DESIGNING TEACHER EVALUATION SYSTEMS
CONTENTS

About the Editors ix
About the Contributors xi
About the Bill & Melinda Gates Foundation xix

CHAPTER 1: Why Measure Effective Teaching? 1
Jeff Archer, Kerri A. Kerr, and Robert C. Pianta

SECTION 1: Using Data for Feedback and Evaluation 7

CHAPTER 2: Grade-Level Variation in Observational Measures of Teacher Effectiveness 9
Kata Mihaly and Daniel F. McCaffrey

CHAPTER 3: Improving Observational Score Quality: Challenges in Observer Thinking 50
Courtney A. Bell, Yi Qi, Andrew J. Croft, Dawn Leusner, Daniel F. McCaffrey, Drew H. Gitomer, and Robert C. Pianta

CHAPTER 4: How Framework for Teaching and Tripod 7Cs Evidence Distinguish Key Components of Effective Teaching 98
Ronald F. Ferguson with Charlotte Danielson

CHAPTER 5: Making Decisions with Imprecise Performance Measures: The Relationship Between Annual Student Achievement Gains and a Teacher’s Career Value Added 144
Douglas O. Staiger and Thomas J. Kane

CHAPTER 6: To What Extent Do Student Perceptions of Classroom Quality Predict Teacher Value Added? 170
Stephen W. Raudenbush and Marshall Jean
SECTION 2: Connecting Evaluation Measures with Student Learning

CHAPTER 7: Combining Classroom Observations and Value Added for the Evaluation and Professional Development of Teachers

Erik A. Ruzek, Christopher A. Hafen, Bridget K. Hamre, and Robert C. Pianta

CHAPTER 8: Classroom Observation and Value-Added Models Give Complementary Information About Quality of Mathematics Teaching

Candace Walkington and Michael Marder

CHAPTER 9: Does the Test Matter? Evaluating Teachers When Tests Differ in Their Sensitivity to Instruction

Morgan S. Polikoff

CHAPTER 10: Understanding Instructional Quality in English Language Arts: Variations in PLATO Scores by Content and Context

Pam Grossman, Julie Cohen, and Lindsay Brown

CHAPTER 11: How Working Conditions Predict Teaching Quality and Student Outcomes

Ronald F. Ferguson with Eric Hirsch


CHAPTER 12: Evaluating Efforts to Minimize Rater Bias in Scoring Classroom Observations

Yoon Soo Park, Jing Chen, and Steven L. Holtzman

CHAPTER 13: Scoring Design Decisions: Reliability and the Length and Focus of Classroom Observations

Jillian N. Joe, Catherine A. McClellan, and Steven L. Holtzman

CHAPTER 14: Assessing Quality Teaching in Science

Susan E. Schultz and Raymond L. Pecheone
CHAPTER 15: Evidence on the Validity of Content Knowledge for Teaching Assessments 493
   Drew H. Gitomer, Geoffrey Phelps, Barbara H. Weren,
   Heather Howell, and Andrew J. Croft

CHAPTER 16: Optimizing Resources to Maximize Student Gains 529
   Catherine A. McClellan, John R. Donoghue, and Robert Pianta

CONCLUSION: Measuring Effective Teaching—The Future Starts Now 583
   Robert Pianta and Kerri A. Kerr

Author Index 591

Subject Index 599
ABOUT THE EDITORS

**Thomas Kane** is the Walter H. Gale Professor of Education at the Harvard Graduate School of Education and faculty director of Harvard’s Center for Education Policy Research. As a deputy director in the K–12 team at the Bill & Melinda Gates Foundation, he directed the Measures of Effective Teaching project. His work has influenced thinking on a range of topics in K–12 and higher education: from measuring teacher effectiveness, to school accountability in the No Child Left Behind Act, to college financial aid, to charter schools, to race-conscious college admissions and measuring the economic payoff to a community from college education. He has been a faculty member at Harvard’s Kennedy School of Government and UCLA’s School of Public Affairs, as well as serving as a senior economist in President Clinton’s Council of Economic Advisers. He has held visiting fellowships at the Brookings Institution and the Hoover Institution at Stanford University.

**Kerri A. Kerr** is an education research and policy consultant. Her work focuses on managing the implementation of large-scale research projects, as well as advising organizations on successful data, accountability, and evaluation strategies. She joined the Measures of Effective Teaching project team at the project’s inception and supported MET’s data collection, analysis, and reporting activities throughout the life of the study. Her prior research focused on school leadership, school transitions, and district-wide scaling of education initiatives. Prior to beginning her career as a consultant, she was a researcher at the RAND Corporation and the chief data and accountability officer at New Leaders for New Schools. A former middle school and high school math teacher, Kerr received a Ph.D. in sociology from Johns Hopkins University.

**Robert Pianta** is the dean of the Curry School of Education, the Novartis US Foundation Professor of Education at the University of Virginia, and founding director of the Center for Advanced Study of Teaching and Learning (CASTL). Pianta’s interests focus on the measurement and production of
effective teaching in classrooms from preschool to high school. He has published more than three hundred scholarly papers and books on these topics. Pianta is senior developer of the Classroom Assessment Scoring System, a method for observational assessment of teacher-student interaction used in many large-scale applications, and is developer of MyTeachingPartner, a web-based coaching and video professional development program. He consults with foundations and with state and federal agencies.
ABOUT THE CONTRIBUTORS

Jeff Archer is a communications consultant specializing in school improvement issues. Previously he covered the teaching profession and school leadership for *Education Week* and documented school improvement practices for the Effective Practice Incentive Community (EPIC), an online knowledge system created by New Leaders for New Schools. For the MET project, Archer helped direct the project’s communication to practitioners.

Courtney A. Bell is a senior research scientist in Educational Testing Service’s Understanding Teaching Quality Center. She completed her doctorate at Michigan State University in curriculum, teaching, and educational policy after earning her B.A. in chemistry at Dartmouth College. A former high school science teacher and teacher educator, Bell’s work focuses on the measurement of teaching and the intersections of policy and practice.

Lindsay Brown is a Ph.D. candidate in curriculum and teacher education at the Stanford Graduate School of Education. She is also an Institute of Education Sciences (IES) fellow, focusing on issues in teacher quality, teacher performance assessment, and professional development. Before joining the Stanford GSE, she worked as a middle school teacher.

Jing Chen is a research project manager at Educational Testing Service. She received her dual Ph.D. degrees in measurement and quantitative methods and in science education from Michigan State University. Her research interests include human and automated constructed response scoring, test development, test equating, standard setting, and learning progression–based assessments.

Julie Cohen is a post-doctoral fellow at the Center to Support Excellence in Teaching at Stanford University. Beginning in 2014, she will be an assistant
professor of curriculum, instruction, and special education at the University of Virginia. Her research focuses on the features of high-quality teaching that cut across content areas and explores how grade level, school settings, and instructional content serve as important contexts for understanding teacher and teaching quality.

Andrew J. Croft is a research associate at the Educational Testing Service. He has provided research support to various projects on teaching quality, educational policy, and the evaluation of teaching. His research interests are teacher preparation, history of education, and teaching and learning in higher education. He received his master’s degree in higher and postsecondary education from Teachers College, Columbia University, in 2011.

Charlotte Danielson, a former economist, is an internationally recognized expert in the area of teacher effectiveness, specializing in the design of teacher evaluation systems that, while ensuring teacher quality, also promote professional learning. She advises state education departments and national ministries and departments of education, both in the United States and overseas. She is in demand as a keynote speaker at national and international conferences and as a policy consultant to legislatures and administrative bodies.

John R. Donoghue is a psychometrician and statistical software developer, and co-owner of ClowderConsulting. He has spent more than twenty years in the testing industry, researching statistical models and creating software for practical implementation. Donoghue received his Ph.D. in quantitative psychology from the University of Southern California. His publications range across IRT, cluster analysis, and multiple comparisons.

Ronald F. Ferguson is the creator of the Tripod Project for School Improvement, the faculty co-chair and director of the Achievement Gap Initiative (AGI) at Harvard University, and faculty co-director of the Pathways to Prosperity Project at the Harvard Graduate School of Education. He has taught at Harvard since 1983. His most recent book is Toward Excellence with Equity: An Emerging Vision for Closing the Achievement Gap, published by Harvard Education Press.
Drew H. Gitomer is the Rose and Nicholas DeMarzo Chair in the Graduate School of Education at Rutgers University. His research centers on the assessment and evaluation of teaching and related policy issues in teaching and teacher education. Prior to coming to Rutgers, Gitomer was a researcher and senior vice president of research at Educational Testing Service in Princeton, New Jersey, where he led the Understanding Teaching Quality Center.

Pam Grossman is the Nomellini-Olivier Professor of Education at the Stanford University School of Education. She has published broadly in the areas of teacher education and professional education, teacher knowledge, and the teaching of English in secondary schools. She is a member of the National Academy of Education and currently serves as the faculty director of the Center to Support Excellence in Teaching at Stanford.

Christopher A. Hafen is a research scientist at the University of Virginia’s Center for Advanced Study of Teaching and Learning and the managing director for the professional development intervention MyTeachingPartner-Secondary. His research interests include understanding the function and development of close relationships in adolescence, with a particular interest in the ways in which contexts drive changes in close relationships over time.

Bridget K. Hamre is an associate research professor and associate director of the University of Virginia’s Center for Advanced Study of Teaching and Learning. Hamre’s areas of expertise include student-teacher relationships and classroom processes that promote positive academic and social development for children. She has also collaborated in the development and testing of several innovative professional development models to support teachers in improving their classroom practices.

Eric Hirsch is the chief external affairs officer for the New Teacher Center, a non-profit dedicated to improving student learning. Hirsch has served as executive director of the Center for Teaching Quality and as education program manager at the National Conference of State Legislatures. He has led fifty statewide and district teaching conditions initiatives over the past decade, collecting more than 1,000,000 educator surveys. Hirsch has worked with