everyone has to be present at the beginning.
- Turn on and check the recording device, if using one.

#### Step 2. The Focus Group Procedure

- Welcome and thank everyone, make introductions, and explain the purpose of the focus group—to gain a range of opinions about the topic.
- Explain the ground rules. The focus group is
  - *Anonymous*: Participants' comments will not be attributed to them individually. They will use an assigned number when making comments and referring to other participants.
  - *Candid*: Honest, constructive information is sought. All comments, positive and negative, are useful in gaining insight about the topic under discussion
  - *Confidential*: Only appropriate parties get the results. Participants should not discuss comments afterwards.
  - *Productive*: Indicate to the participants how the information will be used.
- Index card activity
  - Each participant takes an index card.
  - Participants count off #1, #2, etc. to get their numbers.
  - Each participant puts his or her number in the upper right hand corner of the index card and announces the number each time before speaking.
  - On the index card, each person writes a word or phrase to describe his or her impression of the focus topic. For example, John Holcomb (Chapter 12) might have had participants write a word or phrase to describe their impressions of working in groups on the statistical projects. Below that word or phrase, they are asked to write a number from 1 to 5, where 1 is Low and 5 is High, that describes the level of their satisfaction related to the topic. In Holcomb's case, it could have been their satisfaction with how well their group interacted.

- Once the writing and rating are completed participants are directed to report to the whole group as follows (while the assistant takes notes):
  - “Would a volunteer please read your phrase and rating?” Remind the speaker to state his or her number before speaking.
  - “Please tell why you wrote that phrase and rated the way you did.”
  - Prompt a next volunteer with, “Can we hear from someone else?” (Remind the speaker to state his or her number before speaking.)
  - Continue until all speak or signal an anticipated end with, “Let’s hear from two more.”
- Collect the cards.
- Roundtable activity
  - Going around the circle, divide the group into teams of three or four adjacent participants.
  - Provide each team with a piece of paper with directions to pass the sheet of paper rapidly from person to person with each person jotting down one strength or positive aspect of the intervention or topic under study, saying it aloud while writing it down. They should put their participant numbers next to what they write. An individual can say “Pass” without writing.
  - Working as a team, members then rank-order the top three strengths they identified, with the most important strength first.
  - This listing and ranking is repeated for drawbacks or things that could be improved using the other side of the sheet of paper. The most important thing to improve is placed first.
  - The papers are collected.
- Open-Ended Question(s)
  - The remaining time can be devoted to full group consideration of six to eight open-ended questions. They should be presented in a logical order.
  - As before, participants state their numbers before speaking.
  - Responses can be structured as a round robin, that is, starting at some point in the circle, and going from person to person to get a response. Anyone can pass. This approach ensures that everyone has a chance to speak, and avoids having to solicit volunteers.
- Wrap-up
  - As a closing activity, group members can be invited to say one thing they heard that was really important.
  - Thank the participants and remind them of the confidentiality of what they heard and how the information will be used.

### Step 3. Analyze the Results

- The index card activity produces a set of numbers (ranging from 1 to 5) to average and a collection of short phrases to code or categorize as described in the Coding Data section in this chapter.
- The teams in the roundtable activity have identified and rank-ordered strengths and weaknesses, which can be coded for common themes.
- The recordings or notes taken can be transcribed, if a full record of the discussion is desired, and the transcripts can be coded.

Millis (2004) wrote that a typical transcript entry “might look like this: ‘Student Seven: Unlike Student Two, I felt the evaluation methods were extremely fair because the essays gave us a chance to demonstrate what we actually knew. I found myself integrating ideas in ways I hadn’t thought of before’” (p. 130). She advised paying for a transcription service to produce the transcripts. The appendices in her paper show how to represent the data from the index card and roundtable activities as histograms and color-coded tables.

## Analyzing Qualitative Data

The analysis of qualitative data frequently involves rubrics or coding, neither of which may be familiar to mathematics faculty members. Rubrics are applied to student work such as exams, solutions, papers, projects, or portfolios assigned

Dimensions	Level of Performance Ratings		
	2	1	0
Understanding the Problem	Complete understanding of the problem	Part of the problem misunderstood or misinterpreted	Complete misunderstanding of the problem
Planning a Solution	Plan could have led to a correct solution if implemented properly	Partially correct plan based on part of the problem being interpreted correctly	No attempt, or totally inappropriate plan
Getting an Answer	Correct answer and correct label for the answer	Copying error; computational error; partial answer for a problem with multiple answers	No answer, or wrong answer based on an inappropriate plan

**Table 3.3.** A rubric to evaluate a mathematical problem solving task (Charles, Lester, & O’Daffer, 1987)

in a course, work that is submitted for a grade. The rubric may be an integral part of the grading process or it can be applied separately to measure those aspects of learning that are the focus of the SoTL study. Coding is a method of analyzing textual data such as that gathered from interviews, focus groups, reflective writing assignments or open-ended responses to surveys to determine common themes. Coding schemes can also be applied to written or recorded mathematical work to determine and categorize common mistakes, false starts, or misunderstandings in proofs. We now describe rubrics and coding data in more detail.

## Rubrics

Several of the chapters in this book (see, for example, Chapter 5 by Gretchen Rimmasch and Jim Brandt and Chapter 12 by John Holcomb) make reference to using rubrics to assess student work. A rubric is a guide for evaluating certain dimensions or characteristics of student work. In a SoTL study, these would be the skills or aspects of the learning under investigation. For each dimension, different levels of performance are defined, labeled, and described. A well-crafted rubric provides the criteria by which a task or piece of work will be judged and presents a rating scale with “descriptors” that distinguish among the ratings in the scale. The descriptors clarify the type of work that will be assigned to each level of achievement and help users apply the rubric consistently over time. For example, Charles, Lester, and O’Daffer (1987, p. 30) presented a simple rubric (Table 3.3) for grading problem solving tasks along three dimensions: Understanding the Problem, Planning a Solution, and Getting an Answer.

Rubrics allow a SoTL researcher to reveal specific skills or aspects of learning that an intervention is affecting. For example, if a SoTL inquiry asked whether presenting heuristic strategies, such as solve a simpler problem, collect data to look for a pattern, and work backwards (Pólya, 1957), improved students’ problem solving ability, the rubric in Table 3.3 would help show the effect the instruction was having, in a way that pre- and post-scores on a problem solving test would not. For an example of a rubric derived from the one shown in Table 3.3 applied to a student work sample, see Emenaker (1999).

Some discussions of rubrics within the mathematical community use “rubric” more broadly, to include simple yes/no checklists (Crannell, 1999) or so-called holistic rubrics that in effect combine dimensions and ratings together into a single scale. Table 3.4 contains an example of a holistic rubric for problem solving derived from the California Assessment Program (1989).

Rubrics can be developed for virtually any student work product, presentation, or behavior and, as we have seen, they make possible the assessment of separate aspects of a complex task. An internet search may turn up an existing rubric that can be used or modified. Another option is to create one specifically for the investigation.

### How to Create a Rubric

We begin with a description of the steps involved in creating a rubric.

1. Specify the student learning outcome to be measured with the rubric and the task or piece of work that it will judge.

Score	Descriptor
6	<i>Exemplary response:</i> Gives a complete response with a clear, coherent, unambiguous, and elegant explanation; includes a clear and simplified diagram; communicates effectively to the identified audience; shows understanding of the open-ended problem's mathematical ideas and processes; identifies all the important elements of the problem; may include examples and counterexamples; presents strong supporting arguments.
5	<i>Competent response:</i> Gives a fairly complete response with reasonably clear explanations; may include an appropriate diagram; communicates effectively to the identified audience; shows understanding of the problem's mathematical ideas and processes; identifies the most important elements of the problem; presents solid supporting arguments.
4	<i>Minor flaws but satisfactory:</i> Completes the problem satisfactorily, but the explanation may be muddled; argumentation may be incomplete; diagram may be inappropriate or unclear; understands the underlying mathematical ideas; uses mathematical ideas effectively.
3	<i>Serious flaws but nearly satisfactory:</i> Begins the problem appropriately but may fail to complete or may omit significant parts of the problem; may fail to show full understanding of mathematical ideas and processes; may make major computational errors; may misuse or fail to use mathematical terms; response may reflect an inappropriate strategy for solving the problem.
2	<i>Begins, but fails to complete problem:</i> Explanation is not understandable; diagram may be unclear; shows no understanding of the problem situation; may make major computational errors.
1	<i>Unable to begin effectively:</i> Words do not reflect the problem; drawings misrepresent the problem situation; copies parts of the problem but without attempting a solution; fails to indicate which information is appropriate to the problem.

**Table 3.4.** A holistic-scale rubric for mathematical problem solving ([California Assessment Program, 1989](#))

2. Identify the dimensions of the task or work that the rubric will measure and determine what each includes.
3. For each dimension, decide on the number of performance levels and provide a description for each level. The number of performance levels can vary by dimension, and different weights can be given to each dimension.
4. Test the rubric and, if needed, revise or further develop it.

We now illustrate this process of developing a rubric for a fictitious SoTL study. Suppose as part of a study of a senior capstone course, we pair students as peer coaches for each other's end-of-semester presentation in an attempt to improve their oral presentations. The student learning outcome we wish to measure is how well students have developed the ability to give a mathematical talk and we intend to apply a rubric to the end-of-semester presentation they make on their research project. That completes Step 1.

In the second step of the rubric development process we identify four dimensions of the oral presentation that we want to measure and we provide a detailed description (shown in parentheses) for each dimension:

- Preparation (speaker arrived on time with handouts or materials prepared, seemed familiar with the audio-visual projection system, was appropriately dressed, appeared to have practiced the talk, stayed within the time frame).
- Organization (talk had: an introduction that captured the audience's attention and provided basic information with definitions or examples, as appropriate; a middle that presented data, summarized the method, argument, or proof, and gave the results obtained; and a clear end).
- Knowledge (speaker appeared to understand the mathematical content, was able to explain it clearly without constantly referring to notes, and could answer questions).
- Presentation Skills (presenter spoke clearly and loudly, and made eye contact).

For Step 3, we decide we will start with three performance levels. We can add more later, if finer distinctions are needed. There are many ways to label the levels, for instance, Poor-Acceptable-Good, Emerging-Developing-Advanced, or Limited-Acceptable-Proficient. Numerical values can be 1-2-3, 1-3-5, or something else. We can include sub-dimensions that are rated individually using the same or different scales. For example, for in our Presentation

Presentation Skills Sub-Dimensions	Level of Performance Ratings		
	Needs Improvement (0)	Adequate (1)	Very Good (2)
Eye contact is made	Never or rarely	Some of the time	Most or all of the time
Speaks loudly and clearly	Never or rarely	Some of the time	Most or all of the time

**Table 3.5.** Rubric for the presentation skills dimension of giving a mathematical talk

Skills dimension, for the two sub-dimensions related to speaking loudly and clearly and maintaining eye contact we could use very simple descriptors, Never or rarely (with a value of 0 points), Some of the time (1 point), and Most or all of the time (2 points). If we choose the labels Needs Improvement, Adequate, and Very Good for the performance levels, we could represent this portion of the rubric by the matrix shown in Table 3.5.

On the other hand, our more complex Preparation dimension needs more elaborate descriptors. Keeping the labels Needs Improvement, Adequate, and Very Good, we select 1-3-5 for numerical ratings. This will allow the possibility of assigning intermediate values 2 or 4 in performances that fall between the descriptors. As shown in Table 3.6, we use our detailed description for the Preparation dimension to write the descriptors. A zero (0) value could be assigned for a case where all five indicators of lack of preparation are present.

To complete the rubric, we would need to construct similar tables for the other two dimensions: Organization and Knowledge. The Presentation Skills dimension has a total value of 4 points, while the Preparation dimension has a total of 5 points. If we feel some dimensions are more important than others, we can assign different weights to the dimensions.

For the final step, testing our rubric, we could try it on talks given in other courses, video recorded talks on the internet, or on colloquium speakers who come to the department. We would ask students to comment on how clear and understandable it is as a guide for their work. We would try to recruit colleagues to review or pilot test it and give feedback. Gathering input from other faculty members or students at earlier steps in the development of the rubric is also worth considering. If several people are going to use the same rubric in a given study, it is essential to practice applying it to the same presentation and discuss the results to arrive at a common understanding so that it can be employed consistently and fairly. When applied to student work, in this case to qualitative data in the form of oral presentation of a research project, the rubric enables us to assess the performance along several dimensions, and produces a numerical rating for each dimension and for the presentation as a whole.

Dimension	Level of Performance Ratings		
	Needs Improvement (1)	Adequate (3)	Very Good (5)
Preparation	Lack of preparation indicated by at least two of the following: confusion at the beginning over materials leads to delayed start; speaker struggles with projection system or presentation software due to lack of familiarity or incompatibility of equipment; length of talk departs from time allotted by a significant amount; speaker reads most of talk from a script or notes; unprofessional dress or demeanor detracts from speaker's credibility.	Talk starts on time; little, if any, problem with the projection system that could have been detected in advance; talk ends approximately on time; speaker is able to give talk with only occasional reference to notes; dress or demeanor does not detract from speaker's credibility.	Talk starts on time; no problem with the projection system that could have been detected in advance; talk is well-paced to end on time; speaker is able to give talk with little, if any, reference to notes; dress and demeanor enhance speaker's credibility.

**Table 3.6.** Rubric for the preparation dimension of giving a mathematical talk

Applying the rubric to presentations given in a senior capstone course, first in a semester when no peer coaching system was in place and later when there was one, would generate data on the value of the peer coaching. Of course, this should not be the only data gathered for our hypothetical study. Also, whether the students have access to the rubric prior to the presentation should be considered, because rubrics can have significant pedagogical value. They clarify the instructor's expectations for an assignment or project, allowing students to assess their own work. Rubrics enable the instructor to provide nuanced feedback on projects and presentations. See, for example, Dewar, Larson, and Zachariah (2011, pp. 626–629). For a more in-depth discussion of rubrics, consult *Introduction to Rubrics* (Stevens, 2005). A project by the American Association of Colleges and Universities, *VALUE: Valid Assessment of Learning in Undergraduate Education*, developed fifteen rubrics during an eighteen-month iterative process, including rubrics for written communication, oral communication, quantitative literacy, teamwork, and problem solving. It may be instructive to look at them. (See [www.aacu.org/value/index.cfm](http://www.aacu.org/value/index.cfm).)

We now turn our attention to a method to distill and categorize information from textual data found in transcripts, reflective writing assignments, or open-ended responses to survey questions.

## Coding Data (Content Analysis)

Data collection methods such as think-alouds, focus groups, and open-ended questions in surveys produce textual data, typically a lot of it. Meaning is extracted from verbal data by using labels (codes) to identify common themes that reoccur throughout a data set. The codes make it possible to categorize and organize information as it relates to the research question of the SoTL study. This content analysis allows the researcher to discover patterns that would be difficult to detect by reading alone. The resulting codes help the SoTL scholars to communicate findings to others. The biggest concern when a researcher is new to coding data is how to find the right interpretation. But, in practice, there can be many right interpretations. The key is to explain whatever interpretation is put forth, typically with textual samples, so that others can understand how it was developed (Auerbach & Silverstein, 2003, p. 32).

Reading the raw text, let alone analyzing it and drawing conclusions about patterns, can seem an overwhelming task. (To analyze non-numerical, non-verbal data such as graphs, diagrams, or concept maps, we would identify characteristics of interest and describe them in words yielding verbal data that could then be coded.) We provide an overview and several examples and then outline the steps in the coding process.

The categories or codes that we use can be “predetermined” or allowed to “emerge” from the data. Predetermined categories, sometimes called *a priori* categories, are those suggested by the research question or a published theory. For example, in mathematics, a research question might be “Does using graphing calculators cause students to think more geometrically?” Students could be given similar mathematical tasks to perform first without and then with graphing calculators, or before and after using graphing calculators as a tool. The investigation might use a think-aloud methodology. The data would be analyzed and categorized according to whether it contains markers of *geometric* or *algebraic* thinking.

In Chapter 13, Derek Bruff describes two coding schemes that he developed, both based on the revised Bloom's taxonomy (Anderson & Krathwohl, 2001). He used the six categories—remember, understand, apply, analyze, evaluate, and create—from the taxonomy's cognitive process dimension to code the questions he had asked students on pre-reading quizzes. To analyze and describe the difficulties his students reported having, he coded their responses to this open-ended question: “Please write about the part of the reading that you found most difficult. Try to be as specific as possible. If you did not find the reading difficult, write about the part you found most interesting.” This time his codes—factual, conceptual, procedural, and metacognitive—were derived from the knowledge dimension of Anderson and Krathwohl's (2001) revised Bloom's taxonomy.

In both situations, the researcher began coding knowing, or at least thinking he knew, exactly what to look for. In the first example, it was markers of geometric thinking versus indications of algebraic approaches. In the second, Bruff relied on categories drawn from a well-known taxonomy. Sometimes predetermined categories can be as simple as positive, negative, and neutral. This coding scheme served John Holcomb's purposes in Chapter 12 to ascertain students' perceptions (reported in an open-ended survey question) of the effectiveness of working in groups on realistic projects to learn statistics or to develop communication skills.

Sometimes the researcher has no idea what categories or themes to look for. She has to let the data suggest the themes. This is called using emergent or inductive coding. My study of students' views of mathematics provides an example



Category	Example
Numbers (including computation)	the study of numbers
Listing of Topics (could include numbers)	algebra, pictures, numbers, everything encompasses math
Applications	a “mental gymnastic” that helps you solve real-world and theoretical problems
Pattern, Proof, or Logic	the search for and the study of patterns
Structure, Abstraction, or Generalization	the analysis of abstract systems
Other	a language or everything is mathematics

**Table 3.7.** Emergent categories and example responses to *What is mathematics?*

of this type of coding (Dewar, 2008). The study investigated how undergraduate STEM students’ understandings of mathematics compare to an expert view of mathematics, and whether a single course can enhance future teachers’ views of mathematics. I gathered written responses to “What is mathematics?” first from 55 mathematics and computer science students and 16 mathematics faculty members, and later from seven future teachers at the beginning and end of a particular course. Curtis Bennett and I independently read the responses several times and were able to identify many obvious repeated themes, such as that mathematics was about numbers or it involved making logical arguments. After comparing notes, we settled upon six emergent categories and provided an example each category, as shown in Table 3.7. After another colleague was able to use the information in Table 3.7 to code the original data set with similar results, we were confident that others could understand the coding scheme. The emergent categories became predetermined categories for coding the responses of the seven future teachers.

After completing the coding, I found the work of Schwab (1964) on classifying the disciplines. I noticed that our six emergent categories each aligned with one of the four bases (content boundaries, skills and habits employed by practitioners, modes of inquiry, and purposes or outcomes for the disciplinary work) used by Schwab to classify disciplines. Specifically, both Numbers and Listing of Topics delineated content boundaries, while the Pattern, Proof, or Logic category described modes of mathematical inquiry, in particular, how mathematicians formulate and determine the truth of conjectures. The category Structure, Abstraction, or Generalization corresponded to skills or habits of mind practiced by mathematicians. The Applications category gave a purpose for doing mathematics. The alignment of the categories with Schwab’s theoretical framework provided theoretical support for the appropriateness of the emergent categories.

### Process for Coding Data

We now present a step-by-step process for coding derived from Auerbach and Silverstein (2003), Taylor-Powell and Renner (2003), and the Visible Knowledge Project at Georgetown University (see [blogs.commonsgorgetown.edu/vkp/2009/01/16/coding-data1](https://blogs.commonsgorgetown.edu/vkp/2009/01/16/coding-data1)). The process is reasonable for coding ten to 50 pages of textual data manually. For larger data sets, consider using specialized software that is available for the content analysis of qualitative data.

1. Put the data into a single (Word or Excel) file. Assign an identifying name or number to each respondent’s data.
2. Read and reread the file to become familiar with the data.
3. Keeping the research question in mind, select segments of text that address it.
4. Label in the margin why the text was selected by hand or using Word’s comment feature. This is called “memoing.” The codes or categories will begin to emerge. For a simple example, see *Analyzing Qualitative Data* (Taylor-Powell & Renner, 2003, p. 4).
5. Keep a master list of labels (codes) generated in Step 4.

That completes the initial coding. The next phase, focused coding, includes the steps:

6. Review, revise, and reorganize the codes. Eliminate less useful codes, combine smaller categories into larger ones, and subdivide categories with a large number of segments.
7. Examine the codes for related ideas to identify larger themes or to develop theory.

8. For advanced work, the final step would be to build a hierarchical diagram, a relational chart, or a network diagram. (For examples, see pp. 6–10 of [www.southalabama.edu/coe/bset/johnson/lectures/lec17.pdf](http://www.southalabama.edu/coe/bset/johnson/lectures/lec17.pdf).)

### Assessing the Quality of Content Analysis

We begin with a summary of how the quality of quantitative research is measured. Quantitative research is judged by its reliability, validity, and generalizability. A reliable instrument will yield consistent results, or as John Holcomb writes (Chapter 12, p. 122), “According to Reynolds, Livingston, and Willson (2006), ‘in simplest terms, in the context of measurement reliability refers to consistency or stability of assessment results.’” In Chapter 17, O’Hanlon and her colleagues describe the work they did to check the reliability of their survey.

Validity is the extent to which an instrument measures what it claims to measure. Citing Reynolds, Livingston, and Willson (2006), Holcomb (Chapter 12, p. 122) tells us, “content validity ‘involves how adequately the test samples the content area of the identified construct. . . Content validity is typically based on professional judgments about the appropriateness of the test content.’” It is possible for an assessment tool to be consistent but lack validity. In Chapter 15, Stephen Szydlak discusses how he checked the validity of his survey instrument.

If a finding from a study conducted on a sample population can be applied to the population at large, the result is said to be generalizable. A well-designed, well-executed study that uses the gold standard of random assignment of subjects to an experimental and a control group is considered to yield generalizable results.

What are the corresponding standards for qualitative research? According to Auerbach and Silverstein (2003, pp. 77–87), qualitative data analysis must be *justifiable*. This means it has the following three characteristics:

1. The steps used to arrive at the interpretation are *transparent*. This does not mean others would end up with the same interpretation, but that they understand how the researcher obtained his or her interpretation.
2. The themes that were developed are *communicable*. This means that others can understand the themes even though they may not agree with them.
3. The themes (and constructed theories, if any) tell a *coherent* story.

Being justifiable is the qualitative researcher’s counterpart to reliability and validity in quantitative analysis.

Another measure of quality is whether larger, more abstract themes (that is, constructed theories) are *transferable* to other populations or situations. Small SoTL studies usually don’t involve developing larger constructs, so transferability may not come into play. Being transferable is the qualitative researcher’s counterpart to generalizability in quantitative analysis. One additional measure of the quality of a content analysis of qualitative data is *inter-coder reliability*. This refers to another person being able to code with similar results using your categories. The accepted standard is to achieve at least 80% agreement.

The content analysis employed in the *What is Mathematics?* study illustrates several of these concepts. Because I was able to apply the codes developed from STEM majors and mathematics faculty members to a group of future teachers, most of whom were preparing to teach elementary school (hence non-STEM majors), it provides an example of transferability of the codes (but not a larger theory) to a different population. My ability to correlate the categories with the work of Schwab (1964) showed that they merge with a larger theory. I also tested inter-coder reliability by having a colleague successfully code the original data set using the categories and examples. In Chapter 15, Szydlak also analyzes student responses to the question “What is mathematics?” with similar, but not identical, categories emerging. This demonstrates the possibility of more than one “right” interpretation. More examples of coding data appear in other chapters. For example, in Chapter 7 Lynn Gieger and her colleagues describe their coding process in detail and cite additional references.

## References

- Adhikari, A., & Nolan, D. (2002). But what good came of it at last? How to assess the value of undergraduate research. *Notices of the AMS*, 49(10), 1252–1257. Retrieved from [www.ams.org/notices/200210/comm-nolan.pdf](http://www.ams.org/notices/200210/comm-nolan.pdf)
- Alexander, P. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, 32(8), 10–14.

- Anderson, L., & Krathwohl, D. A. (2001). *Taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.
- Auerbach, C., & Silverstein, L. (2003). *Qualitative data: An introduction to coding and analysis*. New York, NY: New York University Press.
- Bell, P., & Volckmann, D. (2011). Knowledge surveys in general chemistry: Confidence, overconfidence, and performance. *Journal of Chemical Education*, 88(11), 1469–1476.
- Bennett, C., & Dewar, J. (2007). Developing and applying a taxonomy for mathematical knowledge-expertise. *Electronic Proceedings for the Tenth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education Conference on Research in Undergraduate Mathematics Education*, San Diego CA. Retrieved from [sigma.maa.org/rume/crume2007/eproc.html](http://sigma.maa.org/rume/crume2007/eproc.html)
- Bennett, C., & Dewar, J. (2013). SoTL and interdisciplinary encounters in the study of students' understanding of mathematical proof. In K. McKinney (Ed.), *The scholarship of teaching and learning in and across the disciplines* (pp. 54–73). Bloomington, IN: Indiana University Press.
- Berliner, D. (2002). Educational research: The hardest science of all. *Educational Researcher*, 31(8), 18–20.
- Bowers, N., Brandon, M., & Hill, C. (2005). The use of a knowledge survey as an indicator of student learning in an introductory biology course. *Cell Biology Education*, 4, 311–322.
- Brase, C. H., & Brase, C. P. (2013). *Understanding Basic Statistics* (6th ed.). Boston, MA: Brooks/Cole.
- California Assessment Program. (1989). *A question of thinking: A first look at students' performance on open-ended questions in mathematics*. Sacramento, CA: California State Department of Education.
- Charles, R., Lester, F., & O'Daffer, P. (1987). *How to evaluate progress in problem solving*. Reston, VA: National Council of Teachers of Mathematics.
- Crannell, A. (1999). Assessing expository mathematics: Grading journals, essays, and other vagaries. In B. Gold, S. Z. Keith, & W. A. Marion (Eds.), *Assessment Practices in Undergraduate Mathematics*, MAA Notes Series #49. Washington, DC: Mathematical Association of America. Retrieved from [www.maa.org/sites/default/files/pdf/ebooks/pdf/NTE49.pdf](http://www.maa.org/sites/default/files/pdf/ebooks/pdf/NTE49.pdf)
- Cook, T.D., & M.R. Payne. (2002). Objecting to the objections to using random assignment in educational research. In F. Mosteller & R. Boruch (Eds.), *Evidence matters: Randomized trials in education research* (pp. 150–178). Washington, DC: Brookings Institution Press.
- Dewar, J. (2006). Increasing math majors' success and confidence though problem solving and writing. In F. Rosamond & L. Copes (Eds.), *Educational transformations: The influences of Stephen I. Brown* (pp. 101–121). Bloomington, IN: AuthorHouse.
- Dewar, J. (2008). What is mathematics: Student and faculty views. *Electronic Proceedings for the Eleventh Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education Conference on Research in Undergraduate Mathematics Education*, San Diego, CA. Retrieved from [mathed.asu.edu/crume2008/Proceedings/Proceedings.html](http://mathed.asu.edu/crume2008/Proceedings/Proceedings.html)
- Dewar, J. (2010, January). Using a knowledge survey for course and program level assessment in mathematics. Paper presented at the 2010 Joint Mathematics Meetings, San Francisco, CA. Abstract retrieved from [jointmathematicsm meetings.org/meetings/national/jmm/1056-f1-443.pdf](http://jointmathematicsm meetings.org/meetings/national/jmm/1056-f1-443.pdf)
- Dewar, J., Larson, S., & Zachariah, T. (2011). Group projects and civic engagement in a quantitative literacy course. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 21(7): 606–637.
- Ehrlinger, J., Johnson, K., Banner, M., Kruger, J., & Dunning, D. (2008). Why the unskilled are unaware: Further exploration of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, 105, 98–121.

- Emenaker, C. (1999). Assessing modeling projects in calculus and precalculus: Two approaches. In B. Gold, S. Z. Keith, & W. A. Marion (Eds.), *Assessment practices in undergraduate mathematics*, MAA Notes Series #49. Washington, DC: Mathematical Association of America. Retrieved from [www.maa.org/sites/default/files/pdf/ebooks/pdf/NTE49.pdf](http://www.maa.org/sites/default/files/pdf/ebooks/pdf/NTE49.pdf)
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Grauerholz, L., & Main, E. (2013). Fallacies of SoTL: Rethinking how we conduct our research. In K. McKinney (Ed.), *The scholarship of teaching and learning in and across the disciplines* (pp. 152–168). Bloomington, IN: Indiana University Press.
- Gravetter, F. J., & Wallnau, L. B. (2011). *Essentials of statistics for the behavioral sciences* (7th ed.). Belmont, CA: Wadsworth/Cengage Learning.
- Hake, R. (1998) Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66, 64–74.
- Hutchings, P. (2000). *Opening lines: Approaches to the scholarship of teaching and learning*. Menlo Park, CA: The Carnegie Foundation for the Advancement of Teaching.
- Hutchings, P. (2002). *Ethics of inquiry: Issues in the scholarship of teaching and learning*. Menlo Park, CA: The Carnegie Foundation for the Advancement of Teaching.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessment. *Journal of Personality and Social Psychology*, 77, 1121–1134.
- Light, R., Singer, J., & Willett, J. (1990). *By design: Planning research on higher education*. Cambridge MA: Harvard University Press.
- Millis, B. (2004). A versatile interactive focus group protocol for qualitative assessment. In C. M. Wehlburg & S. Chadwick-Blossey (Eds.), *To improve the academy* (Vol. 22, pp. 125–141). San Francisco, CA: Jossey-Bass.
- Nuhfer, E., & Knipp, D. (2003). The knowledge survey: A tool for all reasons. In C. Wehlburg & S. Chadwick-Blossey (Eds.), *To improve the academy* (Vol. 21, pp. 59–78). San Francisco, CA: Jossey-Bass.
- Nuhfer, E., & Knipp, D. (2005). Re: The use of a knowledge survey as an indicator of student learning in an introductory biology course. *Cell Biology Education*, 5, 313–316.
- Pólya, G. (1957). *How to solve it*. Garden City, NY: Doubleday.
- Poole, G. (2013). Square one: What is research? In K. McKinney (Ed.), *The scholarship of teaching and learning in and across the disciplines* (pp. 135–151). Bloomington, IN: Indiana University Press.
- Ramsey, F., & Schafer, D. (2013). *The statistical sleuth: A course in methods of data analysis* (3rd ed.). Boston, MA: Brooks/Cole.
- Reynolds, C. R., Livingston, R. B., & Wilson, V. (2006). *Measurement and assessment in education*. Boston, MA: Pearson Education Inc.
- Schwab, J. (1964). Structure of the disciplines. In G. W. Ford & L. Pugno (Eds.), *The structure of knowledge and the curriculum* (pp. 6–30). Skokie, IL: Rand McNally.
- Shavelson, R.J., & Huang, L. (2003). Responding responsibly to the frenzy to assess learning in higher education. *Change*, 35(1), 10–19.
- Shavelson, R. J., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academies Press.
- Stevens, D. (2005). *Introduction to rubrics*. Sterling, VA: Stylus.
- Stewart, J., & Stewart, G. (2010). Correcting the normalized gain for guessing. *The Physics Teacher*, 48(3), 194–196.

- Suskie, L. A. (1996). *Questionnaire survey research: What works* (2nd ed.). Tallahassee, FL: The Association for Institutional Research.
- Taylor-Powell, E. (2002). *Program development and evaluation, Focus group interviews, Quick tips #5*. Madison, WI: University of Wisconsin-Extension. Retrieved from [www.uwex.edu/ces/pdande/resources/pdf/Tipsheet5.pdf](http://www.uwex.edu/ces/pdande/resources/pdf/Tipsheet5.pdf)
- Taylor-Powell, E., & Renner, M. (2003). *Analyzing qualitative data*. Madison, WI: University of Wisconsin Extension. Retrieved from [learningstore.uwex.edu/pdf/g3658-12.pdf](http://learningstore.uwex.edu/pdf/g3658-12.pdf)
- Trochim, W. (2006a). Inferential statistics. In *Research methods knowledge base*. Retrieved from [www.socialresearchmethods.net/kb/statinf.php](http://www.socialresearchmethods.net/kb/statinf.php)
- Trochim, W. (2006b). Types of data. In *Research methods knowledge base*. Retrieved from [www.socialresearchmethods.net/kb/datatype.php](http://www.socialresearchmethods.net/kb/datatype.php)
- Wiggins, G. (1998). *Educative assessment*. San Francisco, CA: Jossey-Bass.
- Wirth, K., & Perkins, D. (2005). *Knowledge surveys: An indispensable course design and assessment tool*. Paper presented at the Innovations in the Scholarship of Teaching and Learning at Liberal Arts Colleges, St. Olaf, MN. Retrieved from [www.macalester.edu/geology/wirth/WirthPerkinsKS.pdf](http://www.macalester.edu/geology/wirth/WirthPerkinsKS.pdf)
- Working Group on Statistics in Mathematics Education Research. (2007). *Using statistics effectively in mathematics education research: A report from a series of workshops organized by the American Statistical Association with funding from the National Science Foundation*. Alexandria, VA: The American Statistical Association. Retrieved from [www.amstat.org/education/pdfs/UsingStatisticsEffectivelyinMathEdResearch.pdf](http://www.amstat.org/education/pdfs/UsingStatisticsEffectivelyinMathEdResearch.pdf)

# 4

## Resources for Pursuing SoTL and Going Public

Jacqueline M. Dewar  
*Loyola Marymount University*

### Introduction

This chapter, which closes Part I, offers additional resources and advice for completing a SoTL project. These include the need to obtain human subjects clearance in order to publish the results of the study, why it is a good idea to find collaborators for doing SoTL, and where to find them. Suggestions are offered for other sources of support and possible venues for dissemination. As in the previous chapters, the work of the authors in Part II provides examples.

### Human Subjects Considerations

At the outset of a SoTL investigation, if the goal is to publish the results, then human subjects issues will arise, which will be unfamiliar to many mathematicians. According to United States Federal Guidelines, a human subject is a person about whom an investigator (whether professional or student) conducting research obtains data through intervention or interaction with the individual or identifiable private information (32 CFR 219.102.f). Because of past abuses of human subjects in medical trials in populations such as prisoners or minorities in the armed forces, the federal government has developed procedures requiring informed consent for human subjects research (U.S. Department of Health and Human Services, n.d.). Special rules for obtaining informed consent apply to any subject under the age of 18, a situation that can be encountered in SoTL studies of first-year college courses.

Because SoTL publications may involve making the work of our students public, we must follow institutional guidelines for working with human subjects. Most colleges and universities have a committee or group, often called an Institutional Review Board (IRB) or a Human Subjects Review Board, charged with ensuring that the federal guidelines are observed.

Human subjects researchers are expected to inform their subjects of the risks of their involvement in the study and obtain written consent for their participation. Studies that involve little or no physical or emotional risk to the subject, and will not reveal anything about the subject's behavior that would be damaging if it became public, may be exempted from obtaining written consent. Since SoTL research rarely involves a risk to students, it is common to apply to the review board for an exemption. Whether the researcher thinks the study qualifies as exempt or not, she should contact the campus review board and follow its guidelines and procedures. A number of contributors discuss

their IRB applications. For example, Stephen Szydlak (Chapter 15) gives details about his experience with the process, and he includes a copy of the Informed Consent Form used in his study.

As part of the IRB process, the researcher will likely be asked to describe the study's goals, methods, the intended subjects, and what safeguards will be put into place to ensure their safety and privacy. Typically, the researcher must demonstrate familiarity with the federal requirements for protecting of human subjects, usually by completing an online tutorial program. Certification programs in research ethics include *Protecting Human Research Participants*, available free through the National Institutes of Health ([phrp.nihtraining.com](http://phrp.nihtraining.com)), and *Protection of Human Subjects* ([www.citiprogram.org](http://www.citiprogram.org)), offered as a subscription service by the Collaborative Institutional Training Initiative. The time required to complete the tutorials varies from one to several hours. Once obtained, the certifications are usually good for a number of years. How long a certification is considered valid and which tutorial programs are recognized will vary from campus to campus.

In theory, the IRB process must be completed before collecting any data, but if there are existing data, they may be usable. Review boards may grant permission to work with data collected before applying for permission to work with human subjects if the data were gathered as part of normal educational practices. Pam Crawford (Chapter 9) and Derek Bruff (Chapter 13) each mention obtaining permission to use existing data.

Because policies vary, it is essential to become familiar with the procedures and requirements of the local review board. Colleagues on campus in education or psychology are likely to have experience with the process. Asking them for advice is one way to start a collaboration.

## Finding Collaborators and Support

SoTL lends itself to collaboration both formal and informal. Collaborators provide new insights or access to expertise or strengths beyond that of a single scholar (Bennett & Dewar, 2004), just as they do in traditional mathematical research. However, those doing SoTL can find helpful collaborators outside mathematics. Six of the fifteen chapters in Part II describe SoTL projects that involved collaborative work. In Chapter 7 Lynn Gieger and her colleagues at Oglethorpe College attest to the benefits of having participants from multiple disciplines (mathematics, psychology, and biology). They found having multiple perspectives helpful in narrowing the focus of their study, and collectively they had a breadth of experience in qualitative and quantitative research design and analysis. The collaboration between mathematics instructors and academic support staff at Duke University reported in Chapter 6 led to the submission of a grant proposal to the National Science Foundation. Mike Axtel and William Turner (Chapter 14) report on a project that began when both were at the same institution but was completed when they were at different institutions. Chapter 17 describes how mathematics educators and mathematicians at Illinois State University joined forces to study a curricular change in their mathematics teacher preparation program. Chapters 5 and 18 give other examples of collaborative work. Rann Bar-On and his colleagues (Chapter 6) advise: "Key to the success of this approach is finding a group of collaborators who share a common problem and common goals" (p. 60). One contributor to this volume, Rikki Wagstrom (Chapter 19), remarks that she wished she had found a collaborator to share the workload of her project.

As the projects described in Part II of this volume show, collaborators can be found on campus in the mathematics department, in other departments (especially psychology, sociology, and education), or in academic support units. The latter include academic resource centers, teaching or faculty development centers, institutional assessment units, and instructional technology departments. Off-campus collaborators can be found at conference sessions or workshops focused on mathematics teaching, mathematics education, scholarship of teaching and learning, liberal education, civic engagement, assessment, or teacher preparation.

Professional organizations that promote and support SoTL include the International Alliance of Teacher Scholars ([www.iats.com](http://www.iats.com)), the International Society for the Exploration of Teaching and Learning ([www.isetl.org](http://www.isetl.org)), and the International Society for the Scholarship of Teaching and Learning ([www.issotl.org](http://www.issotl.org)). Each holds annual conferences where colleagues from many disciplines gather to discuss and present scholarly work related to teaching and learning. The multi-disciplinary gatherings afford good opportunities for meeting potential collaborators. A link to a list of SoTL conferences appears in the next section.

Faculty members interested in SoTL work often ask about funding sources to support it. Many campuses provide small amounts of funding for course or curricular development, experimenting with innovative teaching methods,

or incorporating technology into courses. These may not be labeled as SoTL grants, but including SoTL as part of the assessment plan should make an application more competitive. External funding for larger projects involving mathematics teaching and learning maybe be available through the American Educational Research Association ([www.aera.net](http://www.aera.net)), the Association for Institutional Research ([www.airweb.org](http://www.airweb.org)), the Department of Education ([www.ed.gov](http://www.ed.gov)), the Environmental Education grants from the Environmental Protection Agency ([www.epa.gov/education/grants/index.html](http://www.epa.gov/education/grants/index.html)), Learn and Serve America, a program of the Corporation for National and Community Service ([www.learnandserve.gov](http://www.learnandserve.gov)), and various funding opportunities administered by the Mathematical Association of America ([www.maa.org](http://www.maa.org)) and the National Science Foundation ([www.nsf.gov](http://www.nsf.gov)). Private foundations that offer funding to improve teaching and learning in mathematics or closely related fields include the Ford Foundation, the Henry Luce Foundation, and the Spencer Foundation. Many campuses have an office to assist faculty members in seeking external funding that can help identify additional external funding sources.

## Going Public with Results

The view of SoTL we present calls for SoTL researchers to submit their findings to peer review and make them public for others in the academy to build upon. Across the academy going public typically involves making conference presentations, possibly with publication in the conference proceedings, or publishing articles in scholarly journals, chapters in edited volumes, or entire books.

### Conferences

Fortunately, many opportunities exist to present the results of SoTL studies in contributed paper sessions at the annual Joint Mathematics Meetings (JMM). In each year since 2007 there also has been a session specifically on the Scholarship of Teaching and Learning in Mathematics at the JMM. Twice (in 2008 and 2009) these have been special sessions jointly sponsored by the American Mathematical Society and the Mathematical Association of America. In 2010 there was an MAA invited paper session co-sponsored by the AMS, and in other years there was an MAA contributed paper session. SoTL work can also appear in RUME sessions at the JMM and at the annual RUME conference.

There are also a number of conferences that focus on scholarship of teaching and learning across all disciplines. The Center for Excellence in Teaching and Learning (CETL) at Kennesaw State University maintains a list of these venues at [cetl.kennesaw.edu/teaching-conferences-directory](http://cetl.kennesaw.edu/teaching-conferences-directory). The list is organized by discipline (including mathematics) and by topic (including scholarship of teaching and learning).

### Journals

Most mathematicians undertaking a SoTL project are aware that the journal *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies* provides a peer-reviewed outlet for SoTL work in mathematics. As Dewar and Bennett (2010) noted, other journals such as the *College Mathematics Journal* and *Mathematics Magazine* typically publish articles focused on mathematical content for teaching rather than on teaching practice. In the *American Mathematical Monthly* the effect of a 1992 editorial change removing the “Teaching of Mathematics” section was to end publication of virtually all articles discussing pedagogy, even though the editor had written:

... articles on the Teaching of Mathematics are not separated from articles on mathematics ... The mathematical principle motivating these changes is the belief that mathematics ought to be viewed as a unified field, both horizontally and vertically. Articles on the mathematics of computers belong next to articles on Riemann surfaces; comments on teaching Calculus ought to be read with as much enthusiasm as comments on the representations of Lie groups ... (Ewing, 1992, p. 2).

Kennesaw State’s CETL provides another useful resource at [cetl.kennesaw.edu/teaching-journals-directory](http://cetl.kennesaw.edu/teaching-journals-directory), one that identifies journals that publish SoTL. It lists disciplinary specific journals (including mathematics) and general or interdisciplinary journals on teaching, most with a link to the journal’s website. SIGMAA-RUME (the Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics) maintains a list of journals that publish papers on mathematics education at [sigmaa.maa.org/rume/journals.html](http://sigmaa.maa.org/rume/journals.html), last updated



on November 29, 2011. The RUME webpage contains a link to another rich resource, a wikispace of journals that publish mathematics education ([mathedjournals.wikispaces.com](http://mathedjournals.wikispaces.com)). Many of these journals will be unfamiliar to teaching mathematicians, so choosing an appropriate journal can be a challenge.

The fifteen projects presented in Part II have all been made public in some form. Three have appeared in print in journals or conference proceedings. Several more are being prepared for journal articles. In Chapter 15, Stephen Szydlik discusses selecting an appropriate publication venue and gives the reasons why he ruled out certain journals. In addition to Szydlik's comments the following guidelines may be useful when choosing a journal. First, identify the target audience. Is it college mathematics professors, faculty members who prepare future teachers, or researchers in mathematics education? Would faculty members in other disciplines be interested in the study? This may help decide between a mathematically focused journal and a more general journal that publishes SoTL. Is it important that the article counts for a tenure or promotion decision? Will a publication in a SoTL journal count? What journals were included in the literature review for the study? They are potential publication venues. Does the journal publish research or practice or both? Journals with a research focus may have a higher standard for evidence and require the work to follow from a theoretical framework. Visit the websites of potential journals and read the instructions or guidelines for authors. Will the article meet the length limitations? Other factors to consider (or not) are acceptance rates, times to publication, and the reputation of the journal. When choosing the journal that seems to be the best fit, also identify second- and third-choice journals. Should the paper be rejected by the first-choice journal, it can be resubmitted to the second-choice journal. Adopt the style of writing that appears in the selected journal and follow instructions carefully when preparing the manuscript. Before submitting it, recruit some local reviewers and make revisions based on their comments.

Writing and publishing an article is more difficult than making a conference presentation. It is also more satisfying and will provide a permanent record of the work, one that can be a foundation for future improvements in teaching and learning by others.

## References

- Bennett, C., & Dewar, J. (2004, December). Carnegie Chronicle: What you really ought to know about collaboration on a SoTL project. *National Teaching and Learning Forum*, 14(1), 4–6. Retrieved from [www.ntlf.com/issues/v14n1/v14n1.pdf](http://www.ntlf.com/issues/v14n1/v14n1.pdf)
- Dewar, J., & Bennett, C. (2010). Situating SoTL within the disciplines: Mathematics in the United States as a case study. *International Journal for the Scholarship of Teaching and Learning*, 4(1). Retrieved from [digitalcommons.georgiasouthern.edu/ij-sotl/vol4/iss1/14](http://digitalcommons.georgiasouthern.edu/ij-sotl/vol4/iss1/14)
- Ewing, J. (1992). Comments. *American Mathematical Monthly*, 99(1), 2.
- U.S. Department of Health and Human Services. (n.d.). OHRP (Office for Human Research Protections) homepage. Retrieved from [www.hhs.gov/ohrp](http://www.hhs.gov/ohrp)



# **Illustrations of SoTL Work in Mathematics**

**Theme 1: Experiments with  
Approaches to Teaching**



# 5

## Assessing the Effectiveness of Classroom Visual Cues

Gretchen Rimmasch and Jim Brandt  
*Southern Utah University*

### Editors' Commentary

This chapter illuminates some of the research design issues discussed in Chapter 3. It shows how the authors developed and piloted a novel intervention, visual cues, during one semester and fully implemented and assessed it in another. The methodology involved two sections of the same course, taught by the same instructor, one as an experimental group and the other as control group. Because using visual cues to assist with some computational skills was specific and limited in scope, many of the concerns mentioned about using control groups in Chapter 3 did not arise. The study involved similar interventions with visual cues in two different settings (first in remedial math, then in calculus). For calculus, the authors discovered that they had to alter their rubric to capture the information they wanted.

### Introduction

The large number of students who must take remedial mathematics courses is a problem for many colleges and universities. The courses must be staffed, and students who place into remedial courses are less likely to successfully complete a degree (Attewell, Lavin, Domina, & Levey, 2006). Remediation has been a problem for decades (see, for instance, Handel & Williams, 2011; Patterson & Sallee, 1986; Weiss & Nguyen, 1998). In many lower division mathematics courses the focus of instruction is often on procedural skills. Students should view these courses as an opportunity to improve their skills, but they often see them as hurdles that do not count towards degree completion. The material in remedial courses is often familiar to students, which can make it more difficult for students to expend the time and effort required to correct previous misconceptions. Conversations about our frustrations with these courses led us to consider what we could do to be more effective in teaching procedural skills. The resulting investigation took the form of a *What works?* study.

We began with an intermediate algebra course, as it involves procedural skills that students have encountered previously. In investigating a *What works?* question, there are many possible approaches (Patterson & Sallee, 1986). One would be to move much of the course online and have students work at their own pace in a computer lab. Another would be to change the grading scheme to a mastery approach with a sequence of gateway exams. Or, the course could be made more interactive with frequent group work and student presentations. In considering our options, we chose a less ambitious approach that focused on trying to be more effective in our lectures.

We decided to focus on lectures related to a single topic in a lower division mathematics course. This was based on advice received at a summer faculty workshop on the integration of technology into teaching and learning. The workshop organizers urged participants not to revise their courses completely, as faculty who attempted overall course revisions often felt overwhelmed or disappointed in the results. They suggested that participants identify one topic and think creatively about how to teach it more effectively. This gives instructors the opportunity to experiment with a portion of a course, to identify approaches that work for them, and, based on those experiences, to revise other portions of the course.

The topic we decided to experiment with was the teaching and learning of exponent rules in an intermediate algebra course. Common student errors in simplifying exponential expressions include applying the wrong exponent rule and performing an exponent rule incorrectly. Rather than simply emphasizing the rules verbally and in writing, we wanted to give students a tool to help them to identify and correctly apply the exponent rules. Since many students indicate that they are visual learners (Felder & Brent, 2005), we began thinking about visual cues that would help students to recognize and correctly apply exponent rules. Our hope was that the visual cues would serve as a visual mnemonic aid.

In searching for previous mathematics education research, we found that Kirshner (1989) had observed that the visual appearance of algebraic expressions played an important role in students' manipulation of them. We were encouraged by accounts of classroom success when using visual tools. For instance, Uygur & Özdaş (2007) reported success with arrow diagrams in teaching the chain rule. But, we did not find any research regarding teaching exponent rules. However, when searching online databases such as Google Scholar, JSTOR, and ERIC for more general mathematics education research regarding visual cues and classroom teaching tools, we found an article by Rasmussen & Marrongelle (2006) describing pedagogical content tools. They defined a pedagogical content tool as “a device, such as a graph, diagram, equation, or verbal statement, that a teacher intentionally uses to connect to student thinking while moving the mathematical agenda forward” (p. 309). Our agenda was to increase computational skill, and we chose visual cues to connect to student thinking. Finding the definition of a pedagogical content tool helped to situate our project in the broader context of mathematics education research. This proved useful when preparing our results for presentation and publication.

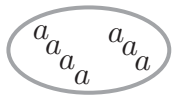
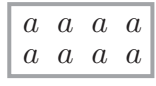
## Exponent Rules Project

### Development of Cues

Current views of learning hold that students construct new knowledge by building on what they already know (Bransford, Brown, & Cocking, 2000). Therefore, we attempted to create visual cues that linked the content with previous mathematical knowledge. We began thinking about cues that might suggest addition of exponents, multiplication of exponents, etc. One of us was teaching a mathematics for elementary education course that discusses various models for arithmetical operations. A model for addition of whole numbers, for instance, involves combining the elements of two disjoint sets into a larger set. Similarly, a model for multiplication of whole numbers involves the area of a rectangle. By using these models as the basis for our visual cues, we hoped to link exponent rules to students' previous mathematical knowledge. The visual cues we eventually settled upon are shown in Figure 5.1 along with the linked idea or example and the associated language. For instance, in the case of a power to a power, we can think of the total number of instances of the base in a rectangular array. Then the appropriate power comes naturally from taking the product of the two powers. To make this connection explicit, we drew a rectangle around the exponents when taking a power to a power. This cue suggested multiplying the length times the width. To reinforce the connection, the associated language “ $m$  by  $n$ ” was used by the instructor.

### Classroom Implementation

To get an idea of the impact of the visual cues on teaching and learning exponent rules, we began with a pilot study where we used the visual cues in one section of intermediate algebra. Our goals in the pilot study were to identify any difficulties in using the visual cues during lecture and to get a sense of student and instructor reactions to them. We found the visual cues to be fairly easy to use during lecture. In many cases, using the cues first required rearranging the exponential expression. For instance, we would rewrite  $a^4b^2a^3$  as  $a^4a^3b^2$  so that the terms involving  $a$  were adjacent

Rule	Visual Cue	Linked Idea or Example	Associated Language
Multiplication with same base	$a^m a^n = a^{m+n}$		"together"
Power to a power	$(a^m)^n = a^{m \cdot n}$		"by"
Product or quotient to a power	$(a \cdot b)^n = a^n \cdot b^n$ $\left(\frac{a}{b}\right)^n = \frac{a^n}{b^n}$	$5(x+2) = 5x + 10$	"distribute"
Negative exponent	$a^{-n} = \frac{1}{a^n}$	$x + 3 = 5$	"move to the other side"
Division with the same base	$\frac{a^m}{a^n} = a^{m-n}$	$\frac{a a a a a}{a a}$	"subtract" or "cancel"

**Figure 5.1.** Exponent rule cues

before circling the exponents 4 and 3 to get  $a^7 b^2$ . This is perhaps an unnecessary step, but it emphasizes the distinction between the different bases in the expression. We were concerned that some students might not like the cues and even wonder why we introduced them. However, no one objected and students seemed to find them a natural part of solving the problem. Both of us felt that our lectures were more effective with the cues. Taking a moment to draw a rectangle around the exponents emphasized the part of the expression we were simplifying and the appropriate rule to apply. This experiment with visual cues encouraged us that they could be effective at improving student performance, and we began planning a more systematic assessment of their impact.

## Assessment

In assessing the impact of the consistent use of visual cues by the instructor, we had two questions. First, did the visual cues improve student performance in simplifying exponential expressions? Second, did they help students to be more confident in working with exponential expressions? After receiving approval from our Institutional Review Board, we administered written assessments to two sections of intermediate algebra. One section was taught using lectures without any visual cues (the control group) and the other section was taught using lecture with consistent use of visual cues (the treatment group). To reduce the differences between sections, the same instructor taught both using the same examples and homework exercises. Because participants had seen exponent rules in previous courses, we administered our assessment before discussing them and again approximately one week after the exam involving exponent rules. These pre- and post-assessments allowed us to collect data on the gain in students' performance and to control for differences between students in the two sections.

The pre- and post-assessment instruments used the same questions for comparison. Each began with several questions involving student confidence and ended with nine exponential expressions to simplify. The questions involving confidence used a five-point Likert scale, where students were asked to rate the degree to which they agreed or disagreed with statements such as "I feel confident simplifying exponential expressions such as  $a^4(ab^{-3})^2$ " or "I find exponents and exponent rules confusing." The simplification questions involved exponential expressions such as  $3x^{-2}x^7$  and  $(m^2n^3)^4$ . All the expressions were similar to those discussed in class and those assigned in the homework.

After collecting the pre- and post-assessments, we entered the data for each student into a spreadsheet. For each simplification question, we recorded whether a participant's answer was correct or incorrect and whether any cue was used to aid in simplifying. To measure learning gains, we computed the difference between post- and pre-assessment scores. To analyze the data, we began with some descriptive statistics regarding student gains. There was little difference between the control group and the treatment group regarding confidence in working with exponent rules. However, in terms of simplifying exponential expressions, the group with visual cues seemed to be more successful. To verify this, we performed a  $t$ -test on the overall gain on all simplification questions comparing the two independent samples (the control group with 34 participants and the treatment group with 37 participants). There was a positive and statistically significant impact on student performance ( $p = .034$ ). Students in the treatment group were more likely to use a type of cue to help in simplifying the expressions ( $p = .026$ ). We presented these findings at the joint meetings (Brandt & Rimmasch, 2011, January) and published them in the *Utah Mathematics Teacher* journal (Brandt & Rimmasch, 2011). These findings led us to believe that instructor use of visual cues could be helpful in teaching procedural skills in other courses.

## Derivative Rules Project

Following our positive experiences with exponent rules, we wondered what other topics could be aided by using visual cues. Would the cues be helpful in classes where the content was less familiar? Would they be useful in classes where the computational procedures were much more involved? Would the cues be useful in classes where the content was more theoretical? We decided to look at how visual cues could be used in a calculus class. This is a class that is computational, but less familiar to the students. For many introductory calculus students, the concepts of limits, derivatives, and integrals are entirely new. These concepts still involve procedural skills, and many students have difficulty mastering those skills. Thus, we thought it would be an ideal place to experiment with visual cues. We decided to consider only one calculus topic and see if the cues could improve student performance. Rather than overloading students with cues for individual functions, we chose to focus on the three rules that involve combinations of functions: the product rule, the quotient rule, and the chain rule.

### Development of Cues

Thinking about the problems our students had with the derivative rules, we realized there were two major issues. The students had to recognize which rule to apply, and they needed to remember and correctly apply it. We wanted to create cues that would help them do both. Initially we hoped to develop cues that linked to previous mathematical knowledge, as we had for the exponent rules. Because the derivative rules are more complex, we settled on a tool that emphasized both the underlying rule and the process for computing the derivative. As shown in the sequence of cues in Figure 5.2, we first drew a polygon around the product or quotient to emphasize the overall form of the function. Then we added the arrows and  $d$ s clockwise around the outside of the polygon, to indicate the steps in the product rule. We used the language of “leave” one function and “differentiate” the other function. We chose a quadrilateral for the product rule

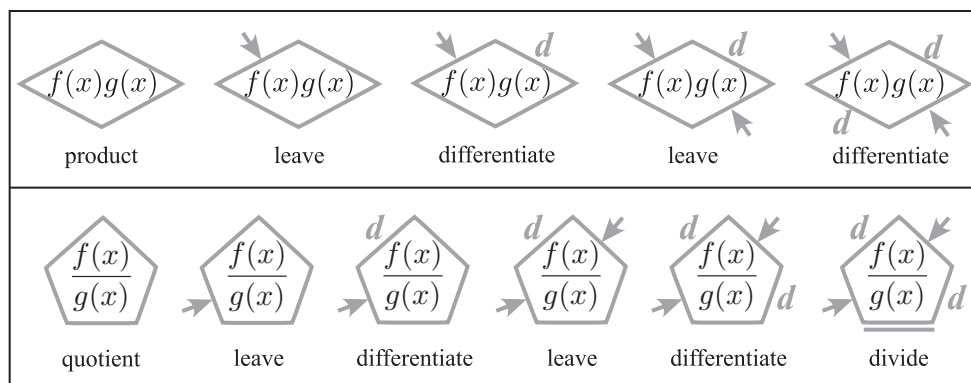
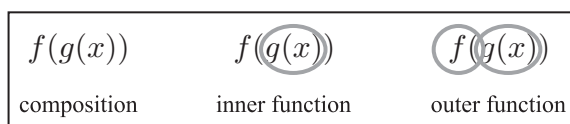


Figure 5.2. Product and quotient rule cues



**Figure 5.3.** Chain rule cue

because there are four steps involved and a pentagon for the quotient rule to reflect its five steps. Although the visual cues were not directly related to multiplication or division, we hoped they would help students recognize the form of the function, remember the derivative rules, and apply them correctly.

As the chain rule is not similar to the product and quotient rules, we wanted a cue that was different. Composition is a multi-step process in the same way that a chain has multiple links. In making this analogy we drew a link around each function (step) in the composition, forming a chain of interconnected links. This emphasized that the different parts of the function were connected to each other. We then described the process as “chaining from the inside out” followed by “differentiating from the outside in.” So, as shown in Figure 5.3, we would begin by circling the innermost function, emphasizing the initial step in the composition. We then drew a circle around the next function, but passing through the original circle to indicate the links of the chain. For example, for  $f(x) = \sin(x^2)$  we would first circle the  $x^2$  and then the sine, with the two circles linked to each other. When differentiating, we would begin with the last link and move inwards. In the example, we would first differentiate the sine, yielding  $\cos(x^2)$  (the link with the  $x^2$  was not involved at that point, so the  $x^2$  stayed). Then we moved to the next link, the  $x^2$ , and differentiated it to get  $2x$ , resulting in  $\cos(x^2)2x$ .

## Classroom Implementation

As with the exponent rules, we did an initial pilot study with one section of calculus to see if students would react favorably to the cues and if the professor felt comfortable using them. In class, we derived the rules in the traditional way, and then simply told the students that we were going to use visual cues to help them remember the rules. We found the cues to be a simple addition to classroom derivative computations. For example, when faced with the derivative of a product, we would enclose the product in a quadrilateral, verbally indicate that we needed to use the product rule, and proceeded to “leave” and “differentiate” the component functions following the pattern in Figure 5.2. While algebra students rarely used the cues after the section on exponents, we found that many calculus students used the derivative cues consistently throughout the rest of the semester. As instructors we felt that the cues were especially helpful when discussing two types of problems. First, the links in the chain helped emphasize each individual function involved in the composition. This was especially beneficial in chain rules with more than two functions, as the number of links indicated the number of times we needed to differentiate. Second, we found the cues to be helpful in problems that required multiple derivative rules. For example, if there was a product in the numerator of a quotient, the quadrilateral for the product was naturally smaller and inside of the pentagon for the quotient. This made it clear that the product rule happened inside of the quotient rule. Following our positive experience with the pilot study, we initiated a more systematic assessment of the effectiveness of the derivative cues.

## Assessment

In assessing the derivative cues, we had the same two questions as with the exponent rules:

- Did the use of visual cues improve student performance in calculating derivatives?
- Did they increase student confidence in their ability to differentiate?

We again received approval from our Institutional Review Board, and administered written assessments to two first semester calculus classes. One section was taught using the cues (a treatment group with 26 participants) and the other section was taught without the cues (a control group with 45 participants). The same instructor taught both sections, using the same classroom examples. Creating assessments instruments to compare student improvement in computing derivatives was a more challenging task than it had been for the exponent rules since many of the students had not computed derivatives previously. We considered using an existing calculus readiness exam as a pre-assessment, but decided that we did not need all the topics it covered. To learn the product, quotient, and chain rules, a prerequisite



skill is the ability to recognize whether a function is a product, quotient, composition, or none of these. Therefore, we developed a pre-assessment that asked students to classify functions as a product, quotient, composition, or none of these. This allowed us to determine if the control and treatment groups were on about the same level when we started discussing the derivative rules. For the post-assessment, we asked students how confident they felt about taking derivatives, knowing which rule to use, remembering the rules, and correctly applying them. Following the questions involving confidence, we asked them to compute the derivatives of six functions and to use implicit differentiation on an equation relating  $x$  and  $y$  to find  $\frac{dy}{dx}$ .

To assess the impact of the cues on student performance, we initially compared the post-assessment scores of the two groups based on whether their derivatives were correct or incorrect. On each question, the treatment group performed at the same or a higher level, but the difference was not statistically significant. However, the treatment group performed worse on the pre-assessment, suggesting that they had improved more than the control group. Also, both groups made many mistakes when computing derivatives, not all of which were related to the product, quotient, or chain rules. For example, one of the functions to differentiate on the post-test was  $f(x) = (3x^2 + 7x) \sin x$ . Many students answered with the correct answer of  $f'(x) = (6x + 7)(\sin x) + (\cos x)(3x^2 + 7x)$ . However, some students seemed to know the product rule, but had difficulty with the derivatives of the functions in the product, giving answers like  $f'(x) = (6x + 7)(\sin x) + (-\cos x)(3x^2 + 7x)$  and  $f'(x) = (6x + 7x)(\sin x) + (\cos x)(3x^2 + 7x)$ . Other students seemed to know the derivative of each function in the product, but had difficulty with the product rule, giving answers like  $f'(x) = (6x + 7)(\cos x)$  and  $f'(x) = (6x + 7)(\sin x) - (\cos x)(3x^2 + 7x)$ .

Since the cues were for performing the product, quotient, and chain rules, we reassessed students' responses using a rubric that focused on a student's recognition and use of the appropriate rule. Each student response was categorized according to three factors: recognizing that a particular derivative rule was required, knowing the required derivative rule, and correctly computing the derivatives of the component functions. For example, for  $f(x) = (3x^2 + 7x) \sin x$ , while  $f'(x) = (6x + 7)(\sin x) + (-\cos x)(3x^2 + 7x)$  is not the correct derivative, the student clearly recognizes that  $f(x)$  is a product and knows the product rule. On the other hand, we were not sure how to assess a response like  $f'(x) = (6x + 7)(\sin x) - (\cos x)(3x^2 + 7x)$ . The student clearly recognized the need for a rule, but it is unclear which rule was intended. Did the student mean to use the product rule and get it wrong, or the quotient rule and get it wrong? Since we couldn't decide what the student was thinking, we gave no credit for this response in terms of recognizing the required rule.

The same four functions appeared on both the pre- and post-assessments. On the pre-assessment, students were asked to classify the functions as a product, quotient, composition, or none of these. As part of the post-assessment, students were asked to compute the derivatives of the same functions, requiring use of either the product rule, quotient rule, chain rule, or none of these. Therefore, we were able to measure their improvement in recognizing the form of various functions. Comparing the gain in the student recognition between the control group and the treatment group, we found that the treatment group, with the visual cues, outperformed the control group. Using a  $t$ -test with two independent samples, the difference between the groups approached statistical significance ( $p = .072$ ).

In our study of the derivative cues, revising the rubric used for assessing student responses significantly affected our results and conclusions. From this we learned the importance of clearly identifying the key focus of the experiment. Our focus was not on success in computing derivatives, but rather on the ability to recognize and apply the product rule, quotient rule, and chain rule. Thus our assessment needed to isolate the impact the cues had on those rules alone. Our initial classification of responses as correct or incorrect did not accomplish this. Fortunately, we were able to modify our rubric because we had used identical functions on the pre- and the post-assessments. Our willingness to consider a subset of the questions and evaluate them in a different manner allowed us to detect if teaching with visual cues affected student recognition of the form of a function. Our experience demonstrates that careful planning in developing the assessment tools and flexibility while investigating the data are essential when analyzing the impact of an experiment.

## Discussion

Rasmussen & Marrongelle (2006) argue that a mathematics “teacher has the obligation of enculturating the students into the discourse and conventional representational forms of the broader community” (p. 395). Computational fluency is part of the culture of mathematics, so teachers need to help students develop their procedural skills. Our research

indicates that visual cues can play an effective role in teaching and learning them. We are not arguing that we should use cues in teaching every mathematical topic; adding a second layer to existing mathematical notation is not the goal. However, by drawing attention to portions of an expression, visual cues can aid classroom discussion and help students attain greater computational fluency. We enjoyed creating the visual cues and implementing them in the classroom. We encourage teachers to identify student difficulties, and to think creatively about pedagogical content tools that might help improve student understanding.

## Future Research

For future research we could continue to investigate our cues. For example, we could see if visual cues have any long term impact on student performance. Do students who learn exponent rules using visual cues in intermediate algebra continue their success at simplifying exponential expressions in future courses? We could also compare the effectiveness of our visual cues with other mnemonic aids. Is it possible that different mnemonic aids work better? Is there a best cue or mnemonic for a task, and why? Another direction would be to develop and investigate visual cues for different mathematical skills. For example, cues might be helpful in teaching logarithm rules in college algebra. We could research the use of visual cues in more theoretical upper division math courses. In an introduction to proof course, for instance, students often have difficulties writing proofs. If an instructor consistently used a visual cue reminding students of the process involved in proof by mathematical induction, would it improve the structure of students' induction proofs?

## Professional Impact

This project has affected our professional work in several ways. First, our institution uses the Boyer model of scholarship in assessing faculty scholarly activity. The scholarship of teaching and learning is an explicit category of scholarship in it, so the project and its related presentations and publications counted as scholarly activity in our annual reviews. Second, the project has led to additional contact and interaction with colleagues at our institution and at other institutions. After seeing our results, several faculty colleagues now use visual cues in teaching exponent rules. Several of our students who are prospective secondary mathematics school teachers use them in tutoring undergraduate students and plan on using them when teaching secondary school. Their experience with visual cues prompted additional conversations about strategies for teaching other mathematics topics. The initial project motivated us to work together on other investigations involving undergraduate mathematics education. Although our graduate degrees are in mathematics, we have become more interested in mathematics education and now have several publications in undergraduate mathematics education. While mathematics education research is often theoretical, our research projects started with practical questions focused on ways to improve mathematics teaching and learning. As a result, our research focus has allowed us to continue to grow as mathematics teachers while engaging in scholarly activity.

## References

- Attewell, A., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *The Journal of Higher Education, 75*(5), 886–924.
- Brandt, J. & Rimmasch, G. (2011, January). *Using visual cues in teaching computational skills*. Paper presented at the MAA Session on the Scholarship of Teaching and Learning in Collegiate Mathematics, I, at the Joint Mathematics Meetings, New Orleans, LA. Abstract retrieved from [jointmathematicsm meetings.org/meetings/national/jmm/1067-v1-1516.pdf](http://jointmathematicsm meetings.org/meetings/national/jmm/1067-v1-1516.pdf)
- Brandt, J., & Rimmasch, G. (2011). Cue up procedural skills. *Utah Mathematics Teacher, 4*, 8–12.
- Bransford, J., Brown, A., & Cocking R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: The National Academy Press.
- Felder, R., & Brent, R. (2005). Understanding student differences. *Journal of Engineering Education, 94*(1), 57–72.

- Handel, S., & Williams, R. (2011). Reimagining remediation. *Change*, 43(2), 28–33. doi:10.1080/00091383.2011.556921
- Kirshner, D. (1989). The visual syntax of algebra. *Journal for Research in Mathematics Education*, 20(3), 274–287.
- Patterson, D., & Sallee, T. (1986). Successful remedial mathematics programs: Why they work. *The American Mathematical Monthly*, 93(9), 724–727.
- Rasmussen, C., & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics in instruction. *Journal for Research in Mathematics Education*, 37(5), 388–420.
- Weiss, K., & Nguyen, T. (1998, March 27). Almost half of Cal state freshmen lack skills. *The Los Angeles Times*. Retrieved from [articles.latimes.com/1998/mar/27/news/mn-34354](http://articles.latimes.com/1998/mar/27/news/mn-34354)
- Uygur, T., & Özdaş, A. (2007). The effect of arrow diagrams on achievement in applying the chain rule. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 17(2), 131–147.

# 6

## **(Re)discovering SoTL Through a Fundamental Challenge: Helping Students Transition to College Calculus**

Rann Bar-On, Jack Bookman, Benjamin Cooke, Donna Hall, and Sarah Schott  
*Duke University*

### **Editors' Commentary**

In this chapter Rann Bar-On, Jack Bookman, Benjamin Cooke, Donna Hall, and Sarah Schott describe how one faculty member's attempt to improve student success in a special freshmen calculus sequence for underprepared students evolved into scholarship of teaching and learning. Key to this progression was collaboration with academic support professionals and non-tenure track faculty. Thoughtful discussions, a few trial interventions, and examining the research literature enabled the group to move from reflection and experimentation to scholarly teaching and then to the scholarship of teaching and learning. After several years of collaborative effort, a grant application to further develop, study, and share the results of this work was submitted to the NSF.

### **Introduction**

Through the efforts of the Carnegie Foundation and others, many faculty are introduced to SoTL through conferences, workshops at their institution, colleagues, or journal articles devoted to SoTL (Hutchings, 2010). As discussed in this volume, these activities have struck a chord with, and given voice to, the scholarly and intellectual interests of many faculty in higher education. The growth of SoTL has provided validation and motivation for faculty to develop SoTL projects. In this chapter, we describe a different introduction to SoTL, one that is more unintentional and less self-conscious. We will discuss how a small group of faculty faced with an instructional problem gradually adopted an increasingly scholarly approach to addressing it. Instead of scholars creating solutions to problems, in this case a problem created scholars.

We were all exposed to SoTL work early in our careers. We mostly set it aside as we pursued our teaching. We rediscovered it when the need arose, and we wanted to stand on the shoulders of those who had attempted to address the problems we encountered. We believe that our experience can be instructive for those who are new to SoTL and for those who have had some involvement with it. Thus we came to our title: “(Re)discovering.”

Our involvement with SoTL was motivated by the goal of improving our teaching. One faculty member perceived a problem and attempted to fix it. But because of heavy teaching loads and little interest by other faculty members, his

was an isolated effort that likely had minimal impact on student success. Subsequently, a collaboration with academic support professionals was initiated, one that fostered thoughtful discussions and experimentation. With the hiring of a new faculty member interested in systemically improving teaching and student success in our calculus courses, the collaboration with the Academic Resource Center (ARC) became more thoughtful. We held regular meetings where we discussed the possibility of applying research literature to the problem. This was a critical turning point as our teaching evolved from good, to reflective, to scholarly. As data were collected and analyzed, the discussions included more people. As further questions were raised and solutions proposed, our work began to look more like scholarship of teaching and learning. It now included reviewing the literature, examining data, developing research questions, and creating research proposals. As a result of the collaboration, the group has submitted a grant application to NSF to further develop, study, and share the results of our work.

Our story was more complicated and less smooth than just described. In this chapter we explain how our work developed and present a model of how faculty whose primary professional responsibilities involve teaching and who have heavy teaching loads can develop a scholarship program for teaching and learning. Key to the success of this approach is finding a group of collaborators who share a common problem and common goals.

## Background—A Stubborn Pedagogical Problem

Before telling how our scholarship of teaching and learning developed, we describe the problem and its context. At Duke and elsewhere, one of the chief obstacles for student retention in the Science, Technology, Engineering, and Mathematics (STEM) disciplines is the difficulty of transitioning from high school to college math and science classes. This is especially true for first-generation and mathematically underprepared students. David Bressoud, past President of the Mathematical Association of America (2009–2010) and former member of the AP Calculus Development Committee (1999–2005), has studied and written extensively about the transition from high school to college mathematics. He wrote in the *Chronicle of Higher Education* (Bressoud, 2010):

There is some indication that the pressure to offer calculus in high school is pushing underprepared students prematurely into the subject. The National Education Longitudinal Study of the high-school class of 1992 reported that of every 10 students who studied calculus in high school, three had to take precalculus when they got to college. Three others did not continue with any calculus in college. No more recent data are available, but the phenomenal growth of students taking calculus in high school suggests that the problem has not abated.

Bressoud suggested that if we are to meet the challenges created by the explosive growth of high-school calculus, it is imperative that we re-examine first-year college mathematics. His observations are consistent with the views of many mathematics faculty members, especially those at selective colleges and universities where students enter college with aspirations to become medical doctors, engineers, or scientists and become discouraged because of their experiences in first-year mathematics courses.

Students entering colleges and universities who are underprepared are less likely to graduate with degrees in the STEM disciplines (Arcidiacono, Acejo, Fang, & Spenner, 2011). At Duke, entering students who want to major in the STEM disciplines are likely to start their math sequence in Math 25L Laboratory Calculus with Functions I. The course is recommended for students with a math SAT score below 680. Within this course, some of the students have a level of ability and preparation significantly below the average 25L student.

Math 25L and 26L, Laboratory Calculus with Functions I and II, were introduced at Duke in 1995. Prior to this, Duke offered a precalculus class for students who were not prepared for calculus. Although students found this class straightforward, it did not adequately prepare them for calculus. Math 25L and 26L were instituted in order to bridge this gap: to allow weakly prepared students to take a class that reinforces algebra and function skills while covering Calculus I over two semesters. The two courses interweave a review of precalculus material with calculus. The approach is to introduce a class of functions (e.g., linear, power, polynomial, rational, exponential, logarithmic, or trigonometric) followed by a study of its properties, at first descriptively, and later using calculus tools. More advanced applications, such as related rates, optimization, and differential equations, are covered toward the end of each semester, though preparatory work for them is undertaken earlier.

Week	4	5	6	7	8	9	10	11	12	13	14	15
Attendance	16	10	7	8	5	2	4	6	3	3	0	1

**Table 6.1.** Combined weekly attendance figures for the two hour-long algebra review sessions for Math 25L in fall 2008

## Story

Our support of first-year students in Math 25L was started as a collaboration between mathematics faculty, who aimed to improve mathematical thinking and competence in undergraduates, and the ARC, whose mission is to support students during the transition from high school to college.

The theory of self-regulated learning (Bandura, 1997; Dembo & Seli, 2008; Zimmerman & Schunk, 2001) underpins the work of the ARC. The center focuses on helping students become independent learners through one-on-one meetings, programs run in collaboration with faculty, and other types of academic support. Beginning students often need to improve metacognitive skills, such as planning, goal setting, using appropriate resources and knowledge, budgeting time, self-monitoring, and self-assessing, in order to become independent learners (Schraw, Crippen, & Hartley, 2006), especially when faced with the higher volume of course material in college.

### Fall 2008—Hastily Organized Algebra Support

In the fall 2008 semester, Jack Bookman took over the leadership of Math 25L. During the second week of class, he introduced an algebra quiz based on the algebra review from the course lab manual. A large portion of the class failed. As a result, Bookman and Benjamin Cooke, a former graduate student in the math department who became an instructor in the ARC, organized algebra review sessions to be run in conjunction with the course. The hour-long review sessions run by Cooke were offered twice a week as a voluntary option for students who had failed the algebra quiz.

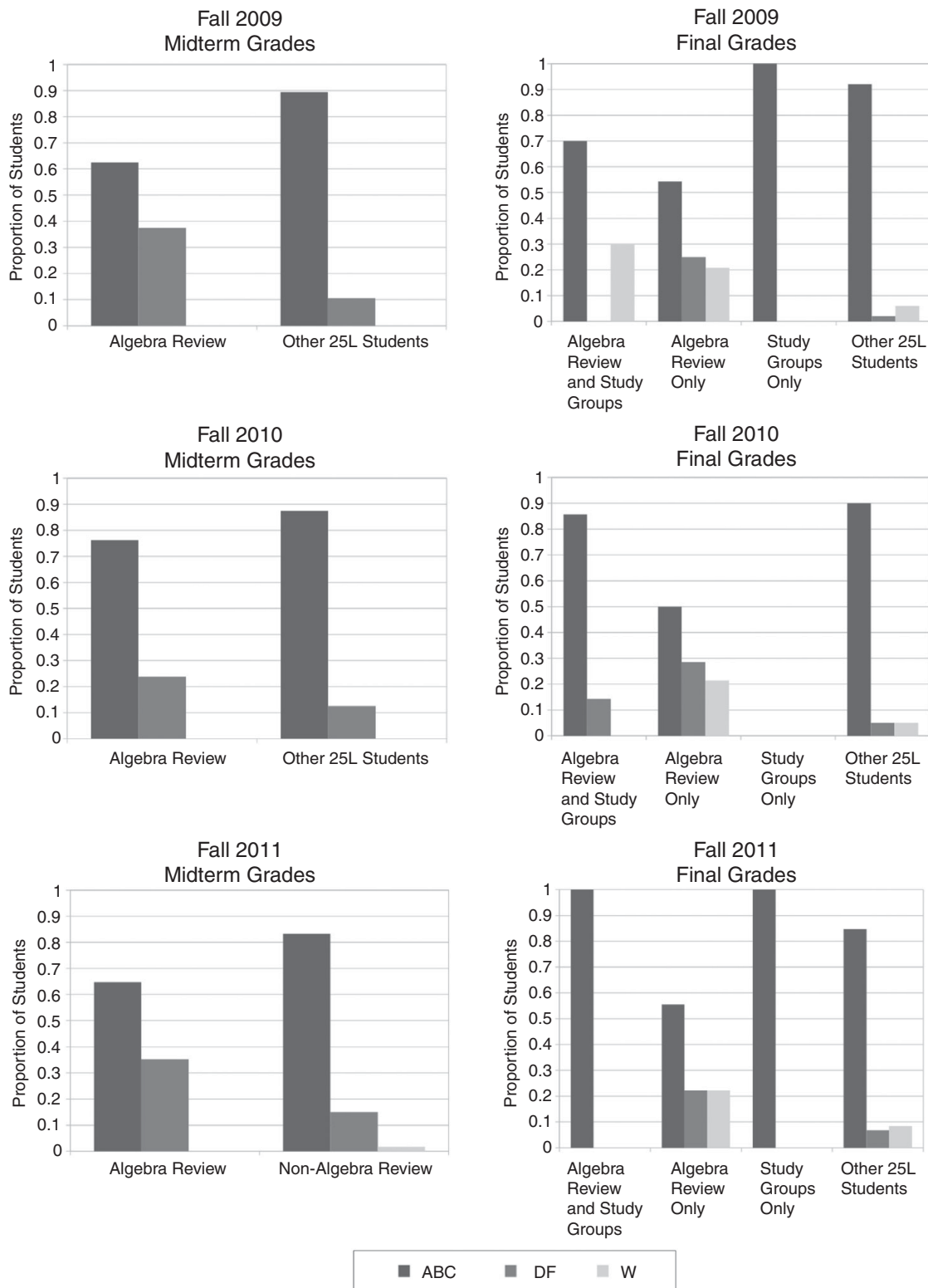
Because of the ad hoc nature of the algebra review, the sessions did not start until the fourth week of the semester and ran through the fifteenth week. As seen in Table 6.1, student participation was sporadic. Feedback from participating students suggested that material reviewed in the support sessions should be more closely tied to material being covered in the course.

### Fall 2009—Mandatory Algebra Review and Voluntary Study Group Sessions

In fall 2009, Rann Bar-On, a newly hired lecturer and recent PhD graduate of the Duke mathematics department, took over leadership of Math 25L. To encourage more student use of Cooke's algebra review sessions, Bar-On made participation mandatory for those students who failed the algebra quiz given in the first week of class. The algebra review sessions were abbreviated to four one-hour sessions held in weeks 2 through 5. The sessions reviewed algebraic topics and demonstrated the algebra skills needed for that week's course content. No follow-up algebra assessment was given. Instead one problem involving considerable algebra on the first test was used to attempt to assess the effect of the algebra sessions. Students in the algebra review had lower midterm grades on average than the rest of the class. This is shown in the top row of Figure 6.1.

After the algebra review sessions concluded in week 5, all students in the course were offered the opportunity to join weekly study groups conducted by Cooke, and some students in the algebra review sessions and some from the rest of the class participated. Students from the algebra review sessions who chose not to join the study groups were more likely to earn a D, F, or W (DFW) for their final grade than those who participated in the study groups, i.e., attended more than one session, as shown in the top row of Figure 6.1.

Selection bias may have influenced these results. Students who chose to attend weekly study sessions may have been using other support services more frequently or making other good academic choices about their study habits and how they used their time. So we cannot claim that attending the study sessions caused students to improve their grades. The data do show that students who entered the course less prepared were more likely to succeed if in their first semester they attended the weekly algebra review and the study group sessions.



**Figure 6.1.** Comparison of grades (ABC, DF, and W) of students participating in algebra review sessions and study groups over three fall semesters: 2009 (top row), 2010 (middle row), and 2011 (bottom row).

The bar graphs in the left column compare midterm grade frequencies of Math 25L students who participated in the algebra review groups (70 students) with those that did not (197 students).

The bar graphs in the right column compare final grade frequencies of four groups of students from left to right: (1) those attending the algebra review and study groups (25 students), (2) those attending the algebra review but not study groups (47 students), (3) those attending study groups but not algebra review (13 students), and (4) those attending neither the algebra review sessions nor study groups (189 students). (Midterm grades were not available for upper class students with a grade of C or higher.)

## Fall 2010—Better Assessment

To improve the assessment of the algebra review sessions, an extra week was added to provide time to administer a quiz at the end. To encourage attendance at review sessions, students' scores on the quiz replaced their lowest lab scores. The hope was to have students choose to attend the sessions and see them as helpful support rather than mandatory remediation.

Follow-up hour-long study group sessions offered twice a week by Cooke ran from week 8 through week 15 of the semester, and were attended primarily by those students from the algebra review sessions. The students who entered the course with weaker preparation and attended both the algebra review sessions and the study sessions all earned grades of D or better in the course. However, those students who attended the algebra review sessions but did not continue with the study groups had a higher rate of DFW grades than the rest of the class, as shown in the middle row of Figure 6.1.

In fall 2010, Bar-On and Cooke met weekly to monitor student progress and to discuss students in danger of failing 25L. These students were referred to the ARC for individual academic skills appointments to work on time management, problem solving approaches, or other academic skills. Even students who consistently attended the algebra review sessions made errors involving key algebraic skills, such as using function notation incorrectly, on tests and quizzes. This continued to be true in the next math course, Math 26L Laboratory Calculus with Functions II, also taught by Bar-On.

## Spring 2011—Learning from the Educational Literature

In an attempt to understand this evidence by drawing on the mathematics education literature, Bar-On and Cooke hypothesized that covariational reasoning (Carlson, Jacobs, Coe, Larsen, & Hsu, 2002), the ability to examine and represent how two covarying quantities change together, is a skill that students in Math 25L and 26L are slow to develop. To measure covariational reasoning, we undertook a study of students' precalculus skills in a variety of Duke laboratory calculus classes after getting permission to use the Precalculus Concept Assessment (PCA) from Marilyn Carlson at Arizona State University (Carlson, Oehrtman, & Engelke, 2010). While we found that students in all Duke calculus classes do better on this assessment than the precalculus high school students studied by Carlson and her colleagues (2002; 2010), we discovered a significant gap in abilities between students in 25L and 26L and those in our higher level Calculus I and II classes, Math 31L and 32L. The data supported the intuition that students in 25L and 26L have weaker precalculus skills than those students whose math SAT scores qualified them for more traditional first-year calculus classes. A summary of the data can be found in Table 6.2.

		25L	26L	32L	Carlson		25L	26L	32L	Carlson
<b>Overall</b>	mean	0.53	0.62	0.78	0.41					
	st. dev.	0.15	0.15	0.13						
	#students	39	38	32						
<b>Reasoning Abilities</b>						<b>Understand growth rate of function types</b>				
Process view of a function		0.49	0.58	0.76	0.35	Linear	0.40	0.47	0.64	0.33
Covariational Reasoning		0.53	0.67	0.81	0.5	Exponential	0.31	0.42	0.72	0.17
Computational Ability		0.58	0.59	0.77	0.42	Rational	0.42	0.67	0.69	0.59
						General non-linear	0.60	0.68	0.9	0.46
<b>Understandings</b>						<b>Understand Function Representations</b>				
Function Evaluation		0.76	0.82	0.91	0.52	Graphical	0.55	0.64	0.82	0.31
Rate of Change		0.42	0.45	0.69	0.26	Algebraic	0.49	0.55	0.71	0.41
Function Composition		0.60	0.73	0.85	0.41	Numerical	0.54	0.74	0.84	0.4
Function Inverse		0.45	0.58	0.69	0.31	Contextual	0.43	0.51	0.75	0.41

**Table 6.2.** Summary of PCA results, spring 2011. Running  $t$ -tests between the overall mean scores, we obtained the following results: 25L is statistically significantly different from 26L, 25L is statistically significantly different from 32L, and 26L is statistically significantly different from 32L;  $p < .01$  in all cases



Week	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Total Class Size
Fall 2009	30	31	28	28	<i>11</i>	<i>8</i>	<i>14</i>	<i>6</i>	<i>13</i>	5	6	5	0	0	96
Fall 2010	18	19	17	15	15	5	6	7	5	6	3	6	6	7	101
Fall 2011	13	15	16	16	16	7	8	9	6	8	0	8	5	9	77

**Table 6.3.** Combined weekly attendance figures for the hour-long algebra review sessions and the hour-long study group sessions (in italics) for fall 2009, fall 2010, and fall 2011

The PCA was administered to students midway through the semester. As can be seen in Table 6.2, there were statistically significant differences in conceptual understanding between students in 25L and 26L and those in the higher level classes. Math 25L and 26L were designed to prepare students for Math 32L. The data, as well as the instructors' collective experience with these classes, indicated that students coming out of 25L and 26L were not as well prepared for the rigors of a second semester calculus course as those coming from the one-semester Math 31L class.

### Fall 2011—Using What We Learned

In fall 2011, the structure of the support sessions did not change significantly, but the PCA was used in conjunction with the algebra quiz to identify students who might benefit from extra algebra support. To encourage attendance at the algebra sessions, the follow-up algebra quiz score replaced the lowest quiz grade. Also an extra emphasis was placed on the concept of a function, including working with function notation, solving system of equations, graphing polynomials and functions, shifting graphs of functions, and finding function inverses. Algebra review sessions ran from weeks 2 through 6, and follow-up study group sessions ran from weeks 7 through 15. All the students who attended both the algebra review sessions and the study group sessions passed the class with a grade of C or better, while those students who attended the algebra review sessions but not the study sessions earned DFW grades with a higher frequency than the rest of the class, as seen in the bottom row of Figure 6.1.

Attendance data for the algebra review sessions and the study sessions are summarized in Table 6.3. The total class size for Math 25L appears in the last column.

After collaborating for three years to support students entering college calculus in the fall semester of their first year, assessing the impact of the algebra review and study sessions throughout the semester, and using the literature to inform our understanding of the mathematical development of precalculus and calculus students, we felt that the course needed further changes. Although students identified as needing extra support who consistently attended the algebra review and the study groups seemed to be passing the course, the performance of the students who did not continue to attend the study groups suffered. Furthermore, even after administering the PCA and obtaining evidence that Math 25L and 26L students' covariational reasoning lagged behind students in higher level calculus courses, we had no plan to bridge this gap.

We needed to develop a further curriculum change to strengthen students' mathematical understanding while supporting first-year students making the transition to collegiate level work. These curriculum changes will be modeled on a successful collaboration between the ARC and the chemistry department at Duke (Hall, Curtin-Soydan, & Canelas, 2013). We have jointly submitted a grant application to the NSF to support and assess the next phase of our work.

### Outcomes and Future Directions

In this chapter, we have presented an alternative pathway to SoTL. A focus on what seemed to be a localized problem evolved into a scholarship of teaching and learning project. We consulted with the IRB for permission to publish the data in this paper, for which the IRB initially encouraged us to submit an exemption application. When it became clear to the IRB that we were reporting internal assessment of an ARC collaboration, the IRB allowed us to submit for publication without IRB exemption. However, because the IRB process varies from campus to campus, SoTL researchers should always consult their local boards. We broadened the scope of our efforts from focusing on our own

classrooms to reading the research literature in the field and going public with the results of our work. This evolving SoTL project continues to prompt additional adjustments and improvements to Duke's curriculum and its student support services.

The study group program will be extended to Math 26L in spring semesters. Algebra review sessions will also be organized for Math 25L in spring semesters, and study groups established for both classes in their off-semester format (25L in the spring, and 26L in the fall). In addition, a tutorial course on learning theory will be offered to undergraduate lab TAs. That these improvements would become the focus of an NSF grant application was an unforeseen outcome of the collaborative effort when it began.

It is worth noting that this endeavor was undertaken by regular rank, non-tenure track teaching faculty with teaching loads that consisted of four courses per semester. Thus, it is reasonable to believe that these efforts can be imitated at universities around the country where faculty members have heavy teaching loads. Although our focus was on the introductory calculus course at Duke University, retention in the STEM disciplines is a problem faced by many colleges and universities. We hope that the story of how we fell into SoTL simply by striving to become more effective educators will inspire faculty members elsewhere to engage in SoTL.

## References

- Arcidiacono, P., Acejo, E., Fang, H., & Spenner, K. (2011). *Does affirmative action lead to mismatch?* Retrieved from [www.seaphe.org/working-papers/](http://www.seaphe.org/working-papers/)
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Bressoud, D. (2010, January 17). The rocky transition from high-school calculus. *Chronicle of Higher Education*. Retrieved from [chronicle.com/article/High-School-Calculus-The/63533/](http://chronicle.com/article/High-School-Calculus-The/63533/)
- Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, 33(5), 352–378.
- Carlson, M., Oehrtman, M., & Engelke, N. (2010). The Precalculus Concept Assessment: A tool for assessing students reasoning abilities and understandings. *Cognition and Instruction*, 28(2), 113–145.
- Dembo, M.H., & Seli, H. (2008). *Motivation and learning strategies for college success: A self-management approach*. Hillside, NJ: Lawrence Erlbaum Associates.
- Hall, D.M., Curtin-Soydan, A.J., & Canelas, D.A. (2013). The science advancement through group engagement (SAGE) program: Leveling the playing field and increasing retention in science. *Journal of Chemical Education*, 91(1), 37–47.
- Hutchings, P. (2010). *Opening doors to faculty involvement in assessment* (NILOA Occasional Paper No. 4). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.
- Schraw, G., Crippen, K.J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36(1), 111–139.
- Zimmerman, B.J., & Schunk, D.H. (2001). *Self-regulated learning and academic achievement, theoretical perspectives*. Hillside, NJ: Lawrence Erlbaum Associates.



# 7

## A Quantitative and Qualitative Comparison of Homework Structures in a Multivariable Calculus Class

Lynn Gieger, John Nardo, Karen Schmeichel, and Leah Zinner  
*Oglethorpe University*

### Editors' Commentary

This chapter describes the work of an interdisciplinary team, Lynn Gieger (Mathematics Education), John Nardo (Mathematics and Computer Science), Karen Schmeichel (Biology), and Leah Zinner (Psychology), to improve the effectiveness of homework in a multivariate calculus class by using an online homework system. Although at the beginning of the investigation, the majority of the team members were not very familiar with qualitative methods, they found the qualitative data particularly useful for providing context and depth to the mixed quantitative results they obtained. Their study highlights the merit of a mixed method approach, particularly when statistically significant results are not obtained from quantitative data. The authors also testify to the value they found in working collaboratively with colleagues from other disciplines.

### Introduction

A conference workshop on SoTL inspired four novices from diverse disciplines (mathematics, education, biology, and psychology) to bring research tools to bear on pedagogical issues. This chapter demonstrates how to start a SoTL project even if the investigator is not fully trained in SoTL methods.

Our calculus sequence serves many purposes. It is both a service we provide to other client disciplines, and also a pipeline into our major that we must nurture. Because Oglethorpe University is a small (approximately 825 undergraduates), private, selective liberal arts college, it has traditionally had only one section of calculus students per academic year. One of the authors, John Nardo, has taught the calculus sequence every other year for the last 12 years. Like many mathematics faculty, he was concerned with his students' entry-level skills and the amount of practice they devoted to mastering new skills. He believed that the structure of homework assignments might have an impact. In broad terms, he wondered how to make "homework," i.e., the practice taking place outside of class, truly effective. Were there better ways, perhaps taking advantage of recently developed learning tools, to encourage practice and mastery?

These questions were on Nardo's mind in the summer of 2010 when several Oglethorpe faculty members attended the 2010 SENCER (Science Education for New Civic Engagements and Responsibilities) Summer Institute. As an

NSF-funded curriculum dissemination initiative, SENCER supports the development of STEM pedagogies that encourage student engagement in “capacious issues” that are conceptually complex and socially relevant (see [www.sencer.net](http://www.sencer.net)). SENCER enlists a wide network of faculty across disciplines and institutions who are interested in the kinds of pedagogical questions found in this volume. SENCER supports approaches aimed at student success and encourages faculty to perform research on pedagogical tools.

For the past five years, Oglethorpe University’s SENCER working group, comprised of faculty from the mathematics, psychology, chemistry, physics, and biology departments, has regularly attended the SENCER Summer Institute (SSI). In 2010, three of us attended a SoTL workshop at the Institute hoping to learn more about the practice of SoTL. Nardo chose this workshop to determine whether a SoTL investigatory approach might be applicable to his calculus homework problem.

The workshop leader, Matthew Fisher, an Associate Professor of Chemistry at St. Vincent’s College and a 2005-2006 Carnegie scholar, encouraged us to use our classrooms as laboratories (or field sites) and to design experiments using the same methodologies that we employ in our disciplines. We came to appreciate the two phases of a SoTL experimental design. The first phase consists of defining a goal. Because it is tempting to be overambitious with project goals, we were urged to refine our goals and warned that the process might take several iterations. Once the question is finalized, the second phase can begin: determining the nature of the experiment to be performed and what types of evidence to collect.

Although quantitative data collection is the norm in most STEM disciplines, SoTL projects can be informed by measurements that are descriptive or qualitative in nature. The process of analyzing qualitative data, as modeled in the SSI workshop, involves coding responses by mining the data (derived from student satisfaction surveys, interviews, focus groups, and reflective writing) for common words or ideas and then revising the codes to reflect the major themes and attitudes of the students.

The focused yet flexible approach of a SoTL investigation seemed perfect for investigating the calculus homework problem observed by Nardo. Moreover, we thought that adopting the SoTL perspective would have broad appeal with colleagues at our home institution, and perhaps elsewhere.

## Formulating a Plan

During the three-hour SoTL workshop, Nardo and his two colleagues discussed how to address his homework problem and study the results of an intervention. A fourth Oglethorpe colleague joined them for the drive home. While in transit, they (the four authors) set about refining the question to be investigated regarding the calculus homework project and then determining a suitable experimental design. The benefits of having participants from several disciplines (mathematics, mathematics education, psychology, and biology) were apparent. The group had a breadth of experience in qualitative and quantitative research design and analysis and was well equipped to try to answer this question using several approaches of inquiry. The biologist and the psychologist brought skill and experience with experimental design, and having a social scientist brought expertise in the design of survey questions and in dealing with the required Institutional Review Board policies for conducting studies on human subjects. The mathematics education team member knew qualitative analysis methodology and provided guidance at this early stage and throughout the study. Moreover, we found collaborating across disciplines to be energizing, fun, and morale boosting.

By the end of the trip, the team had defined a question, determined the experimental design, and begun work on developing the assessment tools. While the SoTL workshop had served as a catalyst for this burst of productivity, we benefitted from other factors. Being away at SSI meant that we were free from the demands of the classroom and other on-campus duties. We were also able to capitalize on our comparable professional expectations, our common interest in student success, and our shared experience of the intimate teaching environment afforded by our home institution.

## The Homework Problem Revisited

Many studies have demonstrated the link between homework and achievement (for a meta-analysis, see [Cooper, Robinson, & Patall, 2006](#)), and specifically, the link between frequency of homework and mathematics achievement gains ([Trautwein, Köller, Schmitz, & Baumert, 2002](#)). Although college students have the time to complete homework, they often choose not to ([Cerrito & Levi, 1999](#)). Thus, if there were a method for ensuring students would complete

homework more frequently, gains in achievement might follow. In the first two courses of the calculus sequence, students had been assigned traditional textbook-based, written homework. At the end of each week, students were to submit full solutions to three textbook problems assigned by the instructor. Students reported (in informal conversations and in course evaluations) doing this graded homework at the last minute, if at all, and they were not overly concerned with completing the assignment since it counted for only 10% of the course grade. Even more worrisome was what this attitude implied for the bulk of homework, which was ungraded and thus did not factor into the course grade at all. How much time were students spending on homework? How much practice and skill development was happening?

An online homework system might solve these problems. Interest in online homework systems and their effectiveness has been growing (Demirci, 2007; Roth, Ivanchenko, & Record, 2008). Research has found that online homework does not negatively impact students' achievement (LaRose, 2010) and that students perceive it positively (Hauk & Segalla, 2005; Zerr, 2007). To investigate the efficacy of online homework, the 10% homework portion of the next semester's Calculus III grade was changed to include online homework, while the other 90% of the course grade (projects, in-class tests, take-home tests, and final examination) was the same as the previous two semesters.

The online homework was administered through WebAssign<sup>®</sup> (an online homework system developed at North Carolina State University). At each class meeting, three online homework problems were due from the previous class meeting's material. The problems assigned were algorithmically generated and had random coefficients; even after multiple attempts, a student was not likely to be given the exact same problem twice. At the end of the week, students submitted one full solution in writing, choosing any of the nine problems from the last week's work, for feedback and grading.

## Study Design: A Mixed Approach

A factor that affected our experimental approach was the fact that we teach at a small liberal arts college. Because only one section of third-semester calculus is offered each year, a design that required having a section to serve as a control group was not possible. Since the research was being conducted with students who expect personalized, high-quality learning experiences, we did not consider it ethical or reasonable to perform studies with a classical control and treatment strategy. This made us think carefully about what sources of data were available to us. We understood that quantitative data would be useful for assessing changes in academic performance, attitudes, and behaviors. However, because of our close relationship with our students, we had an opportunity to gain more detail about how the change in homework affected their attitudes and motivation. This led us to consider collecting qualitative data in addition to quantitative data from surveys and course grades. As has been described by Patton (1990), a quantitative approach lends itself to effective statistical aggregation of data from many people, whereas qualitative analysis of data collected from only a few subjects leads to "an increased understanding of the cases and situations studied" (p. 14). Thus, including a qualitative method—we chose a focus group—seemed appropriate for a study of the small cohort of Oglethorpe calculus students. Because our literature review indicated that qualitative methodology was seldom used in studies at larger institutions, we would be bringing something new to the study of the problem.

## Research Questions

The multidisciplinary perspectives of the research team led us to formulate questions that were designed to be answered both from a hypothesis testing and qualitative analysis approach. We asked:

- How does the use of online homework affect student attitudes, motivation, and performance?
- Will it lead students to study more frequently, be more engaged and responsible for their learning, and lead to improved homework grades, test grades, or course grades?

## Method

### Survey and Focus Group Participants

The subjects in our study were the 17 students enrolled in Calculus III. They were taking the course as a requirement for a mathematics major, mathematics minor, or pre-engineering program. Fifteen students completed a survey on the

first day of class, and sixteen students completed one on the last day of class. Thirteen students completed a paired set of pre- and post-surveys. Using stratified random sampling (the division of a population into smaller groups prior to random selection) based on the course grade in Calculus II and gender, six students were selected for focus group interviews. All six students participated in the focus group at the beginning of the semester; five participated at the end of the semester.

## Survey and Focus Group Materials

Surveys and focus groups were held in the first week of the new class to investigate student views of the previous homework system before work with the new system could cloud those memories.

### Survey Design

The beginning-of-semester survey included 11 questions that asked students to reflect on their habits, skills, and attitudes during their Calculus II course (see Appendix A). For example, students responded on a scale of 1 (Strongly Disagree) to 5 (Strongly Agree) to statements such as “I am skillful in solving calculus problems” and “Homework contributed to my mastery of calculus.” Students also reported how much time they spent on certain study activities, including doing homework and preparing for exams. The survey included an open-ended question that asked students to share anything else about their homework experience. The end-of-semester survey included an additional question asking students to compare the previous graded homework model with the new online graded homework model (see Appendix B). A faculty member, who was not the course instructor, administered the survey. The instructor was not present during the survey administration. Students were assured that the instructor would not see their responses until after final course grades were submitted.

### Focus Group Interviews

We reviewed the qualitative and quantitative data from the beginning-of-semester survey to help us shape the interview protocol for the first focus group. This consisted of an informal conversational interview in which students were encouraged to speak freely about their thoughts and opinions regarding calculus homework. We also prepared several discussion prompts, some of which were drawn directly from the survey. They were on two topics: why the students made specific choices regarding homework, and why they believed the instructor made specific choices regarding homework. For example, we said, “Describe the conditions under which you usually did your homework” and “In your opinion, why did [the instructor] only grade some of the homework he assigned?” We also asked students about some of the results from the survey to clarify why certain patterns emerged. For example, on average, students reported attempting only 58% of the ungraded problems assigned from the text. We asked students why they thought this was the case.

Constant comparative analysis (Corbin & Strauss, 1990) was used to develop a more structured interview protocol for the end-of-semester focus group (see the Analysis and Results section for more detail). The questions fell into three categories: (1) the students’ interpretation of the instructor’s behaviors (“Why do you believe [the instructor] made the change to homework?”), (2) students’ logistic behaviors, that is, what they did (“Did you tend to work on homework problems sooner after class as a result of the new system?”), and (3) students’ motivational behaviors, or why they did what they did (“Were you more or less likely under the new system to seek confirmation from [the instructor] now that there was an additional authority [the website]?”). The same research team member led both focus groups, and the course instructor was not present for either session. Students were assured that the instructor would not see their responses, and they used pseudonyms when responding. Each focus group session lasted about one hour.

## Analysis and Results

We compared students’ grades on offline homework in Calculus II to their grades on online homework in Calculus III. Not surprisingly, students received higher letter grades on online graded homework than on offline graded homework ( $p < .001$ ). When asked to compare online and offline homework (end-of-semester survey, question #12), 64% of

students reported that online homework was more helpful, 29% stated that offline homework was more helpful, and 6% said they were equally helpful.

Students also estimated how much time they spent on graded and ungraded homework each week. The average (self-reported) time spent on graded homework went up only about 12 minutes per week, and the average time spent on ungraded homework went down about 70 minutes per week. Neither difference was statistically significant. We also examined how the change in homework structure might have affected course grades in Calculus III. There was no significant change in average course grade in comparison to Calculus II, nor in comparison to the Calculus III course taught by the same instructor, with the same textbook, in the previous year.

The results provided us with important and useful information, but due to our small sample size, our statistical power to detect significant differences was limited. We turned to qualitative analysis to gain further insight.

For the novice qualitative researcher, we want to clarify that analyzing interviews requires more than reporting a collection of quotations. Our analysis used a modified version of constant comparative analysis (Corbin & Strauss, 1990; Creswell, 1998; Glaser & Strauss, 1967). An essential feature of this approach is that collection, design, and analysis are intertwined. In a pure constant comparative analysis, “categories, codes, and related questions . . . emerge from the data, at which point these questions [are] taken back to the participants for further discussion . . .” (Gieger, 2002). The cycling process repeats until the themes and categories are saturated, that is, until participants felt that the discussion on a particular topic had run its course and they had said all they wanted to say on the issue. We modified this process by having only one iteration, between the first and second focus group interviews. This iteration was why the interview protocol was more structured for the second focus group. This structure resulted from the analysis of the transcript of the first focus group and the answers to the open-ended question on the beginning-of-semester survey. The three authors who did not teach the class all participated in the coding process. This helped organize and synthesize the data and limited subjective bias by including coders from different disciplinary perspectives who were not present at the focus group interviews.

In pure constant comparative analysis, coding has three phases: open coding, axial coding, and selective coding. Open coding is relatively informal, with each coder applying summary titles (for example, “study time”) to larger quotations from the qualitative data. Once the coders come together to share their informal coding results, codes can be either eliminated (because they lack sufficient support in the data) or collapsed (because two coders used similar terms to describe a phenomenon). Axial coding takes the codes produced in the open coding process and combines them into larger categories. The purpose of this stage is to identify central phenomena and to continue to mine the data for the conditions that contributed to them. The final phase, selective coding, attempts to integrate the categories that emerged from the axial coding process into a story line that outlines a theory (Gieger, 2002). Creswell (1998) or Patton (1990) provide a more complete discussion of qualitative design and analysis techniques.

As hoped, the qualitative data allowed us to gain a deeper understanding of the limited quantitative results from the survey. For example, though there was no significant difference in reported time spent on homework during the week, the focus group interviews indicated that the distribution of time had changed. In the words of one student, “I was actually doing calculus throughout the week—which I did not do last semester.”

We learned that the change in study habits was largely due to the convenience and accessibility of the online homework system from comments like these two:

I actually did more problems that I did not have to do because it was just more convenient for me.

It was more convenient to be able to open up the textbook on the computer, read the exact section it was in, instead of having to go back to the room, get the textbook, and figure out everything.

Although the difference was not statistically significant (most likely due to the small sample size), the quantitative data indicated a nontrivial drop in the average number of minutes spent on ungraded textbook homework problems per week. The qualitative data indicated that students were replacing those ungraded textbook problems with ungraded, algorithmically generated practice problems on WebAssign<sup>®</sup> because of the immediate feedback they received from the computerized system:

It’s not like doing ungraded homework in the book. You could work it out, and instead of feeling stuck, I got help from the computer right there.



As a result, students reported seeking out less feedback from the instructor in the course, for example:

Last semester, I used to go to my professor and just drop him a lot of problems. This semester, I didn't have to go.

The most surprising result, however, was that while they saw the benefit of having another resource for feedback and instruction, students were concerned about how this might affect their personal connection to the instructor:

[The instructor] is such a fabulous resource that I think if going to him becomes unnecessary. . . I don't know about that.

and

I definitely think it cut out the professor-student relationship.

In sum, though students saw many benefits to the online homework structure (convenience, immediacy of feedback, and change in study habits), they came at the cost of what they perceived as a loss of personal mentorship.

## Discussion

The SoTL approach performed in this study demonstrates the successful and rapid implementation of a study whose data, although limited in scope, can inform an instructor about instructional changes aimed at bolstering student success. Specifically, the study allowed the instructor to assess, with confidence, the advantages and disadvantages of the new homework structure.

### Direct Implications for Homework in the Calculus Sequence

The qualitative data revealed two striking findings. First, students were working on homework more often. Unlike the old system, in which students did homework once a week right before it was due, students were practicing problems throughout the week. As discovered in the focus groups, when they had a few free moments between classes, they would fire up the laptop to log some practice time. Thus, we were successful in one of our primary objectives: students were now working on their calculus skills throughout the week.

Second, the more students used the online system to practice homework problems and obtain assistance and feedback, the less importance they placed on an office visit. One focus group member said how sad it was to lose that contact. The instructor had assumed that students would still make use of office hours and appointments as they had previously. He learned that he needed to increase his student outreach efforts to compensate for the shift resulting from online homework. As encouraged as we were by the changes in student behaviors due to the instant feedback inherent in online homework systems, we discovered a need to find new ways to encourage the one-on-one relationship between faculty and students that small teaching-oriented institutions promote and value.

### Direct Implications for the Instructor

The major challenge for the instructor in this study was how difficult it was for him to give up control and knowledge of the data for the duration of data collection. He saw the homework, project, and test grades as they were generated or assigned, but he did not see any of the surveys or focus group data until after the students' course grades were submitted. He found his curiosity about the data and the temptation to find some snippets of information hard to drive away.

### General Implications for the Researchers

There are numerous ways in which a SoTL project can affect the work of the investigators as well as the greater teaching and learning community at small liberal arts schools like Oglethorpe University. Below we have listed examples relevant to both students (with respect to success initiatives) and faculty (with respect to professional development).

**Improved Assessment of Individual Classes**

SoTL investigations encourage evidence-based analysis of teaching practices at the classroom level. As instructors revise syllabi at the beginning of a semester, small changes can be introduced to help improve instruction. Viewing the classroom as a site for study can allow a faculty member to cast a fresh eye on classes that may have become routine after repeatedly teaching the same course. Our calculus homework study shows that at our institution, if an instructor initiates a study designed to improve course quality, students likely will be willing participants.

**Improved Review of Major or Program Goals**

The SoTL approach can go hand-in-hand with the major or program assessment reviews of most institutions. Results of SoTL studies, when shared with the department, can lay the groundwork for similar studies in other courses. For example, the results of our study were shared with colleagues in the mathematics department and increased their awareness of the utility of alternate teaching methods.

**Improved Evidence for Self-Evaluations**

When faculty members evaluate themselves and submit applications for tenure and promotion, it is helpful to have a portfolio of materials that address the development of teaching skills. A SoTL study such as the one reported here provides a faculty member with an example that demonstrates a willingness to reevaluate teaching pedagogies and skills.

**Increased Awareness among Faculty of Requirements in Other Disciplines**

Our multidisciplinary research team, which occurred somewhat by chance, was critical to the successful implementation of this study. An unanticipated byproduct of our group discussions was a better understanding of homework expectations in other disciplines. Awareness of what is required of STEM students in each of the contributing disciplines can inform future syllabi designs and promote effective articulation of anticipated goals and skills. Moreover, we discovered that concerns regarding instructor grading load and student fatigue (to the point of inaction) are shared throughout STEM disciplines where content mastery is achieved in large part through iterative problem solving. Discussing the data and results of studies such as ours with colleagues on campus can increase awareness of effective classroom practices. Discussions of ways to improve student success across disciplines are particularly relevant as institutional financial constraints continue to emphasize the importance of student retention.

**Exploration of New Collaborative Research Goals**

Though faculty members at small schools like Oglethorpe are trained in graduate school to pursue rigorous research agendas, their teaching demands leave little time for applying for outside research funding, let alone for performing the actual research. Typically, faculty members involved in STEM research require a research infrastructure that cannot be financially supported by small institutions. By engaging in SoTL, they can utilize basic skills of experimental design, implementation, and analysis to pursue research into effective teaching methodologies. This study demonstrates how expertise in a relatively small number of critical areas can support a SoTL investigation. A SoTL project is an inexpensive alternative for research-motivated faculty members and has the additional benefit of drawing from a relatively abundant experimental subject pool.

**Future Directions and Concluding Remarks**

Our calculus homework study is a good example of how SoTL can be used to evaluate small changes in course design that may benefit both the student and the instructor. However, the entire experience might have been improved in a number of ways. Although our project was managed effectively by dividing responsibility among the authors, our biggest limitation was time. We would have liked a bit more time to review and analyze the data collectively and to explore its points of intersection with previous SoTL work. As it was, deadlines seemed always to be imminent, allowing limited time for reflection and goal setting. An investigation would benefit from scheduling a retreat of some type to afford the same type of productive seclusion as the original SSI conference. The authors could seek financial support for such a retreat through faculty development resources on campus. The retreat would foster discussions of “what types of questions on our campus are SoTL-worthy?” and “how much SoTL at a small liberal arts school is too

much (i.e., with regard to cost versus benefits in the classroom)?” If nothing else, the project can lead to additional scholarly interactions on campus that raise awareness of effective SoTL research practices at other small liberal arts institutions and beyond.

Setting the time and resource constraints aside, this chapter shows how a small ember of an idea grew into a study that introduced four faculty members to SoTL work, pushed them into developing new research skills, and culminated in a national conference presentation and this book chapter. Our experience shows how participating in one national, regional, or state meeting either dedicated to SoTL or having a SoTL workshop with a small number of peers can lead to a worthy project.

## References

- Cerrito, P. B., & Levi, I. (1999). An investigation of student habits in mathematics courses. *College Student Journal*, 33(4), 584–588.
- Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research*, 76(1), 1–62.
- Corbin, J., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13, 3–21.
- Creswell, J. W. (1998). *Qualitative inquiry and research design*. Newbury Park, CA: Sage.
- Demirci, N. (2007). University students’ perceptions of web-based vs. paper-based homework in a general physics course. *Eurasia Journal of Mathematics, Science and Technology Education*, 3(1), 29–34.
- Gieger, J. L. (2002). *A model of academic choice for mathematically talented college women* (unpublished doctoral dissertation). University of Georgia, Athens, GA.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago, IL: Aldine.
- Hauk, S., & Segalla, A. (2005). Student perceptions of the web-based homework program WeBWorK in moderate enrollment college algebra classes. *Journal of Computers in Mathematics and Science Teaching*, 24(3), 229–253.
- LaRose, P. G. (2010). The impact of implementing web homework in second-semester calculus. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 20(8), 664–83.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.
- Roth, V., Ivanchenko, V., & Record, N. (2008). Evaluating student response to WeBWorK, a web-based homework delivery and grading system. *Computers & Education*, 50(4), 1462–82.
- Trautwein, U., Köller, O., Schmitz, B., & Baumert, J. (2002). Do homework assignments enhance achievement? A multilevel analysis in 7th-grade mathematics. *Contemporary Educational Psychology*, 27(1), 26–50.
- Zerr, R. (2007). A quantitative and qualitative analysis of the effectiveness of online homework in first-semester calculus. *Journal of Computers in Mathematics and Science Teaching*, 26(1), 55–73.

## Appendix A

This research project concerns your homework experiences from Calculus I and II last year.

Recall that there were two different types of homework:

- **Graded** homework was approximately once per week and was submitted to Dr. Nardo for grading both the “answer” and the supporting solution.
- Ungraded homework was posted online for every section of the textbook, but it was not submitted for grading.

Please answer questions #1-6 with a whole number on a scale of 1 (Not at All) to 5 (Definitely). Please answer question #7-10 with a whole number estimate.

1. <b>Graded</b> homework contributed to my mastery of calculus.	
2. <b>Graded</b> homework helped me prepare for exams.	
3. <u>Ungraded</u> homework contributed to my mastery of calculus.	
4. <u>Ungraded</u> homework helped me prepare for exams.	
5. I am skillful in solving calculus problems.	
6. I am confident in explaining calculus problems to my peers.	
7. How many minutes per week did you typically spend on <b>graded</b> homework?	
8. What was your typical letter grade on a <b>graded</b> homework assignment?	
9. How many minutes per week did you typically spend on <u>ungraded</u> homework?	
10. What percentage of a typical <u>ungraded</u> homework did you attempt?	

11. Thinking of the other classes you were taking last academic year, which of the items below best applies to your experience in Calculus I/II?
- Overwhelmingly more time spent on calculus
  - More time on calculus than on other classes
  - Same on calculus as on other classes
  - Less on calculus than on other classes
  - Overwhelmingly less time spent on calculus

Please share anything else about your homework experiences in Calculus I/II last year.

Please be specific about whether you’re discussing **graded** or ungraded homework for each item.

## Appendix B

This research project concerns your homework experiences from Calculus III this year compared with Calculus I and II last year.

Please answer questions #1-6 with a whole number on a scale of 1 (Not at All) to 5 (Definitely).

Please answer question #7-10 with a whole number estimate or a letter grade estimate, as requested.

1. <b>Online graded</b> homework contributed to my mastery of calculus.	
2. <b>Online graded</b> homework helped me prepare for exams.	
3. <u>Textbook ungraded</u> homework contributed to my mastery of calculus.	
4. <u>Textbook ungraded</u> homework helped me prepare for exams.	
5. I am skillful in solving calculus problems.	
6. I am confident in explaining calculus problems to my peers.	
7. How many minutes per week did you typically spend on <b>online graded</b> homework?	
8. What was your typical letter grade on an <b>online graded</b> homework assignment?	
9. How many minutes per week did you typically spend on <u>textbook ungraded</u> homework?	
10. What percentage of a typical <u>textbook ungraded</u> homework did you attempt?	

11. Thinking of the other classes you were taking this semester, which of the items below best applies to your experience in Calculus III?

- A. Overwhelmingly more time spent on calculus
- B. More time on calculus than on other classes
- C. Same on calculus as on other classes
- D. Less on calculus than on other classes
- E. Overwhelmingly less time spent on calculus

12. How would you compare the online graded homework (in which you had multiple attempts for each problem) and the previous graded homework model from last year?

- A. Online homework was much more helpful
- B. Online homework was slightly more helpful
- C. Both types of homework were equally helpful
- D. The previous graded homework model was slightly more helpful
- E. The previous graded homework model was much more helpful

Please share anything else about your homework experiences in Calculus III. Please be specific about whether you're discussing **graded** or ungraded homework for each item.

# 8

## Playing Games to Teach Mathematics

Edwin Herman

*University of Wisconsin-Stevens Point*

### Editors' Commentary

Edwin Herman's project was prompted by a desire to incorporate games as a learning device. This chapter details how the project unfolded in stages, because as he gathered evidence he kept refining his question. The first time he gathered evidence regarding whether the students enjoyed the activity. Next, he sought evidence of learning. Because of his knowledge of statistical methods he relied mostly on quantitative evidence. During the last iteration of the course, he became more interested in why game play might be having an effect, leading him to consider qualitative measures. Even readers having neither the interest nor the time to incorporate games in their teaching can learn about the process and the benefits of SoTL from this chapter.

### Introduction

Think back to your earliest investigations into mathematical research. Not graduate school – long before that, when you were getting your first taste of exploring mathematics. For me, it was in high school – the idea of fractional derivatives (as opposed to partial derivatives). I had learned how to take a first derivative, and a second one, and so forth, but what about a *half* derivative? I worked mightily on the problem, struggling to define a formula for a half derivative that worked. Lacking the background needed to answer the question, I was ultimately unsuccessful in discovering a formula (or, rather, re-discovering it – papers and books have been published on fractional derivatives, though I did not find them until years later, in college). Although I lacked the tools and sophistication necessary to answer my question, the attempt taught me a lot about mathematical research and made me a better mathematician in the process.

When you start to embark on a project into the Scholarship of Teaching and Learning, it is worth keeping in mind the idea that the process of doing the research will teach you as you explore. As you progress, you will discover better ways to ask questions – even better questions to investigate. You may or may not succeed in answering your original question in a satisfying way, but you will become a better researcher, and a better teacher, because of it.

One of my first forays into SoTL involved playing board games in the classroom. I love board games. To me they are full of strategy, decision-making, and fun. Naively, I thought that students would pay more attention to a fun activity that involved mathematics in a competitive setting, and I thought this might make them learn it better. I looked for research to support my ideas about games. While writing this chapter I discovered a wealth of information on the use of games and simulations in the classroom (e.g., [Hertel & Millis, 2002](#)); unfortunately, my initial searches were not very fruitful. I found some examples of modified versions of Trivial Pursuit being used in middle and elementary

grades (Dunn, 1989), but little else. It was only later, after I had completed the first part of my project, that I thought to broaden my search to include journals in other disciplines and non-educational journals devoted to game analysis. In spite of my initial lack of success at finding background research, I decided to try games in the classroom and see what happened.

## Game Design

My first step was to design a game for a section of Mathematics for the Social and Management Sciences – without a game, it would be impossible to study the effects of game play! The course, designed primarily for business majors, covers financial mathematics, matrices, a little differential calculus, and some linear programming. I wanted to use the game throughout the course, not merely to illustrate one aspect. Although a simulation (modeling a real life phenomenon) could have explored some of these topics in the course, it was difficult to see how I could incorporate all the topics.

Additionally, there was the issue of rules. As Corbell (1999) advised, “the game itself must be so constructed that the participants can spend as little time as possible mastering the machinery of the game and as much time as possible using it to learn” (p. 177). To accomplish this, I turned to popular games. Other researchers have done likewise, such as Schoen (1996) and Zirkel (1975). In each case, the author took a well-known game and modified it to suit his needs. As a template, I used Trivial Pursuit, where teams compete to answer questions. Trivial Pursuit is played on a board with different paths, allowing teams multiple options for their next question. I deemed this unnecessary so the game became linear. Instead of acquiring pieces, teams would need to get as many questions correct as possible. I modified the game’s rules to keep one team from monopolizing the play by limiting to two the number of turns in a row a team could take. I also borrowed an “all play” feature from the game Pictionary (where some game spaces challenge all teams, not just the team that moved there). Using these ideas, a simple drawing program and some poster boards, I created four copies of a quizzing game in which teams had to answer questions from five different categories, with a total of 36 question cards per category. See Figure 8.1.

I decided to split students into teams for several reasons. As Hatch (2005) pointed out, “a game [should] allow a pupil to hide until he feels confident” (Pupil experience section, para. 5). Making weaker students try to answer questions by themselves while others watched could easily create an uncomfortable atmosphere and negate any benefits of the game. Additionally, I wanted the students to talk about the mathematics. Having them compete as teams encouraged students to talk with their partners about the questions. This strategy had the benefit of allowing the other teams to learn from listening to their opponents’ discussion. Most of the time I limited the size of teams to two, with up to four teams per board.

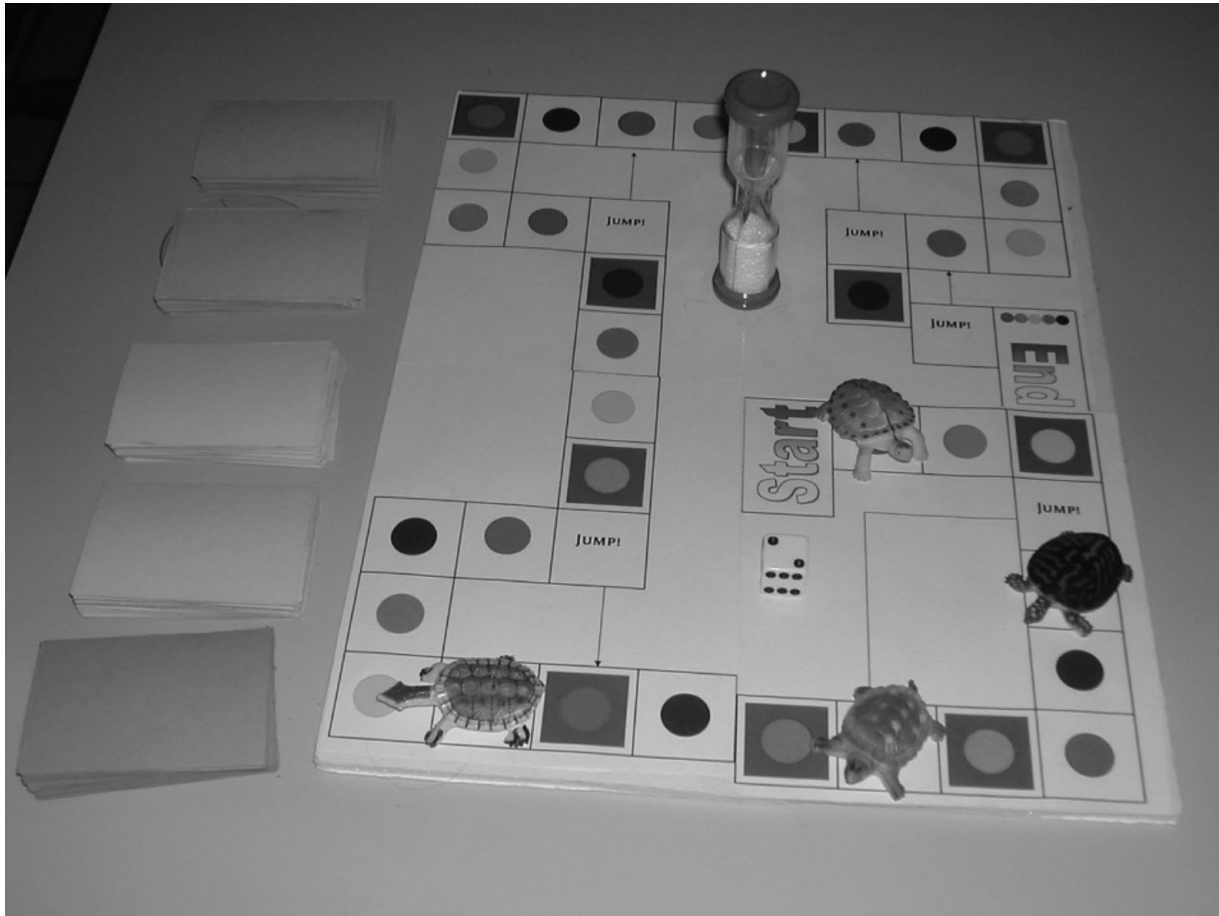
I was also concerned about student ability. Having the best students play against the worst could be frustrating for both. I decided to track victories and assign opponents based on past performance. That way, the more competitive students would have more competitive opponents, while the weaker students would likewise play against more compatible teams. I found that students on the fourth (weakest) board tended to help each other more, even helping those on other teams.

Game play would progress throughout an entire class period. At the end of the period, I collected the scores. To help ensure attention, I made game play a small part of the course grade. I was a little concerned about using a competitive activity as a grade component, but no student complained about it.

Research is divided on the use of competition in the classroom. Lundy (1991) found that competition can have a negative effect on cognitive learning, but others disagree. Berry (2003) asserted that “competition can increase motivation, produce lots of energy, and offer a welcome change of pace” (para. 3) assuming it is not overdone. Rodger (2007) found that women do better in cooperative versus competitive conditions, while Sabato (1989) recommended cooperative learning groups in a competitive setting. In essence, that is what I used: cooperative teams competing with each other.

## Using the Game in Class

As every instructor knows, there is seldom enough time in the term or semester to cover everything. Because of this, I was concerned about devoting too much time to game play – since students rarely looked at future topics, the game did



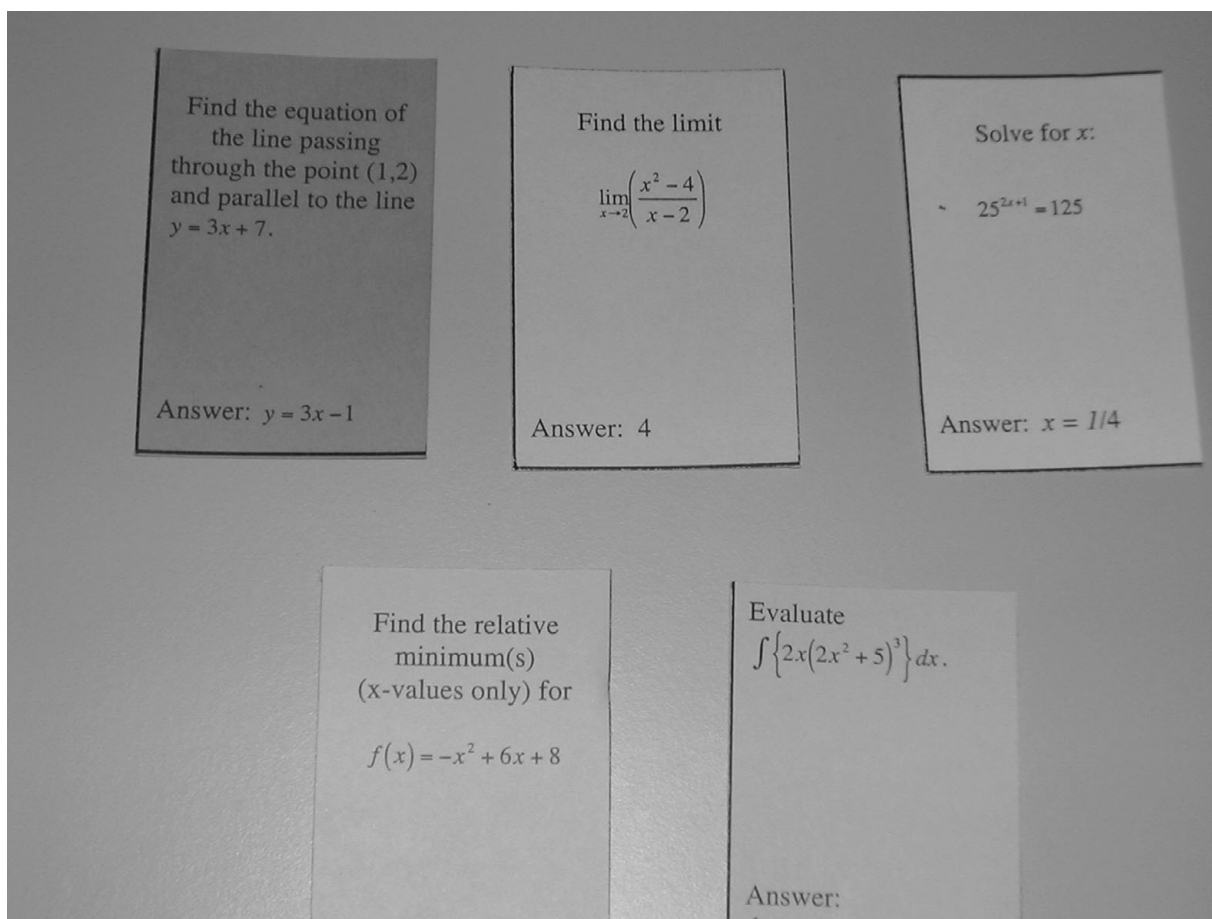
**Figure 8.1.** The game board with timer, die, team markers, and question cards

not present new material at all. On the other hand, if we only played a few times, students would view it as a novelty, not a teaching tool. As [Corbell \(1999\)](#) put it, “if an instrument is to be efficient, it must be given a full opportunity, and therefore enough time, to show its possibilities” (p. 171). In the end I settled for eight days of game playing, starting shortly after the first exam. This might not be long enough to show a discernible effect, but I hoped that it would. In the spring of 2004, however, my goal was to see if the game could be successfully incorporated into a course at all.

It was difficult watching students ask questions without interfering, but I eventually got used to it. The students loved the game; their responses on an end of semester survey were overwhelmingly positive. Out of 30 respondents, 18 recommended “very much” that I use the game in future classes (the other choices were “some” and “not at all” with 11 and one respondents, respectively). As one student put it, “it added a different element to the class – made it more interesting – [and I] got to work with different people.” Although the responses were given anonymously, the one negative comment that I received came from a student who spoke to me at length about the game. After talking with her, though, it became clear that her dislike of the game stemmed from a general feeling of exclusion by her fellow classmates (which had nothing to do with the game) rather than a dislike for the game. Overall, the verdict was positive: students enjoyed the game. As illustrated by the student comment in the preceding paragraph, game play was perceived as a social and fun activity. My next question was whether this translated into better learning.

[Corbell \(1999\)](#) pointed out that there is a “deep-seated prejudice that . . . games are not serious” (p. 163). Play is too often seen as childish, as opposed to the mature concept of learning. Yet play is a necessary step in the understanding of new material ([Holton, 2001](#)). When used successfully, games “generate enthusiasm” ([Ernest, 1986](#)), causing students to “become strongly motivated . . . [and] immerse themselves in the activity” (p. 2). The fun that students feel can “provide a responsive environment” ([Cruikshank, 1980](#)) for learning because games are “psychologically engaging” (p. 77).





**Figure 8.2.** Example questions used in applied calculus course

They give students real and immediate feedback about their decisions. Hatch (2005) argued that this encourages students to perform at a higher level than they otherwise would, gaining more practice than an ordinary textbook assignment would give.

Does this motivation and extra energy translate to better learning? The results are murky. Gremmen and Potters (1997) discovered that the efficacy of gaming is more often supported by subjective indications than objective ones. Freitas (2006) noted that “the attainment of specified learning outcomes is largely unproven through the use of games” (p. 346). In mathematics, however, the results are better. Randel, Morris, Wetzel, and Whitehill (1992, p. 261) stated that “math is the subject area with the greatest percentage of results favoring games” among those they studied. This may be due to the nature of mathematics. In mathematics, just as in game play, participants must create conjectures and devise strategies for solving the problem at hand (Holton, Ahmed, Williams, & Hill, 2001). Ascher (2001) likewise noted that “playing games of strategy and solving mathematics problems share several important features. . . . The challenge [in both cases] is to select an action . . . [at each stage] of a chain of steps. . . . By playing games of strategy, students can begin to develop . . . modes of thought” (p. 96) useful in mathematical work. Thus, playing games may actually improve ability in mathematics.

Unfortunately, most of this research tended to concentrate on short-term games, used to illustrate one topic rather than a whole course. What happens when a game is used throughout the semester, like the one I used in my course?

In the fall of 2004, I taught two demographically similar sections of applied calculus (for business or life sciences majors). I randomly chose one section to play the game (creating new question cards appropriate for the course – see Figure 8.2) while the other section would not play, thus it would serve as a control for my experiment. I thought that this would be an ideal way to examine the game and see if the benefits of using the game outweighed the drawbacks of using it for eight class periods (out of the 50 available). I planned to teach the sections as similarly as possible and then

Section	Exam 1	Exam 2	Exam 3	Exam 4	Exam 5
ONE	1 day	2 days	–	–	2 days
TWO	1 day	–	2 days	2 days	–

**Table 8.1.** Schedule of game play by section

compare the scores from the final exams. If one section had significantly higher scores, that would be an indication that it benefitted from the game (or lack thereof).

Using student finals as an assessment method, I found no significant difference between the sections (Herman, 2006). Student evaluations, however, told a different story. Students from the game playing section rated both the course and the instructor more favorably. Other authors (Azriel, 2005; Sulzman, 2004) attempted similar experiments to determine the effects of games on their classroom and had similar results. Exam scores between a game playing section and a control group showed no significant difference, yet student response to the games was significantly positive. Although this does not demonstrate that classroom games improve learning, it suggests that the time spent on games does not harm student learning and may improve student enjoyment of the course.

After my experiment ended, I realized that there were several ways that its design could have been improved. Final exam scores were not necessarily the best assessment tool. Because college students typically take many finals within a short period of time (the week of finals), final exam scores can be more a reflection of student schedules than of how much they have learned in a class. Additionally, the length of time between playing the game and the final exam was very long. While some authors have reported lasting effects from games (Gremmen & Potters, 1997), it is likely that shorter-term effects would be more dramatic. A third difficulty with my experiment was the existence of two lurking variables, student ability and section dynamics, each of which can create a large amount of variance. For example, a few really good students can skew the results. I resolved to modify my experimental design to address these issues.

## Multivariable Calculus

My next opportunity to compare two sections of a course came in the fall of 2005 in multivariable calculus, typically taken by mathematics and natural science majors. The course covers five chapters of material – parametric and polar curves, vectors, vector functions, partial derivatives, and integration – making the choice of five different topics for the game questions easy to implement. To avoid relying on the final exam to assess the game’s success, I decided to use chapter exams as my primary assessment tool. To avoid the lurking variables of section dynamics and student ability, I decided to alternate game play between sections. One section used the game for certain chapters while the other section did it for others. (This type of study is called a quasi-experimental or switching replications design.) To familiarize students with the game, each section played it once while we covered the first chapter. During subsequent chapters, a section either played the game on two days (separated by at least a week) or did a worksheet covering similar material. The schedule was as shown in Table 8.1.

I was uncertain whether the more mathematically mature students in this course would enjoy the game as much as other students or whether they would consider it childish. As the semester progressed, I noticed some differences. The multivariable students were more comfortable modifying the rules to suit their needs. For example, several groups changed the game so that they used only questions from the current chapter. Others removed the competitive element, using the game as a group review for all players. One student would pick up a card and read it, and then the others would work out the problem, regardless of which team they were on.

I decided not to interfere, letting the students use the game as they saw best, because games work best when “the participant is . . . unrestricted in his or her approach to learning” (Corbell, 1999, p. 175). As Lundy (1991) noted, different students take different approaches to games – some are competitive, while others treat them merely as opportunities to practice skills. Lundy also observed that some students treat a game as a break from real learning and thus gain little from it. There were fewer of these students in multivariable calculus than in my other courses. Being more mature, the students in multivariable calculus were more willing to take an active role in their learning. This agrees with Lundy’s assessment that mature students “take . . . games more seriously and approach . . . [them] as learning devices” (p. 184) rather than as competitions.

## Statistical Model

To analyze the results of my experiment, I used repeated measures analysis of variance (ANOVA), using student, exam number, and whether or not the game was played as factors for the model. Essentially, the repeated measures model took the data points and used them to compute average scores for each student (over all exams), for each of the exams (over all students), and an average difference in exam scores for using the game versus not using it. The model then calculated the amount of natural variation inherent in each of the listed categories. By comparing this variation with the difference between averages, the model determined how likely it was that differences between averages could simply be due to natural variation rather than being a significant effect.

Initially I thought the best model would be multiplicative:

$$SCORE_{ijk} = (STUDENT_i)(EXAM_j)(GAME?_k)(1 + Error_{ijk}),$$

but the error terms (the parts used to determine variation within each category) did not turn out to be normally distributed as required by the model. In the end, the best model was an additive one:

$$SCORE_{ijk} = STUDENT_i + EXAM_j + GAME?_k + Error_{ijk}$$

(where  $GAME?_k$  = some value  $G$  if the section used the game, or = 0 if it did not). Because each section used the game for the first exam, I decided to drop those scores from final analysis (including them changed the results very little). The real question, then, was what was the value of  $G$ , and was it significantly greater than zero? If so, this would show that the students' exam scores were positively affected by usage of the game.

The results of the ANOVA (Analysis of Variance) were as follows:  $t = 1.03$ , with a  $p$ -value of 0.153, and  $G = +1.33$ . One or two extremely poor exam scores by some students were partially to blame for the high  $p$ -value; removing test scores below 50% yielded a  $p$ -value of 0.135 instead. The  $p$ -value shows how likely it is that differences are due to random chance rather than a real effect. While 0.135 is fairly small, tradition uses a cut-off of 0.05 or possibly 0.10. As in my previous experiment, therefore, there was no statistically significant difference between the exam scores of those who played the game and those who did not.

Because of the small number of students involved in the study (fewer than 25 per section), however, natural variability in performance makes the traditional cut-offs difficult to obtain, especially when many other factors contribute to the exam scores (such as study time, sleep, and work load from other courses). My conclusion, therefore, is that the game may have improved student scores, on average by 1.33 points out of 100. Student evaluations painted a clearer picture. At the end of the semester, I asked eight questions about the game and its use; six were rated significantly different from neutral (with  $p$ -values ranging from 0.000 to 0.045). The students believed the game helped them meet other students in the class, they enjoyed playing it, and they recommended using it in future classes. While the game helped them review old material, the students did not think that it helped them prepare for future topics. They were uncertain whether it should be played more or less often or whether it should affect grades. Individual comments support these ratings:

- [It was a] good way of reviewing in a fun way
- The material that hasn't been covered should not be played . . . we don't know how to do the problems . . .
- It would be more helpful to play the game chapter by chapter . . . not every chapter every time we play . . .
- The game's really cool.

It is interesting that the students perceived a benefit from the game even though their scores were not dramatically improved. Three reasons come to mind. The first is related to the Hawthorne effect: that the game is new to them (and fun) and therefore more memorable than other methods. Because they remember the game more than other instructional methods, students are more likely to attribute learning to it. Unfortunately, this misattribution appears to be fairly common (Gremmen & Potters, 1997), making student self-reporting of learning effectiveness a dubious measure. A second possible reason for the perceived benefit is that (as a whole) the students enjoyed the course, disposing them to rate individual aspects of it (such as the game) favorably.

On the other hand, it is possible that the students did, in fact, benefit from the game, more strongly than the exam scores would indicate. The game, whether due to its novelty or to its social nature (Ernest, 1986) or some other aspect,

may have encouraged or empowered the students enough to have an effect throughout their course. For example, by helping students meet each other, game play may have facilitated the development of study groups for the course. It may also have encouraged a more participatory atmosphere in class, or the quiz-answer methodology of the game may have influenced student study methods positively. Certainly, game play can build positive attitudes (Bragg, 2006) and encourage mathematical discussion (Oldfield, 1991). If one of these benefits were present, the effect would likely be seen throughout all chapters, not merely in the ones in which game play took place. To study these possibilities, a different experimental design would be necessary, as well as additional data (for example, carefully worded surveys or records of class participation rates).

## Conclusions and Reflections on the Process

Although I had concerns that more mature students would reject game play as silly, the classes as a whole accepted the game. The students enjoyed the game enough to recommend its use in future classes. Game use, I think, helped bring the students together and fostered a more communal learning environment. I believe that students in game playing sections participated more in class discussions and seemed less afraid of asking questions in class than students in similar courses.

Student feedback was, overall, quite positive, but not every comment was favorable. One student from my Mathematics for the Social and Management Sciences section wrote that “Some days it felt like a good review, while other days the other groups didn’t care.” This probably applies to every teaching method, but it could have been revealing to investigate this aspect more thoroughly, through the use of case studies or more detailed recording of how students played the game in class. In retrospect, I wish I had gathered more detailed data.

Like many mathematicians, the methodology used in my early SoTL investigations relied heavily on numerical statistics. A  $p$ -value, after all, is not subjective – it is a definitive measure of the correctness of a hypothesis. Unfortunately, as McKinney (2007) pointed out, due to practical and ethical considerations we rarely have the luxury of designing a perfect experiment when we deal with the classroom. Subjects are not randomly chosen, many lurking variables are overlooked, and small sample sizes make statistically significant results hard to obtain.

Even ignoring these considerations, there is a wealth of data beyond simple statistics available to classroom researchers. You are more than the designer of the experiment – you are present to witness it unfold. Sometime unexpected benefits (or drawbacks) emerge. In my case, I noticed that students seemed to participate more in class because of the game, something I had not anticipated. Perhaps this was because game play encouraged the students to make stronger social connections with each other, as suggested by student comments and my observations during game play. If I had gathered a broader range of data, perhaps I would have been able to address these new ideas more directly. My advice is this: When designing a classroom study, plan to gather more data than you think you will need or use. It might become useful later.

Beyond data-gathering issues, finding background literature and publishing results can be a challenge. Unlike most mathematical research, the Scholarship of Teaching and Learning cuts across disciplinary boundaries, even appearing in non-educational professional journals. Because of this, it can be tricky finding the right place to publish results. Some of my results, for example, appear in a mathematics journal (Herman, 2006), while others have been submitted to a professional gaming journal.

Challenges aside, SoTL research can be extremely rewarding. As mentioned at the start of this chapter, the process of doing SoTL research can be even more important than the results obtained. Framing and researching a question and designing a project encourages the researcher to experiment within the classroom, much as a painter experiments with styles on a canvas. As the project progresses, the question (or questions) become more refined, more interesting, and the answers can both inform and improve the way you teach.

For me, examining how games affected learning helped me become more aware of what causes students to participate in class and connect with their peers, and my teaching style changed because of it. I now place a greater emphasis on active group learning methods, giving students a bigger role in creating a learning environment. SoTL also gave me the confidence to try other ideas. Since this project ended, I have explored the effectiveness of online quizzes and discussion groups, and the use of written papers to assess critical thinking in statistics courses. Prior to doing SoTL research, I would never have considered incorporating many of these methods into my teaching, nor would I have had

the tools necessary to decide if they were beneficial. In significant ways, SoTL research has made me more confident and a better teacher.

## References

- Ascher, M. (2001). Learning with games of strategy from Mongolia. *Teaching Children Mathematics*, 8(2), 96–99.
- Azriel, J. A., Erthal, M. J., & Starr, E. (2005). Answers, questions, and deceptions: what is the role of games in business education? *Journal of Education for Business*, 81(1), 9–13.
- Berry, R. L. (2003). Creating cooperative classrooms. *Education Digest*, 69(2), 39–42.
- Bragg, L. A. (2006). Hey, I'm learning this. *Australian Primary Mathematics Classroom*, 11(4), 4–7.
- Corbell, P. (1999). Learning from the children: Practical and theoretical reflections on playing and learning. *Simulation & Gaming*, 30(2), 163–180.
- Cruikshank, D. R., & Telfer, R. (1980, Winter). Classroom games and simulations. *Theory into Practice*, 19(1), 75–80.
- Dunn, P. (1989). *Intermediate math trivial pursuit*. Columbus, OH: Good Apple, Inc.
- Ernest, P. (1986). Games: A rationale for their use in the teaching of mathematics in school. *Mathematics in School*, 15(1), 2–5.
- Freitas, S. I. (2006). Using games and simulations for supporting learning. *Learning, Media and Technology*, 31(4), 343–358.
- Gremmen, H., & Potters, J. (1997). Assessing the efficacy of gaming in economic education. *Journal of Economic Education*, 28(4), 291–303.
- Hatch, G. (2005, September). Using games in the classroom. *NRICH: enriching mathematics*. Retrieved from [nrich.maths.org/2928](http://nrich.maths.org/2928)
- Herman, E. (2006, December). Playing games to learn college mathematics. *Connecticut Mathematics Journal*, 1–8.
- Hertel, J. P., & Millis, B. (2002). *Using simulations to promote learning in higher education: An introduction*. Sterling, VA: Stylus Publishing.
- Holton, D., Ahmed, A., Williams, H., & Hill, C. (2001). On the importance of mathematical play. *International Journal of Mathematical Education in Science and Technology*, 32(3), 401–415.
- Klein, J. D., & Freitag, E. (1991). Effects of using an instructional game on motivation and performance. *Journal of Educational Research*, 84(5), 303–308.
- Kraus, W. H. (1982). The use of problem-solving heuristics in the playing of games involving mathematics. *Journal for Research in Mathematics Education*, 13(3), 172–182.
- Lundy, J. (1991). Cognitive learning from games: student approaches to business games. *Studies in Higher Education*, 16(2), 179–188.
- McKinney, K. (2007). *Enhancing learning through the scholarship of teaching and learning*. Bolton, MA: Anker Publishing Company.
- Oldfield, B. J. (1991). Games in the learning of mathematics part 2; Games to stimulate mathematical discussion. *Mathematics in School*, 20(2), 7–9.
- Randel, J. M., Morris, B. A., Wetzel, C. D., & Whitehill, B. V. (1992). The effectiveness of games for educational purposes: A review of recent research. *Simulation & Gaming: An International Journal*, 23(3), 261–276.
- Rodger, S., Murray, H. G., & Cummings, A. L. (2007). Gender differences in cooperative learning with university students. *Alberta Journal of Educational Research*, 53(2), 157–173.

- Sabato, G. (1989). Cooperation and [competition] unleash creative potential. *Social Studies Review*, 28(3), 103–109.
- Schoen, L. M. (1996). Mnemopoly: Board games and mnemonics. *Teaching of Psychology*, 23(1), 30–32.
- Sulzman, E. W. (2004). Games in an introductory soil science course: A novel approach for increasing student involvement with course material. *Journal of Natural Resources and Life Sciences Education*, 33, 98–101.
- Zirkel, G. (1975). Bingo! or how I hit the jackpot for my statistics class. *MATYC Journal*, 9(2), 34–37.



# 9

## Investigating How Students' Linking Historical Events to Developments in Mathematics Changed Engagement in a History of Mathematics Course

Pam Crawford  
*Jacksonville University*

### Editors' Commentary

The author of this chapter, Pam Crawford, holds a doctorate in mathematics with a concentration in teaching collegiate mathematics. The dissertation she wrote gave her experience in conducting research in undergraduate mathematics education (RUME). More recently, she participated in her university's SoTL scholars program, undertaking an investigation prompted by frustrations encountered repeatedly when teaching a history of mathematics course. The mathematics majors in the course were reluctant to engage in historical thinking. She tried an intervention and describes how her study of its effect is an example of SoTL (and not RUME) work, thereby illuminating some of the distinctions between the two.

### Introduction

How often, when thinking about a course you teach, have you wondered what would be the outcome on student learning if you changed some aspect? After making a change, how do you investigate whether or not it improved student understanding and performance? What follows is my tale of researching a change I made in my teaching in a History of Mathematics course.

I hold a Ph.D. in mathematics with a concentration in teaching collegiate mathematics from Western Michigan University. The degree gives me a background in Research in Undergraduate Mathematics Education (RUME), in which researchers apply standard educational research techniques to investigate “the nature of mathematical thinking, teaching, and learning” and try “to use such understandings to improve mathematics instruction” at the collegiate level (Schoenfeld, 2000, p. 641). RUME is an acknowledged research discipline in mathematics with established research procedures and terminology. Scholarship of Teaching and Learning (SoTL), as defined in Chapter 1, is “the intellectual work that faculty members do when they use their disciplinary knowledge . . . to investigate a question about their students' learning (and their own teaching), submit their findings to peer review, and make them public for others in



the academy to build upon.” Clearly SoTL and RUME are related by their concern with investigating student learning. My mathematics education training and my experience with SoTL allow me to comment on their similarities and differences. I begin my chapter by mentioning some of the characteristics of each, so that in the discussion of my SoTL project, I can highlight several distinctions between the two endeavors.

## Similarities and Differences between SoTL and RUME

In Chapter 1, Jacqueline Dewar and Curtis Bennett depict a common space shared by SoTL investigators, RUME researchers, and others interested in the teaching and learning of mathematics by using an equilateral triangle with vertices labeled “Teaching Tips,” “SoTL,” and “RUME.” They explain that a Teaching Tip is a description of a teaching method that an instructor deems “successful” and that students “liked.” They say that when an instructor begins to systematically gather evidence from students that indicates how the method affected their learning, the instructor is moving toward SoTL. The RUME vertex represents the field of mathematics education research where research methodologies, theoretical frameworks, empirical studies, and reproducible results assume greater importance. These aspects influence the questions asked and the methods used to investigate them. SoTL questions typically arise from classroom practice and may connect to nonmathematical topics such as service learning, civic engagement, or student voice in the classroom. In contrast, RUME is primarily concerned with the theory of how people learn mathematics.

Both SoTL and RUME studies follow the same general outline: develop research question(s), conduct a literature review, perform a qualitative or quantitative study (or both), and analyze the data to reach a conclusion. RUME studies use social science research methodology and RUME researchers often have had doctoral-level training in mathematics education, giving them a formal background in theories of mathematical learning.

On the other hand, SoTL studies tend to be the result of disciplinary faculty “going meta” (Hutchings & Shulman, 1999, p. 13) to form and systematically examine questions connected to student learning. Their research questions consider such aspects of student learning as: When do you know student learning is occurring? What conditions favor student learning? How can you deepen student learning? Often a SoTL investigation begins with an instructor’s desire to improve some aspect of teaching and, as scholarship, should end with review by peers and dissemination of the results. Along the way, SoTL researchers seek to triangulate their evidence in support of their conclusions.

My own background and experience can serve to illustrate some of these differences. The introduction to my RUME dissertation (Crawford, 1998), “Fostering Reflective Thinking in First-Semester Calculus Students,” included a section explaining my theoretical framework of constructivism. I wrote about how constructivism has emerged as a research paradigm in mathematics education and how constructivists view understanding as the building of mental frameworks from already existing pieces. That is, previously built frameworks become the content in subsequent constructions (Ernest, 1996; Schoenfeld, 1992; Weaver, 1987). I noted the emphasis in Flavell’s (1976) statement concerning “perhaps *the* central emphasis in learning and development, namely, how and under what conditions the individual assembles, coordinates, or integrates his already existing knowledge and skills into new functional organizations” (p. 231). I concluded the segment concerning constructivism with

An individual’s knowledge is transformed as new ways are created to observe and organize experiences. New mental representations are constructed through different but related mental processes such as encoding, combining, or comparing concepts (Davidson, Deuser & Sternberg, 1994). An important educational goal should be to help students and instructors understand this constantly evolving constructed nature of knowledge (Novak, 1996, p. 7).

The discussion of constructivism was necessary to inform my reader of the theoretical perspective I held while designing my study, conducting my research, and writing about my results.

On the other hand, the nature of knowledge was not a factor in my SoTL History of Mathematics study. I was not operating from a theoretical framework though I was conducting a scholarly study of students’ reactions to changes in teaching methodology. In a RUME study, one determines a theoretical framework and attempts to control (or describes in great detail) all sorts of factors so that other researchers could replicate the study. In my SoTL study, I made careful observations and collected data so that I could analyze my students’ reactions; I did so without emphasizing a theoretical framework or being concerned about replicating the study.

In a RUME study, the theoretical perspective influences the choice of background literature. The RUME researcher desires to situate the study in the current literature and, through the study, extend that body of knowledge. In my SoTL study, I used my review of literature to explain some of the traits I was observing in my students prior to altering my teaching technique and what traits I wanted to develop in my students.

A RUME study frequently begins with a small pilot study to help a researcher work out minor difficulties in materials such as assignment directions, wording of surveys, examination problems, and the like. A pilot study can also help a RUME researcher hone in on the essence of the research question, which is needed to produce a replicable study. Pilot studies may or may not be part of a SoTL study. Prior to teaching the undergraduate course that was the focus of my SoTL study, I was fortunate to be able to do a pilot study in a graduate-level history of mathematics seminar.

SoTL and RUME studies both deal with human subjects. Institutions receiving federal research funding are required by the federal government to have research on human subjects reviewed and approved by an Institutional Review Board (IRB). If the researcher wants to disseminate the results of a study involving students, whether SoTL or RUME, approval by the institution's IRB should be obtained before any data is collected. My background in RUME included training in IRB documentation. IRB requirements can vary greatly from institution to institution, so investigators contemplating their first SoTL study should contact their local board. It has been my experience at a large research-oriented institution and at a small liberal arts institution that IRB Chairs can be very helpful when going through the process.

## Background

Jacksonville University (JU) is a small liberal arts university of about 2,500 undergraduate and 500 graduate students, located in Jacksonville, Florida. We have approximately 25 mathematics majors and 15 mathematics minors in the department, almost none of whom are involved in teacher education. Our upper-level mathematics classes are small with usually eight to 15 students. Our graduates enter careers ranging from military service to employment by local banking or insurance agencies or they go on to earn advanced degrees in such areas as mathematics, finance, or medicine.

As an undergraduate, I double majored in mathematics and history, making history of mathematics one of my academic interests. Our departmental goals for our speaking-intensive history of mathematics course include that students should gain an appreciation for the history of mathematics and are able to

- describe significant theorems and list achievements of famous mathematicians from major mathematical periods,
- understand the role of women in mathematical history, and
- demonstrate communication skills, through oral and written presentations.

I find the history of mathematics fascinating. Yet, each time I taught this course, although for the most part my students achieved the goals of the course, they did not engage in the study of history.

In 2004, I was selected to participate in the first cohort of JU Scholars supported by our Center for Teaching and Learning. My project involved investigating student learning in the upper-level History of Mathematics course required of our junior and senior mathematics majors. Other upper-level mathematics requirements include courses such as real analysis, complex analysis, geometry, and abstract algebra, which focus on mathematical topics. Students are more apprehensive about the history of mathematics course than the other upper-level mathematics courses.

Much of the apprehension is due to the nature of the course, which is very different from other mathematics courses. Although most mathematics courses involve class discussion of mathematics problems and of the mathematical techniques necessary to solve them, traditionally, upper-level mathematics courses are not thought of as participation courses. To do well in them, students need the ability to understand concepts, manipulate symbols, internalize formulas, construct mathematical proofs, and use technology in problem solving.

Those are not necessarily the abilities that best serve my students in the history of mathematics course, particularly since this course also satisfies the speaking intensive graduation requirement set by our university. Mathematical ideas do not occur in a vacuum. In this course, we discuss the development of mathematical techniques and also mathematical philosophical questions such as What influence did the restrictions imposed by the Catholic Church in seventeenth century Italy have on the subsequent shift of mathematical activity to France, Germany, and Great Britain? Classroom

discussion centers on important episodes, problems, and discoveries in mathematics, with emphasis on the historical and social contexts in which they occurred. As [Hetrick and Saxe \(2004\)](#) wrote

Mathematics is alive, dynamic, and still evolving. Its development influences and is influenced by history, culture, and science. Depending upon the prevailing spirit of the times, various factors either encouraged or discouraged the development of new mathematical “truths.” (p. 9)

But, students tend to think of mathematics as static, something that is unchanging and has always been ([Hetrick & Saxe, 2004](#)). When they confront the history of mathematics, the need for integrative learning makes students uneasy. They are asked to investigate connections between developments in history, culture, and science and those in mathematics, that is, to see the big picture. To do this, students are expected to follow themes and areas in mathematics as they develop through time, to understand applied problems that often drove the development of mathematics, and to study solutions to problems using ancient techniques, along with the limitations of those techniques. Requiring them to write a term paper (in a mathematics class!) and make an oral presentation of their paper using technology only adds to students' trepidation. As students like to say, “I'm a mathematics major. If I wanted to write, I would have majored in English!”

I expect students to read two textbooks in both my graduate and undergraduate history of mathematics classes – *Math through the ages* ([Berlinghoff & Gouvêa, 2004](#)) and *Journey through genius* ([Dunham, 1991](#)). Just as art history students are expected to be knowledgeable about great works of art and the lives of the great artists, in his book Dunham relates not only his choices for the 12 great theorems of mathematics but also the lives of their creators, and the theorems' impact on mathematics. This book was written for a nonmathematical audience and should be enjoyed by my undergraduate and graduate students, yet they resist reading the book.

Their resistance might stem from their difficulty in thinking like a historian. Surprisingly, my students fail to be curious about the development of mathematical ideas. They are bothered that there is often no clear resolution to discussion questions such as: Why do we use  $m$  for slope? If the history of mathematics were rerun, would ideas such as prime numbers reappear? Was mathematics discovered or invented (that is, does mathematics exist on its own for mathematicians to discover, in the same way as physicists discover the laws of physics, or do mathematicians invent mathematics, much like a writer invents a novel)? The notion that questions can lack definitive answers is foreign to the thinking skills my students developed as mathematics majors.

## The SoTL Study

As a result, I decided to use a different teaching technique. In my study, I investigated changes in student learning in my spring 2005 undergraduate history of mathematics course when I used more guided discovery, a method in which the instructor acts as a facilitator to guide students' learning, and less lecturing. This approach was modeled on one that a colleague, Marilyn Repsher, used successfully in her graduate-level seminar in the history of mathematics. She required each student to write short papers (two to three pages) based on the investigations of the same six historical events with regard to the events' historical links to the development of mathematics. My students had access to ample resources in the JU library and on the many history of mathematics websites on the Internet. Students were expected to share their findings in class.

The events I asked my students to investigate were

1. The building of the Great Pyramid at Gizeh around 2600 BC,
2. Plato founding the Academy in 387 BC,
3. Syracuse falling to the Romans in 212 BC,
4. Toledo falling to the Christians in 1085 AD,
5. Hitler invading Poland in 1939 AD, and
6. The Soviet Union launching Sputnik in 1957 AD.

Although these events may not appear to have had any impact on mathematics, rather surprisingly they did. For instance, the building requirements of the ancient Egyptians led to the development and use of practical geometry.

Plato made no important mathematical discoveries but his belief that mathematics provides the finest training for the mind was influential in the development of the subject. A Roman soldier killed Archimedes, one of history's greatest mathematicians, during the fall of Syracuse. In more recent history, the fall of Toledo reintroduced European mathematicians to the great works of Greek mathematicians as these manuscripts had long been lost to most Europeans but survived in Arabic translations. When Hitler's invasion of Poland was imminent, Polish cryptographers shared their work on decoding German Enigma machine-encrypted messages with British and French code breakers, which led to Alan Turing's development of the Bombe, an electro-mechanical decrypting machine. The launch of Sputnik and competition with the Soviet Union spurred the growth in mathematical and scientific achievements in the last half of the 20th century.

## Framing the Question

For my study, the research question was "Does the guided discovery teaching technique of having students investigate links between historical events and developments in mathematics improve students' engagement with historical material when compared with my previous teaching methods?"

## Gathering the Data

My measures for determining the success of the guided discovery method came from several sources. Most prominent was whether class discussions improved, both from my viewpoint and from that of the students. Pre- and post-course surveys completed by the students provided information on changes in students' views on and appreciation of mathematical history. The pre-course survey asked:

1. Why should mathematics majors be required to study the history of mathematics?
2. What types of specific things do you expect to learn in this course regarding the history of mathematics?
3. Do you anticipate that this course will be beneficial to you in your career? Why or why not?

Although our upper-level mathematics courses usually enroll eight to 15 students, my history of mathematics course during the semester of the study was quite small, having only four students. Three of them completed a post-course survey at the beginning of the next semester that asked:

1. Of all the things you learned in History of Mathematics, what were the top three?
2. History of Mathematics is a required course for the mathematics major. Do you agree or disagree with this decision?
3. How has learning about the History of Mathematics affected your view of mathematics?
4. How has learning about the History of Mathematics affected the way you view your current mathematics courses?
5. Name three important current events. For each of these events, describe how you think that event will/will not influence the development of mathematics.

By reusing some essay examination questions from prior years, I was able to obtain another type of evidence. I could compare the guided discovery students' levels of performance on midterm and final examinations, particularly on the essay questions. In addition, I included the following question on the final examination:

Explain how completing this course [History of Mathematics] has changed/not changed the way you view:

1. mathematics,
2. the development of mathematics, and
3. the people who participated in the historical development of mathematics.

## Findings

Student responses to the beginning-of-course anonymous surveys were predictable. Students stated that they expected to learn about the origins of various items in mathematics, which they thought was a good thing. Also, they expected to learn names and dates that were important in the development of mathematics, and they did not necessarily think that the course would be beneficial to them in their careers.

These responses contrast with those to the post-course surveys. To the question “Of all the things you learned in History of Mathematics, what were the top three?” one student responded, “The general relationship between the understanding of math and the success of a society” (personal communication, fall 2005). This same student wrote, in response to the survey question concerning the History of Mathematics as a required course for mathematics majors, “It makes you think differently” (personal communication, fall 2005). All the students surveyed believed History of Mathematics should be required for mathematics majors, particularly since the course helped them appreciate problems and techniques studied in other mathematics courses. As one student stated, “Mathematics isn’t just numbers anymore” (personal communication, fall 2005).

In response to the fifth question on the post-course survey, about current events and their impact on the development of mathematics, all students mentioned the war in Iraq as one that would influence the development of mathematics. One student wrote that the war in Iraq would focus resources on developing “mathematics and technology to improve military intelligence allowing for safer resolutions to conflicts” (personal communication, fall 2005). Another cited the location of the conflict as an influence since the war had destroyed many mathematical artifacts housed in Baghdad, such as cuneiform tablets and ancient relics from the House of Wisdom (a Baghdad site dating to 833 AD).

All the students mentioned the recent (in 2005) hurricanes in Florida as another current event that would influence the development of mathematics. They felt that there would be increased interest by scientists to use mathematics, particularly statistics, to improve forecasting and predict hurricane behavior. (Students completed the post-course survey during the fall 2005 semester, that is, during the 2005 hurricane season, which ran from June 1 to November 30 – a range of dates known to almost every Floridian.)

The post-course survey yielded a wide variety in the students’ responses. No two students connected the historical events with mathematics in the same way, which led to interesting and varied class discussions. The in-class discussion that occurred after students wrote short essays on the effect of Hitler invading Poland in 1939 provides more evidence of this variety. Some students focused on the development of the Enigma machine and cryptography by Polish scientists, which then was shared with the Allies. Another student referred to the rise of Hitler, which led to the start of World War II after Hitler invaded Poland in 1939, and how the rise of Hitler resulted in the scattering of mathematicians, particularly Jewish mathematicians, from the universities in Europe to universities in England and the United States, which strengthened the mathematics departments in the United States. Individually and collectively, students were engaging in historical thinking and making connections on their own.

In the student evaluation comments from this undergraduate History of Mathematics class, students wrote that they really enjoyed the class, particularly the class discussions, and that the class “was a nice change of pace from a typical math course” (personal communication, spring 2005). On final examination questions that were repeated from previous years students performed equally well. When asked to explain how completing the course changed, or did not change, the way they viewed mathematics and the development of mathematics, one student wrote “This class inevitably changes perceptions in mathematics. At least that is the way I see it. . . . I consider myself much better rounded in the topic after so much open discussion about the development” (personal communication, spring 2005). Another student responded to the same item by writing,

This course has drastically changed the way I view the development of mathematics. It has always been one of those things where you wonder where it all came from but never really had the drive to really investigate it. I know that’s how it was for me anyway. This class let me see the steps taken to arrive at modern thinking and thought process. It provided me with more respect for the development process, as well as giving me a lot to think about when I’m forced into my thoughts during long car rides (personal communication, spring 2005).

My students became curious about the development of mathematical ideas, as evidenced by several of the statements above. Using a nonmathematical historical event as a seed for students’ investigations into mathematics clearly engaged my history of mathematics students fully in the historical material. As one student commented on his final examination,

“I have told many younger mathematics majors about this class. . . . If you have some sort of love for [mathematics], you certainly should be interested in its history” (personal communication, spring 2005).

## Further Study

A number of improvements could be made to this study. The enrollment in my spring 2005 history of mathematics course was particularly small. Although the data collected for my project came from multiple sources, with only four students, I did not have convincing evidence from which to draw conclusions. I would like to repeat the study with a larger number of students. In a future study I would administer the post-course survey at the end of the semester instead of the beginning of the next semester. Doing that should result in a better response rate. My belief that class discussions improved as a result of the guided discovery teaching technique was based on my perceptions of previous history of mathematics classes’ discussions as compared to the class under study. My data would be strengthened by finding a method to measure the quality and quantity of student discussions.

A question I would like to investigate further is “What is it about their association with these historical events that increases students’ understanding of and appreciation for the relevant mathematics?” Also, I am interested in studying how students’ curiosity in mathematics and their ability to tie together their other mathematics courses can be stimulated through instructors’ inclusion of similar nonmathematical historical events and discussions into course material in other mathematics courses.

I have continued investigating student learning when I used guided discovery in subsequent undergraduate and graduate history of mathematics classes. I have experimented with other questions in my investigation such as How did the defeat of Al Gore in the 2000 presidential election affect the development of mathematics? – a question particularly apropos to a history of mathematics course at a Florida university. I try to expand my list of events to present each year’s students with fresh material, allowing them to make new (to them, and sometimes to me) associations between mathematics and history.

## Broader Significance

The results of my study have been of interest to several audiences. Department of Mathematics faculty members here at Jacksonville University have been curious about ways to foster student discussions of the significance of results presented in courses, though most mathematics courses do not contain as much discussion as my speaking-intensive history of mathematics class. I presented preliminary results of my study, “Student Engagement in the History of Mathematics,” at the spring 2005 Jacksonville University Faculty Symposium for Research and Scholarship. My presentation helped disseminate across campus an example of Scholarship of Teaching and Learning research and demonstrated one of the types of research being performed by faculty members in the department of mathematics. Several faculty members from other disciplines mentioned that they enjoyed hearing about mathematical research they could understand. Some of them later requested more information as to how they could use guided discovery in their courses.

I have had several opportunities to present my paper and to continue my involvement in other aspects of SoTL. At American Mathematical Society/Mathematical Association of America (AMS/MAA) Joint Meetings, I was a co-organizer of SoTL contributed paper sessions in 2011, 2012, and 2013, and of an MAA-AMS SoTL invited paper session in 2010. In addition, I served as a facilitator for MAA minicourses on SoTL in 2007, 2008, and 2009. Not all my presentations have been at SoTL venues; I presented at MathFest in 2007 in a contributed paper session on “Teaching a History of Mathematics Course.”

Although I received tenure prior to becoming involved in SoTL, my selection as a Jacksonville University Scholar for my initial SoTL project and my subsequent opportunities for involvement in SoTL contributed to my promotion to full professor. Other faculty members at JU have used SoTL studies successfully as part of their tenure or promotion portfolios, too.

## Recommendations

For those contemplating an investigation of a course that is offered periodically (like my history of mathematics course), reusing some questions from previous examinations can be helpful to the investigation. In my study, being

able to compare student responses on essay questions to those of previous students gave me data to use when evaluating the new teaching technique.

The strategy just described might mean that the research study would involve the use of previously collected data (e.g., data from earlier semesters such as students' examination responses or course grades). Although the data may have been collected prior to the research study by the SoTL researcher without any intention at that time of presentation, publication, or other dissemination, IRB approval is required before analysis of the data. Because I intended to publish and present the results of my study, I filed the necessary forms and obtained IRB approval prior to the start of my course. That approval included permission to use previous student responses to essay questions. At JU, the IRB required me to obtain permission from the students in my guided discovery course to include them in my study by asking them to sign a consent form that had been approved by the board. As previously mentioned, IRB requirements vary from institution to institution, so boards at another institution may not require obtaining written consent from students. The key message about IRB for beginning SoTL researchers is to become familiar with the local IRB process and allow sufficient lead-time to meet its requirements.

Earlier I discussed the necessity for a theoretical framework in the design of a RUME study, something that is not required in a SoTL study. This difference between SoTL and RUME can be a guide when considering venues for presenting SoTL work. For instance, the *16th Annual Conference on Research in Undergraduate Mathematics Education Announcement and Call for Proposals* included

The conference is a forum for researchers in collegiate mathematics education to share results of research addressing issues pertinent to the learning and teaching of undergraduate mathematics. The conference is organized around the following themes: results of current research, contemporary theoretical perspectives and research paradigms, and innovative methodologies and analytic approaches as they pertain to the study of undergraduate mathematics education. The program will include plenary addresses, contributed paper sessions, and preliminary paper sessions (Special Interest Group, 2012).

Even without knowing the source of this call for papers, the phrase “contemporary theoretical perspectives” would be a clue that, most likely, SoTL papers would not be appropriate for this conference. One final thought: As Schoenfeld (2000) stated, “findings are rarely definitive; they are usually suggestive. Evidence is not on the order of proof, but is cumulative, moving towards conclusions that can be considered to be beyond a reasonable doubt” (p. 649). Mathematics education researchers “should not look for definitive answers but for ideas they can use” (p. 649). The same holds true for SoTL investigations.

## References

- Berlinghoff, W., & Gouvêa, F. (2004). *Math through the ages* (expanded ed.). Washington DC: Mathematical Association of America and Oxton House.
- Crawford, P. (1998). Fostering reflective thinking in first-year calculus students. *ProQuest, UMI Dissertations Publishing* (UMI No. 9835498).
- Davidson, J., Deuser, R., & Sternberg, R. (1994). The role of metacognition in problem solving. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 207–226). Cambridge, MA: MIT Press.
- Dunham, W. (1991). *Journey through genius: The great theorems of mathematics*. New York, NY: Penguin Books.
- Ernest, P. (1996). Varieties of constructivism: A framework for comparison. In L. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 335–350). Mahwah, NJ: Lawrence Erlbaum.
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–235). Hillsdale, NJ: Lawrence Erlbaum.
- Hetrick, N. L., & Saxe, G. W. (2004, Spring). The human side of mathematics: Do our students ever get a chance to see it? *Virginia Mathematics Teacher*, 30(2), 8–10.

- Hutchings, P., & Shulman, L. (1999). The scholarship of teaching: New elaborations, new developments. *Change*, (31)5, 11–15.
- Novak, J. D. (1996). Concept mapping: A tool for improving science teaching and learning. In D. F. Treagust, R. Duit, & B. J. Fraser (Eds.), *Improving teaching and learning in science and mathematics*, (pp. 32–43). London, Teachers College Press.
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. Grouws (Ed.), *Handbook for research in mathematics teaching and learning* (pp. 334–370). New York: Macmillan.
- Schoenfeld, A. H. (2000). Purposes and methods of research in mathematics education. *Notices of the American Mathematical Society*, 47(6), 641–649.
- Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education (SIGMAA on RUME). (2012, August 28) *16th Annual Conference on Research in Undergraduate Mathematics Education Announcement and Call for Proposals*. Retrieved from [sigmaa.maa.org/rume/Site/News.html](http://sigmaa.maa.org/rume/Site/News.html)
- Weaver, J. F. (1987). What research says: The learning of mathematics. *School Science and Mathematics*, 87, 66–70.







# **Illustrations of SoTL Work in Mathematics**

**Theme 2: Crafting Learning Experiences  
around Real-World Data or  
Civic Engagement**



# 10

## Using SoTL to Assess the Outcomes of Teaching Statistics Through Civic Engagement

Cindy Kaus  
*Metropolitan State University*

### Editors' Commentary

In this chapter Cindy Kaus discusses a SoTL project that grew out of her involvement with a national initiative to incorporate civic engagement into the teaching of science and mathematics. She called upon SoTL to provide assessment for the effectiveness of her course redesign. The chapter considers a common problem in doing SoTL, namely encountering difficulties in getting comparison data from control groups taught by other faculty members even when they are willing to assist. The author also describes the professional connections and benefits that accrued to her from employing SoTL to investigate student learning.

### Introduction

The relationship between successful mathematics course completion and degree attainment in higher education is significant (Adelman, 2006, 2009). Hence, the low graduation rates in higher education institutions (Callan, 2008) indicate a need for a more effective and engaging mathematics curriculum. In addition, as more mathematically underprepared students enter higher education (Parsad & Lewis, 2003), engaging students and increasing retention rates in mathematics courses become greater challenges. Low completion rates in general education mathematics courses at Metropolitan State University, a comprehensive public university in St. Paul, Minnesota, led to an investigation of how civic engagement could be used in a statistics course to increase retention and students' interest and confidence in their ability to do mathematics.

Metropolitan State University is an urban institution serving the Twin Cities of Minnesota. The student population is the most diverse in the higher education system in Minnesota. Students of color make up 34% of its student body. The university was founded on principles of connecting higher education with surrounding communities. The idea to incorporate community-based projects into an introductory statistics course emerged as a result of the university's founding principles and also from attending the 2006 SENCER Summer Institute. SENCER, which stands for Science Education for New Civic Engagements and Responsibilities, is an NSF-funded curriculum and dissemination project. It aims to improve collegiate instruction in STEM fields by promoting the teaching of science and mathematics through complex, real world problems. At the institute, SENCER presented convincing research to support its approach to engaging students in mathematics and science. An independent evaluation of the SENCER project

(Weston, Seymour, & Thiry, 2006) found that students who completed a SENCER course learn more, are able to relate learning to real world problems, and are more interested in the STEM disciplines. The evaluation also indicated that students in SENCER courses made significant gains toward their learning objectives and became more confident in their ability to distinguish between science and pseudoscience.

## Stat 201 at Metropolitan State University

Stat 201 is a general education course taught at Metropolitan State University. Approximately 13 sections of Stat 201 are offered each semester and each section enrolls 32 students. The students' majors include but are not limited to mathematics, nursing, psychology, and business. Their skill levels are as varied as their majors. For many non-traditional students, their last mathematics class could have been 5, 10, or even 15 years ago. On the other hand, each class usually has some very able students with strong mathematical skills. Although some of the students take the course in their freshman or sophomore years, a significant number of them are upperclassmen who have postponed fulfilling this general education requirement because of mathematics anxiety. The course includes descriptive statistics and inferential statistics through hypothesis tests, confidence intervals, and  $t$ -tests. Typically, the course is taught in a traditional lecture style with an emphasis on using the graphing calculator to aid in conceptualization, computation, and organization of data. It meets one day each week for 3 hours and 20 minutes. The varied interests and skill sets of the students and the long class periods made Stat 201 an ideal course for community-based projects.

## Research Question Prompted by the Need to Engage and Retain Students

The goal of incorporating community-based projects into a general education requirement was to provide the students with the tools to see the importance of statistics in their disciplines and personal lives. From this goal came the research questions:

- Will incorporating civic engagement into a statistics course increase students' confidence in their ability to do and communicate statistics?
- Will incorporating civic engagement into a high stakes general education statistics course increase retention?

## How Did I Incorporate Civic Engagement into a Statistics Course?

To engage students, the course was revised to focus on and highlight semester-long, community-based group projects. The idea for this came from a 2006 SENCER model course, Quantitative Literacy Through Community-Based Projects (Dewar, Larson, & Zachariah, 2011). The first day in Stat 201 is now devoted to talking about the projects, brainstorming ideas for topics, and discussing what will be involved in successfully completing this portion of the course. Early on, students learn that a community-based project is one that analyzes data related to a local civic, social, or environmental issue and that will aid some person or organization in the community to make well-informed decisions. The project has five stages:

**Stage 1** Brainstorm topics for the projects and do background research.

**Stage 2** Form groups based on student topic interest and submit a proposal.

**Stage 3** Gather data.

**Stage 4** Write a first draft.

**Stage 5** Write the final paper, develop a poster presentation, and write an action letter to a person or organization that could benefit from the results of their project.

By sharing the results of their projects with the community, the students had engaged in service learning. Examples of student-generated project topics are

- Domestic Abuse in Downtown Minneapolis: Gender, Race, and Language (action letter sent to Court Watch, a non-profit program intended to monitor court proceedings in domestic violence),

- Mercury in Metro Lakes and Fish Consumption Advisories (action letter sent to the American Hmong Partnership with an advisory translated into Hmong), and
- Does Where You Live Affect the Price You Pay for Prescription Drugs? (action letter sent to a local senior citizen group).

To emphasize the applicability of statistics, data from civic, social, or environmental issues were incorporated into lecture examples and in-class group work. This showed the students how statistics can be used to understand these issues and make more informed decisions.

## Collecting Evidence on Attitudes and Retention

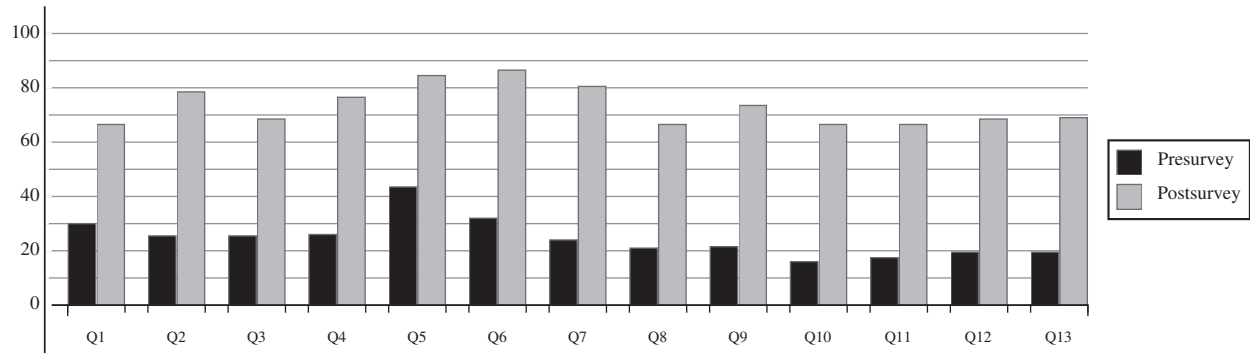
Two survey instruments provided evidence about student interest and confidence in their ability to do and communicate statistics. The first of these, the SALG (Student Assessment of Learning Gains), was developed by SENCER to measure student confidence in doing, communicating, and understanding STEM content. This customizable online instrument is available at [salgsite.org](http://salgsite.org) and the website provides access to hundreds of faculty-generated surveys. I revised existing SALG questions for mathematics and science courses for this investigation of incorporating civic engagement into a statistics course.

Each semester, students in Stat 201 with community-based projects completed the customized online pre- and post-SALG. To ensure a high percentage of participation, in the spring 2007 and fall 2008 semesters I had students complete the pre- and post-surveys in class, yielding 100% response rates. In the fall 2009 semester, students completed the surveys outside of the regularly scheduled class time. They did not receive course credit for completion of the survey. This resulted in a 75% student response rate on the pre-SALG and only 25% participation on the post-SALG. It seems that to get a better response rate, the surveys should be conducted during class time or students should receive credit. To compare student confidence between students who had the course with a strong civic engagement component and students who had been taught statistics in a more traditional manner, the surveys were administered to the other six sections of Stat 201 during the spring of 2007. The response rates were very low, less than 10%. In the fall of 2008, the SALG was again administered to students in the sections of Stat 201 that did not have a strong civic engagement component. To increase student participation, instructors were contacted weeks before the survey was administered and the students received a brief presentation on the first day of class about the survey. Still, the participation rates were so disappointing on the pre-SALG that the students were not asked to complete the post-SALG.

The results of the SALG were used to determine how students' confidence in their ability to do, understand, and communicate statistics had changed over the semester. Students were given thirteen statements related to confidence along with the response choices: Not Confident, A Little Confident, Somewhat Confident, Highly Confident, or Extremely Confident.

The 13 statements from the SALG were: I am confident that I can . . .

1. Discuss statistical concepts with my friends or family.
2. Think critically about statistical-related findings I read or hear about in the media.
3. Make an argument using statistical evidence.
4. Determine the difference between appropriate and inappropriate use of statistics.
5. Interpret tables and graphs.
6. Understand statistical concepts commonly found in books, newspapers and journals.
7. Find journal articles, or statistical data using library or internet databases.
8. Extract main points from a statistical report and develop a coherent summary.
9. Give a presentation using statistics to my class.
10. Describe how statistics is used in analyzing civic or environmental problems.
11. Pose questions that can be addressed by collecting and evaluating statistical evidence.
12. Organize a systematic search for relevant data to answer a question.
13. Write reports using logical reasoning or data as evidence.



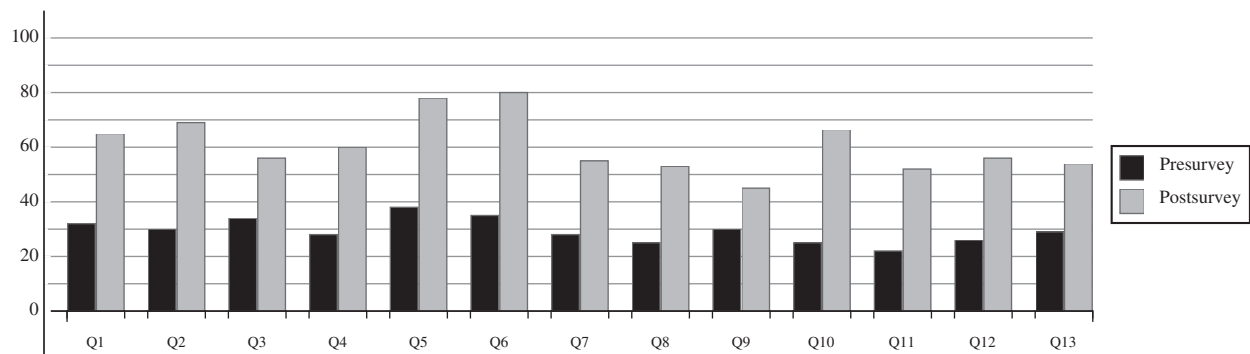
**Figure 10.1.** SALG results, statistics with community-based projects.

The cumulative results of the pre- and post-SALG for all sections taught with community-based projects show large gains in students' confidence in their ability to understand statistics and its applications (see Figure 10.1). Analyzing the data from the spring 2007 and fall 2008 semesters, students who had completed the community-based projects showed a highly significant ( $p < .005$ ) increase in confidence by the end of the semester.

To have a comparison group for the study, I taught one section of Stat 201 in the spring semester of 2010 without community-based projects. Although the course was taught without projects, it included examples in lectures, homework problems, and in-class group work problems with a theme of civic, social, and environmental issues. So, this was not entirely a traditionally taught statistics course and students were given opportunities to see how statistics could be used to better understand civic issues. The cumulative results of the pre- and post-SALG for the one section of Stat 201 taught without community-based projects are shown in Figure 10.2. Although the graph shows statistically significant ( $p < .05$ ) gains for students in their confidence, they are not large as those for the students who had completed the course with community-based projects.

In addition to the thirteen statements about student confidence in the survey, students answered the open-ended question: "Which course activity helped you learn the most?" As shown in Table 10.1, 48% of the students felt that it was the group projects, while students in the section without projects felt that the lectures were the most beneficial. For those students who took the course with community-based projects, 3% reported on the pre-SALG and 35% reported on the post-SALG that they were interested in taking additional mathematics or statistics courses in the future. For students in the course without the projects, their interest in taking additional math or statistics classes increased from 6% to 12%.

The second survey used in the study was the Metropolitan State University Institutional Improvement Questionnaire (IIQ). It is designed to evaluate both courses and instructors and is administered in all classes at the end of each semester. Therefore there is no difficulty obtaining student participation. Of the 22 Likert scale items, the two of most interest to this study were "The course stimulated student thinking" and "How much have you learned in this course?" There was no statistically significant difference in the answers when I taught the course with or without



**Figure 10.2.** SALG results, statistics without community-based projects.

Course Activity Helped Most	% of CBP Students	% of non-CBP Students
Group project	48%	0%
Homework	16%	24%
Lecture	13%	39%
No response or other	13%	11%
Examples done by teacher in class	6%	14%
Reviews for exams	2%	7%
Studying on my own	2%	5%

**Table 10.1.** Post-SALG results, with and without community-based projects (CBP)

community-based projects. Also pertinent to this study were the answers to the open-ended questions “What do you like most about this course?” and “What do you like least about this course?” Thirty-six percent of students answered that the group projects were the part of the course that they liked the most. However, of the 36% favoring the group projects, almost half stated that the group projects were also what they liked the least about the course. Students who did not like the projects said that they enjoyed learning statistics through a community issue, but they did not enjoy working in groups and having to depend on the availability and performance of their group members.

To determine the impact of the engaged learning pedagogy on retention, student grades were analyzed over four semesters. During the first three semesters, five sections of Stat 201 with community-based projects were taught and in the fourth semester one section of Stat 201 was taught without community-based projects. To measure retention, course grades of C– or better were taken as indicators of success and D, F, or W were taken as failure. Some majors require a C– or better in Stat 201 for entrance into their programs, so students would need to re-take the course with a grade of D or lower. The course without community-based projects had a failure rate of 23% while the course with community-based projects had a failure rate of 18%. Although these results show an increase in retention in the course with community-based projects, the findings were not statistically significant, with  $p = .15$ .

Anecdotal evidence provided additional support for changes in students’ views of the value of learning statistics. For the first time in my 11 years of teaching mathematics at the college level, students in a general education course were asking me for letters of recommendation for future employment or graduate school. During the first semester of teaching Stat 201 with community-based projects, six students requested letters of recommendation. When I inquired why they were requesting them from a statistics instructor, they said that they had seen how important statistics was and thought that a letter of recommendation from a statistics professor would carry a lot of weight with an employer or a graduate committee. Another student who was applying for a job in the biostatistics department at the Mayo Clinic in Rochester, Minnesota, asked for a letter of recommendation and if I would serve as a telephone reference. When I received the phone call, I told the director of the Mayo Clinic’s biostatistics program that the student had taken only one introductory course in statistics. She explained that the applicant was very enthusiastic about the project he had completed in my course and had described it in detail during his interview. As a result, he had been asked back to present his findings in a more formal setting. The student is currently employed in the biostatistics department of the Mayo Clinic.

## Conclusions

The study described in this chapter provides compelling evidence of the positive effect of incorporating civic engagement into a statistics course on student interest and confidence in their ability to do and understand statistics. Results from the pre- and post-SALG showed that students from the classes with community-based projects had statistically significant ( $p < .005$ ) higher gains in their confidence than students from the class without community-based projects. The average increase in student confidence for the thirteen statements was 48.5 percentage points for students in the course with community-based projects (Figure 10.1), more than the increase of 31.3 percentage points for students in the course without community-based projects (Figure 10.2). Students in the course with community-based projects overwhelmingly rated the projects as the single activity that contributed most to their learning in the class. They also



had a larger increase in their desire to take more mathematics and statistics courses in the future. Because my section of Stat 201 without community-based projects was used as the comparison group, the students in it had seen applications of statistics to civic issues through lecture examples and in-class group work. This could have led to their increased confidence. For the experiment it might have been better to omit discussion of civic issues in the course without community-based projects. However this was difficult for me to do since my experience with incorporating civic engagement into statistics had been so positive. This brings up an ethical question that arises in any type of research with a control group, namely, should the researcher deny a treatment when there is evidence that it is beneficial? It would have been preferable to have had pre- and post-SALG data from another instructor's course. However, getting an acceptable response rate from students in other sections proved to be too difficult.

The results from the items "The course stimulated student thinking" and "How much have you learned in the course?" on the IIQ showed no significant difference between the courses. However, there are other factors that may have contributed to these results. I had never taught a statistics course before my first section of statistics with community-based projects and the section that I taught without community-based projects was taught three years later. My teaching skills and my familiarity with topics that the students struggled with may have increased significantly. The lack of a significant difference in the answers on these two questions could be attributed to my increased experience with the material. The short answer questions on the IIQ "What did you like most about the course?" and "What did you like least about the course?" gave valuable information. Students in the course with community-based projects listed the projects as both what they liked the most and what they liked the least. The explanations of their choices were enlightening. The students thought they learned a lot from the projects but they didn't like working in groups and having to depend on other students' contributions for their grade.

Student retention was evaluated by analyzing the number of withdrawals and grades of D or F in the course. Incorporating community-based projects into the statistics course resulted in a decrease in grades of D, F, or W from 23% in the traditionally taught course to 18% in the course with community-based projects. This was not statistically significant, so the effect on retention of students is not clear from this study. To get better evidence, further work should be done.

## Next Steps

There are many directions that a continuation and refinement of this study could take. By expanding the study to include other mathematics courses taught with civic engagement components, additional evidence of the impact on student attitudes, confidence, and retention could be obtained. Conducting entrance and exit interviews with students could also provide more information about their attitudes. A longitudinal study could investigate the impact of civic engagement on future academic and professional choices. Another direction to take would be to examine the effect that incorporating civic engagement into mathematics courses has on student populations that are underrepresented in the STEM disciplines. Recent studies (Seymour & Hewitt, 2000) have shown that students from underrepresented groups in the STEM disciplines may be opting out of STEM majors and careers because they do not see the positive societal effects of STEM education or the STEM professions.

## Personal and Professional Impacts

The results of this work have encouraged me to continue incorporating civic engagement into mathematics courses and to investigate its effects. I have integrated civic engagement projects and activities into college algebra, calculus, and differential equations courses. I have also worked with mathematics faculty at local community and private colleges to develop curriculum and assessments to study student learning and attitudes in mathematics.

Incorporating civic engagement into mathematics courses and documenting its effects has had a positive impact on my professional career. The work aligns well with Metropolitan State University's mission, so the administration and my faculty peers valued both the curriculum development and the investigation that followed. The work was an influential factor in my gaining tenure and promotion to associate professor. It led to several grants funded by the Minnesota State Colleges and Universities system, and in 2013 to my being named co-PI on *Engaging Mathematics: Building a National Community of Practice*, a three-year initiative supported by the National Science Foundation through its

TUES-II program. It also played a role in my successful application for a Fulbright to work at the University of Seychelles in Africa from January through June 2014.

The results of the study have been made public at many conferences: the Joint Mathematics Meetings (2009), the SENCER Summer Institutes (2008, 2009, 2011), the Community College National Center for Civic Engagement Hand in Hand Service-Learning and College Completion Conference (2011), Developing a Good Heart in STEM: The First Summit on Incorporating Social Justice and Service-Learning into the STEM Curriculum (2009), and the Minnesota State Colleges and Universities Realizing Student Potential Conferences (2007, 2008). Presenting the work for a broad range of audiences has resulted in opportunities for collaborations. I co-organized a session, “Teaching Mathematics and Statistics through Current Civic Issues,” at the 2009 MAA MathFest and hosted the 2009 SENCER Midwest Symposium on “Teaching Quantitative Reasoning through Civic Issues” at Metropolitan State University.

This study and its results have provided me with opportunities to further my professional development in the scholarship of teaching and learning. The positive impact of incorporating civic engagement into the mathematics curriculum has resulted in more math faculty members, including adjuncts, seeking out ways to connect the mathematics that they are teaching to civic issues. Most importantly, it is clear that students’ confidence in their ability to do and communicate statistics increased significantly when a strong civic engagement component was added to the curriculum.

## References

- Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education. Retrieved from [www.ed.gov/rschstat/research/pubs/toolboxrevisit/index.html](http://www.ed.gov/rschstat/research/pubs/toolboxrevisit/index.html)
- Adelman, C. (2009). *Answers in the toolbox: Academic intensity, attendance patterns, and bachelor’s degree attainment*. Washington, DC: U.S. Department of Education. Retrieved from [www2.ed.gov/pubs/Toolbox/toolbox.html](http://www2.ed.gov/pubs/Toolbox/toolbox.html)
- Callan, P. (2008). *The 2008 national report card: Modest improvements, persistent disparities, eroding global competitiveness*. San Jose, CA: The National Center for Public Policy and Higher Education. Retrieved from [measuringup2008.highereducation.org/print/NCPPEMUNationalRpt.pdf](http://measuringup2008.highereducation.org/print/NCPPEMUNationalRpt.pdf)
- Dewar, J., Larson, S., & Zachariah, T. (2011). Group projects and civic engagement in a qualitative literacy course. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 21(7), 606–637.
- Parsad, B., & Lewis L. (2003). *Remedial education at degree-granting postsecondary institutions in fall 2000*. Washington, DC: National Center for Education Statistics. Retrieved from [www.eric.ed.gov/PDFS/ED482370.pdf](http://www.eric.ed.gov/PDFS/ED482370.pdf)
- Seymour, E., & Hewitt, N. (2000). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.
- Weston, T., Seymour, E., & Thiry, H. (2006). *Evaluation of science education for new civic engagements and responsibilities (SENCER) project*. Retrieved from [www.sencер.net/Assessment/pdfs/FINAL.REPORT\\_SENCER\\_12.21\\_06.pdf](http://www.sencер.net/Assessment/pdfs/FINAL.REPORT_SENCER_12.21_06.pdf)



# 11

## A Pedagogical Odyssey

Michael C. Burke  
*College of San Mateo*

### Editors' Commentary

Michael Burke's odyssey in this inspirational chapter could be characterized as a "vision of the possible" investigation, initiated because he wanted to try something unusual. He began with a desire to help his students gain a deeper understanding of the concept of a function. He also wanted them to encounter genuine applications of mathematics, ones that were truly interdisciplinary. He thought that asking his students to write about this would help them clarify their thinking. As his experiment unfolded, to understand what was happening and to refine what he was trying to achieve, he first used reflective practice, and later SoTL. This chapter underscores the usefulness of observation and data collection on two levels. In aiming to teach his students the value of observation and data collection, the author discovers that this is exactly what he must do as their teacher – observe and collect data. The author writes in an engaging style and with passion about his journey to the discovery that the scholarship of teaching and learning enables him to understand what is happening in his classroom.

### Introduction

*"I've gone to find myself. If I come home before I return, keep me here."*

My title is a bit Homeric, which is perhaps a little highbrow for me, so I thought to follow it with a quote from a t-shirt. I saw the above quote in 2006, written on a t-shirt worn by a young man walking across the Charles Bridge in Prague. If we attempt to decipher its literal meaning, we see that it falls somewhere between puzzling and incoherent. It is for this reason that it seems an apt expression of the sentiments of a young American traveling in Europe, particularly one who finds himself in Prague; it expresses, in a humorous way I think, a sense of massive confusion about who and where he is, about here and there, and about time. Prague frequently has this effect on young Americans. When young people embark on such a journey, they often experience feelings of dislocation and confusion, question their identity, and, ultimately, return home changed, with a different view of the world.

I, too, have been on a journey, a pedagogical journey rather than a physical one. Although I never encountered confusion on the scale described above, it is fair to say that I embarked on my journey with no clear sense of my ultimate destination; in fact, I began my journey without realizing it at all. As I traveled, each stage of my pedagogical exploration yielded fresh questions, and I set off in a modified direction, toward a new destination. Now, after fifteen years, I have perhaps returned home, and I find myself a changed teacher, with a different view of my discipline, a different view of how it fits into the college and the world, and a different view of what and how we should be teaching.

## Embarking on My Journey

I teach at the College of San Mateo, a suburban community college on the San Francisco peninsula. We have a healthy transfer program; the most popular transfer destinations for our students are the University of California (Berkeley and Davis campuses), and the California State University System (particularly San Francisco and San Jose State Universities). The two courses I will be describing, calculus and precalculus, are standard transferable courses taken by students who successfully make the transition to four-year colleges. I used standard textbooks for the courses. But when I embarked on my journey, I decided to augment the standard curriculum in a dramatic way.

I began innocently enough in the mid-1990s. I was thinking about the Rule of Three popularized by the Harvard Calculus Reform Project (Hughes-Hallett et al., 2009), the idea that we should routinely examine functions from three different perspectives: graphically (as a graph), numerically (as a table of values), and symbolically (as an equation or formula). As I thought about the Rule of Three, I realized that a spreadsheet is the perfect tool for the study of functions because users of a spreadsheet can use a formula to build a table, and then, from the table, they can use the spreadsheet to quickly construct an accurate graph. The spreadsheet lets students move from one perspective of a function to another, and thus view a function from all three perspectives. So I began my thinking for this work a number of years ago with the idea that using a spreadsheet should help my students come to a deeper understanding of the mathematical concept of a function.

At the same time, I had three other pedagogical ideas. First, I wanted my students, for motivational purposes, to see some genuine applications of the mathematics they were studying. As Lynn A. Steen (2004) argued, “for students, contexts create meaning. Yet all too often, mathematics students fail to see the relevance of their studies” (p. 24). Second, I wanted to teach through interdisciplinary problems, so that my students would see that knowledge is not constrained by artificial disciplinary boundaries; thus I was interested in what now would be called “integrated learning” (AACU, 2005; Huber & Hutchings, 2004; Kezar & Elrod, 2012). Finally, inspired by the writing across the curriculum movement, I believed that asking my students to write about mathematics would help them clarify their mathematical thoughts (Meier & Rishel, 1998). So I began designing data-based integrative writing assignments that incorporated all these ideas. The assignments required using a spreadsheet to view functions from all three perspectives, the presence of data ensured that the students would see genuine applications, the assignments were interdisciplinary in nature, and the final product was a written paper. My goal in all this was simply to teach a better, more interesting mathematics course.

As I began teaching with these integrative assignments, a funny thing happened. I became interested in helping my students explore non-mathematical issues. How do we decide what is really true? Can we decide for ourselves, or do we have to rely on experts? What is the role of data and evidence in making decisions? Where do reliable data come from, and how do we treat and interpret the data? What kinds of conclusions can we draw from a set of data? Given an issue, what is the interplay between data, preconceptions, opinion, and belief? What is the proper role of science in making judgments and decisions, and in public policy? These questions are subtle, difficult, and important. Moreover, they are rarely addressed in college classrooms; science majors may be asked to deal with these issues, but I suspect that the majority of our students graduate from college without ever giving serious consideration to questions such as these. But, in fact, teaching our students how to ask, and answer, these questions is essential to the teaching of critical thinking, a goal that nearly all of our colleges have endorsed.

Over the years, I constructed a number of data-based integrative writing assignments for my students in calculus and precalculus courses. Topics addressed included the size of the Pacific salmon run, global warming (based on data about carbon dioxide levels), the clarity of Lake Tahoe, a historical look at the population of Ireland, radiocarbon dating, global warming again (based on data about the size of the Arctic ice cap), nuclear waste, and world population. Each assignment presented the students with a data set. Students were to construct a mathematical model (linear or exponential) describing the trend of the data set, and then to implement the model by using a spreadsheet to construct a table and a graph showing the given data and the model projections. Next, they were to write a paper in which they presented their model (with a table and a graph), used it to make some projections (usually extrapolations into the future), and finally, based on the projections, drew a conclusion about the underlying issue.

The sequence of assignments was designed to lead the students through a series of modeling issues that inevitably arise. First, just the idea of a mathematical model is difficult for students to understand. How do we construct mathematical descriptions of natural or man-made phenomena, and how are they helpful to us? Next, there is the

question of what conclusions we can draw. The presence of data and a model lend an air of concreteness to the assignment that students are not accustomed to; grappling with the data forces the students to think. They have to think very carefully to draw appropriate conclusions and, for many, this is a new experience. And there is the possibility of the breakdown of the mathematical model. How can a model break down? What are the consequences? Can the breakdown be anticipated? Does the model and its anticipated breakdown suggest courses of action that we should take? Finally, there is the reconciliation of the mathematical work the students have done with their preconceptions, and with additional readings on the topic. Additional readings are often contradictory. The students' mathematical work frequently suggests conclusions that are at odds with their preconceptions. How do the students sort all of this out? These are difficult issues, deep, in some ways, and fascinating.

My students found the assignments extraordinarily difficult, and I heard, over and over again, comments to the effect that, "I have never been asked to do anything like this before." Although I intended the assignments to be interesting supplements to my courses, discussion of the non-mathematical issues threatened to hijack my mathematics courses. Students found these issues compelling, and they wanted to think and talk about them virtually every day.

## The Scholarship of Teaching and Learning

But what about me? What kinds of demands did teaching in this fashion place on me? First, obviously, I had to learn how to grade written papers. I had to learn how to respond gently to outlandish ideas, expressed in the papers and in class, and how to give careful feedback without being judgmental. I had to learn how to refrain from providing "the answer." The data-based writing assignments are intended as an opportunity for exploration for my students, an opportunity for them to begin to learn how to work with and interpret data and for hypothesizing, discussion, and speculation; this process is short-circuited if I quickly offer a "correct" answer. I had to learn to never offer answers. Finally, I had to learn how to evaluate the new approach to teaching calculus and precalculus. Early on, I had the sense, from class discussion and the written papers, that I was offering my students a rich educational experience, but I could not document it; I did not know how to analyze the results of the work I was doing. I was neither able to demonstrate to others that this approach was effective, nor was I able to identify difficulties so that I could improve the course. I was voyaging in the dark.

Some years into my voyage, I was fortunate to be able to spend some time at the Carnegie Foundation for the Advancement of Teaching, first as a participant in the Integrative Learning Project (see [Huber et al., 2007](#)) and later as a Carnegie Scholar. At Carnegie, I was exposed to the ideas of the scholarship of teaching and learning, and adapted some of them for my own use. My research question as a Carnegie Scholar took the form characterized by Lee Shulman of the Carnegie Foundation as a "vision of the possible" (elsewhere referred to in this book as a *What could be?* question). It centered on the question "What would a mathematics course that used data-based integrative assignments look like?"

My research question was not tightly focused, and it was not the type of question that could be answered in a year or two by an experiment. My investigation went on for many years and had a meandering quality to it; it was an odyssey, if you will. As [Huber and Hutchings \(2005\)](#) point out, such an investigation is not unusual among practitioners of the scholarship of teaching and learning. Perhaps for this reason, my research methods were not elaborate (and I think, for my purposes, they did not need to be). There were no involved experimental designs, no comparison studies, no collection of massive amounts of data that required analysis. Instead, I adopted a simple, straightforward approach that relied on three strategies.

First, I began to analyze in a systematic way the written papers. I found that my students were, for the most part, successful at constructing a mathematical model and at using a spreadsheet to implement the model as a table and a graph. They could also use their models to make projections. But my students had a great deal of difficulty when they tried to reach a conclusion. The conclusions in the papers generally fell into one of three categories. First, a few students did reach appropriate conclusions. A second group of a few students reached conclusions that were directly contradicted by their work with the data. It was as though they were oblivious to the mathematical work they had just done, or perhaps they discounted its value completely. But the majority of my students fell into the third category: They offered what was essentially a non-conclusion such as "it might be this, it might be that, only time will tell" or (even worse from my perspective) a conclusion that sounded fatalistic, "whatever is going to happen will happen, there is no way to know what it is, and we can't do anything about it anyway." Their work with the data seemed to have no

impact on their thinking; they did not appear to know how to interpret the work that they had done. I would even go so far as to say that, in some cases, they did not even believe the work that they had done.

This finding was, of course, distressing. To learn more about their difficulties with formulating conclusions, I followed the analysis of the written papers with a second strategy for gathering evidence on student learning. I began to ask my students, when they handed in their papers, to respond to reflection questions. Some were of a generic nature; others were specific to the particular assignment. Examples of these questions appear in the Appendix. Our colleagues in English refer to such questions as “prompts,” an apt term. The goal of the question is to prompt the students to think and to write about their experiences. The prompts should be open ended, and should invite speculation; if a prompt is too specific, the result is a direct answer. I have found that the particular prompts used for reflection are not all that important. If a student ignores the prompt but writes something thoughtful and interesting, it does not bother me a bit. Reflections are important to me because through them I learn about how my students are thinking, and this informs my teaching. Reflections are also important for the students. I want them to examine their own experiences and to observe what was happening with their thinking.

My students told me, through their reflections, about their difficulties, about where they were confused, and about how difficult it was to write a paper whose conclusion was constrained by the work they had done with data. Most of our students are accustomed to writing papers that consist of their opinions only. I learned that they had an extraordinarily difficult time when their preconceptions were not consistent with the work that they had done with the data; in fact, my students reacted to this with extreme discomfort. I began to emphasize the importance of setting aside preconceptions and not drawing conclusions until after work with the data has been completed: of allowing the work with the data to drive the conclusion. The central question, I told them, is “What do the data tell you?”

The third strategy I used was to evaluate my work in these courses by the standard method, exam scores. I focused on performance on the final exam. The courses I was working with (calculus and precalculus) were standard mathematics courses, with standard content and curriculum. The written papers and the student reflections had convinced me that, with the data-based writing assignments, I was offering a valuable and educationally rich experience that is not ordinarily present in calculus and precalculus courses. But it was important that this did not come at the expense of the mathematical material I was supposed to teach. Since I have taught both calculus and precalculus many times, I used the same final exam I had used in the past, and compared scores with those of previous classes. The precalculus students scored slightly lower on the exam than their counterparts from previous semesters, although the difference was within normal semester-to-semester variation between classes; the calculus scores were essentially the same as the scores from earlier semesters (actually slightly higher). From this, I concluded that it is possible to augment the standard mathematics curriculum with writing assignments of this type without compromising mathematics instruction and student learning. The students also had the benefit of using mathematics to explore a collection of important issues.

## An Assignment: The Population of Ireland

One of the assignments, a historical look at the population of Ireland, provides an example of how I used SoTL to analyze my course. The evidence I collected, assignments and student reflections, enabled me to examine the *What is?* question: What are the difficulties students encounter when attempting to complete such assignments? The answers I obtained changed my thinking about the course as I was teaching it.

The Ireland assignment is presented in two parts. In Part 1, the students are given data about the population of Ireland, in millions, for four years (1700, 1750, 1801, and 1841), as shown in Table 11.1.

Population of Ireland, in millions		
Year	Data	Model
1700	2.0	
1750	2.4	
1801	5.0	
1841	8.2	

**Table 11.1.** Population data for Ireland

A graph shows that the data are nicely exponential. The students then choose two data points from which to construct an exponential model that fits the data. The students are directed to use their model to make projections for the population of Ireland in the years 1851, 1901, 1951, and 2001. So, in the first part of this assignment, students derive their model, construct a table and graph, and make the projections; but they are told not to write a conclusion at this time.

By letting  $t$  represent the number of years since 1700 and  $P$  represent the population, most students obtain a model similar to  $P = 2e^{0.010t}$ . They graph the data and model on the same set of axes to check the fit of their model and to make their projections. Usually, students produce a table and graph somewhat like Table 11.2 and Figure 11.1. The graph (Figure 11.1) indicates that the model provides a good fit for the data. Using the graph (or the table, or the equation), students project that the population of Ireland in 2001 should be about 40 million people.

The students are now ready for Part 2 of the assignment. I provide data for the population of Ireland in the years 1851, 1901, 1951, and 2001 (the bold entries in Table 11.3).

The additional data show that the population of Ireland fell sharply in the decade after 1841, fell again by the turn of the century, was relatively constant through the first half of the twentieth century, and grew slowly in the last half of the twentieth century. Students add the data to their graph as shown in Figure 11.2.

Once the students see the more recent data, they realize that their model projections failed spectacularly. The graph in Figure 11.2 illustrates how badly the exponential model failed. The actual population of 5.6 million in 2001 differs

Population of Ireland, in millions		
Year	Data	Model
1700	2.0	2.0
1750	2.4	3.3
1801	5.0	5.5
1841	8.2	8.2
1851		9.1
1901		14.9
1951		24.6
2001		40.6

Table 11.2. Population data and model projections



Figure 11.1. Exponential model



Population of Ireland, in millions		
Year	Data	Model
1700	2.0	2.0
1750	2.4	3.3
1801	5.0	5.5
1841	8.2	8.2
1851	<b>6.5</b>	9.1
1901	<b>4.4</b>	14.9
1951	<b>4.4</b>	24.6
2001	<b>5.6</b>	40.6

**Table 11.3.** Population data, Part 2 data in bold

dramatically from their model projection of 40 million. Their task now, for the written paper, is to grapple with the question of why their model broke down.

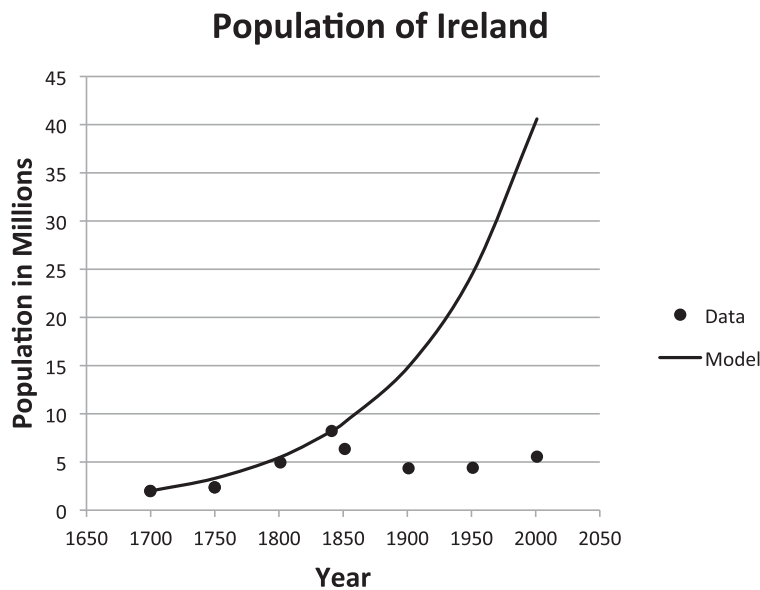
The Ireland assignment introduces my students to the concept of model breakdown. How do they react? Here are a few reflections on the Irish population assignment, offered to illustrate their thinking.

I had a high degree of confidence in my model initially because it fit the data well. Yes, I was surprised at the model’s deviance from the second set of data. I see that math modeling is not so simple. It requires a lot of research into all internal and external factors that can affect the model’s output.

There was some difficulty sticking straight to facts and minimizing my emotions’ influence (especially in the development of the model).

In the assignment, first part and second were totally different [from] each other. Calculating population of Ireland in year 2001 was easy. But hardest part was to see real data and analyze it – what caused the differences? It was really hard for me to organize ideas.

I know Ireland’s history from previous history classes, so I knew the model would be way off. I think that the exponential growth model for what Ireland’s population should have been had nothing gone wrong was an integral part of the paper. It would help readers realize just how tragic the potato famine was.



**Figure 11.2.** Failure of the model after 1841

## Insights from the Data

The student reflections and, to a lesser degree, the written papers, afforded some insights into what my students were thinking and the difficulties they were encountering. First, although they showed that many students come to understand the idea of the breakdown of a mathematical model, there was much confusion about how to deal with it. Before they had data that actually demonstrated the model's failure, students gave almost no thought to anticipating breakdowns of the model. Their comments also indicated that some students had great trouble organizing their thoughts, particularly when their personal emotions and frustrations were involved. Finally, one of the reflections made note of the power of the mathematical analysis. It is one thing to simply state that the Irish potato famine was a great tragedy, but quite another to work through the mathematics of the famine; the mathematical work, the graph, and the dramatic failure of the model all serve to portray powerfully, for my students, the true magnitude of the tragedy.

## Impact and Lessons Learned

In response to these reflections, and similar ones from other assignments, I made substantial changes in the way that I talk about some of the basic issues that arise in a class like this. For example, students who are new to mathematical modeling often believe, and say, that their model proves that a particular outcome will occur, evidently because the projections are supported by algebra that they know to be correct. The Irish population assignment is a good way to disabuse them of this idea.

Beyond that, I now stress the difference between a prediction and a projection. A prediction is an assertion that a particular event will occur. A projection, on the other hand, is a conditional statement, a statement that a particular event will occur provided current trends continue. A prediction that the population of Ireland would be 40 million in 2001 was certainly incorrect. But a projection of a 2001 population of 40 million, provided that the population trends prior to 1841 continued, is not incorrect; the population trends did not continue.

To help deal with the issue of proof, I now frame the question in terms of support versus non-support for the hypothesis, or I ask whether the data and the work they have done with the data is consistent, or perhaps inconsistent, with a hypothesis. Thus, if the issue is the clarity of Lake Tahoe, I will ask if the data support the conclusion that the water quality of the lake is degrading to the point where, at some time in the relatively near future, the water will be so murky that Lake Tahoe will have lost its famed clarity. A conclusion of support for a hypothesis is far short of a claim to have proved it. If the issue is global warming, I will ask whether their work with the data is consistent or inconsistent (or neither!) with the hypothesis that global warming is occurring. The intent here is to provide my students with the language and thinking skills to allow them to draw more careful conclusions, to be able to say, for example, "my analysis of the data supports the hypothesis that global warming is occurring" or "my work with the data is inconsistent with the hypothesis that global warming is occurring." I am trying to teach them how to avoid jumping to extreme (and unwarranted) conclusions such as "I have proved that global warming is occurring" or "I have disproved the claim that global warming is occurring."

A related issue is that of certainty. Many students (and indeed, many adults) are extremely uncomfortable with uncertainty; they prefer absolute certainty. Many take the position that if we are not absolutely certain of a conclusion, then we know nothing at all about a particular phenomenon. But the reality is that we live in an uncertain world, and we need to come to grips with that fact. I am trying to provide my students with a conceptual framework for dealing with uncertainty, with the tools and language to be able formulate conclusions that are helpful and valid, though they fall short of absolute certainty. Consider the terms: projection, consistent, inconsistent – as mathematicians, we should not be surprised that the particular language we choose to use is of critical importance.

To return to the question of the breakdown of a mathematical model, the recognition that a model may break down is only my first goal. I want my students to learn to think about how and when a model will break down. A mathematical model is nothing more than a mathematical description of a trend, and the continuation of a trend depends on other factors. A model will break down when these other factors change. If we can anticipate when that will occur, we can anticipate how and when a model will break down. I tell my students that once they have found a good mathematical model, the most important question to ask about it is the question of how and when it will break down. The reason for this is that often, as in the case of Ireland, the moment of model breakdown is a moment of tragedy. If we can anticipate the breakdown of the model, perhaps there are actions that we can take to avert the tragedy.

## Return to Ithaca (Results)

I stated at the beginning of this chapter that my odyssey has left me a changed teacher with a different view of my discipline, a different view of how my discipline fits into the college and the world, and a different view of what and how we should be teaching. Largely in response to my work with the scholarship of teaching and learning detailed here, my goals as a teacher have changed radically. Initially, I simply wanted to teach better mathematics courses. But now, although I still want to teach mathematics, I also want to engage my students with other issues. I have moved from the desire to embed the mathematics I was teaching into compelling, important, and authentic contexts (via the data-based integrative writing assignments) to fully embracing the complex teaching challenges inherent in teaching through these contexts, contexts that are described as “natural critical learning environments” by Bain (2004, pp. 99–103). I subscribe to the vision for college and university education presented in the Association of American Colleges and Universities publication, *College Learning for the New Global Century* (National Leadership Council, 2007). I see my work as an attempt to fit the teaching of mathematics into that vision. My goal is to teach my students a way of perceiving, of understanding, and of making decisions about the world; I am trying to teach them a scientific worldview, a view that depends upon mathematics as an essential tool, but that does not have mathematics as an end in itself.

Central to a scientific worldview is the recognition of the importance of data and observation. If we are truly interested in the question of global warming, or of the wisdom of using nuclear power, or of the seriousness of the population explosion, it does no good to listen to politicians, or to the opinion makers who litter the airwaves; they operate, largely, in a data-free universe. But these are not political questions. They are questions about the Earth. If we are truly interested in investigating any of these questions, we must ask the Earth. How do we ask the Earth? We engage in the disciplined process of observation, the product of which is data. Only when we have the relevant data can we begin to learn, for ourselves, about any of these issues. Of course, a collection of data, by itself, is only a starting point; once we have the data, we must analyze it, and then interpret it. We can represent this sequence schematically as

observation and data → analysis → interpretation.

As we have seen, the analysis and interpretation of data is non-trivial for students, and my pedagogical voyage has largely been one of discovering that it is important to teach these things, and then learning how to teach them. But it all begins with observation and the collection of data.

While it is true that advanced scientific work requires years of study, the science required to make sense of some of the critical political, social, and environmental issues of the day is easily accessible to our students, and to our citizenry. Straightforward scientific thinking can, and should, be practiced by everyone. Scientific thinking is the disciplined application of common sense principles; it relies on observation and data, and requires a willingness to question, to conjecture, and to test conjectures against the evidence. The data-based integrative writing assignments I have constructed are an attempt to demystify science, to teach scientific thinking, to demonstrate to students that they can think like scientists, and to show them the power of scientific thinking. The integrative assignments illustrate that elementary mathematics, combined with the use of a spreadsheet, offers a powerful tool for the analysis of data. I want our students to understand that they themselves can draw careful conclusions directly from the data, so that they do not need to rely on experts. And I want our students to understand that they don't have to be a scientist to think like one.

## Going Public

I have found these thoughts and results worth sharing. In addition to speaking in the 2010 MAA – AMS Invited Paper Session on the Scholarship of Teaching and Learning (Burke, 2010a), I have presented my insights elsewhere. I wrote a piece for *Carnegie Perspectives* (Burke, n.d.), an online forum for news, views, and perspectives. I also have presented my work (Burke, 2010b) to the Transition Mathematics Project, an initiative designed to help students successfully progress from high school math to college-level math (see [www.transitionmathproject.org](http://www.transitionmathproject.org)). My use of the scholarship of teaching and learning, and most particularly of the systematic gathering of evidence, has enabled me to share with others the knowledge that I gained from my journey.

## Aspirations

Where do I go from here? I find the issues raised by teaching with the data-based integrative writing assignments to be fascinating, compelling, and important. I want to teach a course that is designed around assignments of this type; I would like the data-based integrative writing assignments to be the focus of the course, rather than an interesting supplement. I want to have the time, without the pressures of adhering to a set mathematical curriculum, to fully address the issues I have encountered in response to the assignments, to give the issues the full discussion that they require and deserve.

And I have a final aspiration. When I was a child, I would occasionally visit my father at his office. At age eight or nine, I was puzzled about the meaning of a hand-lettered sign on the wall of his office:

*Observation and Again Observation*

I. P. Pavlov

I have long since come to grips with the sign's meaning, and the sign hangs in my office today. It summarizes one of my primary goals for my students: If I can teach them that, when they want to think about global warming, or using nuclear power, or the population problem, they must begin with careful observation of the world as it is, then I will have achieved this goal. But, the sign also indicates what I must do as a teacher. If I want to make significant changes in what and how I teach, and if I want to understand what is happening in my classroom, then I, too, must begin with observation. And again observation. The scholarship of teaching and learning has provided me with the means to do this.

Any serious look at the world, whether it is a look at an important global issue of the sort my students are investigating or the world of my classroom, must begin with observation. It all begins with the data. I look forward to the day when I see young Americans walking toward me on the Charles Bridge or, more likely, walking toward me across the campus quad, wearing t-shirts that say "Observation and Again Observation."

## References

- AACU (Association of American Colleges and Universities). (2005). Integrative learning. *Peer Review*, 7(4).
- Bain, K. (2004). *What the best college teachers do*. Cambridge, MA: Harvard University Press.
- Burke, M. C. (n.d.). A mathematician's proposal. *Carnegie Perspectives*. Retrieved from [www.carnegiefoundation.org/perspectives/mathematicians-proposal](http://www.carnegiefoundation.org/perspectives/mathematicians-proposal)
- Burke, M. C. (2006). Final project snapshot: Integrative learning in the mathematics classroom. Retrieved from [web.archive.org/web/20080609003449/http://www.cfkeep.org/html/snapshot.php?id=61548260563564](http://web.archive.org/web/20080609003449/http://www.cfkeep.org/html/snapshot.php?id=61548260563564)
- Burke, M. C. (2010a). *Preliminary thoughts on quantitative literacy*. Paper presented at the MAA-AMS Invited Paper Session on the Scholarship of Teaching and Learning in Mathematics, Joint Mathematics Meetings, San Francisco, CA. Abstract retrieved from [jointmathematicsm meetings.org/meetings/national/jmm/1056-be-1510.pdf](http://jointmathematicsm meetings.org/meetings/national/jmm/1056-be-1510.pdf)
- Burke, M. C. (2010b). *Rethinking precollege math: A personnel odyssey* [PowerPoint slides]. Retrieved from Transition Math Project Resources as "Burke Keynote" [transitionmathproject.org/index.php/resources/resources\\_detail/institutes-and-retreats](http://transitionmathproject.org/index.php/resources/resources_detail/institutes-and-retreats)
- Huber, M. T., & Hutchings, P. (2004). *Integrative learning: Mapping the terrain*. Washington, DC: Association of American Colleges and Universities & Palo Alto, CA: Carnegie Foundation for the Advancement of Teaching. Retrieved from [carnegiefoundation.org/sites/default/files/publications/mapping-terrain.pdf](http://carnegiefoundation.org/sites/default/files/publications/mapping-terrain.pdf)
- Huber, M. T., & Hutchings, P. (2005). *The advancement of learning: Building the teaching commons*. San Francisco, CA: Jossey-Bass.
- Huber, M. T., Brown, C., Hutchings, P., Gale, R., Miller, R., & Breen, M. (2007). *Integrative learning: Opportunities to connect*. Retrieved from [gallery.carnegiefoundation.org/ilp](http://gallery.carnegiefoundation.org/ilp)

- Hughes-Hallett, D., Gleason, A. M., McCallum, W. G., Lomen, D. O., Lovelock, D., Tecosky-Feldman, J., Tucker, T. W., Flath, D. E., Thrash, J., Rhea, K. R., Pasquale, A., Gordon, S. P., Quinney, D., & Lock, P. F. (2009). *Calculus: Single variable* (5th ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- Kezar, A., & Elrod, S. (2012). Facilitating interdisciplinary learning: Lessons from Project Kaleidoscope. *Change: The Magazine of Higher Education*, 44(1), 16–25.
- Meier, J., & Rishel, T. (1998). *Writing in the teaching and learning of mathematics* (MAA Notes #48). Washington, DC: The Mathematical Association of America.
- National Leadership Council for Liberal Education and America's Promise. (2007). *College learning for the new global century*. Washington, DC: Association of American Colleges and Universities. Retrieved from [www.aacu.org/leap/documents/GlobalCentury\\_final.pdf](http://www.aacu.org/leap/documents/GlobalCentury_final.pdf)
- Steen, L. A. (2004). *Achieving quantitative literacy: An urgent challenge for higher education* (MAA Notes #62). Washington, DC: Mathematical Association of America.

## Appendix: Reflection Prompts

### Questions Used with Various Assignments

- This assignment required you to integrate mathematical thinking into the paper. What aspects of this were most challenging for you? Please be as specific as possible.
- What aspect of this paper are you most satisfied with, and why? (For example, do you think you have written an interesting introduction? Is the mathematical model clearly explained? etc.) Is this aspect a new development in your writing?
- Did you find this paper easier to write than the other two? Explain why or why not.
- Looking back over the semester, what have you gained from the integration of mathematical modeling and writing into this course? How has it affected your writing? How has it affected your understanding of mathematics?

### Questions Specific to the Irish Population Assignment

- This assignment presented you with a situation in which you were to construct a mathematical model that, in the end, failed rather spectacularly. Did you have a high degree of confidence in your model, initially? Were you surprised when you first saw the second set of data and realized how badly your model projected the population of Ireland throughout the twentieth century? Now that you have completed the assignment, what implications do you see for your work with mathematical models?
- You were required, here, to examine ways in which a mathematical model might break down. In particular, you were asked to think about the conditions underlying the population trend that your model described, and then to identify which of these conditions failed, causing the breakdown of your model. Was this kind of analysis new to you? Was it difficult for you? What strategies helped you to sort through these issues and organize your paper?

# 12

## Presenting Evidence for the Field That Invented the Randomized Clinical Trial

John Holcomb  
*Cleveland State University*

### Editors' Commentary

In this chapter John Holcomb raises difficult questions about the ethics of experimental design when investigating questions about student learning, which led him to forego the traditional approach of using a control group. This decision resulted in difficulties with getting his work published in statistics education journals, despite the work having received extremely positive feedback at conferences. He also discusses the meaning of reliability and validity in measurement and provides an example of a simple scheme for analyzing open-ended responses to surveys.

### Introduction

My SoTL study involved developing data analysis group projects for an introductory statistics course that required students to write technical reports. To determine if students were learning the desired skills with these projects, I designed take-home mid-term and final examinations where students were given individualized data sets and investigative questions. In contemplating how to present evidence of effectiveness, I thought a great deal about the gold standard of evidence in statistics, namely that to conclude that a treatment causes an effect, subjects have to be randomly assigned to a treatment or a control group. In education, this is not always possible. Even when the same instructor teaches both the treatment and control, he can bias the results, consciously or unconsciously. My anxieties about my study lacking a control group were well founded because I experienced difficulty in publishing my results in a statistics education journal. I believe my paper was held to a different standard from that for other articles. This chapter discusses the decisions I made concerning what evidence to collect, the reviewing process, and the contradiction that my work was well received at regional, national, and international conferences, but never was published in a peer-reviewed journal.

### Beginning Ideas

In the late 1990s, when I submitted my application to the Carnegie Academy for the Scholarship of Teaching and Learning (CASTL) program, mathematics departments were developing courses to improve the quantitative literacy of their undergraduate students. These efforts involved fostering student learning of mathematical concepts and communicating them orally and in writing. One course that received a great deal of attention was introductory statistics.

A colleague and I had developed a sequence of six data analysis projects for introductory statistics that utilized real data, computer software, collaboration, and writing. In the projects the students investigated a random sample of

infant births in North Carolina. They applied data analysis techniques learned in class to study relationships among variables such as the birth weight, weeks of gestation, age of the mother, her marital, smoking, and alcohol-consuming status, her weight gain, the gender of the infant, and whether the infant was a singleton birth or one of twins, triplets, etc. Self-selected teams of two to four students submitted a single report for the team to be graded. The project grade depended equally on writing quality and mathematical accuracy. Each team member signed a cover sheet attesting to his or her contribution to the project.

We published an article describing this approach in *The American Statistician (TAS)*, a general readership journal of the American Statistical Association (Holcomb & Ruffer, 2000). The article contained learning objectives for the projects, the rubric we used to evaluate them, and the results of a student survey about whether the students felt the projects helped them learn statistics. The article did not include an assessment of student learning.

During the review process, a referee suggested that we undertake a formal assessment of the impact on student learning of using these projects in introductory statistics. This referee stated that it would be of interest to the mathematical and statistical community. Therefore, the project I proposed for the Carnegie Scholars Program was to make a quantitative and qualitative analysis of the impact of the projects on student learning.

Reading that original CASTL application from over ten years ago makes me realize that there was a significant difference between what I had initially hoped to do and what I did. As a statistician, I was trained to think that the only way to prove that an outcome is caused by a treatment is to implement a randomized experiment. In this setting, this would have involved randomly assigning students to two classes, with a project-based approach in one class and not in the other, having the same instructor teach both sections, and then determining if the results are better in one class than another. Using the taxonomy of SoTL questions (Hutchings, 2000, p. 4) discussed in Chapter 2, I was envisioning a *What works?* investigation. I wanted to collect and present evidence that a project-based instruction method in introductory statistics “worked better” than a class taught with traditional pedagogy.

## Refining the SoTL Question

I have not yet defined what the word “better” meant. In fact, determining this was the first step in my SoTL-study process: I wanted to know if the six-project sequence worked, so I now had to ask, “Do team writing projects using real data lead to individual acquisition of desired skills?” Because I had written learning goals for the statistical projects in the *TAS* article, I had a list of the desired skills. What I had not done was determine how to assess whether students were achieving the learning goals.

In my mind, the two most important learning goals were that students learn how to

1. Perform a thorough summary analysis that included graphics, summary statistics, and their respective interpretations.
2. Formulate statistical hypotheses, formally test them, and then interpret their results, including an understanding of the limitations of those tests.

These are common learning goals; almost every introductory statistics course instructor hopes that students will achieve them. What was a challenge for me in developing this study was how to assess them. I felt, and still feel, that standard examinations in introductory statistics rely too heavily on calculations. Additionally, students are rarely able to perform a complete hypothesis test from start to finish (depicting the data visually, calculating summary statistics, making hypotheses, performing the test, and interpreting results) in a standard exam setting without statistical software.

While wrestling with the question of assessment, I became aware of “authentic assessment.” This concept has been described in various ways. Colvin and Vos (1997) used the term to describe “measuring student performance on tasks that are relevant to the student outside of the school setting” (p. 27). Archbald and Newmann (1988) stated, “A valid assessment system provides information about the particular tasks on which students succeed or fail, but more important, it also presents tasks that are worthwhile, significant, and meaningful—in short, *authentic*” (p. 1). I felt that the homework data analysis projects described in *TAS* were in fact authentic assessments. Thus, it seemed logical that that I needed to test students in a similar manner.

I was now motivated to design take-home components to the mid-term and the final examinations where students would complete a full data analysis on a data set with a structure similar to the homework team projects. In point of fact, I developed 18 individualized data sets and questions and gave each exam to two students presented in a way that suggested each student had received a unique exam.

There were several reasons for adopting an assessment that made individuals accountable. In previous teaching experiences using the project sequences, I felt that some students were not learning the material since they allowed other team members to do most of the work. My hope was that the knowledge that each student would have to complete their own midterm and final examination with a unique data set would motivate all students to participate in the team homework projects.

Because introductory statistics is a required course for many majors, it is reasonable to assume that students will be expected to do statistical analyses in subsequent courses. By assigning individualized data projects, I would have a written record of what students are capable of doing as they left my class. I could use this as evidence in discussions with colleagues from the client departments of the usefulness of the course for their students. The fact that students would complete additional data analysis in subsequent coursework further justified my claim that the study employed an authentic assessment.

I was concerned that students would be tempted to consult their former teammates. Some might argue that giving each student an individual assignment after group work will send a message that group work is really not important. I believe the dual approach is similar to the way in which people are evaluated in the workforce—employees work in teams, but have individual performance reviews. The teamwork is valued, but it is individual performance, within and apart from the team, that determines the pay increase or bonus amount.

My initial thought was to administer the take-home examination once at the end of the semester. After consulting colleagues who were part of the 2000 cohort of CASTL scholars, I modified my plan so as to break up the assignment into two pieces. The first individualized take-home assignment occurred half way through the semester and required students to graphically and numerically summarize data and interpret their results. The final examination had students work on the same data, performing several tests of hypotheses and interpreting the results. This was an excellent suggestion by my colleagues and it shows how consultation with other colleagues, even those outside of the discipline under investigation, can enhance a SoTL study.

I now had two learning assessment tools that I believed would show me if the data analysis homework assignments completed in teams were effective in fostering individual competence in data analysis. I thought I was ready for the randomized study. What I was not ready for was an increase in my own misgivings about administering it.

## Ethical Quandary

As I began to think about initiating the study, I considered how things would appear to a student. I imagined myself as a student enrolled in a traditional section of introductory statistics who was given a data analysis project at midterm. The question, “Would such an assignment be fair?” plagued me. I came to believe that giving students an assignment of this magnitude and of such stakes without providing them sufficient practice, in the form of homework, would be unfair and unethical. It would be analogous to teaching a person new to tennis by drilling them on serving, volleying, and hitting from the baseline and then requiring that person to play a match for the first time expecting him to perform admirably against an experienced player.

I considered making the take-home individualized assignments voluntary for the students in the traditional section, but that bothered me too. I felt that students would not take a voluntary assignment seriously enough for any kind of meaningful performance assessment—especially one for which there was little previous practice.

I wondered who would teach both sections (a project-based course and a traditional course). Having two different instructors introduces a confounding factor, so it seemed I, or someone else, would have to teach both sections. As the instructor of both sections, I could, albeit unconsciously, sway the results in the direction I wanted. The obvious solution was to have a disinterested statistics instructor teach both sections. But then I would need to find such an instructor. At that time, I was a new faculty member at Cleveland State University who did not know faculty members within my own department well. I wondered whether colleagues would be against a project-rich course, consciously or unconsciously, since it was new and different.



These issues forced me to think deeply about my SoTL study. Discussing the challenges with colleagues in the CASTL program led to a chapter about my dilemma being included in *Ethics of Inquiry: Issues in the Scholarship of Teaching and Learning* (Holcomb, 2002). It presents a more complete development of the ethical issues involved and commentary from two SoTL scholars and one of the leaders in statistics education research.

Eventually, I came to envision my study more of what Hutchings (2000, p. 4) describes as a “vision of the possible,” a *What could be?* SoTL study, rather than a *What works?* investigation. What I now set out to do was to present a project-rich course that included individualized assessments to show that team data analysis projects could lead to individual competence in data analysis. To provide evidence, I created and administered 18 individualized mid-term and final take-home examinations using real data.

Using different variables or different observations made the exams individualized. For example, one set of six individualized exams involved data on students’ eating habits in the first year of college. The variables of the study included gender, status as a campus resident or athlete, height and weight at the beginning and end of the spring semester, and fruit, vegetable, and calorie intake. The original study included approximately fifty variables, so it was easy to design six projects that included different subsets of the variables. A second set of six individual examinations came from a study of the taste of meats cooked in taco flavoring. Here participants rated emu, turkey, and beef for appearance, flavor, texture, tenderness, aftertaste, and overall satisfaction. Each exam had students investigate a separate component of these attributes and its relationship to gender, age, income level, and education level of the participant. A set of six other projects came from a study of bone mineral density in elderly subjects screened for osteoporosis. Here the variables were the same for each project (age, gender, bone mineral density, height, weight, calcium intake, treatment), but each individualized exam had a different set of 250 subjects chosen from the 2000 participants. I graded the examinations with the same rubric that was used to grade the team homework projects. I compiled the results to determine what percentage of the students earned a score of 80% or higher on the examinations.

## Additional Learning Goals

In addition to determining if students were learning how to do statistics, I was also interested in two other learning goals: Students would

1. Acquire skills in working with others, and
2. Realize that statistics and statistical ideas have a major impact on their everyday lives.

Assessing these goals was an evolutionary process. I wanted students to answer approximately eight questions after each team homework project. Since I needed Institutional Review Board (IRB) approval for the study, I informed the students that answering the reflection questions was not part of their final grade, and they had the right of not participating in that aspect of the study. The next class day, only four forms were returned. The low response rate occurred even though the form did not request any identifying information. It might have been helpful to explain to the students that providing feedback would benefit future students, but that did not occur to me at the time. After consulting with my colleague who had reviewed the IRB application, we decided to shorten the questions. I then gave students ten minutes of class time to complete the revised reflection questions. Again, they were told filling out the form was voluntary. This time, nearly every student attending class filled out the form. A better approach might be to require students to keep a journal as part of the class assignments. Since the reflection would then be embedded into the coursework, the students would not have the option of refusal.

After group projects two, three, and four, I wanted to inquire: “How did your group complete the project?” and “Did the group interact well?” The first question was addressed by a series of prompts: “Describe how your group completed the assignment. Did you all meet? Did you all do it individually and then discuss? Did you talk via phone or email?” The responses varied greatly, and it was difficult to glean any useful information. The second inquiry was phrased as: “Describe the group interaction. Did your group ‘work successfully’? Did you feel comfortable in your group asking concept or technical questions? Did you feel that any members of the group dominated the discussion and the completion of the assignment?” Again, the format made obtaining clear answers difficult.

After the fourth project, I wanted data on the effectiveness of working in groups on the projects. This time the reflection prompts were: “Describe your overall evaluation of working in groups in and outside of class for this course.

Do you believe that it helped you learn statistics? Do you think it helped you learn communication skills? Did the group provide people with whom you could study or ask questions? Did you ever ask a group member questions outside of class about a topic that confused you?" I should have given space to answer each of the questions, but I did not. They were written together, as shown. Analyzing the responses was challenging because students chose to answer only some of the questions. I decided to tabulate the responses according to the overall tone of the response regarding group work. I sorted the responses into three categories: positive (respondent listed benefits from working with the group), negative (respondent claimed working individually would have been better or they had not benefited from group work), and neutral.

A focus group or an online survey might be a better way to ask questions of this nature. However, even with my crude coding of the data, the results proved enlightening. My perception had been that the groups were not working well, yet I characterized 66% of the responses as positive. I had been considering not using the groups in the future when teaching introductory statistics, but because of this result I have continued to use groups. It seems that I was falling prey to selection bias. Apparently, I was hearing only about the problems with the groups, and not realizing that most groups were working fine and that the majority of students were having a positive experience.

After the take-home midterm examination, I gave the students a five-question "Project Reflection" questionnaire to determine if they felt they were adequately prepared for the midterm. I was concerned that if students were not fully engaging with their team on the homework they would be worried about the individual take-home examinations. Twenty-seven students were in class and returned the questionnaire. My analysis of the results indicated the vast majority of students were not threatened by the exams, indicating they were adequately prepared by completing the team projects.

As a statistician, I was uncomfortable using qualitative tools to assess the secondary goals of the study. These methods, however, proved to be an important component of the investigation. I had been concerned about the student perceptions of group interaction, and the experience of taking an individualized examination to test what they had practiced in groups. It was the qualitative evidence that showed students were working in teams with less conflict than I believed, and that they did not find working on individual exams to be traumatic. The surveys even convinced me to seek funding from the department to purchase better statistical software.

## Going Public

I presented the results of my study at national statistics conferences and at colloquiums in mathematics or statistics departments. The Statistical Education Section of the American Statistical Association sponsors sessions at the Joint Statistical Meetings (JSM). In 2002, I presented the results of my study in a well-attended session at JSM. I was an invited speaker at a sixty-person conference on statistics education in Appleton, Wisconsin sponsored by an NSF-funded grant related to assessment in statistics education. I also presented my work at the First Annual Joint UK and USA Conference on the Scholarship of Teaching and Learning that was co-sponsored by the University of East London and City University of London, London, England, in June 2001. This presentation led to an interview with a journalist for *The Times Education Supplement* followed by an article on my SoTL study (Holcomb, 2001). All of the presentations garnered positive feedback. In addition, I believe that my SoTL investigation and presentations on its results were a major factor in my receiving the 2003 Waller Award from the American Statistical Association for outstanding teaching of introductory statistics by a faculty member in the first ten years of his or her career.

I also submitted an article for publication in the ASA publication *The Journal of Statistics Education (JSE)*. I had served as a referee and an associate editor for the journal for a number of years, so I was familiar with its review process. The journal's mission is to publish manuscripts on innovative approaches to teaching statistics at the K-12, undergraduate, and graduate levels. In my view, many of the papers in this journal were show and tell papers containing little assessment of the innovation's effectiveness. Although throughout my Carnegie experience, other CASTL scholars often expressed concern that they would not be able to publish their work in the journals of their discipline, I was confident that *JSE* would publish mine.

I was shocked when my submission was rejected. Even after eight years, I am still upset. The main criticism was that I needed to perform a randomized study to claim that the project-rich curriculum was effective in teaching students to do data analysis. Although I felt I had clearly indicated that I was not trying to prove anything, but merely show that it is possible for students in a project-rich course to achieve authentic data analysis skills, the reviewers did not seem to

grasp this. The second reason for my surprise and anger was that it seemed to me that the reviewers held my paper to a higher standard than others. In my mind, the mix of quantitative and qualitative methods was a more comprehensive analysis of effectiveness than found in many papers previously published in the journal.

The second major criticism of the paper was that using the mid-term and final examinations as evidence of success was not appropriate because the exams had not been shown to be reliable or content validated. Reliability and validity are concepts within the education testing literature that I was only vaguely familiar with. According to Reynolds, Livingston, and Willson (2006), “in simplest terms, in the context of measurement reliability refers to consistency or stability of assessment results” (p. 86). Also, content validity “involves how adequately the test samples the content area of the identified construct . . . Content validity is typically based on professional judgments about the appropriateness of the test content” (p. 120).

To my knowledge there was no available instrument for statistical data analysis that had been developed and tested for reliability and validity. My view was that assessing the take-home examinations for reliability and validity was a separate study that could take years to complete. Also, the take-home examinations and the grading rubric I used did not lend themselves to the usual means for testing reliability. A very common method of assessing reliability is to use split-half reliability, where a test is split into two equivalent halves that are scored independently. The results of one half of the test are then correlated with results from the other half by calculating the Pearson correlation (Reynolds, Livingston, & Willson, 2006). This was problematic as the questions on the take-home examination differed in point values for each part of the test and the statistical accuracy of the results comprised only 50% of the grade. The writing component graded on the rubric was impossible to split into equal halves.

In my opinion, the dearth of instruments that have been adequately tested for reliability and validity in mathematics and statistics education at the college level is a major barrier for SoTL research and publication. In statistics education, I know of two instruments, the Students Attitudes Toward Statistics (SATS) (Schau, Stevens, Dauphinee, & Del Vecchio, 1995), which measures students’ attitudes toward statistics, and the Comprehensive Assessment of Outcomes (CAOS) test (delMas, Garfield, Ooms, & Chance, 2007) which measures the understanding of statistical concepts for college-level introductory statistics. Neither is appropriate for measuring student understanding and ability to implement data analysis.

After receiving the rejection, I undertook a second study of the data projects. I did not engage in a randomized study for the same ethical reasons as previously stated. I did have a comparison class, not taught by myself, and I surveyed my own students and the comparison class students using the SATS instrument. I also had a colleague evaluate the mid-term and final examinations for appropriateness and to assess student performance. With a colleague using the rubric to evaluate the student performance as well, our goal was to test for reliability by assessing inter-rater reliability. Inter-rater reliability refers to evaluating the degree of agreement when different individuals evaluate subjective material. A correlation is calculated between the scores of the two evaluators (Reynolds, Livingston, & Willson, 2006).

I presented the results on student attitudes in statistics at an invited breakout session at the US Conference on Teaching Statistics in 2007. I never completed a new journal manuscript for submission, fearing rejection this time on the grounds that differences between two classes on the SATS could be explained by the confounding variable of two different instructors. I was also concerned that a new set of reviewers would miss the point that the study was intended to demonstrate it is possible for a course taught with projects to lead to a change in statistical attitudes compared to a traditional class, just as the last set had failed to see the goal of the previous study was to show it is possible for a project-rich course to be effective. This apprehension discouraged my colleague and me from even completing our inter-rater reliability although I believe that approach would have allayed concerns of reliability. The content validity of the take-home examinations would still need to be addressed.

## Final Thoughts

Working on my SoTL study was rewarding and informative. At the course level, I believe I acquired compelling evidence that a project-rich course can lead to student competence in authentic data analysis. I also learned from qualitative evidence that the majority of students found working in teams helpful and that they were adequately prepared for the take-home examinations. Based on the feedback received in this study, I was able to improve the team projects and the take-home examinations. As a result, every time I have taught introductory statistics since the original

SoTL study, I utilized the individual take-home examinations. I am, however, disappointed that I did not succeed in publishing a peer-reviewed manuscript within my discipline on the study.

If I were to do a similar investigation, I would not change much. I would, however, write the manuscript submitted to a statistics education journal differently. I would attempt to frame the paper even more so as a “vision of the possible” course. I might also suggest to the editor reviewers who I think would be open to examining evidence that does not prove causality, but suggests a successful instructional strategy.

Despite the lack of a published paper, the impact on my career from this work has been tremendous. In addition to the Waller Award mentioned earlier, I have been involved in several national efforts in statistics education. I was asked to join the Research Advisory Board for the Consortium for the Advancement of Undergraduate Statistics Education ([www.causeweb.org](http://www.causeweb.org)). This body has advocated for more research into the learning of statistics at the K-16 level. I also served on the advisory board for the Assessment Resource Tools for Improving Statistical Thinking ([app.gen.umn.edu/artist/index.html](http://app.gen.umn.edu/artist/index.html)) project that developed an on-line library of assessment items for introductory statistics and the CAOS instrument described earlier.

The process of undertaking a SoTL study has also made me more respectful of the work of educational researchers, and in particular, mathematics and statistics education researchers. With this new perspective, I have been able to build useful on-campus and off-campus networks that have benefitted my career.

## References

- Archbald, D., & Newmann, F. (1988). *Beyond standardized testing: Assessing authentic academic achievement in the secondary school*. Reston, VA: National Association of Secondary School Principals.
- Colvin, S., & Vos, K. E. (1997). Authentic assessment models for statistics education. In I. Gal & J. B. Garfield, (Eds.), *The assessment challenge in statistics education* (pp. 27–36). Amsterdam, Netherlands: IOS Press.
- delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students’ conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), 28–58.
- Holcomb, J. (2001, July 27). How emu meat can solve your problems. *The Times Higher Education Supplement*, pp. 23.
- Holcomb, J. (2002). The ethics of comparison: A statistician wrestles with the orthodoxy of a control group. In Pat Hutchings (Ed.), *Ethics of inquiry: Issues in the scholarship of teaching and learning* (pp. 19–26). Menlo Park, CA: Carnegie Foundation for the Advancement of Teaching.
- Holcomb, J., & Ruffer, R. (2000). Using a term-long project sequence in introductory statistics. *The American Statistician*, 54(1), 49–53.
- Hutchings, P. (2000). *Opening lines: Approaches to scholarship of teaching and learning*. Menlo Park, CA: Carnegie Foundation for the Advancement of Teaching.
- Reynolds, C., Livingston, R., & Willson, V. (2006). *Measurement and assessment in education*. Boston, MA: Pearson.
- Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement*, 55(5), 868–875.





# **Illustrations of SoTL Work in Mathematics**

**Theme 3: Using Assigned Reading Questions  
to Explore Student Understanding**



# 13

## Conceptual or Computational? Making Sense of Reading Questions in an Inverted Statistics Course

Derek Bruff  
*Vanderbilt University*

### Editors' Commentary

In this chapter Derek Bruff describes his project, trying to get students to read a mathematics textbook more effectively. His report discusses working with existing data, identifying patterns within qualitative data, and trying to approximate a control group experiment within a single course. He also obtained a surprising result that has implications for the teaching of conceptual and computational material in statistics courses. He tells us that his overall plan was inspired by a SoTL project by [Axtell and Turner \(2006\)](#), who have a chapter on that work in this book (Chapter 14). This illustrates a key tenet of SoTL – that work should be made public so that others can build on it. His project also illustrates the interplay between the *What is?* and *What works?* questions in the SoTL taxonomy.

### Introduction

While teaching an introduction to statistics for undergraduate engineering majors, I asked my students to read their textbooks before class. I held them accountable by giving pre-class reading quizzes consisting of two or three open-ended questions about the reading and a “muddiest point” question that asked students what they had found most difficult or confusing in the reading. I wanted my students to come to class familiar enough with the course content so that they could spend class time on activities designed to foster deep learning of the material. After my first semester teaching in this way, I wondered what kind of material my students were able to learn before class from reading their textbook. I also wondered if there were types of pre-class reading quizzes that might help my students learn more from their textbook so that we could progress even further with the material during class. This led me to analyze my students' responses to the pre-class reading questions and eventually to conduct a quasi-control group experiment.

During my investigation I faced the challenges of working with existing data, of identifying patterns in qualitative data, and of conducting something approximating a control group experiment within a single course. My project helped me develop a deeper understanding of the importance of notation in learning mathematics and the integration of conceptual and computational material in a statistics course. It also provided me with experiences that proved useful later in my work with instructors from different disciplines interested in conducting their own SoTL projects.



## Course Context

During 2006 and 2007, I taught a course titled “Probability and Statistics for Engineers” with enrollments of 36 and 56 students, respectively. Almost all the students in the course were engineering majors, many of them juniors or seniors. Most of them had no particular interest in or experience with probability or statistics. The course was a one-semester conceptual and procedural introduction to probability and statistics. I favored depth over breadth, focusing on a few key concepts, distributions, and techniques. I designed the course around the idea that many processes produce variable results that can be quantified to facilitate decision-making. Wiggins and McTighe (2005) referred to a main idea that is meant to persist in the students’ minds long after the course is over as an “enduring understanding.”

I taught the course using what is often called the “inverted” or “flipped” classroom approach. Mazur (2009) noted that in a traditionally taught course, class time is used to transfer information from instructor to students through an instructor’s lecture, leaving students to assimilate that information after class as they struggle with their homework. He argued that assimilation is the harder of the two, and so it should happen during class when other students and the instructor are available to help. The easier phase—the initial transfer of information—can happen before class, as students read their textbooks. This approach inverts the traditional teaching model for mathematics and the sciences. Most of the active learning exercises my students experienced during class were peer instruction activities (Mazur, 1997) facilitated by a classroom response system (Bruff, 2009).

Having students read their textbooks before class is easier said than done. Research has indicated that merely asking students to read before class results in a 20–30% compliance rate (Hobson, 2004). Knowing that, I decided to follow in Mazur’s footsteps and to hold students accountable by giving pre-class reading quizzes. Prior to each class session, I instructed students to read certain pages in their textbooks and complete an online quiz administered through a course management system. Each quiz consisted of two or three open-ended questions about the reading, and a muddiest-point question that asked students to describe one or more difficulties they had encountered with the reading. I graded student responses entirely on effort, not accuracy, since I did not expect them to master the material solely by reading their textbooks. The pre-class reading quizzes comprised 5% of the students’ grades in the course. Since they were due several hours before class started, I was often able to scan the responses and adjust my lesson plans to be more responsive to their learning needs, thereby practicing just-in-time teaching (Novak et al., 1999).

## From *What Works?* to *What Is?*

After teaching this course once using the inverted classroom approach, I felt that I had been able to make good use of class time. Students learned what they could on their own before class from their textbooks, and I spent class time helping them go deeper into the course content. I wondered, however, if I might be able to optimize the process. I had read in McKeachie (1999) that providing students with guiding questions can help them make sense of a reading assignment, but it was not clear to me what kinds of pre-class reading questions would be most effective. My interest in improving the process led to my initial research question:

Research Question 1: What kinds of pre-class reading assignments, including questions about the reading, might help students in an introduction to statistics course for engineering majors learn more from reading their textbooks before class?

While considering how to investigate this question, I ran into a problem. I realized that I would not be able to determine if a different type of pre-class reading assignment helped students to learn more from reading their textbooks without knowing what they were capable of learning from their textbooks on their own. I lacked a way to assess what students learned from reading their textbooks, which meant that I had no way to tell if one kind of pre-class reading assignment worked better than another.

The research question I had identified was what Hutchings (2000) calls a *What works?* question, one that “seek[s] evidence about the relative importance of different [teaching] approaches” (p. 4). These questions are important to investigate, but, as I realized in this project, it is sometimes impossible to answer a *What works?* question without first answering what Hutchings calls a *What is?* question. Hutchings (2000) writes of *What is?* questions:

Here the effort is aimed not so much at proving (or disproving) the effectiveness of a particular approach or intervention but at describing *what it looks like*, what its constituent features might be. Investigations of

this descriptive type might, for instance, look at the dynamics of a class discussion around a difficult topic; they might be efforts to document the varieties of prior knowledge and understanding students bring to a particular topic or aspect of the discipline. (p. 4)

I had been made aware of Hutchings' work while attending the 2006 Carnegie Academy for the Scholarship of Teaching and Learning (CASTL) Summer Institute held at Columbia College Chicago, so I was in a position to deal with this challenge. I formulated a *What is?* question as Research Question 0, the answer to which would enable me to investigate my *What works?* question:

Research Question 0: In an introduction to statistics course for engineering majors, what are students able to learn by reading their textbooks before class?

I learned later, while helping Vanderbilt graduate students conduct SoTL projects as part of their participation in the Teaching Certificate Program my teaching center runs, that it is common for SoTL investigators to begin with a *What works?* question, then shift to a *What is?* question that must be answered first.

Regarding my "What is?" question, I formulated a two-part conjecture that (1) students are most able to master computational procedures (e.g., "Find the probability that an observation of a normal random variable is more than one standard deviation from the mean") through reading their textbooks before class, but that (2) students have more difficulty learning concepts and relationships (e.g., "If the standard deviation of a normal distribution is doubled, how is the graph of the distribution changed?") from their textbooks. Assuming my conjecture to be true, my hypothesis for the *What works?* question (Research Question 1) was that asking students conceptual or "big picture" questions on pre-class reading quizzes would help them learn more from their textbooks than asking them computational questions, since conceptual questions would target the concepts and relationships that students find difficult to learn from textbooks. I thought that students might not be able to fully answer tough conceptual questions on the basis of their pre-class reading, but that having students try would better prepare them to master concepts during class.

Before testing the latter hypothesis, however, I needed to address the former one. So during the summer of 2006 I began exploring data I had available from the prior semester.

## Working with Existing Data

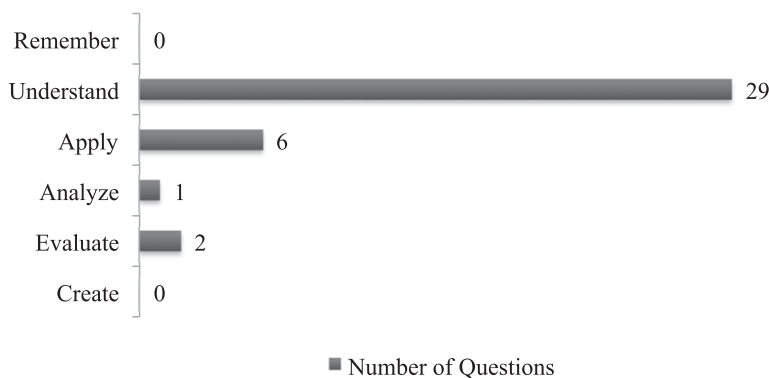
Although I had not set out the previous semester to answer any research question, because of the online quiz format, I had, in fact, captured significant amounts of evidence of student learning. My 36 students had been given 15 pre-class reading quizzes, each consisting of three or four open-ended questions about the reading, yielding close to 1,500 student responses that I could analyze to determine what my students were able to learn by reading their textbooks. (The average completion rate for the quizzes was 78%.) I applied for and received approval from my campus institutional review board to conduct research on the responses, submitting under the "collection or study of existing data" category.

Setting aside my students' responses to the muddiest-point question on each quiz, I began by considering the content questions, giving me a total of 38 questions to analyze. During the semester, I had graded my students' responses to these questions only on effort, but perhaps if I went over their responses again and assessed them on accuracy, I would get a sense of what kinds of questions students were able to answer correctly based on reading alone. Were they, as I had hypothesized, more able to answer computational questions than conceptual ones?

I was familiar with Anderson and Krathwohl's revised Bloom's taxonomy of educational objectives (2001), and I thought its cognitive process dimension (remember, understand, apply, analyze, evaluate, and create) would be useful for categorizing the content questions on the reading quizzes. I coded the questions according to Bloom's six categories. Here are three examples of these questions.

Sample Question 1. Suppose you have seven one-hour television shows recorded on your TiVo and two hours of free time tonight to watch TV. Suppose you wanted to know how many different pairs of shows you can watch tonight. Is this a question about permutations or about combinations?

Sample Question 2. Monty Hall offers you a choice of three doors. Behind two are goats, and behind one is a brand new car. After you make the choice, he opens one of the doors you didn't choose to reveal a goat, and offers you the chance to switch your choice to the one remaining. Should you switch? Why or why not?



**Figure 13.1.** Coding of spring 2006 pre-class reading content questions by Bloom's taxonomy

Sample Question 3. Consider the continuous random variable  $X$ , the weight in pounds of a newborn baby born in the United States during 2006. Suppose that  $X$  can be modeled with a normal distribution with mean  $\mu = 7.57$  and standard deviation  $\sigma = 1.06$ . If the standard deviation were  $\sigma = 1.26$  instead, how would that change the shape of graph of the probability density function of  $X$ ?

Some might code these questions differently, but I coded questions 1 and 3 as “understand” questions since they asked students to demonstrate understanding of particular concepts. I coded question 2 as an “apply” question since it is best answered by applying the technique of decision trees discussed in the textbook. Question 2 is not an easy application question, but I did not feel that it asked students to deal with the relationships among multiple concepts (as an “analyze” question would), nor did it require students to make judgments among multiple options using the standards of the discipline (as an “evaluate” question would).

According to this coding scheme, what kinds of questions had I asked my students on their pre-class reading quizzes? As indicated by Figure 13.1, 29 out of 38 or slightly more than three-quarters of the questions belonged to the “understand” level of Bloom's taxonomy. This reflected my belief that students would need the most help with conceptual material from the readings, so I had asked them questions about concepts to help guide their reading. The distribution of questions was unfortunate for my SoTL investigation; I would not be able to argue that students were more capable of answering one type of pre-class reading question than another because most of the questions were of the same type. One of the drawbacks of working with existing data is that such data are not always rich enough to answer questions of interest. Lesson learned: next time, I would ask a wider variety of types of content questions on pre-class reading quizzes.

At this point, feeling somewhat defeated, I turned to the difficulties my students reported having with the reading assignments through their responses to

Muddiest-Point Question. Please write about the part of the reading that you found most difficult. Try to be as specific as possible. If you did not find the reading difficult, write about the part you found most interesting.

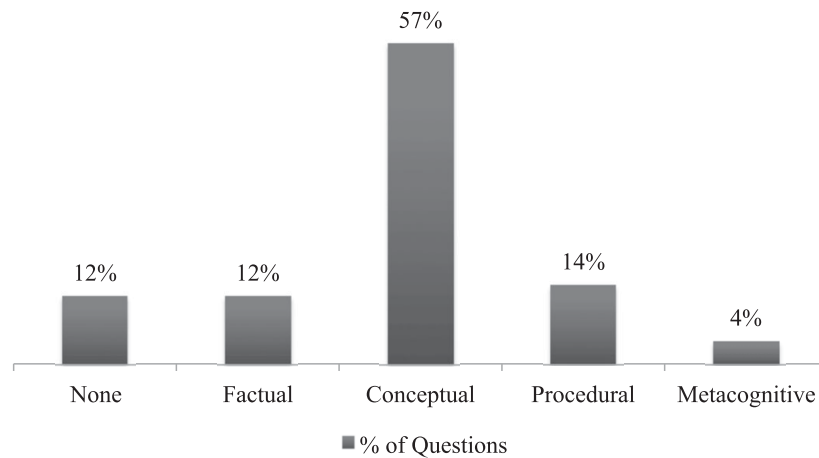
This time I categorized student responses using the knowledge dimension of Anderson and Krathwohl's revised Bloom's taxonomy, which meant determining for each response if the student was struggling with factual, conceptual, procedural, or metacognitive knowledge. Here is an example response in each category:

Factual Knowledge: “The notation in the nonstandard normal distributions was a little confusing. What does  $\phi$  stand for?”

Conceptual Knowledge: “The entire conditional probability section wasn't very intuitive. Examples worked out in class should help.”

Procedural Knowledge: “Determining the rejection region and then finding the corresponding values of  $\alpha$  and  $\beta$  was a difficult process for me to understand simply from the reading.”

Metacognitive Knowledge: “I don't really get this stuff at all.”



**Figure 13.2.** Spring 2006 muddiest points on conditional probability

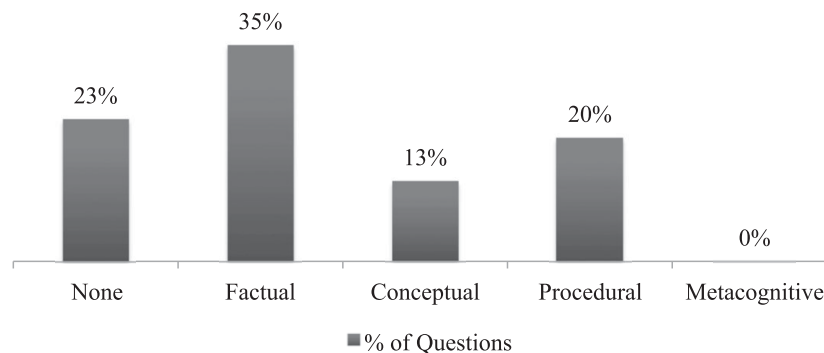
Typically, the student responses that I coded as factual knowledge questions were questions about notation. Coding student responses about concepts and procedures was relatively straightforward. Often educators consider metacognition a relatively high-level learning goal. However, I coded student responses that indicated general confusion as metacognitive. These came from students who did not understand their own learning (or lack thereof) well enough to identify a specific question.

I coded and counted my students' responses to the muddiest-point questions for each of the content topics (e.g., conditional probability, normal distribution, hypothesis testing) addressed in the 15 reading quizzes. This yielded a profile of the kinds of difficulties students had with each topic. Figures 13.2, 13.3, and 13.4 show three profiles.

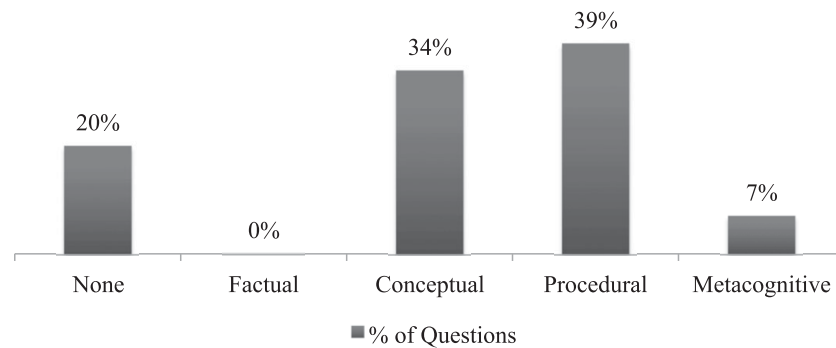
One interesting finding was that for some topics students were confused about factual aspects of the reading. This meant, in most cases, that they did not understand the notation used. (Consider the abundance of Greek letters introduced in a typical textbook section on normal distributions.) This surprised me since students with notational questions could have looked up notational meanings in the textbook as they were doing the readings. Another lesson learned: notation can be a challenge for students, even when they have their textbooks in front of them.

The topic profiles provided an answer to my *What is?* research question. Instead of communicating what students can learn from reading their textbooks, the profiles showed what students struggle to learn. What they struggle to learn varies by topic. Some topics, like conditional probability, pose conceptual challenges to students. Others, like hypothesis testing, are conceptually and procedurally challenging. For some topics, notation trips students up.

The topic profiles suggested a way to address my *What works?* research question. Imagine asking students a series of content questions on a pre-class reading quiz on hypothesis testing that lead to a different profile of muddiest points, perhaps one in which there were few conceptual difficulties identified by students. The shape of the profile would



**Figure 13.3.** Spring 2006 muddiest points on normal distributions



**Figure 13.4.** Spring 2006 muddiest points on hypothesis testing

provide a measure of the effect of different pre-class reading assignments, allowing identification of reading questions that are more or less effective. This was a promising idea, but I found a different way to tackle my *What works?* question in the spring 2007 offering of the course, one that used the results and lessons learned from my analysis of the existing (spring 2006) data.

## A Quasi-Control Group Experiment

Given the homogeneity I had discovered in the kinds of content questions on the pre-class reading quizzes, I modified the quiz format for the spring 2007 semester. Each quiz had four questions: a notational question, a conceptual question, a computational question, and a muddiest-point question asking students to give one question they have about the reading. Although the three content question categories (notational, conceptual, computational) did not align perfectly with either of the cognitive process or knowledge dimensions of Bloom's revised taxonomy, when reviewing the coding I did for the spring 2006 data, it seemed to me that the three categories captured the nature of the course material. Analyzing the accuracy of student responses to the content questions and the nature of the muddiest points identified by students would provide additional data for my *What is?* research question.

To gather evidence for my *What works?* question, I varied the format of the pre-class quizzes late in the semester. For three quizzes, all on readings about linear regression, I divided students into three groups. Group A's quiz had three notational questions and one muddiest-point question. Group B's quiz had three conceptual questions and one muddiest-point question. Group C's quiz had three computational questions and one muddiest-point question. Group assignments rotated so that students received one quiz of each type. This allowed me to run a quasi-control group experiment. This research design met with approval with my campus institutional review board, since it minimized risk to my students' learning and grades.

Here are examples of each type of question:

- Notational Question: In your own words, what is the difference between the meanings of the symbols  $\beta_1$  and  $b_1$ ?
- Computational Question: Suppose you're given the following regressor-response pairs of data: (2, 5.24), (4, 9.99), (6, 15.58), (8, 19.51), and (10, 21.16). Calculate the slope of the line of best fit to these data.
- Conceptual Question: Why is the line of best fit to a set of regressor-response data determined by minimizing the sum of the squares of the residuals instead of the sum of the residuals themselves?

At the start of each of the three class sessions on linear regression, I gave my students an in-class quiz on the relevant textbook material using the classroom response system. Each in-class quiz had one multiple-choice question of each type (notational, conceptual, computational) different from the questions appearing on the pre-class reading quizzes. After answering the in-class quiz questions, students were asked to rate their confidence in their answers. Extra credit was awarded for the in-class quizzes, using the point system in Table 13.1.

This point system was developed by chemistry professor Dennis Jacobs, a 1999–2000 Carnegie scholar (Bruff, 2009). My plan was inspired by a SoTL project by Axtell and Turner (2006) in which in-class quizzes were used to

	Correct Answer	Incorrect Answer
High Confidence	5	0
Medium Confidence	4	1
Low Confidence	3	2

**Table 13.1.** Point system for spring 2007 in-class quizzes

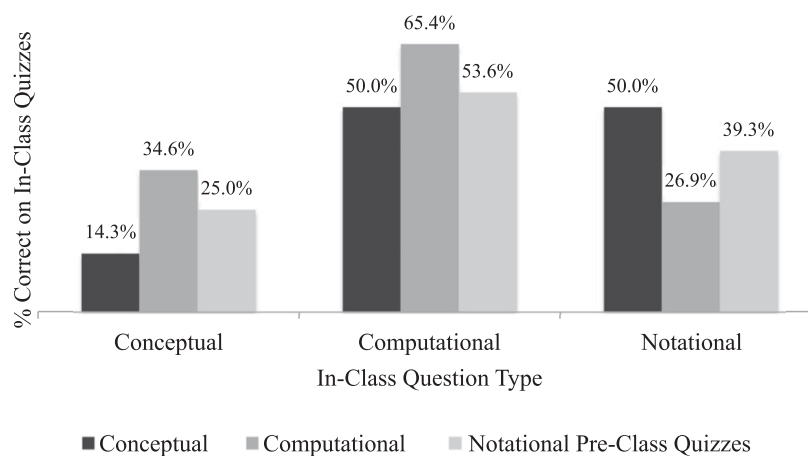
measure knowledge retention from pre-class reading assignments. I learned about Axtell and Turner's project at the 2007 Joint Mathematics Meetings, at which I presented my initial spring 2006 findings. In Axtell and Turner's project, the same or similar questions were asked on pre-class and in-class quizzes. In my study, I planned to use the in-class quizzes to gauge the effectiveness of the three types of pre-class reading questions. The results are summarized in Figure 13.5.

Not surprisingly, students who were asked computational questions before class performed better on computational questions during class the next day than students who were asked conceptual questions the night before. However, the difference was only suggestive ( $p = .13$ ), not significant. What was surprising was that the pre-class computational students also performed better on conceptual questions during class the next day—better than the students who were given conceptual questions the night before! This difference was statistically significant ( $p = .04$ ). It would seem that my practice in the spring 2006 semester of asking students mostly conceptual questions on their pre-class reading quizzes was a poor choice.

I shared these results with my students on the last day of class and asked for their thoughts on why pre-class computational questions might be more effective than pre-class conceptual questions. Here are the reasons they offered:

- Solving computational problems gives students experience with examples of concepts, which helps with learning inductively, which is engineering students' preferred approach to learning.
- It is easy to provide poorly thought out answers to free-response conceptual problems, but one cannot fake one's way through a computational problem. Thus, computational questions challenge students to work harder and thus learn more from the reading.
- The textbook's explanations of concepts were inferior to its explanations of procedures. Thus, students who focused on computational questions before class were able to get more out of the reading.

Pre-class conceptual questions excelled in another area. They prepared students to answer in-class notational questions better than pre-class computational questions did ( $p = .04$ ).



**Figure 13.5.** Performance on in-class quizzes by type of pre-class quiz

To supplement this quasi-control group experiment with qualitative (albeit self-reported) data, I administered an end-of-semester survey to the students asking them about their experiences reading their textbooks and preparing for class. Survey responses were consistent with the above experimental findings. One open-ended survey question asked students to identify what they were able to learn by reading their textbooks before class. About 24% of the students indicated that they were able to learn procedural or computational knowledge in this way, compared with only 15% for factual knowledge and 15% for conceptual knowledge. Also of note is that only 6% of students reported reading their textbooks before class in their science and engineering courses, so my request that they do so was unusual for them.

This second iteration of my SoTL project led me to several important conclusions:

1. The nature of pre-class reading questions matters.
2. Computational questions have a place on pre-class reading quizzes.
3. Conceptual questions might need to wait until class.

## Future Work

I have some ideas for further work. My experimental data indicates that pre-class reading quizzes consisting entirely of computational questions prepare students better for class than quizzes consisting entirely of conceptual questions. Next time I conduct a quasi-control group like this, I will compare all-computational quizzes with quizzes having a mix of question types. It may be that a mix is even more effective. Investigating topics other than linear regression will be important, too, to see if my results generalize to other topics.

My study treated all my students as a single type, but students are, of course, different from one another. Might conceptual pre-class reading questions be more useful for certain types of students, perhaps students of a certain major or gender? This is a question I could partly explore with existing data, and it is one I would investigate in future iterations of this project.

When I presented my project at the 2007 CASTL Summer Institute, the participants in my session and I had a productive conversation about textbooks. Might reading a different textbook yield very different results? Perhaps so, since my students said that the textbook used in the spring 2007 course was weak at explaining concepts. Could my muddiest-point analysis serve as a way to evaluate textbook quality, by identifying textbooks more likely to generate “good” student responses? This is an area for further exploration.

Speaking of “good” responses to muddiest-point questions, when I shared my project in a SoTL contributed paper session at the 2009 Joint Mathematics Meetings, the session moderator, Curtis Bennett, who was also interested in questions students ask about readings in mathematics courses, helped me think about other coding schemes for future work. For example, one scheme might assess how much insight into a student’s thought processes a response provides. This would be useful for just-in-time teaching. It could be employed in grading student responses, perhaps giving one point for a question about the reading, but two points if the student identifies the example in the textbook that triggered his or her question or suggests a possible answer to the question.

A different coding scheme might seek signs of deep learning (Bain, 2004) in student responses. For example, a muddiest-point response that draws a connection between my course and another course might be considered a sign of deep learning, as would a “I wonder if. . .” type of response that could serve as the basis for a student project. (I typically ask students to complete a project at the end of the semester in which they apply statistical techniques to real-world data.)

## Reflections

My SoTL project has paid dividends in my work at the Vanderbilt University Center for Teaching. I consult regularly with faculty members from various disciplines about teaching matters and many times I have shared aspects of my SoTL project with colleagues interested in pre-class reading assignments or the relationship between computational and conceptual learning goals. Although relatively few of my faculty colleagues have been interested in conducting their own SoTL projects, they have been interested in hearing the results of my project.

There has been more interest locally in SoTL among graduate students than among faculty members. The Center for Teaching runs two programs that require participants to conduct small SoTL projects appropriate to their teaching opportunities as graduate students: the Teaching Certificate Program and the Teaching-as-Research (TAR) Fellows program, the latter co-sponsored by the Center for the Integration of Research, Teaching, and Learning (CIRTL, [www.cirtl.net](http://www.cirtl.net)). Having conducted a project of my own has helped me mentor graduate students in these programs. The lessons I learned about shifting from a *What works?* question to a *What is?* question, about coding qualitative data in order to answer a *What is?* question, and about fitting into a single course something approximating a control group experiment are ones I have shared time and again with the graduate students. The success of our TAR Fellows program, which has included many conference presentations by participants as well as two peer-reviewed publications, has led to the program becoming a model for similar programs at other research universities in the CIRTL network.

In my own teaching, I now take more care in crafting the pre-class reading questions that I give my students. In addition to (and sometimes instead of) conceptual questions, I ask notational and computational questions. Also, when considering textbooks for adoption, I am more attentive to how they handle conceptual and computational material. I look for textbooks that scaffold material from the more concrete to the more abstract. I take a similar approach in designing my lesson plans, starting with examples before moving on to conceptual questions tackled through peer instruction.

More generally, the project has led me to question more regularly the assumptions I have about teaching and learning. When my experimental data indicated that computational pre-class reading questions better prepared students for class than conceptual ones, I was surprised. I had expected the opposite result. If I could be wrong about the kinds of pre-class reading questions I should ask my students, I could be wrong about other teaching choices that I make. As a result, I tend to view student work in my courses not only as a chance for students to learn course content and a means by which I evaluate my students, but also as a source of information I can use to understand my students and improve my teaching. I rarely have time to conduct deep analyses of student work during a course, but I make sure to keep my students' work so that I can make more sense of it when I have the time to do so.

Every course I teach presents similar opportunities to look across student work and identify patterns in their learning that can inform my teaching choices. The purpose of teaching is to promote student learning. Asking questions about student learning and attempting to answer them by collecting and systematically analyzing evidence of student learning, particularly qualitative evidence, equips me to fulfill that purpose more effectively.

## References

- Anderson, L., & Krathwohl, D. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.
- Axtell, M., & Turner, W. (2006, May). *Examining the effectiveness of reading questions in introductory university mathematics courses*. Paper presented at the London SoTL 6th Annual International Conference, London, England.
- Bain, K. (2004). *What the best college teachers do*. Cambridge, MA: Harvard University Press.
- Bruff, D. (2009). *Teaching with classroom response systems: Creating active learning environments*. San Francisco, CA: Jossey-Bass.
- Hobson, E. (2004). *Getting students to read: Fourteen tips* (IDEA Paper #40). Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development. Retrieved from [www.theideacenter.org/sites/default/files/Idea\\_Paper\\_40.pdf](http://www.theideacenter.org/sites/default/files/Idea_Paper_40.pdf)
- Hutchings, P. (Ed.). (2000). *Opening lines: Approaches to the scholarship of teaching and learning*. Menlo Park, CA: Carnegie Foundation for the Advancement of Teaching.
- Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.
- Mazur, E. (2009). Farewell, lecture? *Science*, 323(5910), 50–51.



- McKeachie, W. (1999). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (11th ed.). Boston, MA: Houghton Mifflin.
- Novak, G., Patterson, E., Gavrín, A., & Christian, W. (1999). *Just-in-time teaching: Blending active learning with web technology*. Upper Saddle River, NJ: Prentice Hall.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

# 14

## An Investigation Into the Effectiveness of Pre-Class Reading Questions

Mike Axtell  
*University of St. Thomas*

William Turner  
*Wabash College*

### Editors' Commentary

In this chapter Mike Axtell and William Turner describe how they went about undertaking, as novices, a literature review in mathematics education. Their experience revealed to them the critical role that the literature review can play in refining a SoTL research question and how it can aid in designing a study. Readers may want to contrast their study of reading questions with that written by Derek Bruff in the preceding chapter.

### The Backstory

We begin by providing the background of our investigation. We describe what motivated us to use pre-class reading assignments and how over time we developed a system that involved not just assigned readings but reading questions (RQs) as well. We explain why we decided to investigate their use and what it was that we wanted to know, at first.

The process of reading a mathematical text and understanding it is complicated and difficult (Konior, 1993; Reiter, 1998). It is an acquired skill that few, if any, come by naturally. However, it is one that mathematicians must learn. Mathematics conferences and pedagogical publications often contain ideas on how to get, or teach, students to read mathematics (Amick, 1997; Gold, 1998; King, 2001; Ratliff, 1998; Reiter, 1998; Taalman, 1998), and there are hundreds of references on how to help students improve their critical reading skills (Bratina & Lipkin, 2003). The goal is for students to become independent learners capable of teaching themselves, perhaps the ultimate goal of any liberal education program.

If we as teachers are convinced of the need to get students to read a mathematical text, we should then be concerned with determining what our students are gaining, and not gaining, from this task. If our strategies are not leading to desired outcomes, then we should rethink them. This observation led us to collaborate on a SoTL project during the 2005–2006 academic year that focused on student reading.

We had been discussing the need for students to read the text and working on strategies for getting them to do so for several years prior to 2005. Although we wanted students to read and learn independently as part of a liberal education, we faced time constraints. We were both teaching at Wabash College, a four-year liberal arts college located in Indiana, labeled a “more selective” institution by *U.S. News and World Report*. Wabash College has 14-week semesters, and courses meet for 150 minutes per week, either in three 50-minute sessions or two 75-minute sessions. We were teaching the usual freshmen classes and trying to cover at least as much material as schools that had more contact hours for these courses. On a daily basis we were having trouble getting through the material on our syllabus. At the 2004 Joint Mathematics Meetings we attended a special session on getting students to read the text prior to class. Speakers presented their methods for ensuring their students read the text before class, and this seemed like a solution to our problem. Students would come to class knowing some of the material and we could spend our precious class time teaching them the rest.

We spent the next couple of years trying different ways to encourage or require students read the text prior to class. Eventually, we adopted a system where we asked the students a few questions on a section of the text before class. On entering the classroom, the student would receive the answers to the questions and instruction would begin. We were happy with the arrangement and wrote the questions so as to give the student a base level of knowledge on the new material that would provide the background for our lecture. We saved time by not having our students copy definitions or theorems. By having them mimic some basic computational examples in the text, we felt we did not have to spend as much time on elementary examples.

Gradually we realized that we did not know if the pre-class questions were doing anything other than saving instruction time. We assumed the students were entering the classroom with a base level of knowledge since they had answered these questions, but we did not know if the students had retained any of that knowledge from the moment they completed the questions to the moment they sat down in the classroom. Carl Cowen (1991) stressed the need to assess students’ ability to read the text and to retain what they had read. Doing this would enable us to check whether our system was doing what it should and give us feedback we could use to help the students read the text better. Without this information, we were flying blind, and the students’ ability to read and gather information independently might not improve. This motivated us to try to determine what our students were learning from the pre-class reading questions.

## Designing the Study

Many SoTL investigations begin with questions that are too broad and need to be refined or limited to make them researchable. Other difficulties can arise when trying to find a control group or even enough subjects to study. Our study was no exception. What surprised us was how the literature review influenced our question and what data we collected.

Do pre-class reading questions help students learn? This is a vague question, and we had few ideas on how to design a study to answer it. We thought about teaching two sections of a course, one using the questions (as the treatment) and the other not using them (as the control). However, we believed that this made sense only if the two sections were taught in a very similar manner, and professors at Wabash very rarely teach two sections of the same course in a semester. Also we feared the students of one section would resent having to do more work than students in the other if it were taught by the same professor. In a small school, word would quickly spread among the students. So, we decided not to use a control group in the study. Instead, we used two sections – one section of first semester calculus taught by one coauthor (Turner) and one section of precalculus taught by the other (Axtell). Each used pre-class reading questions in the same manner (to be described below) and the same types of data were collected from both sections.

We narrowed our original question “Do pre-class reading questions help students learn the material?” down to “Do students bring knowledge derived from doing the pre-class reading questions into the classroom?” We felt this was the central question whose answer would either justify or invalidate our reason for using reading questions. However, our discussions and our literature search (see below), led us to realize that we were asking different *types* of questions in our pre-class reading questions. We then realized a natural (and more appropriate) question was whether students were more successful at retaining the knowledge gleaned from some *types* of question over others. The final version of our study question was “What types of reading questions best facilitate independent learning, and the retention of the material learned, in our students?”

The process of understanding what we really wanted to study and framing a final research question happened during our literature review and was greatly influenced by what we found. This was our first pedagogical literature review, and it was difficult at first. In mathematics research, there is essentially one database to consult (MathSciNet) and the key words to search for are not hard to figure out. In a pedagogical study, however, there are many databases and the number of seemingly relevant articles can become enormous. We began by exploring the bibliographies of the speakers we heard at the special session on getting students to read mathematics. This was of some use, but our search really picked up speed when we obtained practical information on how to access the pedagogical literature databases efficiently from a doctoral candidate in mathematics education. She recommended we use the databases JSTOR and ERIC.

The search tools in these databases allowed us to search more efficiently for articles using keywords and phrases such as “undergraduate,” “reading for comprehension,” “reading the text,” and so on. The difference between our original searches using the general article search engine provided by the college library versus the JSTOR or ERIC searches cannot be overstated. Most of the articles in our bibliography were found using JSTOR or ERIC. Reading these articles changed the focus and design of our study, as we discuss throughout this article. Scholars who are experiencing frustration with their literature searches might want to consult education or social science colleagues.

We initially viewed the literature review as an onerous task to be completed for form’s sake, but we soon realized that the literature review was playing a central role in shaping our study question and our thoughts for how to design the study. An illustration of this was when our literature review unearthed a paper (Wakefield, 2000) that compared the learning of mathematics to the acquisition of a second language. The paper mentioned the value of repeated exposure to basic vocabulary and techniques of computation. Reading it led to the clarification (in our minds) of two of the three question types (definitional and computational) that we wished to use in our study. Similarly, if learning mathematics is like learning to read (as several papers indicated), then these quotes from the report of the [National Reading Panel \(2000\)](#) prove illuminating:

In fact, a few studies suggest that pre-instruction of vocabulary words facilitates both vocabulary acquisition and comprehension (Ch. 4, p. 22).

. . . high frequency and multiple, repeated exposure to vocabulary material are important for learning gains (Ch. 4, p. 22).

We felt that using definitional reading questions addressed both the frequency of exposure and the pre-instructional recommendations mentioned above. Thus, our literature search was beneficial – the articles we read led us to recognize the types of questions we were asking, and this influenced the research question we wished to answer.

With “What types of reading questions best facilitate independent learning, and the retention of the material learned, in our students?” as the final version of the research question, we settled on three types of reading questions to study:

Definitional: Questions that ask the student to write a definition from the reading.

Computational: Questions that ask the student to mimic an example in the text to solve a similar problem.

Conceptual: Questions that require the student to demonstrate understanding.

One of each type would appear on every pre-class reading question assignment. Here is a typical three-question assignment from a calculus class:

(Definitional) What does it mean to say a function is locally linear?

(Computational) Let  $f(x) = \sin(x)$ . Use  $x = 1.39$  and  $x = 1.41$  to estimate  $f'(1.4)$ .

(Conceptual) Describe how to estimate the derivative of a function by using its graph.

Does this work for every function? If not, what property must the function have?

Students submitted their answers online the evening before the class met. Each answer was scored (0, 0.5, or 1 points) and the scores recorded. These were gathered over the course of a semester from two classes: a precalculus class of 18 students and a calculus class of 27 students. Both classes were comprised mostly of freshmen.

To see whether students had retained the material from their pre-class reading assignment, quizzes were administered at the start of class. They consisted of the same three reading questions that students had answered and submitted the night before. The quizzes were graded in the same manner as the pre-class reading assignments. During the course of the semester, the precalculus class took six reading quizzes, and the calculus class took five. Before class students

did not know if there would be a quiz that day. Thus the quizzes measured the knowledge students derived from the reading questions and brought with them to class the next day.

## Analysis

The most straightforward portion of our study was analyzing our data after the semester was over. Since we had narrowed our study question and organized our data by question type, this process was clear-cut. A full report of the data can be found in our paper (Axtell & Turner, 2007); we provide a partial summary here.

In our initial analysis of the data we were interested in the overall picture. We found that although students did very well on definitional questions on the RQs, they did quite poorly on the exact same questions on the reading quizzes – indicating a lack of knowledge retention. On the other two types of questions, scores on the RQs and the reading quizzes were roughly similar. The drop in definitional scores can be explained by students copying the appropriate definitions out of the text without actively thinking about them. Also, students at this level do not usually realize the importance of paying attention to the exact wording of a mathematical definition.

We next went through the data more carefully, discarding those from any student who didn't turn in an RQ assignment but took the reading quiz (or vice versa). We then analyzed the change in scores between a student's performance (by question type) on the RQ and on the reading quiz. This paired data gave a more accurate view of students' retention of knowledge.

In the precalculus class there were 88 matched data points from students attempting both the definitional reading question before class and the definitional quiz question in class. There was a score decrease of 27 points (out of 88 possible) going from the reading question to the quiz question, a 30.6 percentage point drop in scores for this type of question. The calculus class experienced a 22.9 percentage point decline for definitional questions. The other two question types produced smaller drops (2.6 percentage points for computational and 1.5 for conceptual, when both sections were combined). Overall, both sections had similar data sets.

The analysis showed that students are best able to retain knowledge that required some thought to understand (computational and conceptual questions) and least able to retain casually read and copied material (definitional questions).

## Conclusions of the Study

There is always a danger in trying to say what you hope is true rather than what the data show. This was the case for us when we determined our conclusions. We wanted to say that these pre-class reading questions were helping our students to learn the material better and that we were awesome teachers for setting such a system up. However, our study question and study design were not set up to answer the question of whether our students learned the material better, so we had to stay with our study question.

We were confident in concluding that students can pick up and retain a fair amount of knowledge by reading the text before class. They are good at learning to mimic computational steps to solve problems. They are also capable of retaining conceptual knowledge. However, based on the scores of the conceptual questions on the pre-class reading questions, they are not as good at constructing conceptual knowledge from the text as they are at figuring out how to perform computational tasks. For example, on the RQs in the precalculus class, students scored 12 percentage points better (78.8% versus 66.8%) on computational questions than on conceptual questions. Our study also showed that although students are good at copying a definition from the text, they are poor at recalling it a few hours later.

These conclusions suggest that the pre-class reading questions were achieving part of our goal: namely, ensuring the students arrived in class with a good base level of knowledge about the section to be covered. We were not happy with our students' ability to recall definitions that we thought they should be familiar with. Our literature review convinced us that the learning of definitions is best accomplished by repetition. So we continue to have definitional questions on the RQs and follow them with repetition in class, which we hope will increase students' retention of this knowledge.

## Dissemination Decisions

The number of pedagogical publications in higher education increases every year, as does the number of conferences on pedagogy and SoTL. Thus, the hard part of disseminating one's results is deciding which journal and conference

would be most appropriate. One of the co-authors (Axtell) had heard about the International SoTL conference being held in London and knew it would be a great meeting to attend and present at. The conference's only drawback was that it was a general conference, not focused on mathematics. We submitted our abstract and it was accepted. The presentation led to a submission to and publication in the conference proceedings (Axtell & Turner, 2007). We also gave a presentation at a contributed paper session at the 2007 Joint Mathematics Meetings. Another benefit of doing a literature review is that it will reveal what journals have published articles related to the study topic and hence are potential publication venues.

## Next Steps

We have not continued to study our use of reading questions in the classroom mostly due to one author moving to a new job a few hundred miles away. However, there are related questions that we believe worth pursuing:

- Do RQs help students become better at reading a math text?
- Does student performance on the definitional quiz questions improve as the semester goes on? (indicating that students made progress on learning definitions)
- What other ways are there to improve students' learning and understanding of definitions?

## Personal and Professional Impact

Through this project, we considered the types of reading questions we asked and gathered information about their effectiveness. What we learned from our literature review and the results of the study confirmed our sense that the RQs were doing good things, and we clarified what they were. Namely, the RQs were allowing our students to come to class with some mastery of computational and conceptual ideas. Although students do not master definitions by doing RQs, they allow for repeated exposure to key definitions since the definitions are mentioned again in class. Since repetition is a crucial ingredient of student learning, we dedicated more time to exploring definitions in class and less on computational examples.

Axtell still uses reading questions in his freshmen and sophomore level courses and continues to revise them to improve the in-class delivery of material. He tries to create a series of questions that will replace the first 10 to 15 minutes of a lecture. Passing out (and going over) the solutions to the questions in the first few minutes of class has allowed students to quickly get to the heart of the day's material.

Perhaps the largest impact of the study was in the conversations that occurred across campus. While we did not present our results formally to the campus, Wabash is a small college, and cross-disciplinary pedagogical conversations are common. Our study focused on a mathematics classroom and what types of mathematical questions best help our students, but this is broadly applicable. Professors in all disciplines struggle with how to get students to read the material before class. In addition, SoTL research has blossomed across the college since we did this study. A measure of this is the fact that Wabash College and the Center of Inquiry into the Liberal Arts (an institution housed on Wabash College's campus, see [www.liberalarts.wabash.edu/](http://www.liberalarts.wabash.edu/)) hosted the Third National Conference on Innovations in the Scholarship of Teaching and Learning at Liberal Arts Colleges in March 2009.

This study led us to a number of good outcomes. We learned more about what was happening in our classrooms and how to tailor our activities for desired outcomes. We learned what others were doing through our literature survey and the dissemination of our results at conferences. And our study prompted a series of good conversations at our institution with faculty members in other disciplines. Most importantly, we believe our study helped us to create a better, more intentional, classroom experience for our students.

## References

Amick, L. (1997). Math class – have you seen the preview? *MAA Online Teaching and Learning Exchange*. Retrieved from [www.maa.org/math-class-have-you-seen-the-preview](http://www.maa.org/math-class-have-you-seen-the-preview)

- Axtell, M., & Turner, W. (2007). Examining the effectiveness of reading questions in introductory college mathematics courses. In J. Fanghanel & D. Warren (Eds.), *Proceedings of the International Conference on the Scholarship of Teaching and Learning (2005 and 2006)* (pp. 205–210). London: City University London.
- Bratina, T., & Lipkin, L. (2003). Watch your language! Recommendations to help your students communicate mathematically. *Reading Improvement*, 40, 3–12.
- Cowen, C. (1991). Testing and teaching mathematics reading. *American Mathematical Monthly*, 98(1), 50–53.
- Gold, B. (1998). Requiring student questions on the text. *MAA Online Teaching and Learning Exchange*. Retrieved from [www.maa.org/node/121568](http://www.maa.org/node/121568)
- King, D. (2001). From calculus to topology: Teaching lecture-free seminar courses at all levels of the undergraduate mathematics curriculum. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 11(3), 209–227.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, U.K.: Cambridge University Press.
- National Reading Panel. (2000). *Teaching children to read (Reports of the subgroups)*. Washington, DC: National Institute of Child Health and Human Development. Retrieved from [www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf](http://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf)
- Ratliff, T. (1998). How I (finally) got my calculus students to read the text. *MAA Online Teaching and Learning Exchange*. Retrieved from [www.maa.org/node/121567](http://www.maa.org/node/121567)
- Reiter, A. (1998). Helping undergraduates learn to read mathematics. *MAA Online Teaching and Learning Exchange*. Retrieved from [www.maa.org/node/121566](http://www.maa.org/node/121566)
- Taalman, L. (1998). Problem zero: Getting students to read mathematics. *MAA Online Teaching and Learning Exchange*. Retrieved from [www.maa.org/problem-zero](http://www.maa.org/problem-zero)
- Wakefield, D. (2000). Math as a second language. *The Educational Forum*, 64, 272–279.



# **Illustrations of SoTL Work in Mathematics**

**Theme 4: Exploring Student Understanding of  
the Nature of Mathematics**





# 15

## **Liberal Arts Mathematics Students' Beliefs About the Nature of Mathematics: A Case Study in Survey Research**

Stephen D. Szydlik  
*University of Wisconsin Oshkosh*

### **Editors' Commentary**

Stephen Szydlik's experience teaching a problem-based inquiry seminar for over 10 years led him to question his perceptions about student learning. He lacked empirical evidence about student gains and he couldn't define what he meant by success in his course. This realization led him to a SoTL investigation that required work and reflection in order to frame a researchable question. Particularly worth noting is his description of how he came to define success in the course based on his desired outcomes. His realization that success meant student progress in the higher order activities inherent in doing mathematics led to a study of mathematical beliefs. His discussion of designing survey items about them, and checking for validity and reliability, is another highlight of the chapter. He also provides details of applying to his Institutional Review Board and a detailed discussion of options for publishing his work.

### **Introduction**

Mathematics is a living subject which seeks to understand patterns that permeate both the world around us and the mind within us—*Everybody Counts* (National Research Council, 1989).

What is mathematics? Debated by mathematicians and philosophers of mathematics for hundreds of years, the question is not settled and likely never will be. However, there is little doubt that student conceptions of mathematics are considerably different than those of professional mathematicians.

In this chapter, I describe a Scholarship of Teaching and Learning (SoTL) project that explored the mathematical beliefs of students in two sections of a problem-based general education mathematics course, the problem-based inquiry seminar (PBIS) at the University of Wisconsin Oshkosh. The PBIS course was designed to give students brief but authentic experiences in doing mathematics. Students are asked to collect data, look for patterns, make conjectures, model physical situations, and make logical arguments. Most of the students lack experience with these practices and are initially uncomfortable being asked to do them. Nevertheless, it appears that over the course of a semester, some students change their conceptions of the nature of mathematics. In their communication of mathematics, both formally and informally, these students show signs that they begin to see mathematics as a logical and consistent discipline.

Before initiating my SoTL project, my perceptions about the mathematical beliefs of students in this course were based on anecdotal evidence. Having taught a PBIS course for over a decade, I began to question whether these perceptions were accurate. Although it appeared that the course worked for my students, I had no supporting evidence nor was it clear what I meant by “worked.” At the same time that I began to think about quantifying success in the course, I was accepted into the Wisconsin Teaching Scholars Program. This year-long multidisciplinary program is designed to provide University of Wisconsin System with training and support in conducting introductory SoTL research. Exploring the success of the PBIS course was an obvious project to pursue as part of my participation in the program. Doing this also served a university need, as at the time, our department was looking to strengthen its assessment plan by more systematically examining the success of its courses. So my SoTL project emerged as a natural consequence of a number of factors.

The remainder of the chapter is organized into six sections. The first section provides background on the PBIS course and its students. The second includes a review of some of the relevant literature and the development of the research question. The third describes the methodology of the study, including the evolution of a survey instrument, while the fourth includes some results from the study and directions for further research. The fifth section describes my experience with going public with my SoTL project. The final section includes a reflection on the SoTL process as a whole.

While this chapter concerns a SoTL project that arose in response to the research question, its focus is on describing the experience of doing SoTL research. Conducting SoTL research affords numerous challenges, and I experienced a number of them, particularly in the design and assessment of a research instrument. These issues, which arose as I carried out the project, will be considered in some detail. A more complete treatment of the research results appears in [Szydlik \(2013\)](#).

## Background

At the University of Wisconsin Oshkosh, all students whose major course of study does not require a mathematics course must take a PBIS course. Versions of the course were first developed in the early 1990s in response to uncertainties expressed by faculty about the appropriateness of algebra as a terminal course in mathematics. Although evidence at the time was anecdotal, concerns centered on the high failure rate of general education algebra students and the sense that even successful students were not developing mathematically from their encounter with algebra. In 2005, the PBIS courses were instituted as the general education requirement for most Bachelor of Arts students.

All PBIS courses are limited to 25 students and have a common structure. Their content is to be introduced primarily through problem-solving opportunities, instead of lecture. The classes emphasize discussion and active learning. The course described in this study covers topics found in many collegiate liberal arts mathematics courses today, including the mathematics of voting, apportionment, political power, and elementary graph theory.

The course is structured to provide students with authentic mathematical experiences. By “mathematically authentic,” I mean that throughout the course students are given opportunities to do mathematics in the manner of professional mathematicians. As [Schoenfeld \(1994\)](#) wrote:

Mathematics is an inherently social activity, in which a community of trained practitioners (mathematical scientists) engages in the science of patterns—systematic attempts, based on observation, study, and experimentation, to determine the nature or principles of regularities in systems defined axiomatically or theoretically (“pure math”) or models of systems abstracted from real world objects (“applied math”). The tools of mathematics are abstraction, symbolic representation, and symbolic manipulation. However, being trained in the use of these tools no more means that one thinks mathematically than knowing how to use shop tools makes one a craftsman. Learning to think mathematically means (a) developing a mathematical point of view—valuing the processes of mathematization and abstraction and having the predilection to apply them, and (b) developing competence with the tools of the trade, and using those tools in the service of the goal of understanding structure—mathematical sense-making. (p. 60)

My PBIS course is designed to provide opportunities to engage in these mathematical activities to encourage students to develop their ability to think mathematically and begin to value mathematical processes.

In attempting to put into practice the philosophical spirit of the PBIS courses, I structured the course around carefully selected problems whose solutions generate mathematical conversations that reveal or explore specific content. During a typical class meeting, students worked on a challenging problem in small groups of three or four students for approximately twenty minutes. During this time, I circulated among the groups, observing their activities and providing encouragement and coaching. The class then convened in a semicircle for a discussion of their findings, strategies, solutions, and arguments. Though all were encouraged to speak, students could contribute as much or as little as they wished. I led the discussion from the front of the classroom, asking questions, probing students' responses, introducing notation and terminology, and writing ideas on the board. Throughout the course, students were asked to engage in mathematical behaviors in the sense of Schoenfeld (1994).

While some problems were designed for students to solve completely, others were intended to provide an avenue to discussion of content. They provided students with a context within which we could discuss notation, terminology, or abstract ideas. For example, on the first day of class, students were challenged with the handshake problem, which asks for the total the number of handshakes that take place in a group if everyone shakes hands with everyone else. In their solution of the problem, students were expected to find the correct number of handshakes for an arbitrary number of people and to argue the correctness of their reasoning. Later in the semester, in the context of a package delivery service, students explored the traveling salesman problem. This problem led us to a discussion of Hamiltonian circuits, complete graphs, and approximate algorithms for their solution.

## Framing the Research Question

As I began my SoTL inquiry in the Wisconsin Teaching Scholars program, my goal was to assess the success of my PBIS course. How to do this, and what success meant was not obvious. Upon reflection, I realized that the topics in the course, while engaging and tied to physical realities, were secondary in importance to the mathematical processes that the students encountered. For example, I did not expect my students to long remember the Condorcet criterion for the fairness of a voting system. However, I hoped that when they encounter a universally quantified “if-then” statement, they understand what it means for such a statement to be true or false. While they may not recall that there are  $n!$  ways of arranging  $n$  people from left to right in a photograph, I want them to realize that successful strategies for solving seemingly intractable problems include reducing to a simpler case, collecting data, and looking for patterns. Most importantly, I wanted them to realize that mathematics is a rigorous discipline, requiring precise language and logical arguments, but one that they can make sense of.

According to Steen (1999), “What we assess defines what we value” (p. 1). In my case, the reflection prompted by my SoTL investigation allowed me to clarify what I value in my PBIS class. Although the specific content of the course was engaging, it was the higher order activities inherent in the mathematical process that I most wanted my students to appreciate, or at least to acknowledge. Assessing the success of the course then became a matter of measuring student growth in student awareness and appreciation for mathematical practices. This led me to consider research on mathematical beliefs.

Student mathematical beliefs have been extensively studied. The *Handbook of Research on Mathematics Teaching and Learning* (Grouws, 1992) and its successor, the *Second Handbook of Research on Mathematics Teaching and Learning* (Lester, 2007), contain comprehensive summaries of research in mathematics education and include chapters written by prominent researchers in the field. In these handbooks, Thompson (1992) and Philipp (2007) provided syntheses on the literature that considers mathematical beliefs, much of which focuses on preservice and inservice teachers.

One population whose mathematical beliefs have received relatively little attention is that of college students enrolled in general education mathematics classes. This excludes students such as prospective teachers, business majors, science majors, and students enrolled in developmental mathematics courses. General education courses take on different forms at different institutions. Some require a course (or courses) in algebra, some a quantitative literacy or personal finance class, and some a course in mathematical thinking or problem-solving. Regardless of the type of course offered, however, the size of this student group is large. At my institution (a regional comprehensive university with approximately 12,000 undergraduate students), enrollments in general education mathematics courses made up more than one-third of total student enrollments (2670 general education enrollments out of a 7092 total mathematics student enrollments) in the mathematics department during the year of the study.

Schoenfeld (1992), using the work of Lampert (1990), identified typical student beliefs about the nature of the discipline that are the opposite of mathematicians' beliefs:

- Mathematics problems have one and only one right answer.
- There is only one correct way to solve any mathematics problem—usually the rule the teacher has most recently demonstrated to the class.
- Ordinary students cannot expect to understand mathematics; they expect simply to memorize it, and apply what they have learned mechanically and without understanding.
- Mathematics is a solitary activity, done by individuals in isolation.
- Students who have understood the mathematics they have studied will be able to solve any assigned problem in five minutes or less.
- The mathematics learned in school has little or nothing to do with the real world.
- Formal proof is irrelevant to processes of discovery or invention.

Moreover, conceptions about mathematics are related to learning and can affect student behaviors in the classroom (Garofalo, 1989; Kloosterman, 1996, 2002; Leder & Forgasz, 2002; Schoenfeld, 1985). For example, Garofalo observed that students who believe that mathematics problems can be solved through the direct application of procedures given by the teacher or the text tend to spend their time studying mathematics by memorizing facts and formulas and practicing procedures. Also, students who believe that textbook exercises can be solved only by the methods presented in the textbook may spend their time during an exam trying to remember a method given in the book (or even where in the book the method appeared), rather than trying to reason through a problem. Kroll and Miller (1993) noted that students holding the belief that all mathematics problems can be solved quickly and directly tend to give up quickly if they do not know how to solve a problem. And, students who believe there is just one right way to solve any mathematics problem become dependent on an external authority for verification of their solutions. It is for these reasons that typical beliefs about mathematics can be so disabling for students. Such beliefs contravene the essence of the subject, and those who hold them cannot truly value the discipline.

There is a lack of consistency in the use of the term “mathematical beliefs” in the literature (Leder & Forgasz, 2002; Törner, 2002). Thompson (1992) noted that many authors writing about beliefs do not explicitly define the term and assume that their readers already have a common conception of it. Op ‘T Eynde, DeCorte, and Verschaffel (2002) observed that this is not necessarily the case. Research models of beliefs often include components involving mathematics education, beliefs about the self, and beliefs about social contexts. In this chapter we use the working definition of Op ‘T Eynde, DeCorte, and Verschaffel (2002): “Students’ mathematics-related beliefs are the implicitly or explicitly held subjective conceptions that they hold to be true, that influence their mathematical learning and problem solving” (p. 16). As Schoenfeld (1992) wrote, “Students’ beliefs shape their behavior in ways that have extraordinarily powerful (and often negative) consequences” (p. 356).

Student beliefs about mathematics provided a natural subject for a SoTL inquiry into the effect of the problem-based seminars whose students constituted a relatively understudied population in the domain of beliefs. By measuring beliefs at both the start and the end of the course, I could gather evidence regarding the effect of the class on my students. Thus, my broad investigation into the relative success of the course was resolved into the following two research questions:

1. What do students enrolled in my section of PBIS believe about the nature of mathematics?
2. Do the beliefs change between the beginning and the end of the course?

According to Hutchings' (2000) taxonomy of SoTL investigations, the first question is a *What is?* question, since it involves measuring a student characteristic at a particular moment in time. The second question, since it attempts to measure the effect of a course, is a *What works?* question.

## Collecting Evidence

I chose the most direct method to assess student mathematical beliefs: I asked them. I surveyed the students in two sections of the class at the beginning and at the end of one semester to obtain snapshots of their beliefs so as to measure any changes. Based on the research literature, I designed a paper-and-pencil survey consisting of one open response question and 10 individual five-point Likert scale items. The Likert scale items were adapted primarily from [Lampert \(1990\)](#) and [Schoenfeld's \(1992\)](#) lists of commonly held student beliefs. For example, students were asked for their level of agreement (ranging from Strongly Disagree to Strongly Agree) with the statement, “There is usually only one correct way to solve a mathematics problem.”

In order to quantify the survey data, I categorized each of the 10 statements associated with the Likert items as being either in general agreement or disagreement with view held by the professional mathematical community. Each student response to each survey item was given a score from  $-2$  to  $2$ , with negative scores indicating disagreement with the mathematical community and positive scores indicating agreement. For example, a score of  $2$  was assigned for strongly agreeing with the statement “Mathematics reveals hidden structures that help us understand the world around us,” while strongly agreeing with “There is usually only one correct way to solve a mathematics problem,” received a score of  $-2$ . Adding scores for the ten items allowed me to assign to each participant a mathematical beliefs score between  $-20$  to  $20$ , with a positive score indicating agreement with mathematicians' beliefs.

The survey had a number of advantages. It provided a rapid means of data collection from the sample, and allowed for straightforward quantitative analysis. These were important considerations because I had little additional support for the investigation. Using a survey also allowed for flexibility in design; it was easy to change the survey after piloting it. However, Likert scale surveys are inherently blunt instruments, and the literature identifies a number of their weaknesses. [Ambrose, Philipp, Chauvot, and Clement \(2003\)](#) documented three potential issues, particularly in the context of student beliefs. First, in a written survey, it is impossible to know precisely how a subject interprets the words in the items. Second, in deciding how to respond to an item, the subject often needs to place the question into some personal context. How the subject chooses to contextualize the answer can have a great deal of effect on the response. Finally, Likert scale survey items do not measure how important a belief is to the subject. For example, a respondent may respond “Strongly Agree” to a statement that she has never consciously considered before. In spite of these criticisms, however, Likert scale items remain an important instrument in social science research.

The open response item on the survey asked the respondent to give her best definition of “mathematics” by completing the sentence: Mathematics is . . . . This item allowed students to freely express their conceptions of the discipline. It was the first item on the survey, so students would not be influenced in their definition by the subsequent items. I analyzed the definitions by looking for themes in their responses, counting, for example, the number of students who identified mathematics specifically with formulas. This helped to provide a more nuanced picture of student beliefs.

The items on the final instrument evolved from drafts of the survey that I used in earlier PBIS classes. On the draft versions, I asked students to explain their answers. Their explanations were enlightening and allowed me to identify some of the items in which context and interpretation had an effect on response. For example, one early version included “Ordinary students cannot expect to understand mathematics; they expect simply to memorize it.” Some students found this assertion confusing, as it is really the conjunction of two statements. Students might agree with the first part, but disagree with the second. So, I included only the initial clause in the final survey item. Another draft item, “There is always a rule to follow in solving a mathematics problem,” was revised because of its absolute nature. A student who disagreed with this statement because she believed this was true of most, but not all, mathematics problems would be scored the same as a student who believed that this was true of very few math problems. So, as written, the item did not distinguish between these two views. Feedback from colleagues on the drafts of the instrument was also helpful in its refinement. The final survey instrument appears in Appendix A.

At my institution, all research involving human participants must be reviewed and approved by the university's Institutional Review Board (IRB) before data collection begins. The review ensures that adequate measures are taken to protect those who volunteer to participate in a study. Because my research met conditions demonstrating that it presented minimal risk to participants (it did not involve minors or other protected groups, subjects were not identified, and the subjects' responses were not audio- or video-recorded), the project was classified as “Exempt.” That meant that I was not required to obtain written consent from the subjects in order to use their data. However, I was required to provide

them with a description of my research study containing enough information about the project to allow them to make an informed decision about participating. Appendix B contains the “Informed Consent” form used for my study.

Because my project involved minimal risk to participants, and the structure of the study was straightforward, my IRB proposal received an “expedited” review; that is, the chairperson of the Board inspected and approved it without waiting for the entire Board to convene. As a result, the IRB approval process required only about two weeks. IRB applications that require consideration by the full Board can take significantly longer to be approved. For a more complex project, it is a good idea to begin the IRB approval process far in advance of the planned data collection.

The IRB application requires a detailed description of methodology at the time of application, so the researcher must pay careful attention to research design. This was made clear to me in two ways as I prepared to apply for IRB approval. First, because the application had to include the survey instrument, I had to be certain that the survey I submitted was the final version. This was one reason why I pilot tested several versions of the survey before I constructed the final version. Second, our IRB requires precise descriptions of the participant pool. This prompted me to consider carefully the issues of instrument reliability and validity. Although reliability and validity are crucial considerations in any study of this nature, it was the IRB proposal that caused me to consider them during the initial planning.

Concerns about the reliability and validity of the survey instrument led me to survey other groups besides PBIS students. While general education mathematics students were the focus of the inquiry, I administered the survey to three other groups, to increase my confidence in its reliability. Reliability refers to an instrument’s ability to produce consistent results. First, I surveyed students in a large lecture section of Active Lifestyles, an introductory physical education course that (almost) all undergraduate students are required to take. This sample let me assess the reliability of the instrument. Since most of these students were not taking a mathematics course, and the course they were taking involved little quantitative content, I expected that their mathematical belief scores would not change between the initial and final surveys. A large change in mathematical belief scores from the initial to the final survey in this sample would indicate that the survey is unreliable, while relatively consistent scores would provide some evidence for the reliability of the instrument.

I also wanted to ensure that the instrument had validity, that is, the survey actually measured what it purported to measure, beliefs about the nature of mathematics. For this reason, I administered the survey to mathematics majors and mathematics faculty members. This gave me survey data from individuals with substantially more mathematics experience than PBIS students, and I therefore expected them to have higher mathematical beliefs scores. In particular, I expected that mathematics faculty members would score near the maximum of the scale. A substantial difference between PBIS students and the more mathematically sophisticated majors and faculty would provide evidence for validity. But, if there were little or no difference between these groups and the PBIS students, then either the populations were not different in terms of their mathematical beliefs or the instrument did not adequately detect the difference. The mathematics faculty members were sampled only once rather than in the before and after format used for the other populations.

## Results

The goal of this section is to highlight those results that best illustrate the SoTL process. A more comprehensive discussion of the results may be found in [Szydlik \(2013\)](#).

### Numerical Results

Most of the data collection took place during the spring 2009 semester. Table 15.1 contains the mean aggregate mathematical beliefs scores for the four groups of participants. In each of the samples, the standard deviation was between approximately 3.5 and 4.5. The mean score for the PBIS students on the initial survey was more than a full standard deviation below that of the mathematics majors and approximately three standard deviations below that of the mathematics faculty. In fact, on the initial survey, PBIS students scored lower on average than mathematics faculty members on every item.

The final survey told a dramatically different story. Between the initial and final surveys, student responses shifted just under seven units on average in the direction of alignment with professional mathematicians. This change is highly statistically significant, representing a shift in the sample mean of nearly two standard deviations. Although, the

Population	PBIS Students	Active Lifestyles Students	Mathematics Majors	Mathematics Faculty
Sample Size	32	41	12	18
Sample Mean Initial Survey	1.25	2.12	6.33	13.72
Sample Mean Final Survey	8.22	2.80	8.42	—

**Table 15.1.** Mean aggregate mathematical belief scores for the samples in the study

mathematics majors also moved in the positive direction on the mathematical beliefs scale, the PBIS student sample mean almost matched that of the majors on the final survey.

It is highly unlikely that the change in mathematics belief scores among the PBIS students is due to random chance. Evidence from the physical education students supports this contention. In the sample of 41 Active Lifestyles students who participated in the study, the mean mathematical beliefs score changed from 2.12 to 2.80 between the initial and final surveys. The similar means support the reliability of the instrument because the survey provided relatively consistent measures upon repeated administration to this population. The small positive change in the physical education students' sample mean is also not surprising. Although none of those students were in a surveyed PBIS section, some could have been enrolled in another section of a PBIS course or in some other mathematics course. Experience in those classes, or in some other course with quantitative content could have influenced survey responses.

## Qualitative Results

The survey asked participants to provide a definition of mathematics, and the language and content of many of their definitions overlapped. After identifying the most common themes, I counted how many students either explicitly included or strongly implied each theme in their definitions of mathematics. This simple thematic analysis provided a secondary means of exploring student beliefs, particularly in comparing the initial and final surveys. While almost 62% of PBIS students on the initial survey defined mathematics as involving data, numbers, or formulas, only 29% of students on the final survey did. The percentage of students invoking logic as an aspect of mathematics jumped markedly from the initial to the final survey, going from 14% to 23%. There were also themes on the final survey that were not present at all on the initial survey. The results of that thematic analysis for the PBIS students are included in Table 15.2. (Note that the total in each column is greater than 100% since many definitions involved multiple themes.)

## Discussion

The analysis supports previous research on student beliefs about mathematics. The survey results indicate that general education students do express many of the common beliefs identified by Schoenfeld (1992) that are in contradiction to

Theme	Initial Survey	Final Survey
Problem solving	60%	65%
Data, numbers, or formulas	62%	29%
Logic	14%	23%
Means to understand the world	14%	16%
Procedures, rules, or laws	12%	3%
Patterns	0%	13%
Strategies	0%	10%
Applying previous ideas	0%	10%
Argument or proof	0%	10%

**Table 15.2.** Dominant themes in PBIS definitions of “mathematics”



those of professional mathematicians. At the start of the course the PBIS students viewed mathematics as a collection of facts and procedures, and described doing mathematics as recalling the correct procedure to apply. While students viewed mathematics as substantially involving problem solving, they were just as likely to define the discipline as consisting of numbers, data, and formulas.

This study supports the effectiveness of the PBIS approach in several ways. The expressed beliefs of PBIS students changed substantially over the course of the semester. Although these students had significantly lower scores than the mathematics majors on the initial survey in the study, by the end of the semester, the mean of their mathematical belief scores was nearly the same as that of the mathematics majors. Moreover, the general education students' definitions of mathematics at the end of the semester were richer than at beginning. The PBIS students continued to identify the discipline with problem solving; a notably smaller percentage of students defined it as consisting of data, numbers, formulas, or procedures. More students expressed a value for the logic inherent to the discipline and its power to provide opportunities for rigorous arguments.

There are a number of qualifiers to these results. I was both the researcher and the teacher of the course under study. During the semester, the class had what I deemed authentic mathematical encounters, and we discussed these experiences, using mathematical language. By the end of the semester, students were familiar with many of the common phrases and terminology that we used to describe the mathematical process (e.g., searching for patterns, making arguments, collecting data, etc.) The gains made by the students over the course of the semester could be due to a familiarity with this terminology and a desire to put a correct answer on the survey rather than indicating their true beliefs.

Second, interpretation and context can affect responses in Likert surveys. Although I had hoped to eliminate this issue through the use of multiple pilot surveys, it is still a potential concern for this instrument. This was brought home to me when I analyzed the faculty responses. For example, on Item #6, "Mathematics is as much about patterns as it is about numbers," the PBIS students had higher mean (1.41) on their final survey than mathematics faculty (1.28). One faculty member expressed her concern with the item: "Do you mean only as much about patterns, or do you really mean at least as much about patterns?" Of course, the survey was designed specifically for students and not for faculty members, who read mathematics in a sophisticated way. For example, Item #2, "In mathematics everything goes together in a logical and consistent way," will mean something different to mathematicians familiar with Gödel's incompleteness theorems than it does to students. As these examples demonstrate, it is possible that the students may have interpreted items in the survey in subtle and unforeseen ways.

Interviewing students as part of the SoTL inquiry would have allowed me to inquire about participant responses to ascertain both interpretation and intent. This could have provided insight into the changes in the mathematical beliefs scores by the PBIS students. The student data showed a shift of approximately two standard deviations of the sample mean between the initial and final surveys. However, it is not precisely clear what this meant in terms of belief changes. Did the change reflect a different outlook on the discipline? If so, interviews could also provide insights into the reason, in particular, into which aspects of the class were most responsible.

The results, particularly the change in PBIS student responses, were significant enough to suggest that further research might be fruitful. I intend to explore the mathematical beliefs of general education mathematics students in greater detail, using interviews as a way to delve more deeply. The literature (e.g., [Kagan, 1992](#)) suggests that beliefs are resistant to change, so I also want to follow up with the current study by examining whether the expressed changes in beliefs persist beyond the end of the class.

## Going Public

While the meaning of the term "Scholarship of Teaching and Learning" varies among its practitioners, most definitions include public dissemination of its results as a characteristic (e.g., [Hutchings & Shulman, 1999](#); [Cambridge, 2001](#)). Because scholarship contributes to public discourse, one of the final and most crucial phases of any SoTL project is disseminating the results. This dissemination may take many forms, including conference presentations, teaching workshops, poster sessions, and local or peer-reviewed publication. My experience reflects this diversity of venues for communication, as I shared my research in several ways.

Locally, I shared the research with my colleagues through an assessment report. As noted in the introduction, one reason for investigating the PBIS course was the need to gather data for assessment. While not the focus of the scholarship, obtaining some assessment of the course was a valuable outcome. Our department website promotes our

PBIS courses as offering “the opportunity to develop the ability to distinguish problem solving and critical thinking from exercises and routine thinking and to identify attitudes and beliefs that are conducive to success in challenging situations.” Although my survey was not designed to address this goal directly, the results do provide some support for the success of the course in this area.

I also presented the results of the SoTL project in two different off-campus venues: at a session for SoTL research at a regional conference on diversity, and at a contributed paper session on SoTL at the joint mathematics meetings. The final stage was to submit a paper on the study for peer review and publication. This was challenging for me, and is likely so for many SoTL authors, because we are not trained as educational researchers. Finding an appropriate journal can be particularly difficult in mathematics education, a field that has had professional scholars and dedicated journals for more than 40 years. For example, both the *Journal for Research in Mathematics Education* and *Educational Studies in Mathematics* expect nuanced research studies, carefully grounded in learning theory and theoretical frameworks. They are not appropriate places to submit introductory SoTL work. Other journals such as the *Journal of Mathematics Teacher Education* and *School Science and Mathematics* publish research related to teacher education and elementary and secondary school mathematics, while *Mathematics Teacher* tends to focus on practical concerns of teaching content rather than on formal research studies.

Finding appropriate forums in which to publish SoTL research requires careful consideration by mathematicians. When looking to publish my scholarship, I considered a number of outlets. *Research in Collegiate Mathematics Education*, a continuing series of volumes published by the Conference Board of Mathematical Sciences, seemed appropriate. *The Journal of Mathematical Behavior* might also have been a reasonable choice, since it focuses on mathematical thinking as a social enterprise, and my work fits that category. There are also a number of general SoTL journals that publish this type of research. (For more information on them, see Chapter 4.) Ultimately, *Teaching Mathematics and its Applications*, a journal that promotes the exchange of ideas for improving teaching and learning, seemed most appropriate, and this is where the results of my project were published (Szydlik, 2013). Regardless of the venue, the SoTL inquiry process should have some public dissemination. To be true scholars of teaching and learning requires sharing the results of our investigations with our peers.

## Concluding Remarks

The primary reason for initiating my SoTL project was to examine my PBIS course in a scholarly way. Although the results of the investigation are far from definitive, they have given me evidence that the problem-based inquiry seminar makes a difference with students and their views on mathematics. I hope that the study contributes to the scholarly conversation about student learning.

My inquiry has caused me to think deeply about the course, and to examine my pedagogical practices and their impact on student beliefs. Although I have always considered myself to be a reflective teacher, the study has further heightened my awareness of the classroom environment and its effect on students. Now, when I am leading a discussion, I ask myself whether our discussion is moving my students toward memorizing a procedure or toward deeper conceptual understanding. When I prepare a lesson for my students I ask whether it gives an authentic experience in mathematics: Does it reinforce the view of mathematics as a collection of techniques, or does it encourage students to develop a view of mathematics as a coherent subject that they can make sense of? In this way, my SoTL project has informed my teaching deeply, encouraging me to become more aware of the need to engage my students mathematically. Through this engagement students can come to recognize the beauty and power offered by our discipline.

## Acknowledgements

This work was supported in part through the University of Wisconsin System’s “Wisconsin Teaching Scholars” program.

## References

Ambrose, R., Philipp, R., Chauvot, J., & Clement, L. (2003). A web-based survey to assess prospective elementary school teachers’ beliefs about mathematics and mathematics learning: An alternative to Likert scales. In N. A.

- Pateman, B. J. Dougherty, & J. T. Zilliox (Eds.), *Proceedings of the 2003 Joint Meeting of PME and PMENA* (Vol. 2, pp. 33–39). Honolulu, HI: CRDG, College of Education, University of Hawaii.
- Bell, E. T. (1931). *The queen of the sciences*. Baltimore, MD: The Williams & Wilkins Company.
- Cambridge, B. (2001). Fostering the scholarship of teaching and learning: Communities of practice. In D. Lieberman & C. Wehlburg (Eds.), *To improve the academy* (pp. 3–16). Bolton, MA: Anker.
- Garofalo, J. (1989). Beliefs and their influence on mathematical performance. *Mathematics Teacher*, 82, 502–505.
- Grouws, D. A. (Ed.). (1992). *Handbook for research on mathematics teaching and learning*. New York, NY: Macmillan Publishing Company.
- Hutchings, P. (2000). Approaching the scholarship of teaching and learning. In P. Hutchings (Ed.), *Opening lines: Approaches to the scholarship of teaching and learning* (pp. 1–10). Menlo Park, CA: Carnegie Foundation.
- Hutchings, P., & Shulman, L. (1999). The scholarship of teaching: New elaborations, new developments. *Change*, 31(5), 10–15.
- Kagan, D. M. (1992). Implications of research on teacher belief. *Educational Psychologist*, 27(1), 65–90.
- Kloosterman, P. (1996). Students' beliefs about knowing and learning mathematics: Implications for motivation. In M. Carr (Ed.), *Motivation in mathematics* (pp. 131–156). Cresskill, NJ: Hampton.
- Kloosterman, P. (2002). Beliefs about mathematics and mathematics learning in the secondary school: Measurement and implications for motivation. In G. C. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 247–269). Boston, MA: Kluwer Academic Publishers.
- Kroll, D., & Miller, T. (1993). Insights from research on mathematical problem solving in the middle grades. In D. Owens (Ed.), *Research ideas for the classroom, middle grades mathematics*, National Council of Teachers of Mathematics. Research Interpretation Project (pp. 58–77). New York, NY: Macmillan Publishing Co.
- Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, 27, 29–63.
- Leder, G. C., & Forgasz, H. J. (2002). Measuring mathematical beliefs and their impact on the learning of mathematics: A new approach. In G. C. Leder, E. Pehkonen, & G. Torner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 95–112). Boston, MA: Kluwer Academic Publishers.
- Lester, F. (Ed.). (2007). *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: NCTM, Information Age Publishing.
- National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.
- Op 'T Eynde, P., De Corte, E., & Verschaffel, L. (2002). Framing students' mathematics-related beliefs. In G. C. Leder, E. Pehkonen, & G. Torner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 13–37). Boston, MA: Kluwer Academic Publishers.
- Philipp, R. (2007). Mathematics teachers' beliefs and affect. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257–315). Charlotte, NC: NCTM, Information Age Publishing.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp. 334–370). New York, NY: Macmillan.
- Schoenfeld, A. H. (1994). Reflections on doing and teaching mathematics. In A. H. Schoenfeld, (Ed.), *Mathematical thinking and problem solving* (p. 53–70). Hillsdale, NJ: Erlbaum.

- Steen, L. (1999). Assessing assessment. In B. Gold, S. Keith, & W. Marion, (Eds.), *Assessment practices in undergraduate mathematics* (pp. 1–6). Washington, DC: Mathematical Association of America.
- Szydlik, S. (2013). Beliefs of liberal arts mathematics students regarding the nature of mathematics. *Teaching Mathematics and its Applications*, 32(3), 95–111. doi:10.1093/teamat/hrt002
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. Grouws (Ed.), *Handbook for research on mathematics teaching and learning* (pp. 334–370). New York, NY: Macmillan.
- Törner, G. (2002). Mathematical beliefs—A search for a common ground: Some theoretical considerations on structuring beliefs, some research questions, and some phenomenological observations. In G. C. Leder, E. Pehkonen, & G. Torner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 73–94). Boston, MA: Kluwer Academic Publishers.

## Appendix A

### Mathematical Beliefs Survey

Please give me your best definition of “mathematics” by completing the sentence:

Mathematics is . . .

Now, for each of the following items, please indicate your level of agreement using one of the responses: Strongly Disagree (SD), Disagree (D), Neutral/Not Sure (NS), Agree (A) or Strongly Agree (SA).

Then give a brief explanation why you answered the way you did.

1. To know mathematics means remembering and applying the correct rule or technique to solve a given problem.
2. In mathematics everything goes together in a logical and consistent way.
3. Learning mathematics is mostly memorizing and practicing procedures.
4. Mathematics reveals hidden structures that help us understand the world around us.
5. Ordinary students cannot expect to understand mathematics.
6. Mathematics is as much about patterns as it is about numbers.
7. There is usually only one correct way to solve a mathematics problem.
8. Mathematics is mostly a body of facts and procedures.
9. I understand what it means to make a sound mathematical argument.
10. I am capable of making sound mathematical arguments most of the time.

## Appendix B

### Informed Consent Form Provided to Participants

#### SUMMARY OF RESEARCH PROJECT

The purpose of this summary is to describe the research study, *Beliefs of liberal arts mathematics students about the nature of mathematics*, and to explain the study's scope, aims, and purpose. The goal of this project is to investigate student beliefs about the nature of mathematics and whether those beliefs change during a course that focuses on mathematical problem-solving.

1. The reasonably expected benefits of the project include:
  - a. The benefit to yourself of a better understanding of your own beliefs about the nature of mathematics.
  - b. The benefit to society due to the acquisition of knowledge that may eventually lead to improved teaching of mathematics and a better understanding of how attitudes and feelings about mathematics are formed.
2. The procedures that will be used involve two separate surveys of the participants.

3. It is reasonably foreseeable that you will experience little to no discomfort. The risk of harm that could result from your participation in the project is minimal.
4. There are no alternative procedures that could be used in this study.
5. The expected duration of your participation is approximately 30 minutes, with approximately 15 minutes required for each of the two surveys.
6. Your participation in the study is completely voluntary—you do not have to participate and you can stop at any time. If you refuse to participate now, or withdraw from the study later, it will have no effect on any regular services or benefits available to you at the University of Wisconsin Oshkosh. Your responses to either the survey instrument or the interview will have no effect on your course grade.
7. Any personal information used in this study will be treated confidentially. Information which identifies you as an individual will not be released, without your consent, to anyone for purposes which are not directly related to this research study.
8. If you have any question about this study, or your rights, you may call or write:

Dr. Stephen Szydlik  
220 Swart Hall, Department of Mathematics  
University of Wisconsin Oshkosh  
Oshkosh, WI 54901

Telephone:  
email: [szydliks@uwosh.edu](mailto:szydliks@uwosh.edu)

9. You will be given a copy of this statement, which serves to acknowledge the fact that you have been informed about the project and that you have voluntarily agreed to participate.

This research project has been approved by the University of Wisconsin Oshkosh Institutional Review Board for Protection of Human Participants.

# 16

## The Mathematics of Symmetry and Attitudes towards Mathematics

Blake Mellor

*Loyola Marymount University*

### Editors' Commentary

In this chapter Blake Mellor describes a study of student learning in a mathematics for liberal arts course offered as an alternative to the typical quantitative skills course. His approach to getting baseline data for his study was to teach the traditional quantitative skills course first. As a result of pursuing a SoTL investigation over several semesters, he encountered a number of issues with the use of surveys. Awareness of the lessons learned by Mellor will be useful to those beginning in SoTL. The author presents a different perspective on the personal and professional impact of SoTL. For him, SoTL will not be a path to publication, but will serve as a form of professional development that enables him to revise and improve his courses.

### Introduction

Most college students have a view of mathematics shaped by a high-school curriculum designed to prepare students for calculus. As a result, they view mathematics as primarily a formal exercise in manipulating numbers and algebraic formulas, with applications restricted to the sciences and perhaps finance. Mathematicians, of course, view their subject quite differently. We see mathematics as a creative endeavor, and mathematical thinking as a skill with a wide range of applications. My SoTL investigation emerged from a desire to teach a course for students in the liberal arts that would broaden their view of mathematics and show how it could be applied to the arts and humanities. This long-term study has progressed slowly over the last ten years. In this chapter I describe the genesis of the project, how it has evolved over the past decade, mentioning some of the results, and some of the problems and difficulties I encountered.

### The Genesis and Evolution of the Project

#### In the Beginning

This project began when I was first assigned to teach a mathematics course for liberal arts students. I had the freedom to choose my text, topics, and approach. There are many good texts available, such as [Burger and Starbird \(2005\)](#) and [COMAP \(2006\)](#), but they generally opt for breadth over depth, introducing a variety of interesting or useful topics, rather than exploring one in depth. I wanted a course that would explore a topic deeply, and encourage students to

discover results for themselves. I came across the book *Groups and symmetry: A guide to discovering mathematics* (Farmer, 1996). It covers the classification of symmetric patterns in two dimensions, but is organized as a series of “Tasks” for the students to complete. The emphasis is as much on the process of doing mathematics as on the content.

In spring 2002 I taught a course titled “Mathematics of Symmetry” based on Farmer’s book to a class of 29 students. The course material included the classification of rigid motions in two dimensions (translations, rotations, and reflections) and the classification of the symmetry groups for finite figures, strip (or frieze) patterns, and wallpaper patterns. I was not viewing the course as a site for a SoTL investigation, so I did not collect any data. The students worked through the Tasks in Farmer’s book independently with individual assistance available from me as needed. By the end of the semester students completed projects with in-class presentations on the use of symmetry in some area of the arts or humanities, or creative projects in which they used ideas of symmetry to create their own artwork. The grade was based on journals the students kept of their progress through the text, their projects, and two midterm exams. The student evaluations were generally positive, and I felt that the course had been successful for some students, but I realized I did not have any hard data to decide whether the students’ attitudes towards mathematics had changed. I also thought there were many ways in which the course could be improved.

## Evolution of the Course

Four years later at a new institution, I had another opportunity to offer Mathematics of Symmetry. Inspired by two of my colleagues at Loyola Marymount University (LMU) who were doing work in SoTL, Curtis Bennett and Jacqueline Dewar, I decided to undertake a SoTL study of the class. Since doing research on my teaching was new to me, I wanted to start with data that I could collect and analyze easily. The homework assignments and exams would measure how well students learned the specific content of the course, but it was not clear how to measure their ability to apply any mathematical reasoning skills they gained in the course to other areas. At the very least, it would require considerable effort to develop a reliable instrument. On the other hand, it seemed relatively easy to measure students’ attitudes towards mathematics with an online survey, and if I asked them to respond to statements on a numerical Likert scale (from Strongly Disagree to Strongly Agree), then the analysis would also be relatively easy. So I decided to develop such a survey, with the support of a summer grant from Loyola Marymount’s Center for Teaching Excellence in 2005. I will discuss the survey in more detail in the next section.

To begin my investigation, I taught a section of the usual LMU mathematics course for liberal arts majors, Quantitative Methods for the Modern World, in fall 2005. This course serves the same population as Mathematics of Symmetry, and satisfies the same core curriculum requirement. It is a typical quantitative literacy (QL) course, covering a variety of topics including percentages, descriptive statistics, financial mathematics, and voting theory. It does not look at applications of mathematics to the arts, and most of the mathematics is of the algebraic variety students previously encountered in high school. Teaching the QL course provided an opportunity for me to become more familiar with the types of students who would be taking the Mathematics of Symmetry course. By having the QL class fill out the survey at the beginning and end of the semester, I obtained control group data to use in assessing the effects of the Mathematics of Symmetry course.

In spring 2006, I taught Mathematics of Symmetry for the second time, and had students fill out the attitude survey at the beginning and end of the course. I modified the course by providing more scaffolding for Farmer’s book in the form of lectures and additional worksheets. In addition to submitting their journals as before, I also required students to turn in specific Tasks to be graded. So, rather than proceeding at their own pace, students had deadlines and received feedback throughout the semester. I also scanned the homework all through the semester to be reviewed later to see how students’ reasoning skills changed through the semester. (Five years later, this analysis has not been done, and I’m not sure it ever will be.) Again, students were asked to complete a project on some application of symmetry in the arts or another field, or to do a creative project using ideas of symmetry.

Four years later, in spring 2010, I had another opportunity to teach Mathematics of Symmetry, and I modified the course further. Rather than use Farmer’s book, I wrote notes covering the same material, but with a different organization and emphasis. Also, to make the connections with the arts explicit earlier in the course (rather than just in the projects), I added a short paper early in the course in which students looked at how symmetry was used by a specific artist. I also devoted several classes to lectures on how symmetry appeared in the arts and humanities. As before, I asked students to fill out the pre- and post-attitude surveys.

## Collecting Evidence

The anonymous online pre- and post-surveys served as the primary source of evidence for the project. Considerable work has been done on evaluating students' attitudes towards mathematics. Much of this has involved elementary school students, so I felt it was not applicable to the college student population. As a result, I selected my questions from a variety of sources: the Indiana Mathematics Belief Scales (Kloosterman & Stage, 1992), the Attitudes Toward Mathematics Inventory (Tapia & Marsh, 2004), the Fennema-Sherman Mathematics Attitudes Scales (Mulhern & Rae, 1998), the Maryland Physics Expectations Survey (Redish, Steinberg, & Saul (1998)), questions developed by Alan Schoenfeld (1989), and questions developed by my colleagues Bennett and Dewar (personal communication, 2004). The survey is in the Appendix.

The survey given in 2005 and 2006 consisted of 36 Likert-scale questions and four free-response questions. The survey in 2010 repeated 21 of the same Likert-scale questions and all the free response questions. Eleven additional questions were added to or modified from the earlier instrument. The changes were made because the earlier questions seemed redundant or not useful or because the online survey tool had changed.

In 2005, in Quantitative Methods of the Modern World, 57 took the pre-survey and only 27 took the post-survey. In 2006, in Mathematics of Symmetry, I had 40 respondents at the beginning and 21 at the end. In 2010, it was 49 at the beginning and then 39. A difficulty in comparing the results at the start and end of the semesters is the change in the number of responses, which occurs for several reasons. First, there are students who drop the class after completing the pre-survey, and others who add the class, after the pre-survey has been administered. Second, some students do not respond to the survey, despite repeated reminders and being offered credit. However, at the end of the semester the non-response rate is significantly higher. In 2006 I was able to match respondents at the beginning and end of the semester, but there were only 17 matched pairs. In 2010, due to technical issues, matching was not possible.

In addition, in 2006 I scanned take-home finals along with the all the homework assignments. I also retained some examples of the final projects (those that were not claimed by the students). Finally, I have the results of the university's student course evaluations. While these were not written to gauge students' attitudes towards mathematics, they provide additional information.

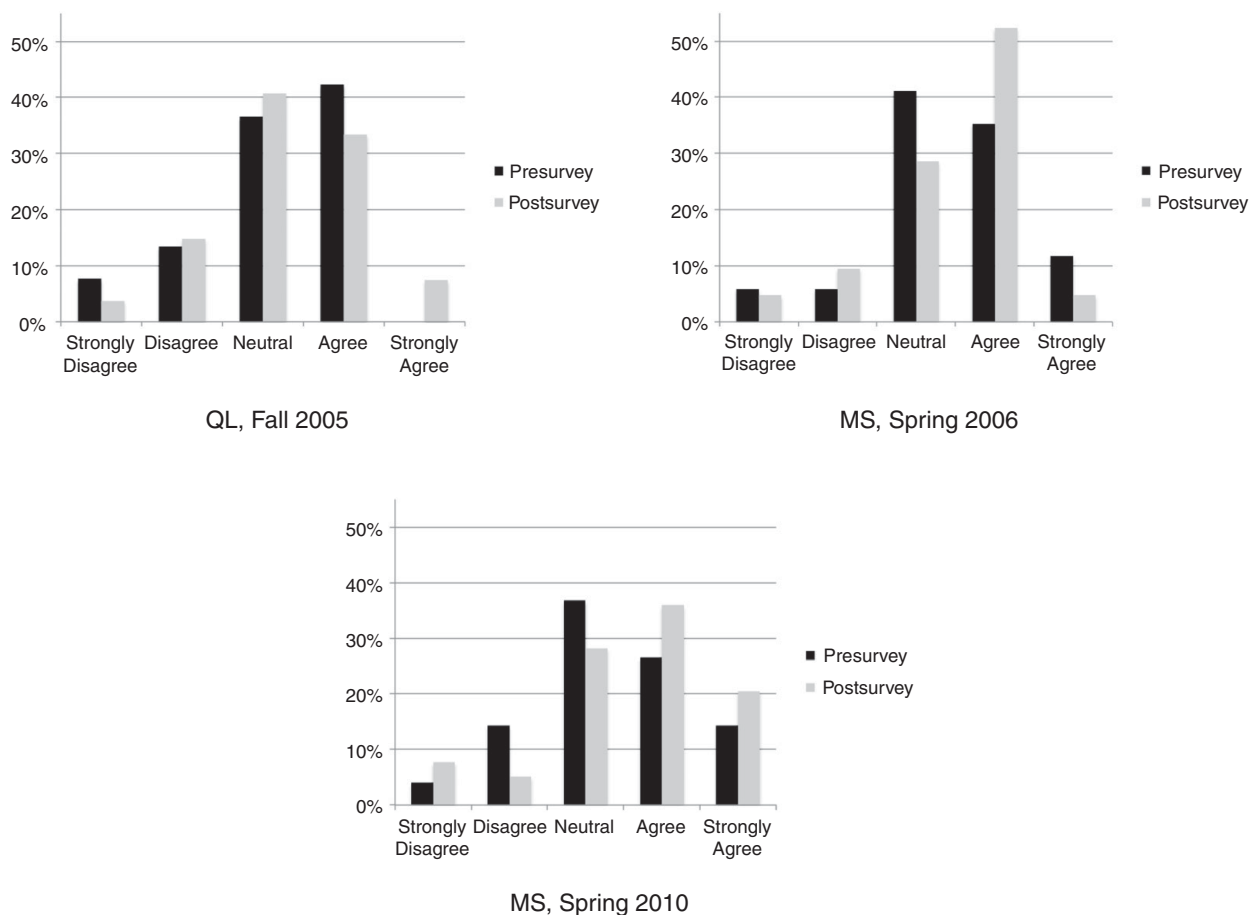
## Analyzing the Evidence

I initially attempted to analyze the data by comparing the means on the pre- and post-surveys and looking for statistically significant differences. This was not very fruitful. The absolute changes in the averages were not large, and the number of responses was so small that the changes were not statistically significant. For example, in 2010 I added a question asking for the level of agreement with the statement "I understand the meaning of symmetry in mathematics." In a course titled Mathematics of Symmetry, if this item doesn't show significant change, nothing will. And yet, while the average on a 5-point scale (from Not at All to A Great Deal) rose 0.8, almost a full point, the  $p$ -value for this item was only 0.18, not even marginally significant at the 0.10 level. For most other items, the  $p$ -values were higher. In 2006, where I was able to look at matched pairs, a few of the items showed marginally significant changes at the 0.10 level, but there were so few of these items that these results could have arisen by chance.

It seems unlikely I will ever have enough data from a single semester to use the difference of means to obtain statistically significant results. Of course, I could combine the results from the same question in different years, but that ignores differences in the course. Even if the syllabus is the same, when the course is offered only once every four years, there are going to be changes in my teaching. Moreover, the course is still under development, so the syllabus has been different each time, and that is likely to continue in the future.

In fact, it can be questionable to combine data from different sections of the same course taught in the same semester. For example, in 2010 I was teaching two sections of Mathematics of Symmetry. Both used the same syllabus and materials, and were taught back-to-back in the same room. However, in the final student evaluations, when asked to rate the "overall effectiveness of instruction," one section gave an average rating of 4.63 out of 5 (considerably above the university average of 4.36), and the other section gave an average rating of 3.52 (in the bottom 10% university-wide). I am still trying to understand the reasons for such a dramatic disparity, but it is clear there were differences between the two sections. Since I had both sections respond to the same online survey, I was unable to separate the sections in the data. One lesson I have learned for the future is to set up different surveys for each section!



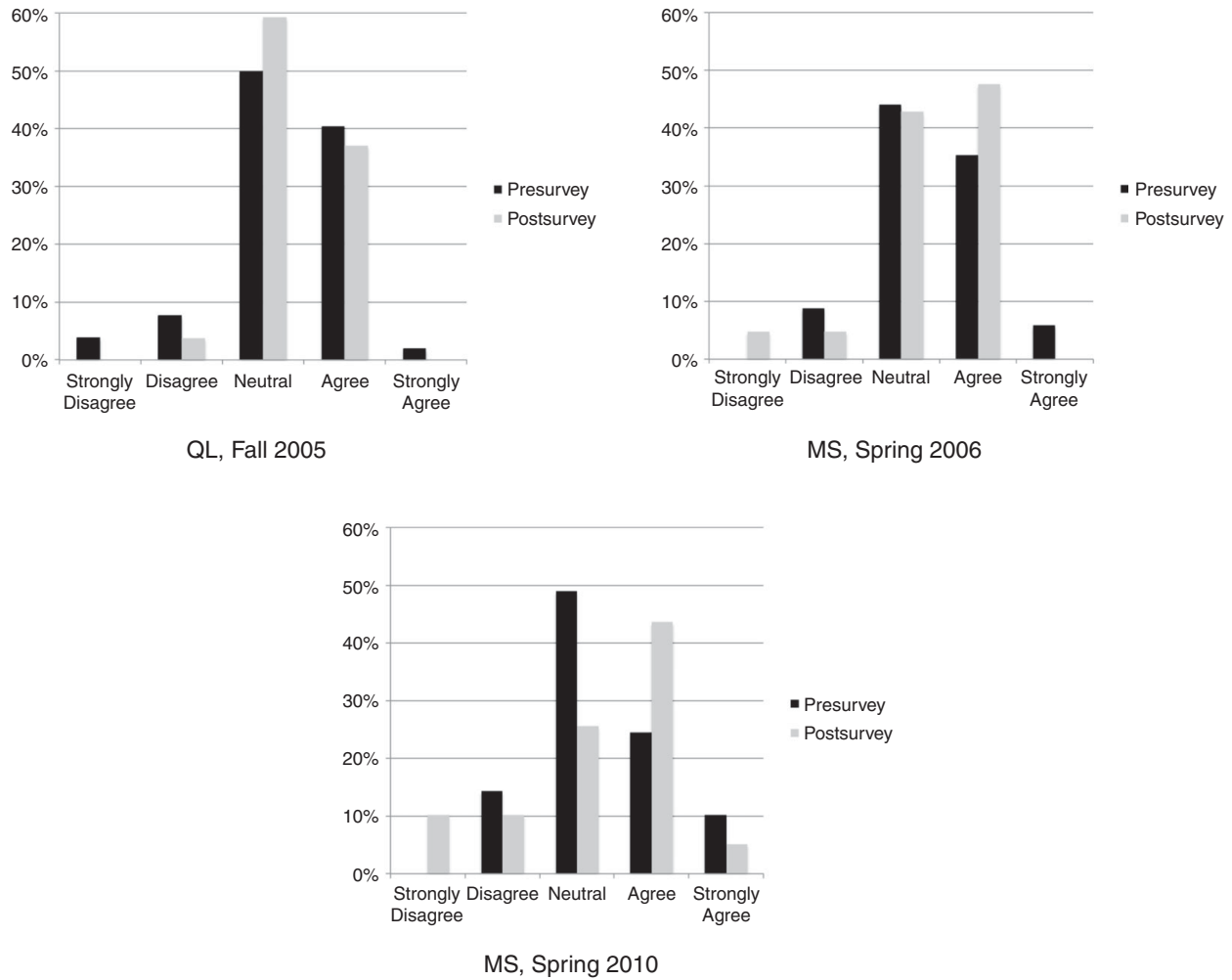


**Figure 16.1.** Mathematics is an important part of a liberal arts curriculum.

Instead of performing statistical tests, I found it more fruitful to look at the data qualitatively, and compare the distributions of responses at the beginning and end of the semester (though this didn't avoid the problem of combining two different sections mentioned before). The Likert-scale questions on the survey were divided into four categories: Confidence, Interest, Usefulness, and Mathematical Thinking. In all four categories, some questions seemed to show an effect, but many more did not. Of course, given enough questions, and considering the relatively small populations, we expect a few of the distributions to show significant changes just by chance. For this reason, I was most interested in questions where the change in the distribution from the beginning to the end of the course was similar in the two versions of the Mathematics of Symmetry (MS) course, and was different from the control group in the QL course, Quantitative Methods for the Modern World. I will discuss five such questions.

There were two related items in the category of Usefulness that showed improvement in both sections of MS and no improvement in QL. The first was "Mathematics is an important part of a liberal arts curriculum," and the second was "Mathematical thinking is useful in the humanities." The results for the three classes are shown in Figures 16.1 and 16.2. In both cases, the percentage of students responding Agree or Strongly Agree dropped slightly in QL but increased by roughly 10 percentage points in MS; these results indicate that MS was more successful than QL in convincing students that mathematics has applications outside of the sciences. Because at most 49% of the MS students agreed on the post-survey that mathematical thinking is useful in the humanities, I feel there is still room for improvement.

The statement, "There are math problems that just can't be solved by following a predetermined sequence of steps," from the Mathematical Thinking category produced dramatically different results in QL and MS. As shown in Figure 16.3, in QL, the percentage of students responding Agree or Strongly Agree dropped by over 30 percentage points. In MS, the percentage responding Agree or Strongly Agree increased an average of 18 percentage points. The distributions illustrate another difficulty with comparing classes in SoTL—it is rare that the students can be considered

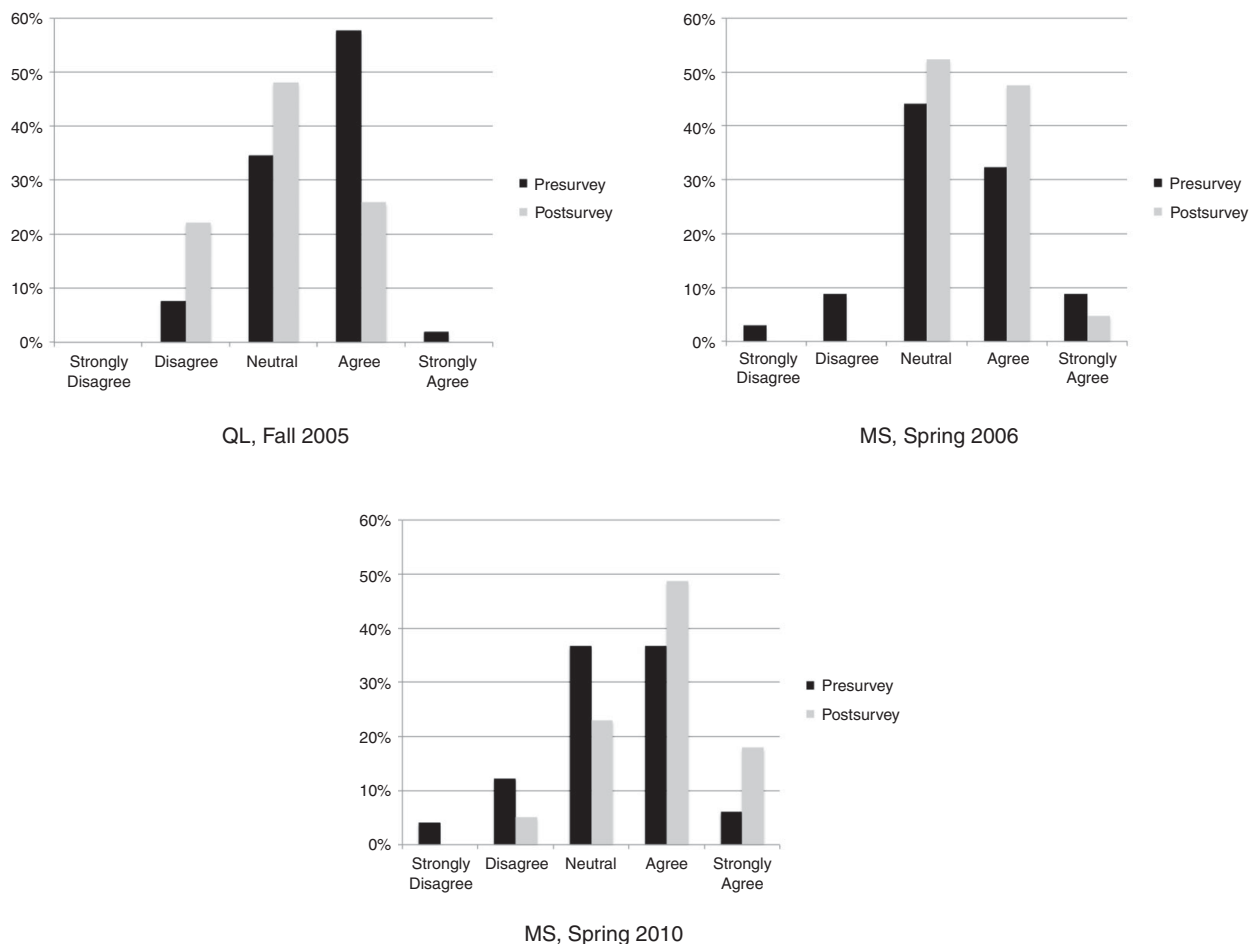


**Figure 16.2.** Mathematical thinking is useful in the humanities.

randomly assigned to the classes. In this case, it is interesting to observe the differences between the classes at both the beginning and end of the semester. Apparently, students taking QL began by believing that mathematics was less formulaic, but that belief was changed by the course. The students taking MS, on the other hand, began by viewing mathematics as more formulaic, and changed their views in the opposite direction. The differences reflect not just the differences in the courses (QL does, in fact, place more emphasis on applying algebraic formulas in prescribed ways), but also in the students who chose to take MS instead. It is possible they chose a course on the mathematics of symmetry in part because they had a formulaic view of mathematics, and were curious about a course that advertised itself as something different.

Another statement in the Mathematical Thinking category, “In mathematics you can be creative and discover things by yourself,” also produced dramatic differences between the classes. In QL, the percentage of students responding Agree or Strongly Agree dropped almost 30 points, while in MS the percentage dropped only 5 points in 2006, and increased 12 points in 2010. The drop in 2006 seems almost negligible, considering the remarkably high percentage of students agreeing with the statement at both the beginning and end of the course. In general, it seems that MS, with its focus on discovery learning and its opportunity for students to do creative projects, was more successful at maintaining or improving the attitude that creativity was part of doing mathematics.

One of the goals of developing MS was to help students with math anxiety. The item that addresses this most directly is “I have a lot of self-confidence when it comes to mathematics.” In 2010, due to changes in the online survey tool, this question was changed to “I am confident in my ability to solve mathematical problems,” and the response choices ranged from Not at All to A Great Deal rather than Strongly Disagree to Strongly Agree. This illustrates



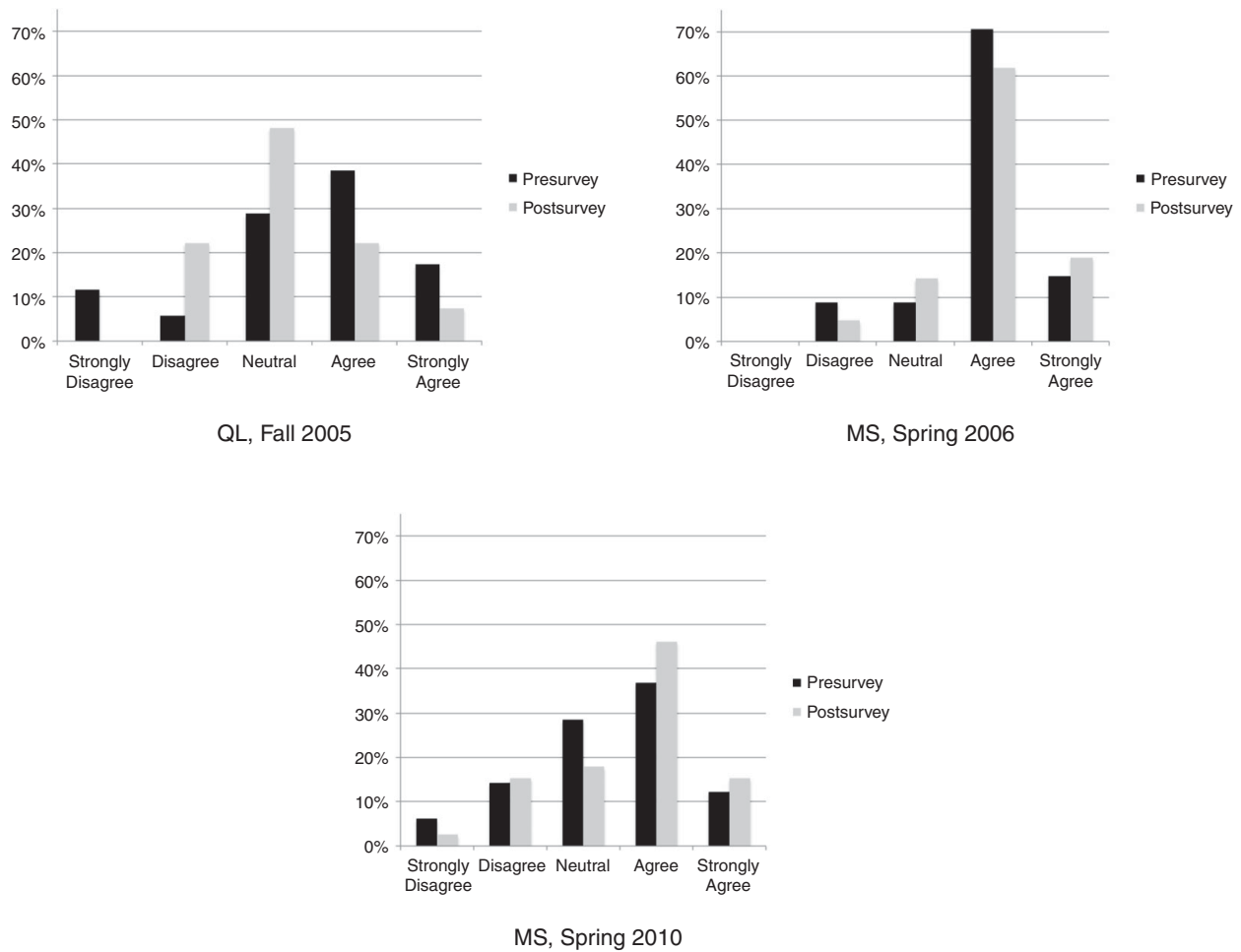
**Figure 16.3.** There are math problems that just can't be solved by following a predetermined sequence of steps.

another issue that frequently arises when doing SoTL work in mathematics. As mathematicians, we are not trained as social scientists, and will frequently make what, in retrospect, seem like obvious errors. In this case, the earlier form of the question should also have been included, to allow for better comparison. However, since each SoTL experiment requires a semester, we try to use what data we have collected despite any errors, and hope to compensate in some reasonable way when doing the analysis.

The results are presented in Figure 16.5. All three classes showed improvement, indicating that both types of courses could have a positive effect on students' confidence. To compare the courses, I used the responses Neutral, Agree, and Strongly Agree in 2005 and 2006, and the responses A Lot and A Great Deal in 2010. This was because the phrasing of the question in 2005 and 2006 (asking for a Lot of Self-Confidence) seemed stronger, and in 2010 the response option Somewhat seemed slightly more negative. The percentage of QL students showing an increase in confidence increased from 40% to 78% (38 percentage points); in MS the percentage increased from 33% to 57% (24 percentage points) in 2006, and from 36% to 56% (20 percentage points) in 2010. The similarity of the numbers between 2006 and 2010 provides justification for how I aggregated the results, and some confidence that the analysis is valid despite the change in the wording of the question. For this item, the more conventional course appeared to have better outcomes, but both courses had a clear (and substantial) positive change.

There were also four free-response questions on the survey:

1. Give at least three terms which describe your feelings when you hear the term "math problem."
2. List three ways you might use mathematics.
3. Briefly describe what you think mathematics is.
4. Briefly describe how (or whether) mathematics fits into a liberal arts curriculum.

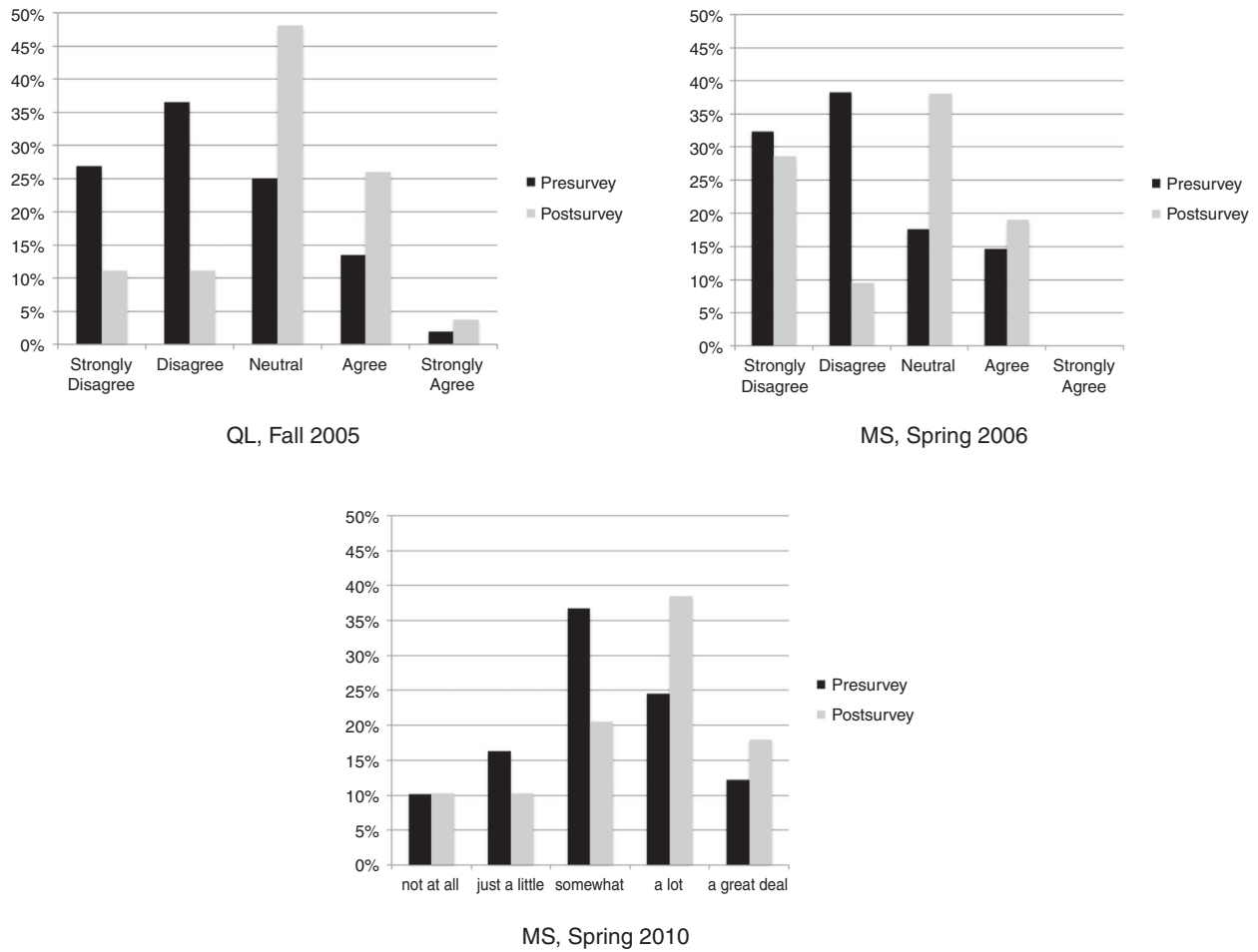


**Figure 16.4.** In mathematics you can be creative and discover things by yourself.

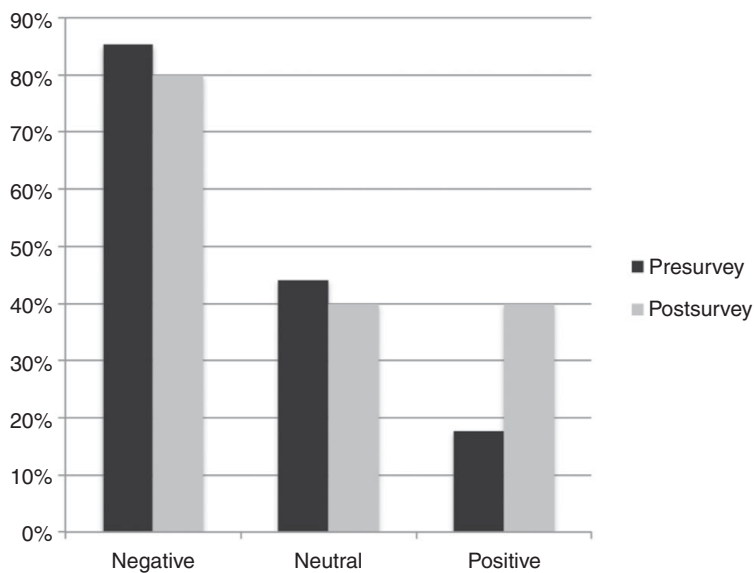
These questions illustrate another of the challenges in doing work in the scholarship of teaching and learning for those of us for whom it is not our primary research interest: How much time do we want to devote to analyzing our data? Answers to free-response questions are likely to provide a more nuanced view of students’ attitudes and beliefs than the Likert-scale questions, but they are more difficult and time-consuming to analyze.

I have managed to do some analysis of the responses for the first three questions for the 2006 MS course. For the first question, I categorized each term offered as Positive, Negative, or Neutral. Positive terms included “excited,” “interested,” “hopeful,” and “optimistic.” Negative terms included “anger,” “panic,” “sadness,” “difficult,” and “fear.” Neutral terms included “challenging,” “thinking,” “numbers,” and “algebra.” There is, of course, subjectivity in my categorization. For example, I decided that “difficult” carried a negative connotation but that “challenging” did not. A student might intend “challenging” to be positive or negative term, so I coded it as Neutral. To quantify the results, one could count the number of terms in each category (with each student contributing multiple terms), or the number of students who contributed a term. I chose the latter, since I am more interested in the students than the terms. The results are shown in Figure 16.6. There were 34 students who took the pre-survey and 21 who took the post-survey. At the end of the course almost as many students were giving negative responses as at the beginning, but the number of students who were supplying a positive response had more than doubled. It is interesting that many students at the end of the course were providing both positive and negative responses, indicating that their view of mathematics may have become more complex.

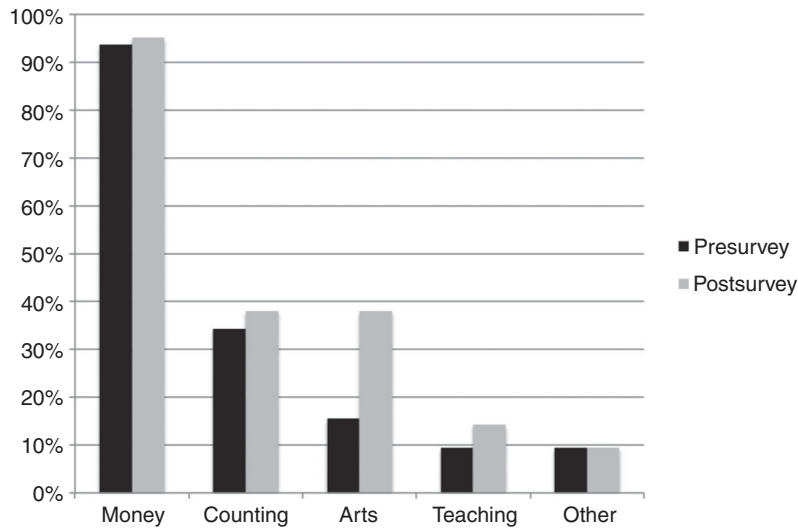
The second question asked students how they might use mathematics. Here, I categorized the responses as Money (including references to finance, taxes, stocks, etc.), Counting (including measurement, cooking, etc.), Arts (including drawing, film, music, etc.), Teaching, and Other. Figure 16.7 shows the percentage of students giving a response in



**Figure 16.5.** I have a lot of self-confidence when it comes to mathematics (2005, 2006) or I am confident in my ability to solve mathematical problems (2010).



**Figure 16.6.** (MS, 2006) Give at least three terms which describe your feelings when you hear the term “math problem.”



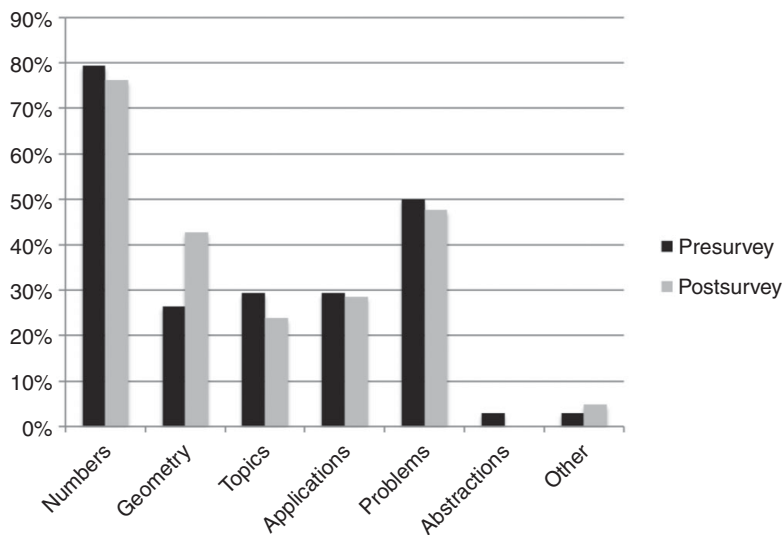
**Figure 16.7.** (MS, 2006) List three ways you might use mathematics.

each category. Financial applications of mathematics were most common, but we can see that over twice as many students listed applications to the arts at the end of the course. This is not surprising, considering the focus on the use of mathematics by artists, and it indicates that the course had some short-term effect.

Analyzing the responses to the third question, describing what mathematics is, was more complex, and I am indebted to Jacqueline Dewar for her assistance. Each response was categorized as referring to Numbers (and equations), Geometry (and shapes and patterns), listing Topics (e.g., “the study of numbers, patterns, formulas, and shapes”), Applications (talking about using mathematics to understand the world), Problem-solving (including references to proofs and critical thinking), Abstraction (talking about mathematics in terms of abstract structures and generalization), or Other. A single response could belong to several categories. For example, a response such as:

I believe that mathematics is the study of numbers, shapes, formulas and patterns and how they relate to the world around us. I think the main goal of math should be to recognize its presence in our world.

would be categorized as Numbers, Geometry (for the references to shapes and patterns), Applications (for “how they relate to the world around us”), and as listing Topics. The results (see Figure 16.8) show that the greatest change from



**Figure 16.8.** (MS, 2006) Briefly describe what you think mathematics is.

the beginning to the end of the course is the increase in the percentage of students who mention geometry, shapes, or patterns as part of their definition of mathematics (from 26% to 43%).

## Conclusions and Next Steps

What conclusions can we draw from the evidence so far? Unfortunately, SoTL rarely provides the kind of clear correct answers or proofs that mathematicians are used to. While some survey items showed meaningful differences between the courses, for many other items the differences were negligible or inconsistent. Some items that seem similar drew different responses, so there is still work to be done on the validity of the instrument. There were some consistent results, but few were surprising. For example, a course explicitly linking mathematics to the arts should cause students to see more connections between mathematics and the arts. But even then, with at most half of the students agreeing that mathematical thinking is useful in the humanities, the effect was not as great as I hoped. That at least some students began to see mathematics as a creative endeavor, and less as a matter of following recipes to solve problems, was encouraging. On the other hand, it was disappointing that there were no significant improvements in many of the items measuring students' interest in mathematics or in learning more mathematics.

Some of these results have been presented in several ways, including a panel presentation on quantitative literacy at the Third Annual ISSOTL Conference (Dewar, 2006), at a workshop on mathematics education via the arts held at the Banff International Research Station (Mellor, 2007a), and as a contributed paper at MathFest 2007 (Mellor, 2007b). The discussions at the conferences provided helpful ideas that I incorporated when I again taught Mathematics of Symmetry in 2010.

There are still more changes I plan to make when I next teach the course. Based on the survey results, I would emphasize the creative aspects of the course and de-emphasize the aspects that are like rote calculation. I plan to include a unit on fractals and self-symmetry, as a natural extension of rigid symmetries, and because the visual beauty of fractals provides a natural connection to the arts. I still believe in the potential of the course to change the way that students view mathematics. The results of my project so far encourage me that this is possible and also indicate how much work there still is to do.

I will also continue to think about what data to collect and how to capture it. It would be interesting to collect and analyze examples of student work to assess whether they are achieving the desired learning outcomes. Though the reality is that I scanned all the students' homework assignments in 2006, I have yet to take another look at them. I plan to retain examples of the projects and exams, but any analysis in the foreseeable future is likely to be restricted to the results of the surveys. I do not plan to make any significant changes to the survey items before next offering the course, but I will be sure to administer it separately to each section.

## Personal and Professional Impact

Over time this project has evolved from asking "what effect does this course have on student attitudes towards mathematics?" to "how can I use the results of my surveys to modify my class to achieve better outcomes?" The project has become less of a research project, with an expected outcome and end date, and more a part of the process by which I revise and continue to develop the course. At some point, when the course is perfected (or at least is much less imperfect than it currently is), I may try to use the data collected over the years to promote the result. But I don't expect this to happen any time soon. In the meantime, working on this project has changed the way I think about assessing and improving all of my classes and has made me a more thoughtful teacher. When I began working in SoTL, I was uncertain whether I wanted to make it a major component of my research. I now realize that the answer is "No." I do not have the expertise (or interest) in analyzing qualitative and quantitative data to make the scholarship of teaching a major component of my scholarship; however, I am interested in making it a larger component of my teaching. So I view the project less as scholarship than as professional development. My goal is not public dissemination of the results, but personal improvement of my teaching. I will present results at conferences and meetings, but to obtain feedback rather than to convince others to follow my lead. If at some point I have developed clear enough results to be worth a more formal scholarly dissemination, that will be a wonderful bonus, but it is no longer the primary motivation or goal of the work.

## Acknowledgement

This research was supported in part by a 2005–6 Faculty Development Grant from the Loyola Marymount University Center for Teaching Excellence.

## References

- Burger, E., & Starbird, M. (2005). *The heart of mathematics: An invitation to effective thinking* (2nd ed.). New York, NY: Key College Publishing.
- COMAP (Consortium for Mathematics and its Applications). (2006). *For all practical purposes: Mathematical literacy in today's world* (7th ed.). New York, NY: W. H. Freeman and Co.
- Dewar, J. (2006, November). *Multiple facets of quantitative literacy*. Panel discussion at the Third Annual ISSOTL Conference, Washington, DC. Abstract and Synthesis retrieved from [www.indiana.edu/~issotl/ISSoTL2006\\_abstractsDG.pdf](http://www.indiana.edu/~issotl/ISSoTL2006_abstractsDG.pdf)
- Farmer, D. (1996). *Groups and symmetry: A guide to discovering mathematics*. Providence, RI: American Mathematical Society.
- Kloosterman, P., & Stage, F. (1992). Measuring beliefs about mathematical problem solving. *School Science and Mathematics, 92*(3), 109–115.
- Mellor, B. (2007a, January). *Mathematics of symmetry: An experimental core mathematics course*. Paper presented at the BIRS Workshop on Innovations in Mathematics Education via the Arts, Banff, Canada.
- Mellor, B. (2007b, August). *How does a course in the mathematics of symmetry affect students in the liberal arts?* Paper presented at MathFest 2007, San Jose, CA.
- Mulhern, F., & Rae, G. (1998). Development of a shortened form of the Fennema-Sherman Mathematics Attitudes Scales. *Educational and Psychological Measurement, 58*, 295–306.
- Redish, E. F., Steinberg, R. N., & Saul, J. M. (1998). Student expectations in introductory physics. *American Journal of Physics, 66*, 212–224
- Schoenfeld, A. (1989). Explorations of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education, 20*, 338–355.
- Tapia, M., & Marsh, G. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly, 8*(2), 1–6.

## Appendix: The Survey

In 2005 and 2006, the survey included the following 36 Likert-scale items. Students were asked to respond Strongly Disagree, Disagree, Neutral, Agree, or Strongly Agree with each item. The questions also used in 2010 are marked with an asterisk.

1. (\*) Mathematics is the study of patterns.
2. (\*) If I can't do a math problem in a few minutes, I probably can't do it at all.
3. (\*) I will use mathematics in many ways as an adult.
4. (\*) In mathematics you can be creative and discover things by yourself.
5. (\*) Mathematics makes me feel uncomfortable.
6. Everything important about mathematics is already known by mathematicians.
7. (\*) Mathematics is mostly facts and procedures that have to be memorized.
8. (\*) I am interested in learning more mathematics.



9. (\*) I believe studying math helps me with problem solving in other areas.
10. (\*) Mathematics is dull and boring.
11. The skills I learn in my math classes don't help me in any other classes.
12. After a conjecture has been proven, it is possible to find a counterexample.
13. (\*) To really solve a math problem, you need to do more than get the right answer, you also need to understand why the solution works.
14. (\*) Mathematics is an important part of a liberal arts curriculum.
15. (\*) To solve math problems you have to be taught the right procedure, or you can't do anything.
16. I have a lot of self-confidence when it comes to mathematics.
17. I see mathematics as a subject I will rarely use in daily life as an adult.
18. (\*) People who do well in mathematics are just naturally good at it.
19. (\*) Mathematics is the study of numbers and equations.
20. (\*) I really like mathematics.
21. I don't care why a mathematical procedure works as long as I can get the correct answer.
22. (\*) Mathematical thinking is useful in the humanities.
23. Mathematics is a very interesting subject.
24. The topics covered in a math class are often unrelated.
25. Math is one of my least interesting classes.
26. Math is a growing field.
27. (\*) The main benefit of math classes is learning how to reason logically.
28. The best way for me to learn mathematics is by solving many problems rather than by carefully analyzing a few in detail.
29. (\*) Trial and error and experimentation are often useful strategies in mathematics.
30. Students majoring in the arts and humanities don't need to know college-level math.
31. If I see five examples where a formula holds, then I am convinced the formula is true.
32. I find I can do hard math problems if I just hang in there.
33. I expect to do fairly well in any math class I take.
34. (\*) I like to solve new problems in mathematics.
35. (\*) The main benefit from studying mathematics is developing the ability to follow procedures.
36. (\*) There are math problems that just can't be solved by following a predetermined sequence of steps.

The following questions were used in 2010. Students were asked to respond Not at All, Just a Little, Somewhat, A Lot, or A Great Deal to each item.

1. I understand the meaning of symmetry in mathematics.
2. I understand the meaning of a mathematical classification.
3. I can classify visual patterns according to their symmetries.
4. I can recognize patterns and make conjectures.
5. I can analyze examples to test conjectures.
6. I can use logical reasoning to prove conjectures.
7. I am confident in my ability to solve mathematical problems.
8. I am interested in mathematics.
9. I am in the habit of using skills learned in math classes in other areas of my life.

10. I am in the habit of using mathematics to inspire creative ideas.
11. I am in the habit of using mathematics to help describe the world around me.

In all three years the survey included the free-response questions:

1. Give at least three terms which describe your feelings when you hear the term “math problem.”
2. List three ways you might use mathematics.
3. Briefly describe what you think mathematics is.
4. Briefly describe how (or whether) mathematics fits into a liberal arts curriculum.



# 17

## Mathematics Research Experiences for Preservice Teachers: Investigating the Impact on Their Beliefs

Wendy A. O'Hanlon, David D. Barker, Cynthia W. Langrall, John A. Dossey,  
Sharon M. McCrone, and Saad I. El-Zanati  
*Illinois State University*

### Editors' Commentary

In this chapter a group at Illinois State University describe how they used the scholarship of teaching and learning to investigate whether having preservice teachers participate in a mathematical research experience for undergraduates (REU) program influenced their beliefs about teaching and learning mathematics. Unable to find an appropriate survey instrument, they developed their own. They explain the organization and content of the survey, tell how they piloted and tested it for reliability, and describe how the results are being used to improve the REU program.

### Introduction

The Scholarship of Teaching and Learning (SoTL) can serve as a critical tool for improving instruction at the post-secondary level. If we, mathematicians and mathematics educators, can systematically explore, share, and reflect on our understanding of quality collegiate instruction, improvement in overall course quality and student learning can follow. The multifaceted nature of student learning presents a challenge for any educational study. One such facet is student beliefs. For example, an effective instructional strategy for teaching an honors calculus course for mathematics majors may not translate to a calculus course designed for business majors because of the different attitudes and beliefs of these two groups. Differences in students' views of the nature of mathematics, the importance of building mathematical understanding, the beauty of mathematics, or the standards for reasoning and proof, all can influence the success of instruction. In this chapter, we describe a SoTL study designed to inform and improve a mathematics education program that engages future secondary teachers in authentic mathematics research experiences. One of the goals of the program, the *REU Site: Mathematics Research Experience for Pre-service and In-service Teachers* (REU), is to change teachers' beliefs about mathematics and about the teaching and learning of mathematics and thus to influence the way they teach.

### Impact of Beliefs on Secondary Mathematics Teachers

Numerous researchers have cited the impact of beliefs on teaching and learning (Ernest, 1989; Philipp, 2007; Schoenfeld, 1985; Silver, 1985; Thompson, 1992). Research on inservice mathematics teachers has shown that classroom

practice, including activities and actions, is tied to a teacher's beliefs (Correa, Perry, Sims, Miller, & Fang, 2008; Lloyd, 2002; Lubinski, & Jaberg, 1997; Wilson & Cooney, 2002). Ebert and Risacher (1996) found that "Teachers' beliefs about how students learn mathematics, their beliefs about mathematics itself, and their knowledge of teaching in general, are likely to affect how they design and teach lessons" (p. 5). Different beliefs about teaching and learning mathematics may cause a teacher to use different instructional practices and can affect how concepts and skills are conveyed through instruction (Ebert & Risacher, 1996). Ball (1990) suggested that teacher preparation programs must address preservice teachers' beliefs about mathematics, as studies have shown they influence the way mathematics is understood. Swars, S. Smith, M. Smith, and Hart (2008) stated, "The pressing goal of teacher education programs should be the facilitation of these changes during program experiences" (p. 47). Unfortunately, research has indicated that traditional mathematics courses do little to challenge or positively change future teachers' beliefs and actions (Brown & Borko, 1992; Brown, Cooney, & Jones, 1990).

Beliefs do not exist in isolation, but comprise a system that is interrelated, often in complex ways (Green, 1971). Beliefs about the nature of mathematics and beliefs about the teaching and learning of mathematics are related. One of the main goals of our REU program has been to address both sets of beliefs concurrently. To assess and improve our program we wanted to know the beliefs our preservice teachers had at the start and how they change over the course of the REU experience. Knowing what our preservice teachers believe would help us to design our program to maximize its benefit. Understanding how they change would enable us to infer influence of the program on their beliefs. For instance, fostering the belief that exploration and making conjectures are valuable components of the learning process can influence the way preservice teachers will teach. Through our SoTL investigation we sought to answer the questions:

- What are the initial beliefs of preservice secondary mathematics teachers about the teaching and learning of mathematics?
- What changes occur in preservice secondary mathematics teachers' beliefs about the teaching and learning of mathematics after their participation in the REU?

## Developing a Beliefs Instrument

To measure beliefs about teaching and learning mathematics, we reviewed the literature to find instruments that could be used for this purpose. However, after an initial trial of several instruments (Day, 1993; Fennema, Carpenter, & Peterson, 1986; Tabachnick, & Zeichner, 1985), we found they did not serve our purposes. Since we were examining the beliefs of secondary mathematics education majors, we wanted to focus on content and pedagogy in the secondary mathematics education program (e.g., instructional strategies, curriculum, student learning, technology, assessment). Also we wanted to reflect the recommendations of NCTM's *Principles and Standards for School Mathematics* (2000). We found that none of the instruments used in our trial were suitable, so we began to develop a teaching and learning beliefs survey (Appendix A).

The first step was to identify teacher beliefs that could affect teacher practice and that could be influenced by a mathematics research experience. To do this we reviewed the recommendations of the Conference Board of Mathematical Sciences (CBMS) for preparing mathematics teachers (CBMS, 2001, 2012) and a bibliography on mathematical habits of mind found in Cuoco, Goldenberg, and Mark (2010). We chose to focus on mathematical habits of mind because of their connection to the process of mathematics research. These habits, important components of our REU program, include finding, describing, and explaining patterns; creating and using representations; generalizing from examples; stating generality in precise language; and expecting mathematics to make sense. Cuoco et al. (2010) stated that, "our teaching experience and our work with teachers convinced us that raising the methods used by mathematicians to the same level of importance as the results of those methods would be a viable organizer for school curriculum" (p. 683). By analyzing the CBMS recommendations and the mathematical habits of mind, we identified these categories about teacher beliefs and attitudes for inclusion in our survey: procedural versus conceptual understanding, teaching through discovery versus presentation; problem solving; reasoning and proof; connections; cognitively guided versus mathematically guided; perseverance; mathematical authority; communication; technology; and, as a single grouping, attitude, confidence, and efficacy.

We then developed sets of items for each category. To produce a more reliable instrument, we used both positive and negative statements. We also included statements to discern differences between preservice teachers' beliefs about their learning, their beliefs about how students learn mathematics, and their beliefs about the teaching of mathematics. We felt it was important to make these distinctions because the research experience could affect the preservice teachers' beliefs about their learning, but not change their beliefs about teaching or their students' learning. Surveys were constructed using a Likert scale ranging from 1 to 5, with 1 representing Strongly Disagree and 5 Strongly Agree. During data analysis, the data from negatively phrased items were reversed so that higher scores represented productive beliefs consistent with those described in *The Mathematical Education of Teachers* (CBMS, 2001, 2012) and in the *Mathematical Practices of the Common Core State Standards* (National Governors Association, 2010).

After a trial of the survey with a group of secondary mathematics education students, we examined the results. We determined clusters of items that correlated with one another to see if those clusters mapped to the belief areas we were investigating, using factor analysis.

We also looked carefully at the data, reread each item for how it might be interpreted by a student, and identified ambiguous statements that contained multiple ideas or allowed for more than one interpretation. It is impossible to determine what a response, such as Agree, means for an ambiguous statement because it can be viewed from more than one perspective. Those items were either rewritten or deleted. Appendix B contains a matrix that shows the assignment of each item in the final form of the survey to one of the belief categories (e.g., problem solving, reasoning and proof, communication) and perspectives (personal learning, student learning, or teaching).

We also examined the percentage of variance explained by each individual item in relationship to the overall survey results. This shows the ability of an item to discriminate among the respondents. If everyone answers an item the same way, the survey designer would remove it, unless knowing about (nearly) unanimous agreement on a certain statement was of particular interest.

Once the final version of the survey was determined, we calculated its reliability using Cronbach's alpha, a standard test for consistency or stability of a test or survey instrument (McDonald, 1999). Research instruments are expected to obtain  $\alpha$ -values of at least 0.70 for acceptable use in research settings. Nunnally (1978) proposed that values of  $\alpha \geq 0.80$  are sufficient unless decisions determining the fates of individuals (e.g., grades, admittance, promotions) are involved. As our value of  $\alpha = 0.80$  meets the stronger criterion, the present instrument appears to be sufficiently reliable for measuring overall beliefs about teaching and learning mathematics.

## Investigating Change in Preservice Teacher Beliefs

To examine changes in beliefs, the preservice teachers in the REU were asked to complete the survey at the beginning and end of the program. We examined differences in pre- and post-mean response scores for changes having an effect size,  $|ES|$ , greater than 0.5. The effect size is a measurement of the difference of the group's average response data at two points in time. While there are many ways of defining the concept (see Hedges & Olkin, 1985), we used the group's average score at the end of the experience minus the group's average score at the beginning divided by a measure of standard deviation of the scores at the beginning, that is,

$$ES = \frac{\bar{x}_{\text{end}} - \bar{x}_{\text{begin}}}{\bar{s}_{\text{begin}}}$$

So items satisfying  $|ES| > 0.5$  have shown a change of at least a half a standard deviation. Table 17.1 shows the twelve items that resulted in changes of that magnitude for participants in the summer 2011 REU.

The first nine statements in Table 17.1 correspond to beliefs about discovery learning, problem solving, procedural versus conceptual understanding, and reasoning and proof. Each pre- to post-mean score change in these categories represented movement toward greater alignment with the vision in the CBMS recommendations (CBMS, 2001, 2012). For example, preservice teachers strengthened their belief (3.83  $\rightarrow$  4.50) that students should be given time to explore mathematical ideas. This strong positive change in belief concerning the role of exploration may influence their use of exploration in their future teaching. Their belief that the teacher should be the sole provider of mathematical verification of argument weakened, a finding that suggests a move to favoring instruction that asks students to provide justification.

Statement	ES	Mean Scores: Pre → Post 1: Strongly Disagree → 5: Strongly Agree
I learn best when I am given time to explore a mathematical idea.	0.80	3.83 → 4.50
Students should engage in problem solving before they master computational procedures and basic concepts.	0.64	3.00 → 3.67
Students should not attempt problem solving until they understand basic concepts.	0.67	2.58 → 3.58
Students should understand computational procedures before they spend time practicing them.	0.61	3.50 → 2.83
The teacher should provide verification for mathematical arguments rather than expecting students to do so.	0.84	2.92 → 2.25
During instruction, teachers should emphasize the interconnections among mathematical ideas.	0.58	4.17 → 4.50
Struggling with mathematical concepts is beneficial to developing understanding.	0.56	4.08 → 4.58
The structure of mathematics as a subject is more important in making instructional decisions than the natural development of students' ideas.	0.65	2.92 → 2.58
If a student is having difficulties solving a problem, then the teacher should tell the student how to solve it.	0.75	2.08 → 1.58
I do not like to be challenged in mathematics class and would prefer the teacher guide me to the correct solution quickly.	1.18	1.33 → 1.92
During class discussions, students should analyze and critique other students' work.	0.51	4.33 → 4.00
Calculators can be used to explore and investigate mathematical concepts.	0.65	4.58 → 4.25

**Table 17.1.** Items with effect size  $|ES| > 0.5$  and corresponding pre- and post-mean score responses for summer 2011 REU students

Beliefs changed about whether or not students should be using and practicing computational procedures before they have developed an understanding of them. Their beliefs about themselves as learners also changed. We interpreted these results as indicating that our program has had a positive influence on certain aspects of our preservice teachers' beliefs.

However, simply identifying what went well was not the ultimate goal of our SoTL investigation. To improve our program and collegiate mathematics instruction in general, it was also important to identify aspects of student beliefs that remain unchanged or that changed in ways other than what we intended. As shown by the last three statements in Table 17.1, not all student beliefs moved in the desired direction. Some of the statements concerning perseverance, mathematical authority, and technology showed changes opposite to what we desired.

There were also three items for which mean scores (shown in parentheses) did not change from the beginning to the end of the program:

- I prefer to learn mathematics topics separately rather than connecting mathematical ideas (1.93);
- It is not necessary for students to understand the interconnections among mathematical ideas as long as they have some understanding of the individual topics (2.08); and
- I feel confident in my abilities to solve mathematics problems (4.42).

The identical pre- and post-mean response scores for the first two items indicated the REU had little impact on preservice teachers' beliefs about the importance of building connections. We conjectured that the REU did not provide experiences that prompted changes in this area. Because connecting mathematical ideas is important for mathematics research and for developing a deep understanding of mathematics we plan to target this belief more directly. We will include a problem-solving situation designed to highlight mathematical connections and incorporate a discussion of their role in mathematics. By adding these two components early in the program, we hope to increase the number of connections preservice teachers make during the REU and to increase their belief that connections are important.

The third item indicates that, on average, our preservice teachers' mathematical confidence remained unchanged. The confidence of some preservice teachers increased while the confidence of others decreased. Given the nature of mathematics research, increasing confidence can be difficult. The choice of problem and the support provided contribute to the likelihood of success, but do not guarantee it. However, our survey revealed that, on average, our preservice teachers both entered and left the program with a high level of confidence. Given that our REU program selected extremely talented preservice teachers from across the nation, this is not surprising. Nevertheless, our SoTL study does raise the question of the impact of mathematics research experiences on an average group of secondary mathematics education majors. In this situation, the issue of confidence may need to be addressed.

## Concluding Remarks

In this chapter, we have described the development of a beliefs survey that was used to study our *REU Site: Mathematics Research Experience for Pre-service and In-service Teachers* at Illinois State University. Incorporating mathematics research experiences into secondary mathematics teacher preparation provides future and practicing teachers the opportunity to learn mathematics from a discovery approach. We believe this will strengthen their knowledge of mathematics and their ability to teach. One purpose of the REU is to change teachers' beliefs about mathematics and hence their beliefs about teaching and learning. We believe that including research experiences in teacher preparation programs can assist future teachers to develop the mathematical capability to provide instruction that is consistent with current recommendations (CBMS, 2001, 2012; NCTM, 1991, 2000; National Governors Association, 2010).

Our SoTL investigation increased our understanding of the incoming beliefs of our preservice teachers and how these beliefs changed. In developing the survey, we began to think more deeply about teacher beliefs. This enabled us to identify areas for improvement and to adjust our program accordingly. However, we are still trying to understand how preservice teachers' beliefs are changing and what might be affecting their belief structures during the REU program. Thus, we are introducing the use of journals for participants to record weekly reflections.

The data we have obtained from our SoTL investigation have been used to inform the development of our program from year-to-year, to assist in our grant reports to the National Science Foundation, and to make presentations at national conferences. Although the design of our study does not enable us to attribute changes in beliefs directly to participation in the REU, the research tools and findings provide valuable insights that can guide us as we continue to work on preservice teacher beliefs. More factors need to be investigated to improve our understanding of how the research experiences for preservice teachers affect their classroom practice. It would be beneficial to follow our REU participants into their classrooms to determine the influence of the REU experience. How often do they use the ideas discussed in the REU? Are the beliefs that were developed in the REU evident in their teaching? Are mathematical habits of mind evident in their planning and execution of instruction? Do they incorporate authentic mathematical experiences into their instruction? We will continue to use the SoTL to revise and adjust our REU program to maximize its benefit. Once data from future SoTL studies help us understand what aspects of doing mathematics research have the greatest impact on preservice teachers, other institutions may be convinced to adopt similar programs.

## Acknowledgment

This material is based upon work supported by the Division of Mathematical Sciences at the National Science Foundation under Award A1063038. This report reflects the views and positions of the authors, and not necessarily the views of the National Science Foundation.



## References

- Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *Elementary School Journal*, 90, 449–466.
- Brown, C. A., & Boriko, H. (1992). Becoming a mathematics teacher. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 209–239). Reston, VA: National Council of Teachers of Mathematics.
- Brown, S. I., Cooney, T. J., & Jones, D. (1990). Mathematics teacher education. In W. R. Houston, M. Haberman, & J. Sikula (Eds.), *Handbook of research on teacher education* (pp. 639–656). New York, NY: Macmillan.
- CBMS (Conference Board of the Mathematical Sciences). (2001). *The Mathematical Education of Teachers*. Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.
- CBMS (Conference Board of the Mathematical Sciences). (2012). *The Mathematical Education of Teachers II*. Providence, RI and Washington, DC: American Mathematical Society and Mathematical Association of America.
- Correa, C. A., Perry, M., Sims, L. M., Miller, K. F., & Fang, G. (2008). Connected and culturally embedded beliefs: Chinese and US teachers talk about how their students best learn mathematics. *Teaching and Teacher Education*, 24, 140–153.
- Cuoco, A., Goldenberg, E. P., & Mark, J. (2010). Organizing a curriculum around mathematical habits of mind. *Mathematics Teacher*, 103(9), 682–688.
- Day, R. P. (1993). *Emerging and evolving themes: Three case studies of preservice mathematics teachers' beliefs*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.
- Ebert, C. L., & Risacher, B. F. (1996, April). "Alternative pathways to teaching"—An investigation of the factors that influence the acquisition of pedagogical content knowledge for traditional and non-traditional teachers. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.
- Ernest, P. (1989). The impact of beliefs on the teaching of mathematics. In P. Ernest (Ed.), *Mathematics teaching: The state of the art* (pp. 249–254). London: Falmer Press.
- Fennema, E., Carpenter, T. P., & Peterson, P. L. (1986). *Mathematics beliefs scales*. Unpublished manuscript, University of Wisconsin, Madison.
- Green, T. F. (1971). *The activities of teaching*. New York, NY: McGraw-Hill.
- Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press.
- Lloyd, G. (2002). Mathematics teachers' beliefs and experiences with innovative curriculum materials. In G. C. Leder, E. Pehkonene, & G. Torner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 149–159). Boston, MA: Kluwer Academic Publishers.
- Lubinski, C. A., & Jaberg, P. A. (1997). Teacher change and mathematics K–4: Developing a theoretical perspective. In B. S. Nelson (Ed.), *Mathematics teachers in transition* (pp. 223–254). Mahwah, NJ: Lawrence Erlbaum.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum Associates.
- NCTM (National Council of Teachers of Mathematics). (1991). *Professional standards for the teaching of mathematics*. Reston, VA: NCTM.
- NCTM (National Council of Teachers of Mathematics). (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Retrieved from [www.corestandards.org/Math/](http://www.corestandards.org/Math/)
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

- Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257–315). Reston, VA: National Council of Teachers of Mathematics.
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, FL: Academic Press.
- Silver, E. A. (1985). Research on teaching mathematical problem-solving: Some underrepresented themes and directions. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 247–266). Hillsdale, NJ: Lawrence Erlbaum.
- Swars, S. L., Smith, S. Z., Smith, M. E., & Hart, L. C. (2009). A longitudinal study of effects of a developmental teacher preparation program on elementary prospective teachers' mathematics beliefs. *Journal of Mathematics Teacher Education*, 12, 47–66.
- Tabachnick, B. R., & Zeichner, K. M. (1985). *The development of teacher perspectives: A final report*. Madison, WI: Wisconsin Center for Education Research.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127–146). New York, NY: Macmillan.
- Wilson, M., & Cooney, T. J. (2002). Mathematics teacher change and development. In G. C. Leder, E. Pehkonene, & G. Torner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 127–147). Boston, MA: Kluwer Academic Publishers.

## Appendix A

### Teaching and Learning Mathematics Beliefs

Please use the following codes to respond to each of the following statements by indicating whether you:

- 5: Strongly Agree with the statement
- 4: Agree with the statement
- 3: are Undecided about the statement
- 2: Disagree with the statement
- 1: Strongly Disagree with the statement

1. It is important for a student to discover mathematics.
2. Students should not attempt problem solving until they understand basic concepts.
3. Students should learn mathematics by developing and investigating mathematical conjectures, arguments, and proofs.
4. It is not necessary for students to understand the interconnections among mathematical ideas as long as they have some understanding of the individual topics.
5. I do not like to be challenged in mathematics class and would prefer the teacher guide me to the correct solution quickly.
6. During class discussions, students should analyze and critique other students' work.
7. Allowing students to discuss mathematics with a partner, or in a group, is an important instructional strategy.
8. Students should not be allowed to use calculators until they have mastered the mathematics concepts.
9. It is important for my students to be confident in their mathematical ability.
10. Students should spend time practicing computational procedures before they are expected to understand those procedures.
11. The teacher should demonstrate how to solve mathematics problems before the students are allowed to solve problems.

12. I rely more on logic than on external authority (e.g., the textbook or a teacher) to determine the soundness of a mathematical argument.
13. Students should understand how mathematical ideas interconnect and build on one another to produce a coherent whole.
14. During class discussions, the teacher should be the authority in terms of whether a student's mathematical conjecture or justification is correct.
15. Calculator use allows students to spend more time on critical thinking.
16. I feel confident in my abilities to solve mathematics problems.
17. I need to practice using a computational procedure before I can understand how that procedure works.
18. I learn mathematics through practicing skills and procedures.
19. Teachers should provide opportunities for students to critique mathematical arguments and discuss their own conjectures.
20. I like to be challenged in mathematics class.
21. During class discussions, students should play a role in determining whether mathematical justifications are valid.
22. Students should use calculators primarily to check their pencil and paper computations.
23. It is important for my students to find mathematics interesting.
24. I learn best when I am given time to explore a mathematical idea.
25. The best way to teach problem solving is to focus on one type of mathematics problem at a time.
26. During instruction, teachers should teach each mathematical idea separately rather than emphasizing the interconnections among ideas.
27. I prefer to not work in groups, as I learn better on my own.
28. Calculator use will cause a decline in basic computational skills.
29. Students should engage in problem solving before they master computational procedures and basic concepts.
30. If a teacher encounters a student who is struggling with a problem, she should ask a question that will help the student find a solution, but should not directly give the answer.
31. Struggling with mathematical concepts is detrimental to developing understanding.
32. Group discussions often lead to tangents, or incorrect mathematics, and should be limited in their use.
33. Calculators can be used to explore and investigate mathematical concepts.
34. My job as a teacher is to make the mathematics interesting and engaging.
35. Students learn mathematics best from teachers' demonstrations and explanations.
36. I rely on external authorities like the textbook, or a teacher, to judge the soundness of a mathematical argument.
37. During instruction, teachers should emphasize the interconnections among mathematical ideas.
38. Struggling with mathematical concepts is beneficial to developing understanding.
39. I feel inadequate in my abilities to solve mathematics problems.
40. Students should understand computational procedures before they spend time practicing them.
41. I learn best when my teacher provides a clear explanation.
42. I primarily use problem solving to apply mathematics rather than to learn mathematics.
43. Students should learn mathematics by studying mathematical arguments and proofs presented by the teacher or shown in the textbook.
44. I prefer to learn mathematics topics separately rather than connecting mathematical ideas.
45. I learn best when I am able to share and discuss my mathematical ideas with peers.
46. Students understand mathematical ideas well only when they can explain ideas in their own words.
47. I learn mathematical ideas when I can see how they interconnect and build on one another.

48. The structure of mathematics as a subject is more important in making instructional decisions than the natural development of students' ideas.
49. The teacher should provide verification for mathematical arguments rather than expecting students to do so.
50. If a student is having difficulties solving a problem, then the teacher should tell the student how to solve it.
51. I need to understand computational procedures before I can use them.
52. Teachers should encourage students to invent ways to solve mathematics problems.
53. I learn mathematics through problem solving.
54. I worry about my abilities to teach mathematics.
55. Only advanced students should use calculators.

## Appendix B

### ISU Teaching and Learning Belief Matrix

Core Idea	Personal Learning	Student Learning	Teaching
Procedures vs. Understanding	Positively worded (Pos): 51 Negatively worded (Neg): 17	Pos: 40 Neg: 10, 18	
Discovery vs. Presenting	Pos: 24 Neg: 41	Pos: 1 Neg: 35	Pos: 52 Neg: 11
Problem Solving	Pos: 53 Neg: 42	Pos: 29 Neg: 2	Neg: 25
Reasoning and Proof	Pos: 12 Neg: 36	Pos: 3 Neg: 43	Pos: 19 Neg: 49
Connections	Pos: 47 Neg: 44	Pos: 13 Neg: 4	Pos: 37 Neg: 26
Cognitively Guided vs. Mathematically Guided (content guided)			Neg: 48
Persevering vs. Immediacy	Pos: 20 Neg: 5	Pos: 38 Neg: 31	Pos: 30 Neg: 50
Mathematical Authority			Pos: 6, 21 Neg: 14
Communication	Pos: 45 Neg: 27	Pos: 46	Pos: 7 Neg: 32
Attitude, Confidence, and Efficacy	Pos: 16 Neg: 39	Pos: 9, 23	Pos: 34 Neg: 54
Technology*	Pos: 15, 33; Neg: 8, 22, 28, 55		

\* The technology statements were not associated with any of the perspectives (personal learning, student learning, or teaching).





# **Illustrations of SoTL Work in Mathematics**

## **Theme 5: Tackling Large Questions**



# 18

## The Question of Transfer: Investigating How Mathematics Contributes to a Liberal Education

Curtis D. Bennett and Jacqueline M. Dewar  
*Loyola Marymount University*

### Editors' Commentary

An opportunity, not a teaching or learning problem, prompted the investigation in this chapter, where all three types of SoTL questions make an appearance. The authors raise questions about evidence: what to gather and from whom, and how to analyze it. They describe unexpected difficulties they encountered. The study, completed in 2004, continues to offer lessons and open new avenues for the authors, their department, their university, and the discipline, especially related to quantitative literacy and first-year-seminar courses.

### Introduction

Appropriate transfer of knowledge, even within the same domain and across remarkably similar contexts, has been demonstrated to be very difficult to achieve (Bransford, 2000; Broussard, 2012; Dufresne, Mestre, Thaden-Koch, Gerace, & Leonard, 2005; Mestre, 2002). Yet, mathematics teachers at all levels often sell the study of mathematics as a good foundation for almost any career (Paulos, 1995), referring to the problem solving and logical thinking skills the major is supposed to develop. Recently, Alan Tucker (2011) wrote, “The value of studying mathematics is perhaps more in its mental training than its content. The wide-ranging accomplishments of math majors speak for themselves” (p. 705). As examples of mathematics majors who found success in other fields, Tucker cited a famous economist (John Maynard Keynes), a biologist (Eric Lander), a hedge fund operator (Jim Simons), and a basketball superstar (Michael Jordan). However, in our SoTL investigation, we investigated the transfer of mathematical skills from an entirely different perspective.

Our inquiry into how the proof writing and problem solving skills gained as a mathematics major transfer beyond the mathematics classroom started when we were selected as 2003–2004 scholars in the Carnegie Academy for the Scholarship of Teaching and Learning (CASTL) program. The suggested theme that year was Liberal Learning and we had applied with paired projects that were going to examine different aspects of the question:

What role does mathematics play in a liberal education?

This question belongs to the *What is?* category in the taxonomy of SoTL questions (discussed in Chapter 2), but we will see later on that our investigation touched on the other types of SoTL questions as well.



As we noted in our application, “The liberal arts are the method by which we seek to discover the truth about the most important matters of human life through reason and reflection” (Agresto, 2002). An important measure of a discipline’s contribution to a liberal education is whether the results of studying the discipline go beyond the classroom and into the person’s life. We reasoned that if we could not find evidence of this for mathematics majors, then we could not expect it to happen with mathematics for liberal arts students who typically take a single low-level course in mathematics or quantitative literacy. On the other hand, if we found evidence of transfer, then we might get some ideas about how to design a mathematics course for non-majors.

## The Subjects of Our Investigation

One author (Dewar) had a treasured piece of student work that suggested a special year-long workshop course for beginning mathematics majors at Loyola Marymount University (LMU) might be a site for this investigation. The course was introduced into the curriculum two decades ago in hopes of increasing retention in the major. It engages students in problem solving, asks them to express their mathematical reasoning both verbally and in writing, and provides information about career opportunities and the relevance of mathematics to modern society. It also attempts to foster a sense of community among a group of 12 to 20 students who, as beginning mathematics majors, might be in different mathematics courses ranging from precalculus to differential equations. Due to its success (Dewar, 2006a), the course still plays a significant role in the curriculum.

A year after taking the course, Dewar’s former student drew upon the problem solving methods he had encountered there to shape a philosophy-of-life term paper for his philosophy course. He reinterpreted mathematical problem solving strategies (e.g., look for a pattern, work backwards, solve a simpler problem, try special cases, and practice persistence) so that he could apply them to problems encountered in family life, in society, and in the service of world peace. He wrote, “My philosophy of life is very basic: Practicing mathematical problem solving can train people to think creatively in order to overcome difficulties to enable communities to live together in a peaceful and happy society.” Thus, the investigation began with a compelling example of how mathematics had contributed to one student’s liberal learning. This transfer of mathematical understanding and methods was truly “a vision of the possible,” and all by itself could have launched a *What could be?* SoTL investigation. By luck, the other author (Bennett) was scheduled to teach the workshop course the following year, so he would study its students.

Meanwhile Dewar, on sabbatical that year, would study a cross-section of undergraduate mathematics majors and one recent graduate (12 students in all), to try to document the evolution of our students’ understanding (and valuing) of proof. Each discipline has its own method for developing, creating, or discovering new knowledge or, in mathematics, new truths. For example, biology has the scientific method, sociology the sociological imagination, and in mathematics we have proof. Most students find learning how to construct a mathematical proof and communicate it to others the hardest part of the mathematics major. In our major, students first encounter proofs in the workshop course, next in a sophomore course devoted to learning techniques of proof, and later in upper division courses. Thus, having one of us focus on students at various stages of learning about mathematical proof, and how that experience might carry over to other areas, made sense to us as part of an investigation of how mathematics contributed to liberal learning. The investigation of proof yielded results unrelated to liberal learning that we reported elsewhere (Bennett & Dewar, 2007, 2013).

## The Question of Transfer: Collecting Evidence

To investigate how mathematics students’ proof and problem solving skills transferred, we needed to obtain students’ perceptions. We surveyed students in the workshop course, conducted interviews with freshmen and senior mathematics majors, and ran a student focus group with six of Dewar’s 12 subjects. To triangulate our data, we also interviewed faculty members from other disciplines, hoping to learn what they saw mathematics majors bringing to their classrooms that other students did not.

We anticipated that our interviews of colleagues from other disciplines would provide us with the best evidence, but we were wrong. Although most instructors talked about how the mathematics students differed from the other students in their classes, we encountered two problems. The first was to find a way to ensure that they were contrasting our students to comparable students from other disciplines. That is, we were concerned that the results we gathered

would be confounded by differences in students' academic backgrounds or abilities unrelated to their learning in mathematics. This was less of a concern for the interviews with instructors teaching undergraduate honors courses, and more for those with the education faculty members who were teaching courses in the credential program where some of their students had received undergraduate degrees elsewhere and, in some cases, quite a few years earlier. Because the differences reported by the education faculty members were connected to experiences in the mathematics major, our concerns about different backgrounds or abilities being confounding factors, were not realized.

Another problem we encountered was that instructors teaching undergraduates in courses such as English, history, and philosophy seemed to bring their perspectives on, and perhaps even prejudices against, mathematics into the discussion. Rather than reporting on the work or attributes of mathematics students, one instructor repeatedly stated that he saw that the student did such and such, but that must have been because the student had a French major in addition to a mathematics major or that the student was doing something because of another course. Another instructor wanted to ascribe any special attributes of reasoning or writing that he had observed to what happened in his class instead of what the students brought to the class. Only the interviews with the two education faculty members were free of such disciplinary-oriented perspectives. Both said that the mathematics students were likely to write more concise and organized papers. They also observed that the mathematics majors paid attention, for good and for ill, to definitions. The attention to definitions was good when it led to precision in their writing. The ill was that they were ill at ease with open-ended assignments, such as observing a class and writing a report. The humanities faculty members had also mentioned that the mathematics students' writing was terse and concise, but viewed this as a less positive characteristic than the education faculty members did.

The student surveys, interviews, and focus group were more straightforward to conduct and gave us qualitative data to analyze. We had expected that students would mention problem solving and thinking skills learned in the discipline being transferred, but not communication skills. We were surprised by how frequently students mentioned transferring communication and writing skills to other classes.

We surveyed the 15 freshmen in the mathematics workshop class at the end of the academic year. One question was how they saw the mathematics transfer to other classes, both mathematical and nonmathematical. Seven specifically commented on how what they learned had transferred to non-mathematics classes. Five of those mentioned writing skills as transferring. For example, one stated, "When writing history papers, I could tell my analysis was more precise." Students also mentioned thinking skills but not as often as writing.

We also conducted interviews with five freshman and five seniors at the end of the spring 2004 term. For the seniors, the first question asked how their view and use of problem solving and proof had evolved over their four years at our institution. Verbal, or qualitative, data gathered like this is rich in information, but requires careful coding, as described in Chapter 3. Rather than mine our data for a few quotes that supported the hypothesis, we searched for patterns or themes. After transcribing the interviews, we looked for topics, comments, or perceptions that showed up repeatedly. But students often used different words to describe the same thing, so on occasion we had to debate whether certain statements reflected similar ideas. Even though writing was not mentioned in our question, seniors talked about it, and how it transferred to other classes. For example, one senior said,

When I write papers, I almost write like a proof. I think of what I know at the beginning and set that up, and then I progress through it and manipulate it until I get my end results or what I want.

Another senior student stated,

I find how I write papers now is very mathematical—very scary.

Freshmen made similar statements during their interviews, but without using the phrase "write papers." For example, one commented about writing:

I tend to explain things more, look deeper into the problem. Put down all the things I know, and if I know this do I know anything that relates to it that I know.

The student is describing approaches to thinking and reasoning that involve writing (because of the phrase "put down"), and that could be applied in all sorts of nonmathematical contexts. However, given the source, a freshman in the workshop class, and knowing the course emphasizes problem solving and communicating mathematical reasoning, we believe that the student is describing using writing as part of the problem solving process.

Because of difficulties like this in interpreting the qualitative data, having multiple data sources echoing a common theme provides more support for that theme or conclusion. In our study, a focus group provided us with a different data source and third set of subjects, one including students at all levels. Participants observed that their writing styles had changed because of their mathematics classes. One junior said:

Yeah, I definitely agree that the problem solving skills and thought processes and writing skills that I learned in [the workshop course] and in all the other math classes I took really helped me to become a better writer in less technical but still technical classes like philosophy and science, and even sometimes English when I'd have to write argumentative papers it would really help.

This was a clear statement about writing. In the focus group, students were asked how the workshop course influenced their writing. We expected that students would mention how they wrote in other mathematics classes, but they discussed how they wrote in non-technical classes. One exception was a freshman student who said that she was not taking any writing classes so that she could not contribute to the discussion. Because the response came from a focus group and not from a yes or no bubble-in survey, we understood the reason for the negative response. This illustrates why a SoTL investigator might want to spend the additional time and effort required to gather and analyze both qualitative and quantitative data.

## Analysis

Did the experience of doing mathematical writing lead mathematics students to adopt, consciously or unconsciously, a more concise technical style in other writing? This question has the form of a *What works?* question, and we used our disciplinary knowledge from mathematics about proof writing in the undergraduate curriculum to answer it. In general, mathematics students were not writing papers in their mathematics classes but were describing their mathematical reasoning by writing informal mathematical arguments as freshmen and, later on, more formal proofs. The format of a mathematical proof is different from that of an essay or research paper. In undergraduate classes, students would not be writing introductions to or summaries of their proofs. While mathematical proofs are, or should be, written in complete sentences, proofs done for homework assignments seldom extend beyond a single paragraph. They often employ a shorthand notation and most proofs require careful use of definitions. These observations support the idea that what was transferring for the students were the skills of making logical connections, writing concisely, and using words very specifically, three characteristics of good mathematical proofs.

## Conclusions

So what can we conclude about mathematics contributing to liberal learning? The evidence we began with, showing that for one student mathematical problem solving became a metaphor and a method for approaching life's problems, was not reflected in evidence gathered later. Instead, many of the 30 or so students in our study indicated that the mathematical reasoning and writing skills that they developed in the mathematics major influenced the way they thought about and approached writing, mathematical or otherwise. Students at all levels of the major said this. In the interviews, education, but not humanities, faculty members agreed with the students and mentioned specific characteristics of mathematics majors' writing that could be connected to their mathematical training, namely organization, conciseness, and attention to definitions (precision). Our data supports the claim that mathematics majors are prized for their problem solving and logical thinking skills, and it suggests more than that. We found that organization, precision, and conciseness in writing skills can be benefits of the major. When combined, our data and analysis suggested that the study of mathematics in the major contributes to students' liberal education through logical thinking and communication skills that are developed through mathematical problem solving, reasoning, and writing proofs. These results were reported at the Third Annual ISSOTL Conference (Dewar, 2006b).

Our study has many limitations. It was conducted on one campus, where the mathematics curriculum has an unusual course for beginning majors that stresses mathematical problem solving and communication. Moreover, students in the course comprised about half of the subjects in the study, and every student in the study had taken or was currently

enrolled in the course. The results should be considered a “vision of the possible.” We believe this vision can be helpful for those concerned with quantitative literacy or first-year-seminar courses, as we discuss in the next section.

## Applications of these Results

Our findings may suggest appropriate goals for quantitative literacy (QL), a topic that is commanding increasing attention in the 21st century. Between 1997 and 2004, the College Board, the MAA, and the National Council for Education and the Disciplines produced four major reports on QL (Madison & Steen, 2003; Steen, 1997, 2001, 2004). In 2004, the MAA approved the formation of a special interest group on QL and the Association of American Colleges and Universities devoted an issue of *Peer Review* to QL (Tritelli, 2004).

We reasoned that, if a goal of the QL movement is to design courses for non-majors so that they will carry the mathematical methods they learn beyond the classroom, then we should know what mathematics majors themselves carry beyond the discipline, and what types of courses would contribute most to the transfer. Our study suggests that those planning QL courses should not focus on calculation or applying computational algorithms. Instead, they should emphasize problem solving, reasoning skills, and argumentation, both written and oral. For more than 15 years, our university has offered a QL course intended for every student whose major does not require mathematics. It includes topics such as number sense, percentages, mathematics of finance, elementary probability and statistics, and voting theory. Over the last few years, our department has tried several new approaches to QL, replacing voting theory with semester-long group research projects with written reports and end of semester presentations, offering special topics courses such as the Mathematics of Symmetry that involve projects or reports, and developing some shorter projects that involve writing. We have taken a scholarly approach and documented the results (Dewar, 2006b; Dewar, Larson, & Zachariah, 2011). A study of the special topics course on symmetry appears as Chapter 16 in this volume.

Our study may also have applications to first-year-seminar courses, which are becoming more common. They are typically small discussion-based courses designed to assist students in their academic and social development and in their transition to college (Hunter & Linder, 2005). Developing critical thinking and writing skills are the two most common academic first-year-seminar goals (National Survey of First-Year Seminars, 2009). We could reconfigure the workshop course into a freshman seminar for non-math majors, as our university core curriculum is being revised to include first-year-seminars. Our findings along with previous assessments of the workshop course would provide a scholarly basis for redesigning the course as a first-year-seminar experience for students for any major. Ideas for a redesign were reported in a contributed paper session at the 2011 MathFest (Dewar, Larson, & Zachariah, 2011).

## Next Steps

What would we do next if we were to continue this study? We could address the limitations of our study by revisiting the original study on a larger scale, and seek additional evidence from other colleagues, both in the humanities and in education, to see if our original results are replicated. We could improve our methodology and think about better evidence to collect. We might develop a non-mathematical writing assignment for our freshman mathematics majors, and think about how to get similar data from our seniors. We should become more familiar with the latest literature on transfer and the latest work on quantitative literacy. We might even seek to communicate directly with experts on transfer or quantitative literacy such as Jose Mestre or Lynn Arthur Steen.

After finding that characteristics of mathematical writing appear in mathematics majors’ writing in other courses, we now wonder if writing skills learned by psychology students in a research methods class or by engineering students learning to write lab reports transfer to other domains. So, instead of revisiting the original study in mathematics, we could seek collaborators in other fields to explore these ideas in parallel. This could be a case where a SoTL study transcends disciplines, as discussed in Chapter 2. Because our study suggests that writing instruction may be effective when students have a specific disciplinary goal in mind that goes beyond learning to write, in any future work we might want to explore connections with the decades-old writing-across-the-curriculum movement.

We could also devise a mathematical problem solving and writing course as a special topics course to meet the QL requirement of the university core. Doing so would provide us with an ideal laboratory to study how mathematical

training contributes to liberal learning. The question could be: Do students who are not mathematics majors develop transferable skills from engaging in mathematical reasoning and learning to express it in writing?

## Personal and Professional Impact

As a result of this project, we gained enormous insights into our students, our teaching, and the difficulties of doing the scholarship of teaching and learning. We had naively hoped that finding evidence of transfer of mathematical skills would be straightforward. However, gathering and interpreting such evidence is complicated and messy. To derive any conclusions, we had to learn about focus groups and social science methodologies for analyzing qualitative data. Even without the depth of training in qualitative data analysis that a doctoral program in mathematics education would provide, we were able to identify a common theme in the data that had been gathered by different methods from multiple sources. By gathering and analyzing our data in a systematic way, we obtained evidence that we used to back up our claims or proposals.

The investigation also gave us new insights. While we always believed that learning mathematics helps students become better problem solvers and more logical, it was a surprise that so many students saw their mathematics major as helping them learn to write. In our teaching, particularly in the QL class, we are now more likely to give writing assignments and ask students to reflect on how mathematical thinking could and should influence their writing. The project also led our department to examine how and where we teach writing throughout our curriculum. We little expected that the effects of our SoTL investigation would have such an effect on our department.

## References

- Agresto, J. (2002, January 14). Do the liberal arts serve a public function? *Liberal Arts Online* 2(1). Retrieved from [www.liberalarts.wabash.edu/lao-2-1-liberal-arts-public/](http://www.liberalarts.wabash.edu/lao-2-1-liberal-arts-public/)
- Bennett, C., & Dewar, J. (2013). Interdisciplinary encounters while using SOTL to examine the evolution of students' understanding of proof. In K. McKinney (Ed.), *SoTL in and across the disciplines*, (pp. 54–73). Bloomington, IN: Indiana University Press.
- Bennett, C., & Dewar, J. (2007). Developing and applying a taxonomy for mathematical knowledge-expertise. *Electronic Proceedings for the Tenth Special Interest Group of the Mathematical Association of America on Research in Undergraduate Mathematics Education Conference on Research in Undergraduate Mathematics Education*, San Diego CA. Retrieved from [sigmaa.maa.org/rume/crume2007/eproc.html](http://sigmaa.maa.org/rume/crume2007/eproc.html)
- Bransford, J. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Bressoud, D. (2012, September 12). Teaching and learning for transference. *Launchings*. Retrieved from [launchings.blogspot.com/2012/08/teaching-and-learning-for-transference.html](http://launchings.blogspot.com/2012/08/teaching-and-learning-for-transference.html)
- Dewar, J. (2006a). Increasing math majors' success and confidence through problem solving and writing. In F. Rosamond and L. Copes (Eds.), *Educational transformations: The influences of Stephen I. Brown*, (pp. 101–121). Bloomington, Indiana: AuthorHouse.
- Dewar, J. (2006b). *Multiple facets of quantitative literacy*. Panel Discussion at the Third Annual ISSOTL Conference, Washington, DC, November 9–12, 2006. Abstract and Synthesis available at [www.indiana.edu/~issotl/ISSoTL2006-abstractsDG.pdf](http://www.indiana.edu/~issotl/ISSoTL2006-abstractsDG.pdf)
- Dewar, J., Larson, S., & Zachariah, T. (2011). Group projects and civic engagement in a quantitative literacy course. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 21(7), 606–637.
- Dewar, J., Larson, S., & Zachariah, T. (2011, August). *Extrapolating from a first year experience for mathematics majors*. Paper presented at 2011 MathFest of the Mathematics Association of America, Louisville, KY.
- Dufresne, R., Mestre, J., Thaden-Koch, T., Gerace, W., & Leonard, W. (2005). Knowledge representation and coordination in the transfer process. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective*, (pp. 155–215). Greenwich, CT: Information Age.

- Hunter, M. S., & Linder, C. W. (2005). First-year seminars. In M. L. Upcraft, J. N. Gardner, & B. O. Barefoot (Eds.), *Challenging and supporting the first-year student: A handbook for improving the first year of college*, (pp. 275–291). San Francisco, CA: Jossey-Bass.
- Madison, B., & Steen, L. (Eds.). (2003). *Quantitative literacy: Why numeracy matters for schools and colleges*. Princeton, NJ: National Council on Education and the Disciplines.
- Mestre, J. P. (2002). Probing adults' conceptual understanding and transfer of learning via problem posing. *Journal of Applied Developmental Psychology*, 23, 9–50.
- National Survey of First-Year Seminars. (2009). Retrieved from [www.sc.edu/fye](http://www.sc.edu/fye)
- Paulos, J. A. (1995). *A mathematician reads the newspaper*. New York, NY: Basic Books.
- Steen, L. (Ed.). (1997). *Why numbers count: Quantitative literacy for tomorrow's America*. New York, NY: The College Board.
- Steen, L. (2001). *Mathematics and democracy: The case for quantitative literacy*. Princeton, NJ: National Council for Education and the Disciplines.
- Steen, L. (2004). *Achieving quantitative literacy: An urgent challenge for higher education*. Washington, DC: Mathematical Association of America.
- Tritelli, D. (Ed.). (2004). Quantitative literacy [Special issue]. *Peer Review*, 6(4).
- Tucker, A. (2011). Revitalizing the 1960 mathematics major. *Notices of the AMS*, 58(5) 704–5.



# 19

## Using SoTL Practices to Drive Curriculum Development

Rikki Wagstrom  
*Metropolitan State University*

### Editors' Commentary

In this chapter, Rikki Wagstrom describes how she applied SoTL processes to aid in the development and evaluation of a new curriculum that integrated civic issues into a prerequisite course for college algebra. Her experience illustrates how it can take a long time to identify and frame an appropriate research question. She describes searching the literature and tells how it led her to a useful model, one that prompted her to change the site of her investigation and revise her research question. She provides insights into the problems that can arise in finding faculty members to teach experimental and control sections, and the tough decisions that have to be made about how much data to collect.

### Introduction

My SoTL journey originated with the 2006 Summer Institute sponsored by SENCER (Science Education for New Civic Engagements and Responsibilities, see [sencer.net](http://sencer.net)). SENCER, a National Science Foundation funded program, supports mathematics and science faculty in creating or modifying curriculum, pedagogies, and courses to integrate civic issues. When I attended the institute, the courses already developed through SENCER were mostly in the sciences; only a few were in mathematics. I returned to Metropolitan State University in St. Paul, Minnesota, inspired to design a SENCER mathematics course and curious about how using the SENCER approach would affect introductory-level courses such as developmental mathematics and college algebra, where low retention rates are a problem.

When I returned for the 2007 SENCER Summer Institute, I participated in a pre-institute workshop on the Scholarship of Teaching and Learning. Workshop participants began to develop a SoTL research question to investigate during the coming year and to report on the following summer. I started with many questions: Would teaching college algebra through a civic issue generate interest in mathematics? Would it improve students' ability to think about and use algebra sensibly? Would it better prepare them for future mathematics courses? To launch my SoTL project, I settled on the question: How does teaching algebra through civic issues affect students' abilities to apply appropriate mathematical arguments or tools to mathematics-related problems arising in their lives?

### Building on the Work of Others

One of the first workshop assignments was to search the literature for research relevant to our questions and to create an annotated bibliography. The workshop leaders gave us a list of databases featuring educational research and titles of particular journals that publish SoTL research. ERIC (Educational Resources Information Center, found



at [www.eric.ed.gov](http://www.eric.ed.gov)) was highly recommended as it contains information pertaining to mathematics education. I also looked at the online SoTL journal *Mountain Rise* (see [mountainrise.wcu.edu/index.php/MtnRise](http://mountainrise.wcu.edu/index.php/MtnRise)) and SoTL case studies and publications available through the Carnegie Foundation.

When exploring new bodies of research, doing literature searches, particularly in databases, can be challenging. I find it helpful to begin with an All Text rather than a Keyword or a Descriptor search. This generally locates records at least tangentially related to my topic of interest. From the relevant records, I can obtain a list of descriptors, author names, and journal titles to use in further queries. As an example, if I want to locate for articles on integrating civic issues into algebra courses in the ERIC database, a search on “civic issues” AND “algebra” using the KW Identifier or DE Descriptor fields produces no results. However, using the TX All Text field yields three records, one of which is the *PRIMUS* article, “Group Projects and Civic Engagement in a Quantitative Literacy Course” (Dewar, Larson, & Zachariah, 2011). This record contains many descriptors, including numeracy, algebra, citizenship education, comparative analysis, and social responsibility. Thus, from one search I can acquire many leads.

My search located a report to the U.S. Department of Education, Office of Postsecondary Education (Schaufele & Zumoff, 1996) on a FIPSE-funded project at Kennesaw State University titled “Earth Algebra.” The principal investigators had developed a curriculum to teach college algebra through the context of global warming, and investigated the impact on students’ views of mathematics, and on their algebraic, data analysis, and mathematical modeling competencies. They reported improvements in students’ views of mathematics and in their performance in data analysis and mathematical modeling, but they found no significant difference between the Earth Algebra students’ algebra skills and those of traditional college algebra students.

I also searched for information about critical thinking and quantitative reasoning to help me define what I would mean by “applying algebra sensibly” or applying “appropriate mathematical arguments or tools.” I found a collection of methods, developed by fellows of the National Institute for Science Education (1999), for assessing student learning in STEM courses at the Field-tested Learning Assessment Guide (FLAG) website ([www.flaguide.org](http://www.flaguide.org)). Their Mathematical Thinking Classroom Assessment Techniques (Ridgway, Swan, & Burkhardt, 2001) provided examples of questions to assess different types of critical thinking and quantitative reasoning skills. Although I did not use any of these questions in my assessment surveys, they were good models.

## Redirecting My Project and Rethinking My Research Question

Although my plan had been to study the impact on student learning of incorporating civic issues into a college algebra course, after my literature search I shifted the site of my study to Math 101 Exploring Functions, the prerequisite for our college algebra course. In addition to the Earth Algebra project, I found other examples of curricula integrating civic issues in college algebra and precalculus, but none for college algebra prerequisite courses. All of these courses typically have low student retention rates. That raised the question whether student learning gains similar to those cited in the Earth Algebra study might be seen below the level of college algebra. Using Earth Algebra as a model, my project became developing a curriculum for our Math 101 course integrating sustainability issues and evaluating its impact on student learning.

I then reconsidered my research question: How does teaching algebra through civic issues affect students’ abilities to apply appropriate mathematical arguments or tools to mathematics-related problems arising in their lives? The phrase “teaching algebra through civic issues” implied that learning algebra would be a by-product of studying a civic issue. I decided that the phrase “integration of civic issues” would better express my objectives. Moreover, the Earth Algebra study suggested that my assessment goals were too narrow. I had no idea what Math 101 students would take away from a civic engagement experience. It seemed to me that this project needed to explore more than students’ ability to apply mathematics. Consequently, my research question evolved into a *What could be?* question: How might integrating civic issues into a college algebra prerequisite course improve quantitative reasoning, algebraic skill development, and student confidence and interest in studying mathematics?

## Initial Planning

My timeline for the SoTL project was: plan and design assessment instruments during fall 2007; develop, pilot test, and revise new curriculum during spring and summer 2008; assess curriculum from fall 2008 through summer 2009; and compile data and prepare for dissemination during fall 2009. In the fall 2007 semester, I used the Earth Algebra study

	Treatment group	Control group
Pre-test	Yes	No
Post-test	Yes	Selected problems
Pre-course Math-SALG	Yes	No
Post-course Math-SALG	Yes	No

**Table 19.1.** Assessments administered in Math 101

as a guide for how to go about collecting evidence. It had been a controlled experiment involving seventeen sections of college algebra offered concurrently in one quarter. Students in the treatment group (sections of college algebra using the Earth Algebra curriculum) and in the control group (sections using the traditional curriculum) completed a battery of pre- and post-assessments, the results of which were analyzed statistically.

Despite my desire to conduct a similar study in Math 101, I was concerned about potential unintended effects of pre-course testing on the students and student-teacher relationships. Math 101 students frequently are extremely anxious about their mathematical ability. When I teach this course I go out of my way to ease their anxiety on the first day of class and set up a positive classroom atmosphere. Administering a battery of tests on the first day would hardly contribute to the positive initial experience I desired, and I feared the consequences might carry through the entire semester. I also worried about how adjunct faculty, who teach most of our Math 101 sections, would view such an assessment plan. Many of them shared my concerns about the unintended consequences of testing students on the first day, and some might interpret the study as an evaluation of their teaching and feel pressured to alter their courses in some way. Consequently, I structured my assessments so as to minimize the invasiveness of the study, even though that meant deviating from the Earth Algebra study design.

According to my revised research question, my learning objectives included improved quantitative reasoning skills, algebraic skill development, and student confidence and interest in studying mathematics. To establish a baseline for algebraic skills, I started with my Math 101 final examination from the previous semester when I taught the course using the traditional curriculum. I extracted a collection of typical algebra problems to create a post-test for sections of Math 101 using the traditional curriculum (control group) and sections using the new curriculum (treatment group). I asked a colleague familiar with Math 101 to review the problems. Initially, I had planned to work with the other Math 101 instructors to create a common set of problems for the post-test. However, Math 101 is taught primarily by part-time instructors hired on a semester-by-semester basis. Consequently, I could not know for certain who would be teaching the course a year later.

I also wanted to examine whether integrating civic issues into the curriculum could affect students' quantitative literacy. Toward that end, I identified related quantitative literacy skills: finding and interpreting rates of change and percentage changes in diverse contexts, performing unit conversions, making plausible estimates, and developing appropriate mathematical models (functions) given underlying assumptions. I developed problems to assess these skills and added them to the post-test. Ultimately, the post-test problems would be part of the final examinations given in the treatment and control groups. I selected a subset of six post-test problems for a pre-test. Because the pre-test included problems emphasizing topics that would not be covered in the sections of Math 101 using the traditional curriculum, only students in the treatment group would take it. I considered writing a shorter version of the pre-test for the control group, but I decided to collect only the most vital data from the control group, namely the results from the post-test. The final examination problems that were used in the pre- and post-comparisons appear in the appendix.

To evaluate students' perceptions of their ability to apply mathematics and their interest in mathematics, I used SENCER's Student Assessment of Learning Gains for mathematics, commonly referred to as the Math-SALG (SENCER, 2012). These pre- and post-online surveys measure students' mathematical confidence in applying mathematics to topics of civic interest and assess their interest in further study of mathematics. Although I originally intended to administer the Math-SALG to all Math 101 students, at the end of summer 2008 I decided not to administer it to the control group. It was a decision I came to regret, but at the time I felt that it was necessary for me to scale back my assessment efforts. I will say more about this later. A summary of the assessments administered in Math 101 sections is shown in Table 19.1.

I thought that this assessment structure would enable me to investigate student learning in the treatment group and provide data for comparing algebraic skills with the control group.

## Piloting the Curriculum and Assessment Instruments

During spring 2008 I developed what would be the first of several iterations of the new curriculum. I had already written the equivalent of a workbook for our traditional sections of Math 101, which I had used for my lectures. The new curriculum would be an adaptation of this workbook organized into two parts. The first would contain core topics such as unit conversion, plausible estimation, rates of change and percent change, linear and exponential behavior, and an introduction to functions. The second part of the workbook would apply these topics to linear, exponential, and logarithmic functions. The chapters introducing algebraic content, like functions, properties of the real numbers, simplifying algebraic expressions, and equation solving, were largely unaltered. Beginning in the fall 2008 semester, the curriculum would include two sustainability projects exploring ecological footprints using unit conversions, plausible estimations, and linear, exponential, and rational functions.

During the spring 2008 semester, I used the new curriculum and the assessment instruments in a section of Math 101. The assessment results provided insights that guided me in revising areas of the curriculum, particularly in the development of quantitative literacy and reasoning skills. Because I had never done a study like this, it was helpful to try out the assessment instruments before undertaking the formal study so as to work out unforeseen problems.

I made two additional revisions to the curriculum during summer 2008, modifying the assessment instruments each time so as to gauge their effectiveness and to track changes in student learning over earlier versions of the curriculum. Since the same (or similar) test questions were used previously, there was some risk of results being skewed by sharing of test information between classes. Because Metropolitan State University is a non-residential university, and most students are working adults who are on campus only when attending a class, the students tend to be isolated. Therefore I was willing to accept what I believed to be a small risk. Additionally, I did not return any of the pre- or post-tests and the use of scratch paper was monitored.

During the spring and summer trials I worked through many challenges that would otherwise have hurt my study. In spring 2008, there were problems collecting data with SENCER's online pre- and post-course Math-SALG. Setting up user accounts was difficult for some students in the course. Several students created more than one user account because they thought their first attempt had not worked, resulting in duplicate pre-course data. After the spring 2008 term, I wrote a hard-copy version of the Math-SALG for administration in class. This was not perfect either because students needed to use an assigned ID number on both the pre- and post-course Math-SALG surveys. Every semester, there would be students who could not locate their ID number for the post-course Math-SALG. This resulted in lost data, but not to the same extent as with the online version.

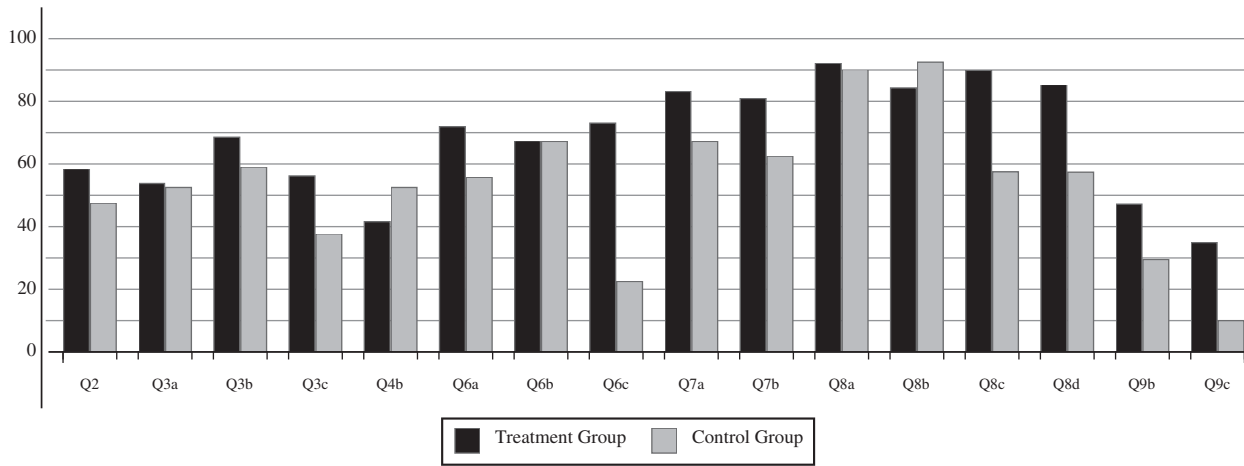
Another challenge I faced during the pilot period was developing a curriculum that met my objectives. During the spring trial, I could see the shortcomings in the curriculum from viewing my students' work. As an illustration, the chapter on unit conversion and plausible estimation did not provide sufficient scaffolding to enable students to develop competency. It took the entire trial period to expand and refine the curriculum to my satisfaction.

A final difficulty was my inexperience in developing assessment tools for particular skills. I learned that student mistakes on multi-part questions can be difficult to interpret. A better approach is to break up a multi-part question into separate problems each with its own context. Also, pre-test questions on material that will be taught in the course can lead students to believe they were registered for a course for which they were underprepared.

## Conducting the Study

During summer 2008, I submitted paperwork to our Human Subjects Review Board (HSRB) in anticipation of going public with data obtained from sections of Math 101 offered the following year. I included the assessment instruments, the cover letter that would be given to students explaining the study, and the voluntary consent form. I needed to describe how student confidentiality would be ensured. I encountered no difficulties in gaining approval.

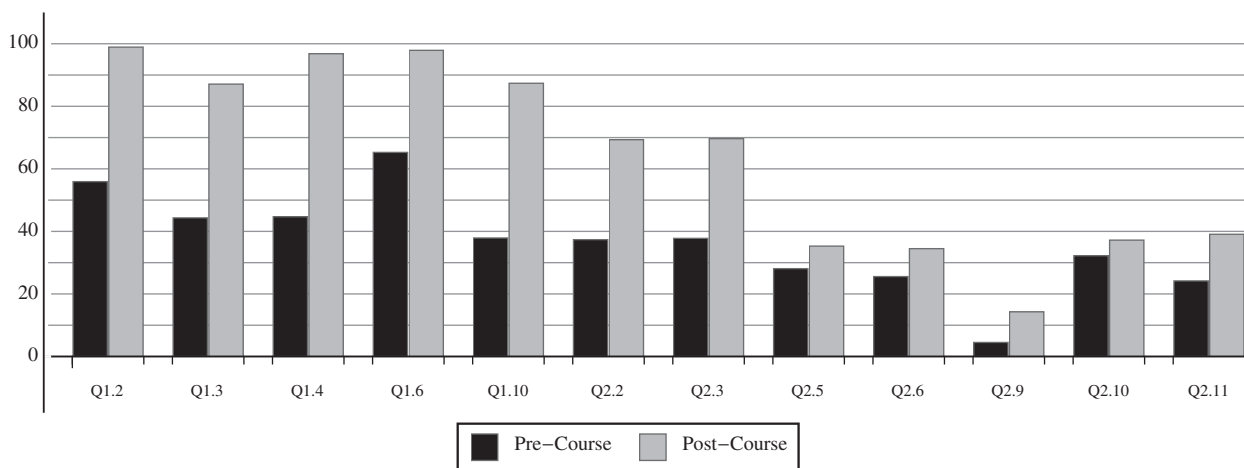
During summer 2008, I also contacted instructors who were scheduled to teach Math 101 in fall 2008 to discuss the study and ask for their participation. The teaching schedule for the spring 2009 semester was not yet determined, so spring 2009 instructors were contacted later. I only contacted instructors who had taught Math 101 several times, or had taught mathematics at an equivalent level. I gave them the post-test problems, and asked them to use all the problems that they would typically include in their final examinations. I believed that this was a fair concession, because they had no part in creating the post-test problems, and I wanted to respect their autonomy.



**Figure 19.1.** Post-test scores of treatment and control groups

Between fall 2008 and summer 2009, 167 students registered in nine sections of Math 101 voluntarily consented to participate in the assessment study and completed the assessments. Six sections (106 students) were in the treatment group and three (61 students) were in the control group. Of the six sections in the treatment group, I taught five and a colleague taught the sixth. The three sections in the control group were taught by two part-time instructors and one full-time faculty member. Two other sections of Math 101 used the new curriculum, but were taught by instructors with less experience. I decided not to administer assessments in these sections because the curriculum incorporated new content that the instructors would need to learn, and it had a different approach to teaching algebra. I did not want to stress them by assessing their students, nor did I want to add faculty experience and expertise to the list of confounding variables in the study.

During the fall 2009 semester, I compiled and summarized the data that I had assembled. The number of student responses for each post-test problem depended on how many instructors incorporated the problem into their final exams. Because instructors were free to select which questions they wanted to include, the number of student responses for each post-test problem varied by question. The sample sizes of responses by students in the treatment group ranged from 89 to 106, while the sample sizes for the control group ranged from 40 to 61. The post-test results for the treatment and control groups and the pre- and post-Math SALG scores for the treatment group are shown in Figures 19.1 and 19.2, respectively. The vertical axis in Figure 19.1 shows the percentage of correct responses to post-test problems and, in Figure 19.2, displays the percentage of responses in the range Somewhat Confident to Extremely Confident or



**Figure 19.2.** Math-SALG Pre- and Post-comparison

Somewhat Interested to Extremely Interested on the Math-SALG. Question numbers are shown below the horizontal axes.

## Findings and Dissemination

Six post-test problems on algebra and quantitative literacy skills were used as a pre-test for the students in the new curriculum. Their average scores on these problems increased 53 percentage points, indicating that the new curriculum was successful in developing algebra skills and improving quantitative reasoning.

Trends in the assessment data in Figure 19.1 suggested that the MATH 101 curriculum integrating civic and environmental issues was at least as effective as a traditional mathematics curriculum at building students' mathematical skills. The Math-SALG data from the course taught with the new curriculum (Figure 19.2) showed that students' confidence in using mathematics and their interest in learning and applying mathematics increased. During fall 2009, I wrote a paper describing the curriculum and reporting the assessment findings. Because it dealt with both mathematics curriculum and civic engagement, finding a journal for publication could have been difficult. Fortunately, through my involvement with SENCER, I was familiar with the peer-reviewed journal *Science Education and Civic Engagement: An International Journal* (SECEIJ). I submitted the paper to SECEIJ and it was accepted and published (Wagstrom, 2010).

## Insights: Looking Back

In contrast to the Earth Algebra study, my SoTL project was neither grant-funded nor collaborative. At the beginning, I did not fully appreciate the consequences that these factors would have on the project. My two-year timeline included one year for curriculum development and design of the assessment instruments and one year for data collection and analysis. The first year was intense. I vastly underestimated how much time and mental energy would be needed to redesign the curriculum repeatedly, design and revise the assessment instruments, and compile and analyze data. External funding would have provided me time for more research on the development of critical thinking and quantitative reasoning skills and time to actually write the curriculum.

I should also have collaborated with at least one colleague on this project. I had to play the roles of teacher, curriculum developer, and evaluator, which was a lot for a study of this nature and size. The demands involved in moving the project forward were so great that I considered abandoning it. Instead, I decided not to administer the Math-SALG in the control group because I needed to let something go, and I thought that the Math-SALG data were less important to the study than the data from the post-test. As I stated earlier, this was a regrettable decision. If I had worked with a collaborator on this project, the additional data probably would have been collected and there could have been additional comparisons between the treatment and control groups.

One of my concerns at the start of this project was the effect of pre-course assessments on class dynamics. Looking back, I think a researcher can successfully mitigate the impacts of pre-course assessment without having to compromise the amount of data being collected by explaining to students in clear terms what it is that the researcher is trying to learn, how the assessment will benefit future students, and how current students can assist in the effort. That the study would not affect students' grades or their relationship with their instructor, and that the intention is to evaluate the curriculum and not the students are other important messages to convey. Similar explanations can also be helpful in gaining the cooperation of the other instructors.

When I began using the new curriculum in my Math 101 sections, I knew there was a distinct chance that it would not be received well. Many of the students either feared or openly disliked mathematics. Initially, they had no textbook, they received only packets of worksheets. The mathematics in the worksheets was often different in approach from what they were accustomed to seeing. They were required to learn sustainability and quantitative literacy content not listed in the course description. The students were aware that more was being expected of them than of students in traditional sections of the course. They knew they were test cases. And yet almost no students dropped at the beginning of the semesters. It is possible that the assessment study may have played a role. Student comments expressed appreciation for the care I took to find out if the curriculum was "doing a good job" and for the opportunity to be a part of the endeavor by being given a voice. In fact, these students nominated me for our university-wide award for teaching excellence.

## Project Benefits: Looking Forward

The student-learning gains observed in my SoTL project enabled me to obtain a Center for Teaching and Learning (CTL) STEM grant from the Minnesota State Colleges and Universities system to expand the curriculum and develop a new course, Math 102 Mathematics of Sustainability. After the summer 2009 semester, Math 101 was replaced with two college algebra prerequisite options for students, one of which was Mathematics of Sustainability. During spring 2009, I collaborated with the two part-time instructors and the faculty colleague who used the new Math 101 curriculum during the previous year to plan the new course. The CTL STEM grant and later sabbatical support afforded me time to revise and expand the integration of sustainability content into the course.

In many respects, the Math 101 study was incomplete. The curriculum I wrote for the course offered students opportunities, in homework exercises and projects, to develop a range of quantitative reasoning skills, including plausible estimation, developing mathematical models (functions) given underlying assumptions, recognizing an inappropriate use of mathematics, interpreting graphs, and making arguments using mathematics. Yet only the first two skills were assessed in the post-test; the other three were never evaluated. The curriculum became more extensive than the assessment instruments (which were written at the start of the project) and I did not revise them accordingly.

The new Mathematics of Sustainability curriculum further expands opportunities for students to develop quantitative reasoning skills and I am interested in doing a more extensive evaluation of its effectiveness. Unfortunately, performing a controlled experiment similar to the Math 101 study to evaluate the effects of Mathematics of Sustainability is not possible. The students enrolling in that course are required to have a higher mathematics placement score than students enrolling in our other college algebra prerequisite course, so the two groups of students represent different populations. However, since the publication of the Math 101 study, I have been contacted by numerous educators expressing interest in using some, or all, of the curriculum in their courses. In the future it might be possible to develop an assessment instrument to supplement the Math-SALG that they could administer to evaluate student learning outcomes. Obtaining information from courses taught at different institutions with differing student populations could give a more complete picture of the effects of civic engagement on student learning.

My SoTL project work has given me a focus for my teaching and my scholarly interests, and has benefitted my career in very tangible ways. For successfully integrating civic issues into Math 101, I was nominated for, and accepted into, the SENCER Leadership Fellows program in 2008. Also, because I work at an institution that values student-centered teaching, civic engagement, and scholarship, the SoTL project factored significantly in my successful application for tenure and promotion.

Like any research, doing SoTL has challenges and rewards. In this chapter I have tried to describe the challenges and lessons learned from my first experience in this field in the hope that other mathematicians new to the Scholarship of Teaching and Learning will find them useful.

## References

- Dewar, J., Larson, S., & Zachariah, T. (2011). Group projects and civic engagement in a quantitative literacy course. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 21(7), 606–637.
- National Institute for Science Education, University of Wisconsin-Madison. (1999). *College Level One*. Retrieved from [archive.wceruw.org/nise/About\\_NISE/NISE\\_Brochure/College\\_Level\\_One1.html](http://archive.wceruw.org/nise/About_NISE/NISE_Brochure/College_Level_One1.html)
- Ridgway, J., Swan, M., & Burkhardt, H. (2001). Assessing mathematical thinking via FLAG. In D. Holton & M. Niss (Eds.), *Teaching and learning mathematics at university level: An ICMI study* (pp. 423–430). Dordrecht: Kluwer Academic Publishers.
- Schaufele, C., & Zumoff, N. (1996). *Executive summary: Earth Algebra P116B10601*. Retrieved from [catalogue.nla.gov.au/Record/5616129](http://catalogue.nla.gov.au/Record/5616129)
- Science Education for New Civic Engagements and Responsibilities. (2012). *The SENCER-SALG*. Retrieved from [www.sencernet.org/Assessment/sencersalg.cfm](http://www.sencernet.org/Assessment/sencersalg.cfm)

Wagstrom, R. (2010). Teaching pre-college algebra mathematics through environmental sustainability: Curriculum development and assessment. *Science Education and Civic Engagement: An International Journal*, 2(2). Retrieved from [www.secej.net](http://www.secej.net)

## Appendix

### Math 101 Post-Test Questions Used in the Study

Problems shown below were embedded in the final examination given to sections of Math 101 using the new curriculum. Problems marked with an asterisk (\*) were also on final examinations of sections of Math 101 using the traditional curriculum. Problems marked with a pound sign (#) were on the pre-test for the sections of Math 101 using the new curriculum.

2. \*The table below is an input/output table for a linear function whose slope equals  $-2/5$ . Complete the table.

Input	0	1	6	7	9
Output		6			

3. \*Determine the equations for each of the following functions.
- An exponential function whose graph passes through the points (0,10) and (1,7).
  - A linear function whose graph is parallel to the graph of  $y = 3x - 9$  and passes through the point (2,1).
  - A linear function whose graph has a  $y$ -intercept equal to  $-4$  and an  $x$ -intercept equal to 6.
4. (a) For the data set given below, determine whether it describes a linear or an exponential function. Work must be shown in order to receive credit.

$x$	1	4	7	10
$y$	17	13	9	5

- (b) \*Determine the equation for the function in (a).
6. \*#Solve the following equations and/or inequalities.
- $2x - (-1 - 4x) = 4x + 2$
  - $-4x + 8 > 10$
  - $\frac{y-4}{2} = y + 1$
7. \*The graph of a function  $w(x)$  is shown below (Figure 19.3). Use it to answer the following questions.

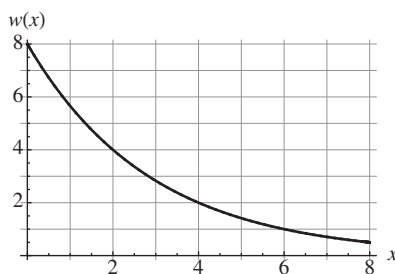


Figure 19.3. Graph of a function  $w(x)$

- Determine the approximate value of  $x$  for which  $w(x) = 3$ .
  - Determine  $w(6)$ .
8. \*Consider the functions  $f(t) = 3t - 1$  and  $h(t) = t^2 + 4$ .
- Determine  $f(5)$ .
  - Determine  $f(x - 3)$  and simplify.
  - Determine  $h(f(1))$ .
  - Determine the value of  $t$  for which  $f(t) = 0$ .

9. In 1970, the world's population was approximately 3.7 billion people. In 2005, the world's population reached approximately 6.5 billion people.
- (a) #By what percentage did the world's population grow during this time period?
  - (b) \*#What was the average rate of change of the world's population during this time period? Remember to give the units!
  - (c) \*Interpret the rate of change in part (b) in practical terms using everyday language. Please use complete sentences.
10. #The estimated cost to the city of San Francisco (to its tax-paying residents) for each plastic bag customers receive at a grocery store is about 17 cents. This includes costs associated with contamination of the recycling stream, collection and disposal costs, and landfill processing. (Source: [www.1BagataTime.com](http://www.1BagataTime.com)) Let's assume the cost per bag to the tax-paying residents of Saint Paul is also approximately 17 cents. The population of Saint Paul is approximately 300,000 people. Suppose the average number of plastic bags a person collects per week is 2 bags. Estimate the total annual cost to the entire city of Saint Paul. You must show work to receive credit.







# **Epilogue**

## **Value and Benefits of SoTL**



# 20

## Synthesis of the Value and Benefits of SoTL Experienced by the Contributors

Curtis D. Bennett and Jacqueline M. Dewar  
*Loyola Marymount University*

### Editors' Commentary

The primary goal for this volume is to provide guidance for mathematics faculty members interested in undertaking a scholarly study of their teaching practice, but a secondary goal is to promote a greater understanding of this work and its value to the mathematics community. In this chapter we reflect on the value of SoTL, generally, and take stock of the outcomes and benefits that accrued to the 25 contributing authors as a result of their scholarly inquiries into teaching and learning.

### Introduction

In 1999, Lee Shulman, then President of the Carnegie Foundation for the Advancement of Teaching, wrote entertainingly and perceptively about what it looks like when learning does not go well (Shulman, 1999). He coined a “taxonomy of pedago-pathology” consisting of *amnesia*, *fantasia*, and *inertia*. According to Shulman, amnesia refers to students forgetting, as a matter of course, what they learned. He joked that sometimes they even forget that they attended some classes. Fantasia denotes persistent misconceptions, where students are unaware that they misunderstand. Finally, inertia signifies that students are unable to use what they learned.

We suggest that teaching possesses similar pathologies. Amnesia is a good label for the many things about our teaching we forget from one semester to the next, things that went well and things that didn't, even when we are teaching the same course. In fact, Shulman called this “pedagogical amnesia” (as cited in Hutchings, 1998, p.17). Without observing and collecting data, we have little evidence or direction for improving and are unlikely to learn from our mistakes. The remainder of the taxonomy also transfers. Fantasia refers to our misconceptions about what students bring to class, think, learn, find difficult, or don't understand. Inertia signifies that we continue to teach as we have in the past independent of whether or not students are learning all that we want them to learn.

The chapters in Part II showed how the scholarship of teaching and learning helps instructors escape these pathologies. Our authors attested to the power of SoTL in addressing them and noted the many professional benefits that can arise from undertaking SoTL projects.

### Amnesia

Undertaking the scholarship of teaching and learning requires us to collect evidence of our students' learning and thinking during class. Doing SoTL pushes us to examine that evidence after the course has ended. This evidence along

with our later analysis of it provides us with an antidote to pedagogical amnesia. Whether we undertake to share our results with others, as many of the authors did, or find ourselves making our evidence, as Blake Mellor described in Chapter 16 (p. 166), “more a part of the process by which I revise and continue to develop” a course, just beginning a SoTL project can help fight pedagogical amnesia.

Throughout Part II, our authors commented on how they found compiling evidence important to improving their classes. The authors of Chapter 7 mentioned that they were able to analyze teaching practices at the classroom level and make small evidence-based changes. In Chapter 8 (p. 83), Edwin Herman noted that although he noticed a change in student participation, his investigation was unable to address these ideas more directly because of a lack of evidence, leading him to advise readers “to gather more data than you think you will need or use,” as it may be useful later. Michael Burke (Chapter 11) and Derek Bruff (Chapter 13) also echoed this theme when they noted that collecting evidence is an important tool for classroom improvement.

Classroom amnesia also results in an inability to share one’s results with colleagues in a meaningful way. As seen in Chapters 6, 8, 9, 11, 14, and 18, SoTL practitioners have been able to promote change by sharing their evidence with others. For example, the Duke University team’s SoTL project (Chapter 6) gave rise to “additional adjustments and improvements to Duke’s curriculum” (p. 65). Likewise, Curtis Bennett and Jacqueline Dewar’s investigation (Chapter 18) contributed to their department’s curriculum review. Communicating SoTL evidence and outcomes can have an even wider impact. Mike Axtell and William Turner (Chapter 14) stated, “Perhaps the largest impact of the study was in the conversations that occurred across campus” (p. 141).

## Fantasia

While collecting and revisiting evidence helps to prevent amnesia, analyzing it helps us learn what we don’t know. Many SoTL projects begin with *What works?* questions, projects that try to analyze whether a pedagogical strategy promotes student learning. Over and over again, however, our authors talk about shifting to *What is?* questions because, as they start to look at their data, they discover the unexpected.

Through such discoveries, SoTL has the potential to change us as teachers. As Michael Burke (Chapter 11) stated, he ended “with a different view of my discipline, a different view of how my discipline fits into the college and the world, and a different view of what and how we should be teaching” (p. 107). While for most authors the changes were not so grand, their analysis of evidence often surprised them and gave them a new skepticism about what they thought they knew. As Derek Bruff (Chapter 13, p. 135) stated, “the project has led me to question more regularly the assumptions I have about teaching and learning,” because he had expected, “the opposite result.” Curtis Bennett and Jacqueline Dewar (Chapter 18) were surprised “that so many students saw their mathematics major as helping them learn to write” (p. 188).

But the analysis of evidence not only allows us to make surprising discoveries, it also makes us more aware as teachers. More aware of what the students bring to the class, more aware of where they encounter difficulties, and more aware of how they think. This new awareness shows up in Chapters 7, 8, 11, 12, 13, 15, 17, and 18. From Edwin Herman (Chapter 8) we heard that he became “more aware of what causes students to participate in class” (p. 83). This awakening to how the students function and learn socially shows up in John Holcomb’s statistics project and Stephen Szydluk’s work with liberal arts students (Chapters 12 and 15).

Michael Burke (Chapter 11) experienced a larger change in his awareness. He learned that his students’ difficulties were at a more basic level than he had thought, and by reading their reflections, he saw his class in a new light. This knowledge led him on his personal “Odyssey” as he called it. For Burke, engaging in the scholarship of teaching and learning was a transformative experience, leading to many changes in his classroom, his teaching, and in his students.

Before examining how SoTL combats inertia, we discuss the professional rewards of doing SoTL.

## Professional Benefits

In Part II, our authors related how practicing the scholarship of teaching and learning led to professional rewards. For some it counted as merit or in their tenure documents, for others it led to recognition, grants, or moving forward in their profession. They noted how the work resulted in professional connections and fostered new interests, opening doors for them and further professionalizing the work of college teaching.

Almost all of the projects described in Part II resulted in conference presentations or publications (presentations in Chapters 5 and 9 through 19 and publications in Chapters 5, 8, 11, 14, 15, and 19). Moreover, Chapters 5, 7, 9, 10, and 19 mention how the SoTL work undertaken at their institutions counted towards merit, tenure, or promotion. In all of those, except Chapter 7, SoTL counted for more than evidence of teaching. While the decision of what counts in the merit process is institutionally specific, by producing a tangible product with evidence that can be examined by peers the scholarship of teaching and learning allows for greater consideration of the intellectual work of teaching and improving student learning.

For several authors SoTL work led to professional recognition beyond their departments or institutions. Rikki Wagstrom's project (Chapter 19) played a role in her acceptance into the SENCER Leadership Fellows program. John Holcomb's project (Chapter 12) was a component of his winning the Waller Education Award given by the American Statistical Association.

Faculty development initiatives on teaching predate the scholarship of teaching and learning by as many as four decades (Sorcenelli, Austin, Eddy, & Beach, 2005). Hutchings, Huber, and Ciccone (2011) took note of their differing but converging histories and how they are coming to find common ground. Our authors attest to the value SoTL as a means of professional development in three ways: enabling them to develop new skills or interests; providing them with new professional connections, collaborators, or networks; and empowering them to take a leadership role in the professional development of others.

In Chapter 5, Gretchen Rimmasch and Jim Brandt told us that they "have become more interested in mathematics education" (p. 57). Many authors implied they learned new skills in collecting and analyzing qualitative evidence. Curtis Bennett and Jacqueline Dewar made this explicit in Chapter 18: "to derive any conclusions, we had to learn about focus groups and social science methodologies for analyzing qualitative data" (p. 188).

All of our twenty-five authors have made new professional connections as a result of doing SoTL. Many of the authors told of new collegial or professional connections on their own campuses and elsewhere. Some connections involved outreach on their part. For example, Cindy Kaus (Chapter 10) co-organized a session, "Teaching Mathematics and Statistics through Current Civic Issues," at the 2009 Mathematical Association of America's MathFest and hosted the 2009 SENCER Midwest Symposium on "Teaching Quantitative Reasoning through Civic Issues" at her institution.

New professional connections can prompt SoTL practitioners to take leadership roles at their institutions or beyond. In Chapter 13 Derek Bruff offered a testimonial to the value of SoTL as preparation for working in faculty development. He wrote, "My SoTL project has paid dividends in my work at the Vanderbilt University Center for Teaching. I consult regularly with faculty from various disciplines about teaching matters and many times I have shared aspects of this SoTL project with colleagues interested in pre-class reading assignments or the relationship between computational and conceptual learning goals" (p. 134). Pam Crawford's work (Chapter 9, p. 93) led to colleagues requesting more information on "how they could use guided discovery," while John Holcomb (Chapter 12) has been appointed to statistics education leadership groups.

## Inertia

Perhaps the most significant benefit of SoTL for faculty members in general, and our authors in particular, is its ability to fight classroom inertia. Faculty members find balancing the demands of teaching, research, and service increasingly challenging in the 21st century. Teaching as we always have in the past, that is, succumbing to inertia in our teaching, is very tempting. Our authors used words like "growth," "fresh," "courage," "change," "seeking," "fascinating," "challenge," and more. Such words show the invigoration and transformation that SoTL scholars experience and often engender in others.

SoTL work invites us to change and improve our teaching. Rather than fretting about our students, we seek to understand them and to change our teaching in response to them. Michael Burke in Chapter 11 encapsulated this when he stated that as a result of his SoTL project his "goals as a teacher have changed radically" (p. 114). But not all change is radical, nor should it be. In Chapter 10, Cindy Kaus explained that her project resulted in more faculty members "seeking out ways to connect the mathematics they are teaching to civic issues" (p. 105). Sometimes small changes arise from projects. John Holcomb (Chapter 12) mentioned seeking departmental funding for better statistical software as a result of his work, while Derek Bruff (Chapter 13, p. 135) spoke of taking "more care."

Beyond fighting the inertia of how we teach, SoTL encourages us to continually think about improvements and be more reflective in our practice. Stephen Szydlak (Chapter 15) wrote that his inquiry caused him “to examine [his] pedagogical practices” (p. 153) and “heightened [his] awareness of the classroom environment” (p. 153), while Gretchen Rimmasch and Jim Brandt (Chapter 5) mentioned how their SoTL work allowed them to “continue to grow” (p. 57). Blake Mellor (Chapter 16) related that he had become a “more thoughtful teacher” (p. 166). Best of all, as Edwin Herman (Chapter 8) stated, SoTL “encourages the researcher to experiment within the classroom” (p. 83) and made him “more confident and a better teacher” (p. 84).

Beyond breaking the inertia of how we teach, SoTL fights the ennui that faculty sometimes face in teaching. Lynn Geiger, John Nardo, Karen Schmeichel, and Leah Zinner (Chapter 7) mentioned that SoTL allows “a faculty member to cast a fresh eye on classes that may have become routine” (p. 73). Furthermore, SoTL has the ability to inspire as it raises “fascinating, compelling, and important” (p. 115) issues as noted by Michael Burke in Chapter 11.

In closing, it is our hope that you, the reader, have found a greater understanding and appreciation for the scholarship of teaching and learning, whether that translates into your placing a greater value on the SoTL work of others, experimenting with a small project, or engaging whole-heartedly in the scholarship of teaching and learning.

## References

- Hutchings, P. (1998). Defining features and significant functions of the course portfolio. In P. Hutchings (Ed.), *The course portfolio: How faculty can examine their teaching to advance practice and improve student learning* (pp. 13–18). Washington, DC: AAHE.
- Hutchings, P., Huber, M., & Ciccone, A. (2011). *Scholarship of teaching and learning reconsidered*. San Francisco, CA: Jossey-Bass.
- Shulman, L. (1999). Taking learning seriously. *Change*, 31(4), 10–17. Retrieved from [www.carnegiefoundation.org/eLibrary/taking-learning-seriously](http://www.carnegiefoundation.org/eLibrary/taking-learning-seriously)
- Sorcenelli, M. D., Austin, A. E., Eddy, P. L., & Beach, A. L. (2005). *Creating the future of faculty development: Learning from the past, understanding the present*. San Francisco, CA: Jossey-Bass.

# Index

- a priori categories, *see* coding data
- AACU (Association of American Colleges and Universities), 114, 187
- active learning, 23, 128, 146
- American Mathematical Society, *see* AMS
- amnesia, xiii, 203, 204
- AMS (American Mathematical Society), 5, 16, 47, 93, 114
- Analysis of Variance, *see* ANOVA
- ANOVA (Analysis of Variance), 82
- assessment
- alignment of, 20, 22, 25, 33, 56, 132
  - authentic, 118, 119
  - of program, 73, 146, 152
  - SoTL and, 9
- Assessment Resource Tools for Improving Statistical Thinking, 123
- assessment tools, 20, 30, 41, 44, 56, 68, 81, 119, 194
- Assessment Resource Tools for Improving Statistical Thinking, 123
  - Attitudes Toward Mathematics Inventory, 159, *see also* assessment tools
  - Comprehensive Assessment of Outcomes (CAOS), 122, 123
  - Fennema-Sherman Mathematics Attitudes Scales, 159, 167
  - Field-tested Learning Assessment Guide (FLAG), 192, 197
  - Indiana Mathematics Belief Scales, 159
  - Maryland Physics Expectations Survey, 159
  - Precalculus Concept Assessment (PCA), 63, 64
  - Student Assessment of Learning Gains (SALG), 101, 102, 103, 104, 193, 194, 195, 196, 197
  - Students Attitudes Toward Statistics (SATS), 122
- Attitudes Toward Mathematics Inventory, 159
- Banchoff, T., 5, 6
- beliefs
- about mathematics, 145–156, 163, 167
  - about teaching and learning mathematics, 171–179
- Bloom's taxonomy, 22, 39, 42, 129, 130, 132, 135
- Boyer model, 57
- Boyer, E., 3, 4, 5, 6, 8
- Bressoud, David, 60
- CAOS, *see* Comprehensive Assessment of Outcomes
- Carlson, M., 7, 63
- Carnegie Academy for the Scholarship of Teaching and Learning (CASTL), xv, 4, 5, 14, 117, 118, 119, 120, 121, 129, 134, 183
- Carnegie Foundation for the Advancement of Teaching, 3, 4, 109, 203
- Carnegie scholar(s), 5
- Carnegie scholars, xv, xvi, 4, 5, 6, 13, 31, 68, 109, 118, 132, 183
- CASTL, *see* Carnegie Academy for the Scholarship of Teaching and Learning
- CBMS, *see* Conference Board of Mathematical Sciences
- Centers for Teaching and Learning, 16, 47, 89, 134, 135, 158, 167, 197
- civic engagement, xiii, xvi, 5, 17, 22, 46, 88, 99, 105
- classroom response system, 128, 132, 135
- clickers, *see* classroom response systems
- coding data, 36, 68, 71, 121, 130, 131, 132, 134, 135, 185
- emergent (inductive) categories, 39–40
  - predetermined (*a priori*) categories, 39–40
- coding data, 39–41, 130
- collaboration, 46–47
- College Learning for the New Global Century* (AACU), 114
- communication skills, 39, 89, 121, 185, 186
- community-based learning, 7, 9, *see also* service learning
- Comprehensive Assessment of Outcomes (CAOS), 122, 123, *see also* assessment tools
- Conference Board of Mathematical Sciences (CBMS), 153, 172, 173
- confounding factor, 119, 122, 185, 195, *see also* lurking variable, *see also* research design
- constant comparative analysis, 70, 71, *see also* coding data
- axial coding, 71
  - open coding, 71
  - selective coding, 71
- constructivism, 88, 94
- content analysis, *see* coding data
- control group, 20–21, 41, 51, 53, 54, 55, 56, 69, 81, 99, 104, 117, 123, 127, 134, 135, 138, 158, 160, 193, 195, 196, *see also* research design
- course(s)
- intermediate algebra, 51, 52, 53, 57
- criterion-referenced evaluation, 22, *see also* research design
- critical thinking, 83, 108, 153, 165, 178, 187, 192, 196
- Cronbach's alpha, 173
- curriculum development, 9, 104, 136, 191–197
- data, *see also* evidence
- qualitative, xiii, xvi, 7, 15, 19, 21, 23, 27, 31, 35–41, 67, 68, 69, 70, 71, 127, 134, 135, 166, 185, 186, 188
  - quantitative, xiii, 7, 15, 19, 21, 23, 24, 27, 67, 68, 69, 70, 71, 166, 186
  - quantitative vs. qualitative, 23–24
  - triangulating, 7, 19, 88, 184



- dissemination, 5, 45, 88, 94, 140, 141, 152, 153, 166, 192, 196  
 journals, xiv, 8, 47–48  
 choosing a journal, 48, 153, 196  
 conferences, 47
- educational research, xv, 6, 17, 42, 87, 191  
 gold standard, 20, 41, 117
- educational research, 19–23
- educational researchers, 20, 123, 153
- effect size, 173–174
- emergent categories, *see* coding data
- ethical concerns, *see* research design, ethical concerns
- evidence, 19–41
- existing data, 46, 94, 127, 129, 130, 132, 134, *see also*  
 Institutional Review Board
- expedited review, 150
- Experimental Design, *see* research design
- Experimental group, *see* research design. Experimental group
- fantasia, xiii, 203, 204
- Fennema-Sherman Mathematics Attitudes Scales, 159, 167,  
*see also* assessment tools
- Fisher, M., 68
- flipped classroom, *see* inverted classroom
- focus group, xv, 20, 27, 31, 39, 43, 44, 68, 69, 72, 121, 184,  
 185, 186, 188, 205
- focus group, 33–35, 70, 71
- generalizability, 41, *see also* quantitative research
- going public, xvi, 47–48, 65, 114, 121, 146, 152, 194, *see also*  
 dissemination
- guided discovery, 90, 91, 93, 94
- Hawthorne effect, 82
- Huber, M., 4, 6, 7, 9
- human subject(s), xvi, 21, 33, 45–46, 68, 89, *see also* IRB  
 definition of, 45
- Human Subjects Review Board, *see* Institutional Review Board
- Hutchings, P., 4, 6, 7, 9, 14, 23
- Indiana Mathematics Belief Scales, 159, *see also* assessment  
 tools
- inductive categories, *see* coding data
- inertia, xiii, 203, 204, 205–206
- Institutional Review Board (IRB), 45, 53, 55, 68, 89, 129, 132,  
 145, 149, 156  
 Exempt Study, 45, 64  
 existing data, 46, 94, 127, 129, 130, 132, 134  
 expedited review, 150  
 informed consent, 45, 46, 150, 155  
 IRB application, 46, 120, 150  
 IRB process, 46, 64, 89, 94
- integrative learning, *see* learning, integrative
- Integrative Learning Project, 109
- intermediate algebra, 51, 52, 53, 57
- inverted classroom, 127, 128
- IRB, *see* Institutional Review Board
- ISSOTL (International Society for the Scholarship of Teaching  
 and Learning), 5, 46, 186
- Jacobs, D., 132
- just-in-time teaching, 128, 134, 136
- knowledge survey, 20, 27–31, 42
- learning  
 community-based, 7, 9  
 integrative, 90, 109, 115  
 self-regulated, 61, 65  
 service, 88, 100, 105
- liberal education, xiii, 46, 116, 137, 138, 183, 184, 186
- Likert scale, *see* survey
- literature review, 16, 48, 69, 88, 137, 138, 139, 140, 141
- literature search, 13, 138, 139, 192  
 ERIC, 16, 17, 52, 139, 191, 192  
 Math Ed Literature Database, 17  
 PsycINFO, 16  
 snowball strategy, 17  
 Web of Science, 17
- literature search, 16–17
- logical thinking, 183, 186
- lurking variable, 81, 83, *see also* confounding factor, *see also*  
 research design
- MAA (Mathematical Association of America), xiv, xv, xvi, 5,  
 6, 7, 17, 47, 60, 93, 105, 114, 187, 205  
 minicourse, xv, xvi, 5, 93
- Making the Connection: Research and Teaching in  
 Undergraduate Mathematics Education*, 7
- margin of error, 22, 27, 29
- Maryland Physics Expectations Survey, 159, *see also*  
 assessment tools
- math anxiety, 23, 24, 100, 161
- Mathematical Association of America, *see* MAA
- mathematical habits of mind, 40, 172, 175, 176
- MET, *see* The Mathematical Education of Teachers
- MET II, *see* The Mathematical Education of Teachers II
- metacognition, 39, 61, 65, 94, 95, 130, 131, 154
- methodology, 20, 31, 32, 39, 51, 83, 88, 146, 150, 187, *see also*  
 research design
- mixed quantitative and qualitative design, *see* research design
- mnemonic, 52, 57
- muddiest point, *see* questions, muddiest point
- National Council of Teachers of Mathematics, *see* NCTM
- National Science Foundation (NSF), 19, 44, 46, 47, 59, 60, 64,  
 65, 68, 99, 104, 121, 175, 191
- NCTM (National Council of Teachers of Mathematics), 42,  
 154, 176, 177
- normalized gain, 23, 43
- NSF, *see* National Science Foundation
- online homework, 24, 25, 67, 69, 70, 71, 72, 76
- p*-value, 54, 56, 63, 102, 103, 133, 159
- pathologies of teaching, xiii

- Pearson correlation, 122
- PCA, *see* Precalculus Concept Assessment
- pedagogical content tool, 52, 57, 58
- peer review, xiv, xv, 3, 4, 6, 8, 13, 19, 87, 152
- pilot study, *see* research design, pilot study
- pre- and post-comparisons, *see* research design, pre- and post-comparisons
- pre-existing data, *see* existing data
- Precalculus Concept Assessment (PCA), 63, 64
- predetermined categories, *see* coding data
- Project Kaleidoscope (AACU), 5, 9, 10
- proof, *see* writing, proof
- qualitative research
- coherent, 41
  - communicable, 41
  - inter-coder reliability, 41
  - inter-rater reliability, 122
  - justifiable, 41
  - quality of, 41
  - transferable, 41
  - transparent, 41
- quantitative research
- generalizability, 41
  - quality of, 41
  - reliability, 41, 117, 122, 145, 150, 151, 171, 173
  - validity, 41, 117, 122, 145, 150, 166
- quasi-control group experiment, *see also* research design
- quasi-experimental design, *see* research design, quasi-experimental
- questions
- muddiest point, 127, 128, 129, 130, 131, 132, 134
  - reading, xiii, 23, 127, 135, 137–141
- randomized study, 42, 117, 118, 119, 121, 122, *see also* research design
- Rasmussen, C., 7
- reading
- questions, xvi
- reading questions, *see* questions, reading
- reflective
- teaching, 9
- reliability, 41, *see also* qualitative research, quantitative research
- inter-coder, 41
  - inter-rater, 122
- remedial mathematics, 51
- research design, 4, 15, 19, 46, 51, 68, 74, 132, 150, *see also* methodology
- confounding factor, 119, 122, 185, 195
  - control group, 20–21, 41, 51, 53, 54, 55, 56, 69, 81, 99, 104, 117, 123, 127, 134, 135, 138, 158, 160, 193, 195, 196
  - ethical concerns, 21, 22, 46, 69, 83, 104, 117, 119–120, 122
  - experimental group, 20, 51
  - gold standard, 20, 41, 117
  - lack of a control group, 20–22
  - lurking variable, 81, 83
  - mixed methods, 67–69
  - pilot study, 52, 55, 89, 194
  - pre- and post-comparisons, 20, 21, 22, 23, 25, 29, 30, 36, 53, 54, 56, 91, 101, 102, 103, 104, 158, 159, 173, 174, 175, 193, 194, 195
  - quasi-control group experiment, 127, 132, 134
  - quasi-experimental design, 20, 21, 81
  - random assignment, 20, 41, 42, 117, 118, 161
  - switching replications design, 20, 81, 132
  - treatment group, 53–54, 55, 56, 21, 193, 195
- research design, 19–23
- Research Experience for Undergraduates, *see* REU
- Research in Mathematics Education, xv
- Research on Undergraduate Mathematics Education, xiv, 3, 6, 42, 87, 94, 95, 188, *see also* RUME
- research question
- framing, xiii, xv, 13, 15, 19, 139, 147–148, 191
  - narrowing, 15–16, 138
  - refining, xiii, 7, 68, 77, 83, 118, 137, 138
- REU (Research Experiences for Undergraduates), 171–175
- Richlin, L., 4
- rubric, xvi, 27, 35–39, 43, 51, 56, 118, 120, 122
- VALUE (Valid Assessment of Learning in Undergraduate Education), 39
- RUME (Research on Undergraduate Mathematics Education), xiv, xv, 3, 5, 6–7, 14, 47, 48, 87–89, 94, 95
- Salem, A., 6
- SALG, *see* Student Assessment of Learning Gains
- sample size, 22, 71, 83, 151, 195
- SATS, *see* Student Attitude Toward Statistics
- Schoenfeld, A., 94, 146, 147, 148, 149, 159
- scholarship
- types of, 3
- Scholarship Assessed* (Glassick, Huber & Maeroff), 8
- scholarship of teaching and learning, *see* SoTL
- Scholarship Reconsidered* (Boyer), 3, 8
- Science Education for New Civic Engagements and Responsibilities (SENCER), 10, 67–68, 99–100, 101, 105, 191, 193, 194, 196, 197, 205
- SENCER, *see* Science Education for New Civic Engagements and Responsibilities
- SENCER Leadership Fellow, 197, 205
- service learning, 88, 100, 105, *see also* community-based learning
- Shulman, L., 4, 109, 203
- SoTL
- and RUME, 6–7, 87–89, 94
  - benefits of, xv, 3, 9, 203, 206
  - counting as research, 8
  - definition of, 3, 4
  - emergence in mathematics, 3, 5–6
  - evaluation of, 8
  - origins of, 3–4
  - starting point of, 13–14
- SoTL and RUME, 6–7, 87–89, 94
- SoTL taxonomy, xv, xvi, 7, 13, 14–15, 22, 23, 118, 127, 148, 183
- theory-building, 14, 32
  - vision of the possible, 14, 107, 109, 120, 123, 184, 187
  - What could be?*, xv, 7, 14–15, 23, 109, 120, 184, 192

- SoTL taxonomy (*cont.*)  
*What is?*, xv, 7, 14–15, 16, 22, 23, 110, 127, 128–129, 131, 132, 135, 148, 183, 204  
*What works?*, xv, 7, 14–15, 16, 20, 23, 51, 118, 120, 127, 128–129, 131, 132, 135, 148, 186, 204
- split-half reliability, 122
- statistically significant, 22–23, 25, 54, 56, 63, 67, 71, 81, 82, 83, 99, 100, 102, 103, 104, 133, 150, 152, 159, 160, 166, 192
- statistics  
 ANOVA, 82  
 Chronbach's alpha, 173  
 course(s), xvi, 21, 83, 99–105, 117–123, 127–129  
 effect size, 173–174  
*p*-value, 159  
 Pearson correlation, 122  
 sample size, 22, 71, 83, 151, 195  
 stratified random sampling, 70  
*t*-test, 22, 54, 56, 63, 100
- Straley, T., 5
- stratified random sampling, 70
- student  
 attitudes, 15, 22, 26, 68, 69, 70, 83, 101, 104, 122, 24, 158, 159, 163, 166, 171  
 competition, 78, 81, 85  
 confidence, 21, 27, 31, 42, 72, 99–105, 116, 132–133, 160, 162, 164, 172, 175, 179, 188, 192, 193, 196  
 cooperation, 78, 84, 85  
 descriptions of mathematics, 40, 145, 157, 162, 165–166  
 interest, 99, 100, 101, 102, 103, 160, 166, 191, 192, 193, 196  
 letters of recommendation, 103  
 motivation, 21, 25, 31, 69, 70, 78, 79, 80, 84, 108, 154  
 retention, 19, 21, 24, 60, 65, 73, 99, 100, 103, 104, 184, 191, 192  
 study group, 61–65, 83  
 voice, 7, 88, 196
- Student Assessment of Learning Gains (SALG), 101, 102, 103, 104, 193, 194, 195, 196, 197, *see also* assessment tools
- Student Attitudes Toward Statistics (SATS), 122, *see also* assessment tools
- student voice, 7, 88
- study group, 61–65, 83
- survey  
 Attitudes Toward Mathematics Inventory, 159  
 design, 25–27  
 Fennema-Sherman Mathematics Attitudes Scales, 159, 167  
 Indiana Mathematics Belief Scales, 159  
 Likert scale, 24, 25, 26, 53, 102, 149, 152, 153, 158, 159, 160, 163, 167, 173
- Maryland Physics Expectations Survey, 159  
 response rate, 26, 93, 101, 104, 120, 159
- sustainability, 192, 194, 196, 197, 198
- switching replications design, *see* research design, switching replications
- t*-test, 22, 54, 56, 63, 100
- taxonomy  
 Bloom's, 22, 39, 42, 129, 130, 132, 135  
 Mathematical Knowledge and Expertise, 32, 42  
 of pedago-pathologies, 203  
 SoTL, xv, xvi, 7, 13, 14–15, 22, 23, 118, 127, 148, 183
- taxonomy of mathematical knowledge and expertise, 32, 42
- taxonomy of SoTL questions, *see* SoTL Taxonomy
- teacher preparation, xvi, 46, 172, 175, 177
- teaching  
 award, 5, 121, 123, 196, 205  
 good, xiii, xv, 3, 4, 16  
 reflective, xiii, 4, 16, 59  
 scholarly, xiii–xiv, xv, 3, 4, 8, 16, 59, 60  
 tips, 6, 88, 136
- Teaching Center, *see* Center for Teaching and Learning
- teaching commons, 6, 7, 115
- teaching problem, *see* SoTL, starting point of  
*The Mathematical Education of Teachers* (MET), 173, 176  
*The Mathematical Education of Teachers II* (MET II), 176
- theory-building, 14, 32
- think-aloud, xvi, 27, 31–33, 39
- transfer of knowledge, 183–188
- treatment group, 21, 53–54, 55, 56, 193, 195, *see also* research design
- triangulation, 7, 19, 88, 184, *see* data, triangulating
- Tucker, A., 183
- validity, 41, 117, 122, 145, 150, 166, *see also* quantitative research
- visual cue, 22, 51–57
- WebAssign*®, 69, 71
- Wieman, C., 22
- Winkel, Brian, xiv
- writing  
 assignments, 27, 36, 108, 109, 110, 114, 115, 188  
 prompts, 110, 116, 120  
 proofs, 31, 57, 183, 185, 186  
 reflective, 21, 23, 27, 36, 39, 68, 120, 121, 175
- Writing Across the Curriculum, 108