Maria Teresa Tatto
Michael C. Rodriguez · Wendy M. Smith
Mark D. Reckase · Kiril Bankov *Editors*

Exploring the Mathematical Education of Teachers Using TEDS-M Data

Springer

Exploring the Mathematical Education of Teachers Using TEDS-M Data

Maria Teresa Tatto • Michael C. Rodriguez
Wendy M. Smith • Mark D. Reckase
Kiril Bankov
Editors

Exploring the Mathematical Education of Teachers Using TEDS-M Data

Springer

Editors
Maria Teresa Tatto
Mary Lou Fulton Teachers College
Arizona State University
Tempe, AZ, USA

Wendy M. Smith
Center for Science, Mathematics
& Computer Education
University of Nebraska
Lincoln, NE, USA

Kiril Bankov
Faculty of Mathematics and Informatics
University of Sofia
Sofia, Bulgaria

Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia, Bulgaria

Michael C. Rodriguez
Department of Educational Psychology
University of Minnesota
Minneapolis, MN, USA

Mark D. Reckase
Michigan State University
East Lansing, MI, USA

ISBN 978-3-319-92143-3 ISBN 978-3-319-92144-0 (eBook)
https://doi.org/10.1007/978-3-319-92144-0

Library of Congress Control Number: 2018949360

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

For more than a century, researchers in mathematics education have had an international dimension to their work. They have attempted to study how mathematics is taught and learned both within and across countries. When, for example, the Commission Internationale de l'Enseignement Mathématique (reincarnated later as the International Commission on Mathematical Instruction) was established in 1908, its first project was to invite reports on the state of mathematics teaching in its member countries. The resulting surveys were used to make cross-country comparisons and stimulate within-country reforms.

During the last half century, international studies of the teaching and learning of mathematics in various countries have grown enormously in scope, number, and quality. Researchers have collected vast amounts of data on course syllabi and textbooks, students' attitudes and performance on mathematics problems, and recordings of instruction in representative samples of mathematics classrooms across multiple countries.

The twenty-first century has seen researchers' attention move beyond the teaching and learning of school mathematics to consider the teachers of that mathematics—what they know and believe about the subject and its teaching as well as how they learn to teach. The Teacher Education and Development Study in Mathematics (TEDS-M) is the largest and best known international study of mathematics teacher preparation. The present book contains reports of exploratory analyses of the TEDS-M data that were primarily produced as the result of a series of workshops to acquaint scholars—especially beginning mathematics education researchers—with the study and the data.

The first section of the book looks at programs to prepare teachers to teach mathematics in primary and lower secondary school. The researchers found considerable variation in programs across and within countries as well as in the backgrounds and preparation of the teacher educators staffing those programs. The TEDS-M data clearly offer a rich selection of program characteristics, and the chapters in the first section only begin to explore the ways in which those characteristics might be related.

The second section focuses on the prospective teachers in the various programs—their knowledge, beliefs, and other characteristics. Again, the researchers found variety across programs and countries, and they were able to discern some provocative patterns connecting what the prospective teachers had learned and what their program had offered them. The cross-sectional nature of the TEDS-M data limited the attention that could be paid to change, but the researchers were nonetheless able to draw some provocative inferences about program outcomes.

The third section addresses methodological issues raised by TEDS-M that concern sampling, instrument development, and validation. The statistical sampling, which involved institutions and populations of prospective teachers and teacher educators in multiple countries, was highly complex, as was the instrument development process. Sophisticated procedures connected to the development of anchor points and to the use of tests of differential item functioning were used to explore how the various assessment instruments were working. Further, the validity of the TEDS-M knowledge assessment was studied through expert judgments of item content against specifications for teacher knowledge. Clearly, this groundbreaking study has made important inroads in the investigation of international comparisons.

In the past few decades, it has become increasingly difficult to recruit participants for empirical studies in education. A researcher in mathematics education, acting alone, cannot hope to acquire a large, let alone a representative, sample of new teachers or of teacher educators from a given country. Studies of representative samples of such groups from more than one country demand the support of international organizations and of national research centers. The TEDS-M data set, therefore, provides a unique resource for researchers in mathematics education. Although the chapters in the present book demonstrate vividly the potential of that resource, they are far from exhausting it. TEDS-M is a gift to you the readers of this book that deserves to be carried forward in the research you will do.

Athens, Georgia Jeremy Kilpatrick

Preface

In this book, we continue to explore the data collected by the Teacher Education and Development Study in Mathematics (TEDS-M) in 2008, and made publicly available in 2012. TEDS-M is the first and only study to date that has been carried out at an international level to explore the outcomes of mathematics teacher education with nationally representative samples of preservice teacher education programs, their educators, and their future teachers at the primary and secondary levels in 17 countries. The study researchers focused on understanding the depth and breadth of mathematics knowledge that future teachers acquire at the end of their preservice programs, essentially making TEDS-M a study of the outcomes of higher education. TEDS-M is the first and only rigorous study where researchers have sought to test a series of hypotheses about what constitutes effective teacher education for future mathematics teachers that had emerged after years of continuous, albeit small-scale, studies of teacher learning. The TEDS-M researchers used a survey methodology to design a comprehensive evaluative research study that could reflect the complexity of the endeavor.

The study was innovative in that the research team developed assessments of future teachers' mathematics content and pedagogical knowledge, among other measures, while collecting and documenting evidence of validity and reliability. This task was only possible with the collaboration of mathematics teacher educators and their students as part of an international and interdisciplinary team of mathematicians, psychometricians, survey and sampling experts, and policy analysts. The TEDS-M International Center received funding from the U.S. National Science Foundation to develop and coordinate the study. Participating countries sought their own funding, and the International Association for the Evaluation of Educational Achievement (IEA) sponsored the study and contributed logistical support.[1] The IEA's network of collaborators, originally founded in 1958, has marked 60 years of

[1] Maria Teresa Tatto then at Michigan State University directed the study. Other organizations involved in the TEDS-M study included ACER, the IEA Secretariat, the IEA-DPC, and Statistics Canada.

exploring different, innovative and rigorous approaches to evaluating educational effectiveness[2] of which this study is an example.

The TEDS-M study was not without challenges. Many factors mediate the ultimate outcomes of teacher education programs. The knowledge that future teachers are able to demonstrate at the end of their preparation in preservice university-based programs is dependent on their previous learning in their primary and secondary schooling, and importantly through their passage in higher education institutions through consecutive or concurrent teacher preparation programs. Further practical experiences in schools through internships or field experiences create yet another level of learning. Evaluating one's own program using an international and comparative assessment creates important opportunities to learn from what other systems are doing and important opportunities for change. As Arthur W. Foshay remarked in the report of the first IEA study ("Educational Achievements of Thirteen-Year-Olds in Twelve Countries"; retrieved from http://www.iea.nl/brief-history-iea):

> If custom and law define what is educationally allowable within a nation, the educational systems beyond one's national boundaries suggest what is educationally possible.

The comparative study of teacher education across different countries also reflects important societal values including conceptions of the knowledge that is important for societies to have, how it is to be acquired, and from whom. Indeed, we believe that the direction a society elects to follow can be identified by the manner in which it treats its teachers and their students.

Although change is typically slow in teacher education, currently we have a combination of stability and fast-paced change resulting from the introduction of alternative routes to preparing teachers, in what seems to be a global movement. It is in this fast-paced environment, where much of the change is driven by market forces rather than by research evidence, that the TEDS-M study provides an important contribution to the field.

The publication of this book serves to celebrate TEDS-M's research findings and methodological advances, to remind us of what is possible and effective in traditional teacher education programs, and to model research that can inform future policy and scholarship. Whereas the critical reader may question the relevance of data that was collected a decade ago, recent research continues to uncover the same problems that TEDS-M highlighted, and the importance of continuing to do the kind of research we undertook.[3] In some cases, the problems have been exacerbated by the proliferation of alternative routes and redefinitions of what counts as a qualified teacher, to the point that scholars have remarked that the project of education in

---

[2] IEA's work originated in 1958 when a group of scholars, educational psychologists, sociologists, and psychometricians met at the UNESCO Institute for Education (UIE) to discuss problems associated with evaluating school effectiveness and student learning. See http://www.iea.nl/brief-history-iea

[3] See for instance Tatto, M.T., Burn, K., Menter, I., Mutton, T., & Thompson, I. (2018). *Learning to teach in England and the United States: The evolution of policy and practice*. Abingdon, England: Routledge.

universities is in jeopardy.[4] Driven both by teacher shortages and political considerations, popular movements away from formal teacher preparation reveal assumptions about the skills and knowledge necessary to teach effectively; we reject those assumptions and find teaching to be an incredibly complex and difficult art. The lack of a vigorous response from teacher educators to defend the worth of their programs continues to leave the field vulnerable to ideological attacks. We hope that the TEDS-M research will serve as a model to create a basis for inquiry-based learning in teacher education programs.

The purpose of this book, then, goes beyond adding to the key research reports produced by the study,[5] and the already rich research literature based on the TEDS-M data since it was published in 2012, and which can be found in numerous research journals and reports, in books, and elsewhere. We began this book with three purposes in mind. First, we wanted to honor our commitment to the U.S. National Science Foundation to not only make the data publicly available but to enable dissemination of the research, and provide access to our complex database, with a focus on junior scholars and advanced doctoral students. After an initial meeting at Michigan State University, U.S., at the end of 2012, we issued a call for proposals to the national and international mathematics research community. We funded a series of meetings and data analysis workshops for these scholars in exchange for the production of high quality publishable articles and chapters. The call resulted in a number of proposals. These were reviewed by our internal team, during a meeting at Michigan State University, in 2013 and those with the highest quality were selected. We held two meetings with these colleagues in July 2014 in Limerick, Ireland, sponsored by the National Centre for Excellence in Mathematics and Science Teaching and Learning (NCE-MSTL), and in July 2016 at the University of Minnesota, U.S. These week-long meetings resulted in the development of 20 paper proposals, which we organized into four symposia proposals for the American Educational Research Association (AERA), Association of Mathematics Teacher Educators (AMTE), the National Council of Teachers of Mathematics (NCTM), and the Comparative International Education Societies Annual Meetings (held in 2015–2016). The symposia were accepted for presentation, and we experienced positive receptions in these conferences. In addition to continue to give visibility to the study, these conferences helped us to advance the work on the papers. We held

[4] See Furlong, J. (2013). Education – An Anatomy of the Discipline: Rescuing the University Project? London: Routledge.

[5] See for instance: Tatto. M.T. (ed.) (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M). Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries: Technical Report*. Amsterdam: International Association for the Evaluation of Student Achievement. Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M. & Reckase, M. (2012). *Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam: International Association for the Evaluation of Student Achievement. Brese, F., & Tatto, M.T. (Eds.) (2012). *User guide for the TEDS-M international database. Amsterdam*, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

a working meeting in January 2017 to check on the progress of our efforts and to finalize plans for publication. We finished writing the book in March 2017. Once the chapters were written, they were reviewed by the book editors, then by expert statisticians in addition to our own, and once approved, the chapters went through three external editors to ensure clarity and meaning. Once submitted to Springer in mid-2017, the book was reviewed anonymously by six peers following Springer's rigorous review process, and finally approved for publication in early 2018. Second, we wanted the book authors to explore not only issues that have not yet been studied using the TEDS-M database, but also to provide more exploration into aspects of the study that had not been published before, such as the curriculum analysis and the level of belief alignment between future teachers and their educators. Third, we wanted to dedicate a section of the book to discuss in more depth the most important methodological advances achieved by the TEDS-M researchers, such as the sampling strategies and the creation of anchor points to give contextual meaning to the assessments results.

Ultimately, the goal of the book is to inform audiences about how use of the TEDS-M study and data: (a) helps to strengthen the knowledge base to address current national priorities such as increasing the number of fully competent mathematics teachers; (b) helps to understand the nature and contributions of preservice teacher education, as a way to inform policies for selection, preparation, induction, and professional development of mathematics teachers; (c) serves as an example of a scientific approach to the study of teacher education and teacher learning in mathematics; and (d) provides concepts, definitions, measurement strategies, indicators and instrumentation to strengthen the research in this field, and the knowledge base of teacher education effectiveness. We offer this book in the hopes that it will contribute to new studies, and more rigorous scholarship in teacher education, and more generally, in higher education.

July 2018 Maria Teresa Tatto

Acknowledgments

We continue to be thankful to the International Association for the Evaluation of Educational Achievement (IEA) for maintaining the TEDS-M database and the website where scholars continue to add work that uses the TEDS-M database and the methods that emerged from the study. We thank the participants in the TEDS-M study including the National Research Centers, teacher educators, and future teachers in Botswana, Canada, Chile, Chinese Taipei, Georgia, Germany, Malaysia, Norway, Oman, Philippines, Poland, Russia, Singapore, Spain, Switzerland, Thailand, and the United States, whose efforts in data collection continue to benefit the field.

We are thankful to Maria Teresa Tatto for hosting the first two workshops in Michigan State University, Paul Conway of the National Centre for Excellence in Mathematics and Science Teaching and Learning for hosting a very productive workshop in Limerick, and Michael Rodriguez for hosting the last workshop at the University of Minnesota.

We are thankful to Lindsay Augustyn from the University of Nebraska-Lincoln, Jeffery Cole from Michigan State University, and Colin Ong-Dean who served as copy editors for this book. They worked tirelessly with the authors and editors to make sure that the grammar was uniform and APA formatting and styles were consistent across chapters. We are thankful to the Michigan State University Center for Statistical Training and Consulting (CSTAT) whose consultants helped review each chapter for the appropriateness, use, and interpretation of the analysis of the data and statistical results.

We are thankful to our current institutions Arizona State University, the University of Minnesota, the University of Nebraska-Lincoln, Michigan State University, and the University of Sofia, for providing support and the gift of time and space that has allowed us to see this project to fruition.

We are thankful to Annemarie Keur, our editor at Springer, for her encouragement and guidance through this process.

Finally, this work could not have been accomplished without the continuous support of the National Science Foundation. We are thankful to program officers James Dietz and Finbarr Sloane for their dedication to this project and to others at NSF for their interest in this cross-national research including Janice Earle, Elizabeth Vanderputten, and Larry Suter.

The TEDS-M workshops and the resulting work that we present in this book were sponsored in part by a grant from the US National Science Foundation NSF REC 0514431 (M.T. Tatto, PI and Executive Director). Any opinions, findings, and conclusions or recommendations expressed in this book are those of the researchers and do not necessarily reflect the views of the National Science Foundation.

Maria Teresa Tatto
Michael C. Rodriguez
Wendy M. Smith
Mark D. Reckase
Kiril Bankov

Contents

1 Introduction: Exploring the Mathematical Education of Teachers Using TEDS-M Data 1
Maria Teresa Tatto

Part I Exploring Different Dimensions of Teacher Education Programs

2 Introduction: Exploring Different Dimensions of Teacher Education Programs in the TEDS-M Study 17
Maria Teresa Tatto and Wendy M. Smith

3 Preparing High Quality Mathematics Primary Teachers: Exploring Program Strategies and Standards in the United States, Russia, Poland, and Chinese Taipei 31
Yadira Peralta and Maria Teresa Tatto

4 The Intended, Implemented, and Achieved Curriculum of Mathematics Teacher Education in the United States 69
Maria Teresa Tatto and Kiril Bankov

5 Developing Diverse Teachers: Analyzing Primary Mathematics Teacher Education Programs Prioritizing Selection of Diverse Future Teachers 135
James Pippin

6 A Comparative International Study of Differences in Beliefs Between Future Teachers and Their Educators 165
Michael C. Rodriguez, Maria Teresa Tatto, José Palma, and Kyle Nickodem

Part II Exploring Future Teacher Characteristics, Knowledge, Beliefs and Opportunities to Learn

7 Introduction: Exploring Future Teacher Characteristics, Knowledge, Beliefs and Opportunities to Learn in the TEDS-M Study 195
Maria Teresa Tatto and Wendy M. Smith

8 The Mathematical Education of Primary Teachers 205
Maria Teresa Tatto

9 How Primary Future Teachers' Knowledge Is Shaped by Teacher Preparation 257
Hong Qian and Peter Youngs

10 Opportunities to Learn Mathematics Pedagogy and Connect Classroom Learning to Practice: A Study of Future Teachers in the United States and Singapore 279
Traci Shizu Kutaka, Wendy M. Smith, and Lorraine M. Males

11 Preparing Primary Mathematics Teachers to Learn to Work with Students from Diverse Backgrounds 311
Elizabeth B. Dyer

12 Differences in Beliefs and Knowledge for Teaching Mathematics: An International Study of Future Teachers 349
Traci Shizu Kutaka, Wendy M. Smith, and Anthony D. Albano

13 Future Teachers' and Teacher Educators' Perceptions of Learning Mathematics Instruction and Relationships to Knowledge 379
Rachel A. Ayieko

14 The Mathematical Education of Secondary Teachers 409
Maria Teresa Tatto

15 An International Study of the Relationship Between Learning to Teach Students from Diverse Backgrounds and Mathematical Knowledge for Teaching in Future Secondary Mathematics Teachers 451
Elizabeth B. Dyer

Part III Methodological Challenges and Strategies

16 Introduction: Methodological Challenges and Strategies in the TEDS-M Study 479
Maria Teresa Tatto and Wendy M. Smith

17 Sampling for TEDS-M 489
Sabine Meinck and Jean Dumais

18 Developing Anchor Points to Enhance the Meaning of the Mathematical and Mathematical Pedagogy Score Scales from the TEDS-M Study 527
Mark D. Reckase

19 Examining Sources of Gender DIF in Mathematics Knowledge of Future Teachers Using Cross-Classified IRT Models 543
Liuhan Sophie Cai and Anthony D. Albano

20 Standing the Test of Time: Validating the TEDS-M Knowledge Assessment Against MET II Expectations 563
Edward A. Silver and Jillian P. Mortimer

Contributors

Editors

Maria Teresa Tatto is Professor in the Division of Educational Leadership and Innovation at Arizona State University, and the Southwest Borderlands Professor of Comparative Education at the Mary Lou Fulton Teachers College. She is the principal investigator for the Teacher Education and Development Study in Mathematics and the First Five Years of Mathematics Teaching Study, both designed to explore the connections between preservice preparation and what is learned on the job during the first years of teaching. She is a former president of the Comparative and International Education Society and studies the effects of educational policy on school systems. She is an Honorary Research Fellow in the Department of Education at the University of Oxford, England.

Michael C. Rodriguez is Professor of Educational Measurement in the Department of Educational Psychology and Campbell Leadership Chair in Education and Human Development at the University of Minnesota. He is a collaborator for the Teacher Education and Development Study in Mathematics and the First Five Years of Mathematics Teaching Study, providing psychometric support for instrument development and analysis. His interests include item writing, multilevel modeling and item response modeling, early literacy, and assessment of social-emotional learning. He is the recipient of the International Reading Association's Harris Research Award.

Wendy M. Smith is a Research Associate Professor and Associate Director for the Center for Science, Mathematics, and Computer Education at the University of Nebraska. She is a collaborator for the Teacher Education and Development Study in Mathematics and the First Five Years of Mathematics Teaching Study, which explore connections between preservice preparation and what is learned during the first years of teaching. She is a mathematics teacher educator and former middle level mathematics teacher. Her research interests include K-16 mathematics education, rural education, teacher leadership, estimating teacher professional development effects on student achievement, and institutional change.

Mark D. Reckase is a University Distinguished Professor Emeritus of Measurement and Quantitative Methods. He specializes in the development of educational and psychological tests, educational policy related to testing, and the psychometric theory that supports the assessment of cognitive skills and content knowledge. In particular, his research is on applications of unidimensional and multidimensional item response theory (IRT) models, computerized adaptive testing (CAT), assessment using performance tasks, standard setting on educational tests, and methods for evaluating the quality of teacher education programs.

Kiril Bankov is Professor of Mathematics Education in the Faculty of Mathematics and Informatics at the University of Sofia, and in the Institute of Mathematics and Informatics at the Bulgarian Academy of Sciences. He is a mathematics coordinator for the Teacher Education and Development Study in Mathematics and the First Five Years of Mathematics Teaching Study, both designed to explore the connections between preservice preparation and what is learned on the job during the first years of teaching. His research interests include the preparation of mathematics teachers, mathematics teaching, assessment in mathematics education, extracurricular work with mathematically gifted students, and mathematics competitions.

Contributors

Anthony D. Albano is an Assistant Professor in Educational Psychology at the University of Nebraska-Lincoln. His primary research interests include multilevel modeling of item-response data, linking and equating, assessment literacy, and technology in testing and test development.

Rachel A. Ayieko is an Assistant Professor in the Department of Instruction and Leadership in Education at Duquesne University. Her research interests focus on international and comparative studies on mathematics teacher education, teacher quality, mathematics classroom practices, and opportunities to learn mathematics in relation to teacher knowledge and students' learning outcomes. Recent publications have appeared in the *Forum for International Research in Education (FIRE)* and the *African Journal of Research in Mathematics, Science and Technology Education (AJRMSTE).*

Liuhan Sophie Cai is a doctoral candidate in the Department of Educational Psychology at the University of Nebraska-Lincoln (UNL). She joined the Quantitative, Qualitative & Psychometric Methods program in 2013 and received her master's degree in 2015 at UNL. Her research interests include educational measurement, differential item functioning, equating, and item response theory.

Jean Dumais joined Statistics Canada in 1982 after having completed his MSc (Math). Jean worked as a survey statistician on topics as diverse as agriculture, health, price indices, census of population, education, transportation, and on general statistical consultation and training. Jean became the chief of the statistical consultation group in 2003, supervising a team of 8 to 10 senior statisticians. Jean worked on TIMSS 1995, on TEDS-M, and on TALIS since its creation; Jean retired from Statistics Canada in June 2017.

Elizabeth B. Dyer is a Research Associate in the STEM Program at WestEd. She received her PhD from the Learning Sciences program at Northwestern University. Before her doctoral studies, she was a research associate at Horizon Research Inc., working on research studies and evaluations of initiatives in STEM education. Her research interests include mathematics teacher learning and instructional improvement.

Traci Shizu Kutaka is a Research Associate at the Marsico Institute for Early Learning and Literacy at the University of Denver. Her research interests include early childhood care and education as well as teacher preparation and professional development. Kutaka has extensive experience conducting quantitative, qualitative, and mixed-methods research, including collecting survey, interview, and observational data; examining the reliability and validity of measurement and assessment tools; and writing articles and reports for both academic and nonacademic audiences.

Lorraine M. Males is an Assistant Professor in the Department of Teaching, Learning, and Teacher Education in the College of Education and Human Science at the University of Nebraska-Lincoln. Her work broadly focuses on preparing secondary mathematics teachers and understanding the influence of curriculum materials on student and teacher learning. Her current project employs the use of eye tracking technology to examine teachers' attention to mathematics curriculum materials, how teachers interpret what they attend to, and the decisions that teachers make to respond to materials when planning and enacting instruction.

Sabine Meinck is Head of both the Research and Analysis Unit and the Sampling Unit of the International Association for the Evaluation of Educational Achievement (IEA). She is responsible for coordinating all research and sampling-related activities undertaken by the IEA, a remit that also includes providing courses designed to build research capacity. She is experienced in all aspects of the survey design and analysis of international large-scale assessments (ILSAs), and has been involved with almost all ILSAs conducted over the last decade. For the Teacher Education and Development Study in Mathematics, she was responsible for sampling, weighting and variance estimation.

Jillian P. Mortimer is a doctoral candidate in the School of Education at the University of Michigan. She holds a Bachelor of Science degree in elementary education from the University of Michigan and a master's degree in curriculum and instruction from the University of Richmond. Her research interests include elementary teacher education in mathematics.

Kyle Nickodem is a doctoral candidate in the Department of Educational Psychology at the University of Minnesota. After earning a bachelor's degree in psychology from the University of Notre Dame, Kyle served 2 years as an AmeriCorps member with the educational nonprofit College Possible. He currently works as a consultant for the Center for Applied Research and Educational Improvement (CAREI) providing data analysis, assessment, and evaluation services for local education agencies and Minnesota school districts. He is a recipient of the Russell W. Burris Fellowship from the University of Minnesota.

José Palma is a doctoral candidate in the Department of Educational Psychology at the University of Minnesota. Before joining the Quantitative Methods in Education program, José was a bilingual teaching assistant in St. Paul Public Schools. He currently provides psychometric support in the development of the Individual Growth and Development Indicators-Español early literacy project (funded through the Institute of Education Sciences). His interests include the assessment of special populations, multilevel measurement, early literacy in dual language learners, achievement gaps, and opportunity to learn. He is the recipient of the Diversity of Views and Experiences Fellowship and the Multicultural Recognition Award from the University of Minnesota.

Yadira Peralta is a doctoral candidate and a Fulbright scholar in the Department of Educational Psychology at the University of Minnesota. She holds a bachelor's degree in applied mathematics and a master's degree in economics from the Center for Research and Teaching in Economics (CIDE) in Mexico. Before joining the Quantitative Methods in Education doctoral program, Yadira worked as a deputy director of statistical analysis, under the educational indicators area at the National Institute for Educational Assessment and Evaluation (INEE) in Mexico. Her research interests include statistical models for correlated data, nonlinear models, and generalized longitudinal models.

James Pippin is a Postdoctoral Research Associate in the Department of Teacher Education at Michigan State University (MSU). As a Graduate Research Assistant in Educational Policy at MSU, he worked on the Teacher Education and Development Study in Mathematics and the First Five Years of Mathematics Teaching Study. Prior to his doctoral work, he taught K-12 English for several years in South Korea and China. His research interests focus on policies and practices that support the preparation, recruitment, development, and retention of effective teachers for marginalized students in the United States and internationally.

Hong Qian is a psychometrician at the National Council of State Boards of Nursing. She earned a dual PhD degree in measurement and quantitative methods, and curriculum, instruction and teacher education at Michigan State University. Her research focuses on international and comparative analysis of teacher preparation programs, how teacher education programs affect teacher knowledge, and how educational policy influences teacher commitment and instruction. Recent publications have appeared in *Journal of Mathematics Teacher Education*, *Journal of Teacher Education*, *Elementary School Journal*, *Journal of Educational Change*, and *Educational Measurement: Issues and Practice.*

Edward A. Silver is the Senior Associate Dean for Research and Graduate Studies and the William A. Brownell Collegiate Professor of Education in the School of Education and Professor of Mathematics in the College of Literature, Science, and the Arts at the University of Michigan. He teaches and advises graduate students in mathematics education and conducts research related to the teaching and learning of mathematics, with particular attention to the study of mathematical thinking, enhancing the knowledge and effectiveness of mathematics teachers, and productive interchange between education research and practice.

Peter Youngs is a Professor in the Department of Curriculum, Instruction, and Special Education at the University of Virginia. His research interests focus on education policy effects on teaching and learning in the core academic subjects including state and district policy related to teacher preparation, induction, and evaluation and their effects on teacher instruction, commitment, and retention. Recent publications have appeared in *Educational Researcher*, *Journal of Teacher Education*, and *Journal of Mathematics Teacher Education.* He received the AERA Division K (Teaching and Teacher Education) Early Career Award and he currently serves as coeditor of *Educational Evaluation and Policy Analysis.*

Abbreviations

ACER	Australian Council for Educational Research
AERA	American Educational Research Association
AIC	Akaike Information Criterion
AMTE	Association of Mathematics Teacher Educators
APA	American Psychological Association
BIC	Bayesian Information Criterion
BRR	Balanced Repeated Replication
CBMS	Conference Board of Mathematical Sciences
CCSS	Common Core State Standards
CI	Confidence Interval
CR	Constructed-Response
DIF	Differential Item Functioning
DPC	Data Processing and Research Centre now: IEA Hamburg
EI	Essential Idea
FT	Future Teacher
FPT	Future Primary Teacher
FST	Future Secondary Teacher
FTQ	Future Teacher Questionnaire
ICC	Intra-class correlation coefficient
IDB	International Data Base
IEA	International Association for the Evaluation of Educational Achievement
IHE	Institutes of Higher Education
ILSA	International Large-Scale Assessments
IPQ	Institutional Program Questionnaire
LCA	Latent Class Analysis
MC	Multiple-Choice
MCK	Mathematical Content Knowledge
MET II	The Mathematical Education of Teachers II
MKT	Mathematical Knowledge for Teaching
MPCK	Mathematical Pedagogical Content Knowledge
MSU	Michigan State University

MT21	Mathematics Teaching in the 21st Century Study
NCTM	National Council of Teachers of Mathematics
NCME	National Council on Measurement in Education
NIE	National Institute of Education
NRC	National Research Coordinator
NSF	National Science Foundation
NYC	New York City
OECD	Organisation for Economic Co-operation and Development
OTL	Opportunities to Learn
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PPS	Probability Proportional to Size
PST	Preservice Teacher
SE	Standard Error
SES	Socioeconomic Status
SMP	Common Core State Standard for Mathematical Practice
SRS	Simple Random Sampling
SSBIC	Sample-Size Adjusted Bayesian Information Criterion
STEM	Science, Technology, Engineering, and Mathematics
TALIS	Teaching and Learning International Survey
TEDS-M	Teacher Education and Development Study in Mathematics
TIMSS	Trends in International Mathematics and Science Study
TP	Teacher Preparation
TPU	Teacher Preparation Unit
TPI	Teacher Preparedness Inventory
USA	United States of America
WinW3S	Windows® Within-School Sampling Software

Chapter 1
Introduction: Exploring the Mathematical Education of Teachers Using TEDS-M Data

Maria Teresa Tatto

Abstract How does teacher education contribute to the learning outcomes of future teachers? Are there programs that are more successful than others in helping teachers learn to teach? How do local and national policy environments contribute to teacher education outcomes? This chapter introduces the book to readers and invites them to explore these questions across a large number of settings. The chapter illustrates why investigating the impact of pre-service teacher education on teachers' learning outcomes is a necessary component to understanding variation in the quality of teachers who enter the field. The chapter also provides an overview of the Teacher Education and Development Study in Mathematics (TEDS-M) a cross-national study of primary and secondary mathematics teacher education sponsored by the International Association for the Evaluation of Educational Achievement (IEA), and funded by the U.S. National Science Foundation and participating countries. The book includes original work that explores new facets of the TEDS-M methodology and data, along with results and policy implications; and illustrates the challenges and possibilities in engaging in systematic research on teacher education. Because we lack models to frame research on teacher education processes and outcomes, the book seeks to provide guidance to future research in this area by

The text in this *Overview* contains shortened and slightly edited versions of text that has appeared in the following publications: Tatto et al. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework*. Amsterdam, the Netherlands: International Association for the Evaluation of Student Achievement, and Tatto et al. (2012). *Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, the Netherlands: International Association for the Evaluation of Student Achievement. Text cited directly or indirectly from those sources will not be made recognizable.

An extensive report on the descriptive findings on the characteristics of teacher education programs and teacher educators in the study can be found in Tatto et al. (2012), and in Tatto (2013). We summarize the key concepts and findings here to orient the reader; the chapters in this part however are original contributions written exclusively for this book.

M. T. Tatto (✉)
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: teresa.tatto@asu.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_1

outlining the methodology followed by the TEDS-M study as well as findings from secondary analyses of the rich TEDS-M database.

Introduction

How does teacher education contribute to the learning outcomes of future teachers? Are there programs that are more successful than others in helping teachers learn to teach? How do local and national policy environments contribute to teacher education outcomes? This book invites readers to explore these questions across a large number of settings. Although these questions seem simple, authoritative answers are hard to find. Recent work in the United States, for example, has tended to focus more on the learning outcomes of pupils of program graduates rather than the learning outcomes of the prospective teachers themselves (e.g., Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Goldhaber, Liddle, & Theobald, 2013; Koedel, Parsons, Podgursky, & Ehlert, 2015). Yet important research has been done documenting that teachers' knowledge and beliefs are key to pupils' achievement and that teachers' previous preparation should be considered an important policy priority (Campbell et al., 2014; Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Hill, Rowan, & Ball, 2005; Metzler & Woessmann, 2012; Wilkins, 2008). Thus, investigating the impact of pre-service teacher education on teachers' learning outcomes is a necessary component to understanding variation in the quality of teachers who enter the field.

This book uses the data collected by the Teacher Education and Development Study in Mathematics (TEDS-M) in 2008. The TEDS-M study is a cross-national study of primary and secondary mathematics teacher education sponsored by the International Association for the Evaluation of Educational Achievement (IEA) and funded by the National Science Foundation and participating countries. TEDS-M focuses on *how teachers are prepared to teach mathematics in primary and lower secondary school.* Consequently, TEDS-M is a study of the variation in the nature and impact of teacher education programs within and across countries.

The purpose of this book is twofold: first, to describe the different phases of the TEDS-M study and showcase original work that explores new facets of the TEDS-M database, along with results and policy implications; and second, to illustrate the challenges and possibilities in engaging in systematic research on teacher education. Because we lack good models to frame research on teacher education processes and outcomes, the book seeks to provide guidance to future research in this area by outlining the methodology followed by the TEDS-M study as well as findings from secondary analyses of the rich TEDS-M database.

The book is organized around the TEDS-M conceptual framework and research questions, and has three parts. *Part I* includes chapters that explore the characteristics of the teacher education programs studied, including the curriculum, the strategies and guidelines that programs use to prepare highly knowledgeable teachers, and the preparation of teachers to meet the needs of diverse learners. This part also includes a study focusing on teacher educators, particularly examining the degree of alignment between the beliefs of teacher educators and future teachers. *Part II* moves to the study of future teachers' beliefs, knowledge, and opportunities to learn. *Part III* includes chapters that address some important methodological issues that arose in TEDS-M and that have not been discussed in depth elsewhere. In particular, chapters in Part III discuss the challenges of creating a common language across settings and countries before undertaking the research; developing rigorous instruments with validity evidence that produce reliable scores, as well as a sampling frame across countries that is sensitive to within-country variation, culture, and norms; and the development of anchor points to convey contextual meaning to the study findings. The last two chapters use TEDS-M data to examine differential item functioning, and to provide validity evidence to support the use and interpretation of TEDS-M assessment results against the expectations included in the CBMS report: The Mathematical Education of Teachers (MET II).[1]

The TEDS-M Framework

The impetus for TEDS-M, conducted in 17 countries, was recognition that teaching in general, and specifically in the so-called STEM subjects, has become more challenging worldwide, as growth in knowledge demands frequent curricular change, and as large numbers of teachers reach retirement age. It also has become increasingly clear that effectively responding to demands for teacher preparation reform will remain difficult while there is lack of consensus on what such reform should encompass. In the absence of empirical data, efforts to reform and improve educational provision in the highly contested STEM arena continue to be undermined by traditional and implicit assumptions. TEDS-M accordingly focused on collecting, from the varied national and cultural settings represented by the participating countries, empirical data that could inform policy and practice related to recruiting and preparing a new generation of teachers capable of teaching increasingly demanding mathematics curricula.

Although future teachers and school systems must place their trust in the numerous and diverse teacher education programs across the world, no comprehensive, authoritative study of the outcomes of teacher education had been carried out at the time that the TEDS-M study took place, and none has been done since. The lack of work in this area made it essential to do a comprehensive study of teacher education's immediate outcomes to identify what knowledge, skills and, dispositions future teachers have close to graduation and when they are declared ready to teach. An important assumption of the TEDS-M study is that the education of teachers is

[1] http://www.cbmsweb.org/the-mathematical-education-of-teachers/

not generic, and that learning to teach occurs within subject contexts. Consequently, TEDS-M is subject-specific and focuses on mathematics teacher education as an area to study.

Two particular purposes underpinned TEDS-M. The first was to identify how the countries participating in TEDS-M prepare teachers to teach mathematics in primary and lower-secondary schools. The second was to study variation in the nature and impact of teacher education programs on future teacher knowledge and beliefs within and across the participating countries. The information collected came from representative samples (within the participating countries) of preservice teacher education programs, their future primary and lower-secondary school teachers, and their teacher educators.

The 17 countries that participated in TEDS-M were Botswana, Canada (four provinces), Chile, Chinese Taipei, Georgia, Germany, Malaysia, Norway, Oman (lower-secondary teacher education only), the Philippines, Poland, the Russian Federation, Singapore, Spain (primary teacher education only), Switzerland (German-speaking cantons), Thailand, and the United States (public institutions only). Across the 17 participating countries, approximately 22,000 future teachers from 751 programs were surveyed and tested. Teaching staff within these programs were also surveyed—close to 5,000 mathematicians, mathematics educators, and general pedagogy educators.

The overall TEDS-M study has three overlapping components:

- Studies of teacher education policy, schooling, and social contexts at the national level;
- Studies of primary and lower secondary mathematics teacher education programs, standards, and expectations for teacher learning; and
- Studies of the mathematics and related teaching knowledge of future primary and lower secondary mathematics teachers.

TEDS-M explored the associations among these components, such as associations among teacher education policies, program practices, and future teacher outcomes as shown in the TEDS-M Conceptual Framework in Fig. 1.1.

Specifically, TEDS-M investigated the following research questions:

1. What are the policies that support primary and secondary teachers' achieved level and depth of mathematics and related teaching knowledge?
2. What learning opportunities, available to prospective primary and secondary mathematics teachers, allow them to attain such knowledge?
3. What level and depth of mathematics and related teaching knowledge have prospective primary and secondary teachers attained by the end of their preservice teacher education?

A common question across these three areas of inquiry concerned cross-national and intra-national variation—specifically, to what extent do teacher education policy, opportunities to learn, and future teachers' mathematics subject and pedagogy knowledge vary across and within countries?

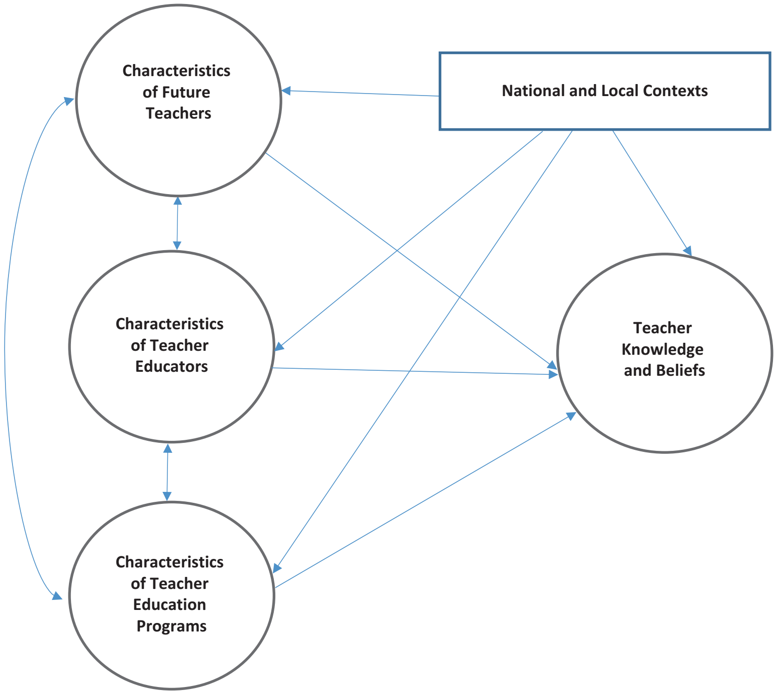


Fig. 1.1 TEDS-M Conceptual Framework

Studying Mathematics Teacher Education: TEDS-M Findings to Date

The main TEDS-M findings are well-documented, both in reports which can be found on the IEA website under the association's complete list of publications available online in the ILSA Gateway (http://www.ilsa-gateway.org/ and search by study 'TEDS-M'), in the ERIC system and in various several articles and special issues.

Although the TEDS-M study has provided and continues to provide new insights into the nature of mathematics teacher education across the participating countries, one of the most important findings for the field of mathematics teacher education, and comparative education more broadly, is the high degree of variation and complexity encountered in the 17 participating teacher education systems. This organizational complexity proved to be more challenging than that encountered in comparative studies of K-12 education within individual countries. Awareness of this complexity led to an understanding that country-by-country comparisons, as done in most international and comparative studies, could be carried out only after efforts to ensure that similar types of teacher education programs were being compared. We discuss these efforts below.

Variation in and Across Countries

The TEDS-M study team did not select countries for participation in the study; rather, countries throughout the world were invited to participate in TEDS-M. The 17 countries that agreed to participate in the study differed with respect to many important geographic, demographic, economic, and educational characteristics. The TEDS-M sample includes very large countries such as the United States of America (U.S.) and the Russian Federation, as well as small countries such as Singapore. These countries vary greatly in financial resources, as measured by per capita income, and in the aggregate size of their economies. In addition, a few have high fertility rates, which lead to rapidly increasing school enrollment, whereas other countries have fertility rates below replacement levels, which could lead to declining school enrollment. Most of the TEDS-M countries have a relatively favorable combination of these interacting characteristics, whereas just a few face serious funding challenges due to growing enrollments. This latter situation is, unfortunately, very widespread outside of the TEDS-M participating countries. TEDS-M is not representative of the world's countries. Instead, it comprises a relatively advantaged, but still diverse subsample from which much can be learned.

Program Variation

The countries that participated in TEDS-M vary in terms of selectivity and status of teachers, and the degree to which teaching mathematics is conceived as needing general or special mathematics preparation. These conceptions of mathematics teaching are reflected in the selectivity of teacher education programs, which is closely related to the supply of beginning teachers: a shortage of candidates who want to be teachers may result in lowering standards of admission and selectivity during and at the end of the programs (as in the United States). In contrast, an oversupply of applicants (as in Chinese Taipei), may lead to tighter admission and more stringent selectivity policy and practices.

TEDS-M provides valuable evidence of diversity in the number, size, and nature of teacher education institutions across the world. The TEDS-M study team surveyed 349 programs that prepare future teachers to teach primary pupils exclusively, 226 programs that prepare future teachers to teach secondary pupils exclusively, and 176 programs that prepare future teachers to teach primary and secondary pupils. The number of institutions that housed these teacher education programs across participating countries ranged from one institution in Singapore that had multiple programs preparing future primary and secondary teachers, to 78 in Poland. The nature of these institutions differs widely within and between countries. Some are Institutes of Higher Education (IHE) such as universities or colleges outside universities; some offer programs only in education; some are comprehensive in the fields of study offered; some offer university degrees; some of these institutions are public and some are private.

The usual way to categorize teacher education programs is according to the design of their opportunities to learn: whether they prepare teachers for primary or secondary schools. However, for TEDS-M, this turned out to be an oversimplification. The terms *primary* and *secondary* do not mean the same thing from country to country. There is no universal agreement on when primary grades end and secondary grades begin. Therefore, instead of relying on an assumed primary-secondary dividing line, TEDS-M constructed a more refined categorization based on a fine-grained analysis of the programs. To ensure that programs with similar purposes and characteristics were being compared across countries, TEDS-M used two organizational variables: grade span (the range of school grades for which teachers in that program were being prepared to teach) and teacher specialization (whether the program was preparing specialist mathematics teachers or generalist teachers). Programs were classified into program-types within countries based on the grade spans for which they prepared teachers, and according to whether they prepared *generalist* teachers or *specialist* teachers of mathematics.

Variation in Opportunities to Learn in Teacher Education Programs

One reason for our effort to classify programs in terms of grade span and specialization is that the resulting groups are likely to have different opportunities to learn (OTL), and the OTL in turn are likely to lead to different knowledge results. TEDS-M found OTL for mathematics, mathematics pedagogy, and general pedagogy depended on the grade level and the curriculum future teachers were expected to teach. For example, programs for future primary teachers gave more coverage to the basic concepts of numbers, measurement, and geometry and less coverage to functions, probability and statistics, calculus, and structure than did programs for lower secondary teachers.

Analogous patterns were also observed among secondary-level teachers. Programs that were intended to prepare teachers to teach higher grades tended to provide, on average, more OTL mathematics than the programs that prepared teachers for the early grades. The findings of this study thus reflect what seems in some countries to be a cultural norm—namely, that teachers who are expected to teach in primary, and especially early primary grades, do not need much mathematics content beyond that included in the primary and secondary school curricula. The pattern among future secondary teachers is generally characterized by more and deeper coverage of mathematics content; however, there was more variability in OTL among those being prepared for the early secondary grades (known in some countries as "middle school") than among those being prepared to teach Grade 11 and above.

Not surprisingly, the countries with programs that provided the most specialized opportunities to learn challenging mathematics had higher scores in the TEDS-M knowledge assessments. In TEDS-M, future primary-level and secondary-level specialists were found in high-achieving countries such as Chinese Taipei, Singapore, and the Russian Federation; these teachers had significantly more OTL university- and school-level mathematics than primary and secondary teachers in others countries. Opportunities to learn more and deeper mathematics seemed to be related to cultural notions of the knowledge needed to teach mathematics in primary and secondary schools. Yet the question of how much content knowledge teachers need to teach effectively is still an issue of much debate.

TEDS-M offers an opportunity to examine how these distinct assumptions play out in practice. If relatively little content knowledge is needed for the early grades, then less emphasis on mathematics preparation and non-specialization can be justified. The key question is whether teachers prepared in this fashion can teach mathematics as effectively as teachers with more extensive and deeper knowledge, such as that more often possessed by specialist teachers. Although TEDS-M does not provide definitive conclusions in this regard (this question requires the study of beginning teachers and their impact on pupils), it is important to confirm that TEDS-M future teachers who will be mathematics specialists in primary schools have higher knowledge scores on average than their generalist counterparts in the same countries.

Variation Among Teacher Educators

To complement its emphasis on the nature and extent of mathematics content and pedagogy offered to future teachers, TEDS-M surveys included questions for teacher educators about themselves, their students, and their programs. Demographic data on teacher educators at the level collected by TEDS-M fills a gap in the literature and is an important contribution of the study. The TEDS-M data on teacher educators provides insight into the variability of teacher educators across the countries studied in a number of other areas. Among the close to 5000 teacher educators surveyed for TEDS-M, the percentage with doctoral degrees in mathematics ranged from 7% in the Philippines to over 60% in Georgia, Chinese Taipei, Poland, and Oman; the percentage with doctoral degrees in mathematics pedagogy ranged from about 7% in the Philippines to 40% in Georgia. Among these teacher educators, the percentage who reported having experience teaching primary or secondary school ranged from about 20% in Oman to 90% in Georgia. All the teacher educators were asked if they considered themselves mathematics specialists. Their responses varied according to whether the respondent was a mathematician teaching mathematics content to future teachers, a mathematics educator teaching mathematics pedagogy, or a teacher educator teaching general pedagogy. Nevertheless, a surprising number among those teaching mathematics content or mathematics pedagogy described themselves as not being specialists: close to 40% in Chile and the Russian Federation, and close to 50% in Chinese Taipei, Malaysia, and the Philippines. In contrast, close to 90% of those educators in Germany, and Oman declared mathematics as their "main specialty," whereas those in Botswana, Georgia, Poland, Singapore, Switzerland, and Thailand ranged from 70% (in Thailand) to 85% (in Georgia).

Variation Among Future Teachers

As with programs and teacher educators, TEDS-M provided important information on the variability in teachers' demographic characteristics within and across countries. Future teachers being prepared to teach at the primary and secondary school levels in the TEDS-M samples were predominantly female, although there were more males at the higher levels and in particular countries. Most of the future teachers that participated in TEDS-M come from well-resourced homes, leaving low-income families underrepresented in every country in one of the largest occupations that has also historically offered an accessible avenue of social mobility. Many reported having access to such possessions as calculators, dictionaries, and DVD players, but not personal computers—now widely considered essential for professional use—especially teachers in less affluent countries such as Georgia, the Philippines, Botswana, and Thailand. A relatively small proportion of the sample of future teachers who answered the test did not speak the official language of their country (which was used in the TEDS-M surveys and tests) at home, indicating that linguistic minorities may be underrepresented in some countries.

In other respects, the self-reports of future teachers were encouraging. Most future teachers described themselves as above average or near the top of their year in academic achievement at the end of upper secondary school. Among the reasons given by future teachers for wishing to become teachers, liking to work with young people and wanting to influence the next generation were particularly important. Many believed that although teaching is a challenging job, they had an aptitude for it.

Variation in the Outcomes of Teacher Education Programs

Whereas diverse approaches are embodied in each of the programs studied in TEDS-M, it could be argued that they represent variations in the search for the optimal balance among plausible OTL the knowledge needed in mathematics teaching (Ball & Bass, 2000; Shulman, 1987). As suggested in initial reports, there is important variation within and across countries in the outcomes measures used by TEDS-M, namely in the assessments of Mathematics and Mathematics Pedagogy Content Knowledge. We summarize these briefly below.

Mathematics and Mathematics Pedagogy Content Knowledge

There is a clear and unmistakable finding regarding the TEDS-M research question about the knowledge attained by future primary and secondary teachers at the endpoint of teacher education: knowledge for teaching mathematics varies considerably among individuals within every country and between countries. The difference in mean mathematics content knowledge (MCK) scores between the highest- and lowest-achieving country in each primary and secondary program group was between 100 and 200 points—one and two standard deviations. This is a substantial difference, comparable to the difference between the 50th to the 96th percentile in the whole group. Differences in mean achievement between countries in the same program group on mathematics pedagogical content knowledge (MPCK) were somewhat smaller, ranging from about 100 to 150 points. So, within each program group, at the end of their teacher preparation programs, future teachers in some countries have substantially greater MCK and MPCK than others.

For each participating country, the results of TEDS-M serve as a baseline for further investigation. For example, content experts may look at the descriptions of the kinds of mathematics and mathematics pedagogy knowledge attained in each program or country and study how changes in OTL may correlate with improved performance. Policymakers may want to investigate ways to encourage more talented secondary school graduates to select teaching as a career, or investigate how teacher preparation programs of the same duration can lead to higher scores on MCK and MPCK. One conclusion that can be drawn from TEDS-M is that goals for improving MCK and MPCK among future teachers should be both ambitious and achievable.

Beliefs

Teachers' actions in the classroom are guided by their beliefs about the nature of teaching and learning, and about the subjects and students they teach. Acknowledging this, the TEDS-M study team gathered data on beliefs from future teachers of mathematics and from the educators charged with the responsibility of preparing them to be teachers. The survey included measures of beliefs about the nature of mathematics (e.g., Mathematics is a set of rules and procedures, Mathematics is a process of inquiry), beliefs about learning mathematics (e.g., by following teacher direction or through student activity), and beliefs about mathematics achievement (e.g., mathematics as a fixed ability). The belief that mathematics is a set of rules and procedures and that it is best learned by following teacher direction have been characterized in the literature as *calculational* and *direct-transmission* (Philipp, 2007; Staub & Stern, 2002). The belief that mathematics is a process of inquiry and that it is best learned by active student involvement is consistent with those described in the same literature as *conceptual* and *cognitive-constructionist*.

Data on beliefs from three groups (future primary teachers, future secondary teachers, and teacher educators) were compared, and, in contrast with the knowledge scales, the differences of substance were not among program groups, but rather among countries. Consequently, the analysis was based on comparisons by country in a way that was not feasible with the knowledge scales. In general, the pattern of beliefs described as a *conceptual* or *cognitive-constructionist* orientation is endorsed by teacher educators and future teachers in all countries, although somewhat more weakly in Georgia. The pattern of beliefs described as *computational* or *direct-transmission* was endorsed by teacher educators and future teachers in Botswana, Georgia, Malaysia, Oman, the Philippines, and Thailand, but not by teacher educators and future teachers in Germany, Norway, and Switzerland. Patterns of responses from several countries (Chile, Chinese Taipei, Poland, the Russian Federation, Singapore, and Spain) were generally consistent with the conceptual orientation, and emphasized the belief that mathematics cannot only be learned by memorizing a series of rules and procedures (*Mathematics as a Set of Rules and Procedures*). The view of *Mathematics as a Fixed Ability* carries with it the implication that mathematics is not for all, that some children cannot and will not succeed in mathematics. This view may have implications for how children are grouped and how they are taught. It is a minority view in all countries surveyed, but still a matter of concern in that it stands in opposition to the apparent international consensus on the need for all children to learn mathematics at a higher level than has generally been the case. This opposition view was supported by future teachers and teacher educators in Botswana, Thailand, Georgia, Malaysia and the Philippines, and rejected in Germany, Switzerland, the United States, and Norway.

There are substantial between-country differences in the extent to which beliefs are held in association with other tendencies. For instance, the program groups within countries endorsing beliefs consistent with a computational orientation are generally among those with lower mean scores on the knowledge tests. However, it would be unwise to generalize from this, for two reasons. First, the sample of countries is quite small. Second, the countries differ greatly from one another both culturally and historically, in ways that may influence both beliefs and knowledge in unknown ways. In some countries scoring high on the MCK and MPCK tests, future teachers endorsed both belief in mathematics as a set of rules and procedures and as a process of inquiry. The TEDS-M findings show that both conceptions, *computational* and *constructivist*, are endorsed in mathematics teacher education, and what is at issue is the appropriate use and balance of each.

Variations in Context and Policy

TEDS-M has shown teachers' careers and working conditions range from those where teachers are carefully selected, well-compensated, and highly regarded to those where there is less selectivity, low salaries, and low status. These careers and conditions are shaped by the two major systems of teacher employment (career-based and position-based) found in the world's public schools, together with various mixed or hybrid models.

Career-based refers to systems where teachers are recruited at a relatively young age to remain in one coherent, clearly organized, public or civil service system throughout their working lives. Teacher education is facilitated by the predictability and stability of careers in these systems. Promotion follows a well-defined path of seniority and other requirements, and teaching assignments follow bureaucratic deployment principles and procedures. Countries able to afford career-based staffing can generally avoid major teacher supply problems and have an advantage in recruiting higher-ability applicants.

Position-based systems take a very different approach to teacher employment. Teachers are not hired into the national civil service or a separate national teacher service. Rather, they are hired into specific teaching positions within an unpredictable career-long progression of assignments. As a result, access is more readily open to applicants of diverse ages and atypical career backgrounds. Movement in and out of teaching to raise children or pursue other opportunities is possible. In these systems, it may be difficult to recruit and retain sufficient numbers of teachers, especially in areas like science and mathematics, where there are attractive opportunities in other occupations.

In short, this distinction between career- and position-based systems has a major impact on teacher education. Since appointment in a career-based system involves a commitment to lifelong employment, such systems are more justified in investing in initial teacher preparation, knowing that the educational system will likely realize the return on this investment throughout the teacher's working life. Often this commitment is made even before the beginner receives any teacher training. In contrast, in position-based systems, such an investment in initial preparation is less justifiable, since the system is based on the assumption that individuals may move in and out of teaching on a relatively short-term basis, and often the graduates of teacher education in such a system never occupy any teaching position at all.

One long-term policy that has increasingly influenced teacher education in a large number of countries worldwide, including those participating in TEDS-M, is to require teachers to have university degrees. Obtaining an all-graduate teaching force, all of whom have higher education degrees (not just diplomas) has been one of the main goals of teacher education policy in many countries over the years and has affected teacher recruitment and the subsequent experience of these teachers once they are employed.

The TEDS-M study team also sought to examine the range of policies affecting teacher education programs, especially those related to accountability concerns, finding great variation in approaches, including the existence of criteria to insure the quality of entrants to teacher education programs, criteria to assess the quality of graduates before they can gain entry to the teaching profession, and accreditation reviews to insure programs' accreditation.

Overall, TEDS-M researchers have found a positive association between the strength of accountability strategies and arrangements and country mean scores in the TEDS-M tests of MCK and MPCK; countries with strong arrangements, such as Chinese Taipei and Singapore, scored highest on these measures. Countries with weaker arrangements, such as Georgia and Chile, tended to score lower on the two measures of future teacher knowledge.

These findings have implications for policymakers concerned with promoting teacher quality. Policies can be designed to cover the full spectrum, from policies designed to make teaching an attractive career to policies for assuring that entrants to the profession have attained high standards of performance. TEDS-M researchers point to the importance of ensuring that policies designed to promote teacher quality are coordinated and mutually supportive. Specifically, TEDS-M provides evidence that countries such as Chinese Taipei and Singapore, that do well on international tests of student achievement such as TIMSS, employ a full range of strategies. They not only ensure high quality of entrants to teacher education, but also have strong systems for reviewing, assessing, and accrediting teacher education providers. They also have strong mechanisms for ensuring that graduates meet high standards of performance before gaining certification and full entry to the profession.

Reform that recognizes these findings is critical. The TEDS-M study team found that all participating teacher education systems were implementing reforms in teacher education, attempting to change their education systems in order to increase the mathematics achievement levels of their students. In the European countries in TEDS-M, changes to entire university systems are underway as a result of the Bologna accord for the creation of a European Higher Education Area. In other countries, such as Malaysia, changes in teacher education toward more advanced levels of education for teachers were precipitated by concerns about the limitations and weaknesses of current mathematics, science, and technology education. Although reform is virtually ubiquitous in the TEDS-M countries, it is important to keep in mind that, as in any cross-sectional study, TEDS-M provides only a snapshot of mathematics teacher preparation in the year 2008–2009, when the data were collected.

TEDS-M's Contribution to the Study of Mathematics Teacher Education

TEDS-M is not only the first large scale comparative international study of teacher education outcomes with representative samples, but in higher education as a whole. Moreover, the surveys were completed with high response rates and coverage of the target populations, in most cases meeting the very high IEA standards for sampling and response rates. In the instances where the IEA standards were not met, the response rates still compared favorably with general experience in higher education surveys, especially surveys in which the targeted participants are all volunteers.

TEDS-M thus lays the foundation for future rigorous national and cross-national research in teacher education, making available a common terminology, sampling methods tailored to teacher education, instruments, and analyses that can be adapted and improved for use in subsequent teacher-education studies, whether they be in mathematics or in other areas. TEDS-M has also served to develop strong research capability within the countries that participated in this study. Finally, the TEDS-M database has continued to contribute to this new line of research by enabling secondary analyses by researchers around the world.

References

Ball, D., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on teaching and learning mathematics* (pp. 83–104). Westport, CT: Ablex.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 416–440.

Campbell, P. F., Nishio, M., Smith, T. M., Clark, L. M., Conant, D. L., Rust, A. H., … Choi, Y. (2014). The relationship between teachers' mathematical content and pedagogical knowledge, teachers' perceptions, and student achievement. *Journal for Research in Mathematics Education, 45*(4), 419–459.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*, 499–531.

Goldhaber, D., Liddle, S., & Theobald, R. (2013). The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review, 34*, 29–44.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Koedel, C., Parsons, E., Podgursky, M., & Ehlert, M. (2015). Teacher preparation programs and teacher quality: Are there real differences across programs? *Education Finance and Policy, 10*(4), 508–534.

Metzler, J., & Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation. *Journal of Development Economics, 99*(2), 486–496.

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). Charlotte, NC: National Council of Teachers of Mathematics & Information Age Publishing.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1.

Staub, F. C., & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology, 94*(2), 344–355.

Tatto, M. T. (Ed.). (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report*. Amsterdam, The Netherlands: IEA.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework. Teacher Education and Development International Study Center, Michigan State University, East Lansing, MI, and IEA.

Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education, 11*(2), 139–164.

Part I
Exploring Different Dimensions of Teacher Education Programs

Chapter 2
Introduction: Exploring Different Dimensions of Teacher Education Programs in the TEDS-M Study

Maria Teresa Tatto and Wendy M. Smith

Abstract This chapter provides an introduction to Part I of the book, which focuses on examining the characteristics of teacher education programs that participated in the Teacher Education and Development Study in Mathematics (TEDS-M), a cross-national study of teacher education programs that prepare future primary and secondary mathematics teachers. The study collected institutional data from 751 programs in 17 countries. In this introduction to Part I, we provide an overview of the characteristics of the teacher education programs studied and of the teacher educators within them, the methods of TEDS-M data collection, and the challenges encountered when collecting program-level data. The overview is followed by an introduction to the chapters in Part I, which explore different aspects of the participating programs. The first chapter looks at the types of strategies used in the pursuit of program quality and their relationship with programs' outcomes. The following

The text in this introduction contains shortened and slightly edited versions of text that has appeared in the following publications: Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M., & Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, the Netherlands: International Association for the Evaluation of Student Achievement. Brese, F., & Tatto, M. T. (Eds.) (2012). *User guide for the TEDS-M international database*. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA). Tatto, M. T. (2013). The Teacher Education and Development Study in Mathematics (TEDS-M). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: technical report*. Amsterdam: International Association for the Evaluation of Student Achievement. Text cited directly or indirectly from those sources will not be made recognizable. An extensive report on the descriptive findings on the characteristics of teacher education programs in the study can be found in Tatto et al., (2012). We summarize the key findings here and reproduce some figures and tables to orient the reader; the chapters in Part I however are original contributions written exclusively for this book.

M. T. Tatto (✉)
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: teresa.tatto@asu.edu

W. M. Smith
University of Nebraska, Lincoln, NE, USA
e-mail: wsmith5@unl.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_2

chapter explores the characteristics of the curriculum of mathematics teacher education in programs preparing primary and secondary teachers in the U.S. The penultimate chapter looks at the extent to which teacher education programs are designed to educate diverse populations, and what opportunities are provided to future teachers to learn how to effectively teach diverse groups of students. The last chapter of Part I investigates the views of teacher educators about mathematics teaching and how these are reflected in the views of their future teachers.

Characteristics of Teacher Education Programs

In the first part of the book, we focus on examining the characteristics of teacher education programs. Within TEDS-M, a key priority was to understand the characteristics of the teacher education programs where future teachers are prepared. In the first four chapters we explore different aspects of the participating programs. Chapter 3 authors look at what types of strategies are used in the pursuit of program quality and their relationship with programs' outcomes. Chapter 4 authors explore the characteristics of the curriculum of mathematics teacher education in programs preparing primary and secondary teachers in the U.S. The Chapter 5 author looks at the extent to which teacher programs are designed to educate diverse populations, and what opportunities are provided to future teachers to learn how to effectively teach diverse groups of students. The last chapter of Part I, Chap. 6, investigates the views of teacher educators about mathematics teaching and learning and how these compare to the views of their student teachers.

In this introduction to Part I, we provide a brief overview of the teacher education programs and the teacher educators within them, the methods of data collection, and the challenges we encountered when collecting program-level data.

Data on Teacher Education Programs

Table 2.1 lists the program-types included in the TEDS-M target population and shows how they differ within and between countries. Although the names of the program-types vary from country to country, the characteristics and purpose of program-types in different countries are often similar: They aim to adequately prepare future teachers to teach the official school curriculum (Tatto et al., 2012, pp. 28–32). Whether and how this is accomplished is the subject of the different chapters in this section.

To understand programs' characteristics, institutional data were collected via the Institutional Program Questionnaire (IPQ) from 751 programs in the 17 TEDS-M countries, including 349 programs that prepare future teachers to teach exclusively at the primary school level, 226 programs that prepare future teachers to teach at the lower secondary school level, and 176 programs that prepare future teachers to

Table 2.1 Organizational characteristics of teacher education program-types in the TEDS-M study

Country	Program-type	Consecutive/ Concurrent	Duration (years)	Grade span	Specialization	Program group	Test administered
Botswana	Diploma in Primary Education	Concurrent	3	1–7	Generalist	3: Primary-lower secondary (grade 10 max)	Primary
	Diploma in Secondary Education, Colleges of Education	Concurrent	3	8–10	Specialist	5: Lower secondary (grade 10 max)	Secondary
	Bachelor of Secondary Education (Science), University of Botswana	Concurrent	4	8–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
Canada Ontario	Primary/Junior	Consecutive	4 + 1	1–6	Generalist	2: Primary (grade 6 max)	NA
Ontario	Junior/Intermediate	Consecutive	4 + 1	4–10	Generalist & Specialist	Both 3 (Primary-lower secondary – grade 10 max) & 5 (Lower secondary – grade 10 max)	NA
Ontario	Intermediate/Senior	Consecutive	4 + 1	7–12	Specialist (in 2 subjects)	6: Upper secondary (up to grade 11 and above)	NA
Quebec	Primary	Concurrent	4	1–6	Generalist	2: Primary (grade 6 max)	NA
Quebec	Secondary	Concurrent	4	7–11	Specialist	6: Upper secondary (up to grade 11 and above)	NA
Nova Scotia	Primary	Consecutive	4 + 2	1–6	Generalist	2: Primary (grade 6 max)	NA
Nova Scotia	Secondary (Junior & Senior)	Consecutive	4 + 2	7–12	Specialist	6: Upper secondary (up to grade 11 and above)	NA
Newfoundland-Labrador	Primary/Elementary	Concurrent	5	1–6	Generalist	2: Primary (grade 6 max)	NA

(continued)

Table 2.1 (continued)

Country	Program-type	Consecutive/ Concurrent	Duration (years)	Grade span	Specialization	Program group	Test administered
Newfoundland-Labrador	Intermediate/Secondary	Consecutive	4 + 1	7–12	Specialist	6: Upper secondary (up to grade 11 and above)	NA
Chile	Generalist	Concurrent	4	1–8	Generalist	BOTH 3 (Primary-lower secondary – grade 10 max) & 5 (Lower secondary – grade 10 max)	Both
	Generalist with further mathematics education	Concurrent	4	5–8	Generalist	5: Lower secondary (grade 10 max)	Secondary
Chinese Taipei	Elementary Teacher Education	Concurrent	4.5	1–6	Generalist	2: Primary (grade 6 max)	Primary
	Secondary Mathematics Teacher Education	Concurrent	4.5	7–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
Georgia	Bachelor of Pedagogy	Concurrent	4	1–4	Generalist	1: Lower primary (grade 4 max)	Primary
	BA Mathematics	Concurrent	3	5–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
	MS Mathematics	Concurrent	5	5–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
	MS Mathematics	Concurrent	5	5–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary

Germany	Teachers for Grades 1–4 with Mathematics as Teaching Subject (Type 1a)	Hybrid of the two	3.5 + 2.0	1–4	Generalist	1: Lower primary (grade 4 max)	Primary
	Teachers for Grades 1–4 without Mathematics as Teaching Subject (Type 1b)	Hybrid of the two	3.5 + 2.0	1–4	Generalist	1: Lower primary (grade 4 max)	Primary
	Teachers of Grades 1–9/10 with Mathematics as Teaching Subject (Type 2a)	Hybrid of the two	3.5 + 2.0	1–9/10	Specialist (in 2 subjects)	BOTH 4 (primary mathematics specialist) & 5 (lower secondary – grade 10 max)	Both
	Teachers for Grades 1–10 without Mathematics as Teaching Subject (Type 2b)	Hybrid of the two	3.5 + 2.0	1–4	Generalist	1: Lower primary (grade 4 max)	Primary
	Teachers for Grades 5/7–9/10 with Mathematics as Teaching Subject (Type 3)	Hybrid of the two	3.5 + 2.0	5/7–9/10	Specialist (in 2 subjects)	5: Lower secondary (grade 10 max)	Secondary
	Teachers for Grades 5/7–12/13 with Mathematics as a Teaching Subject (Type 4)	Hybrid of the two	4.5 + 2.0	5/7–12/13	Specialist (in 2 subjects)	6: Upper secondary (up to grade 11 and above)	Secondary
Malaysia	Bachelor of Education (Primary)	Concurrent	4	1–6	Specialist (in 2 subjects)	4: Primary mathematics specialist	Primary
	Diploma of Education (Mathematics)	Consecutive	4 + 1	1–6	Specialist (in 2 subjects)	4: Primary mathematics specialist	Primary
	Malaysian Teaching Diploma (Mathematics)	Concurrent	3	1–6	Specialist (in 2 subjects)	4: Primary mathematics specialist	Primary
	B.Ed (Mathematics) Secondary	Concurrent	4	7–13	Specialist (in 2 subjects)	6: Upper secondary (up to grade 11 and above)	Secondary
	B.Sc.Ed (Mathematics) Secondary	Concurrent	4	7–13	Specialist (in 2 subjects)	6: Upper secondary (up to grade 11 and above)	Secondary

(continued)

Table 2.1 (continued)

Country	Program-type	Consecutive/ Concurrent	Duration (years)	Grade span	Specialization	Program group	Test administered
Norway	ALU with Mathematics Option	Concurrent	4	1–10	Generalist with extra math	BOTH 3 (Primary-lower secondary – grade 10 max.) & 5 (Lower secondary – grade 10 max)	Both
	ALU without Mathematics Option	Concurrent	4	1–10	Generalist	BOTH 3 (Primary-lower secondary – grade 10 max.) & 5 (Lower secondary – grade 10 max)	Both
	PPU	Consecutive	3 + 1 (or 5 + 1)	8–13	Specialist (in 2 subjects)	6: Upper secondary (up to grade 11 and above)	Secondary
	Master's (MAS)	Concurrent	5	8–13	Specialist (in 2 subjects)	6: Upper secondary (up to grade 11 and above)	Secondary
Oman	Bachelor of Education, University	Concurrent	5	5–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
	Educational Diploma after B.Sc	Consecutive	5 + 1	5–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
	Bachelor of Education, Colleges Of Education	Concurrent	4	5–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
Philippines	Bachelor in Elementary Education	Concurrent	4	1–6	Generalist	2: Primary (grade 6 max)	Primary
	Bachelor in Secondary Education	Concurrent	4	7–10	Specialist	5: Lower secondary (grade 10 max)	Secondary

Poland	B. Ped. Integrated Teaching, First cycle	Concurrent	3	1–3	Generalist	1: Lower primary (grade 4 max)	Primary
	M.A. Integrated Teaching, Long Cycle	Concurrent	5	1–3	Generalist	1: Lower primary (grade 4 max)	Primary
	Mathematics BA First Cycle	Concurrent	3	4–9	Specialist	BOTH 4 (Primary mathematics specialist) & 5 (Lower secondary – grade 10 max)	Both
	Mathematics MA Long Cycle	Concurrent	5	4–12	Specialist	BOTH 4 (Primary mathematics specialist) & 6 (Upper secondary – up to grade 11 and above)	Both
Russian Federation	Primary Teacher Education	Concurrent	5	1–4	Generalist	1: Lower primary (grade 4 max)	Primary
	Teacher of Mathematics	Concurrent	5	5–11	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
Singapore	PGDE, Primary Option C	Consecutive	4 + 1	1–6	Generalist	2: Primary (grade 6 max)	Primary
	BA (Ed) (Pri)	Concurrent	4	1–6	Generalist	2: Primary (grade 6 max)	Primary
	BSc (Ed) (Pri)	Concurrent	4	1–6	Generalist	2: Primary (grade 6 max)	Primary
	Dip Ed, Primary Option A	Concurrent	2	1–6	Specialist (in 2 subjects)	4: Primary mathematics specialist	Primary
	Dip Ed, Primary Option C	Concurrent	2	1–6	Generalist	2: Primary (grade 6 max)	Primary
	PGDE, Primary Option A	Consecutive	4 + 1	1–6	Specialist	4: Primary mathematics specialist	Primary
	PGDE, Lower Secondary	Consecutive	4 + 1	7–8	Specialist (in 2 subjects)	5: Lower secondary (grade 10 max)	Secondary
	PGDE, Secondary	Consecutive	4 + 1	7–12	Specialist (in 2 subjects)	6: Upper secondary (up to grade 11 and above)	Secondary

(continued)

Table 2.1 (continued)

Country	Program-type	Consecutive/ Concurrent	Duration (years)	Grade span	Specialization	Program group	Test administered
Spain	Teacher of Primary Education	Concurrent	3	1–6	Generalist	2: Primary (grade 6 max)	Primary
Switzerland	Teachers for Grades 1–2/3	Concurrent	3	1–2/3	Generalist	1: Lower primary (grade 4 max)	Primary
	Teachers for Primary School (Grades 1–6)	Concurrent	3	1–6	Generalist	2: Primary (grade 6 max)	Primary
	Teachers for Primary School (Grades 3–6)	Concurrent	3	3–6	Generalist	2: Primary (grade 6 max)	Primary
	Teachers for Secondary School (Grades 7–9)	Concurrent	4.5	7–9	Generalist, some specialization	5: Lower secondary (grade 10 max)	Secondary
Thailand	Bachelor of Education	Concurrent	5	1–12	Specialist	BOTH 4 (Primary mathematics specialist) & 6 (Upper secondary – up to grade 11 and above)	Both
	Graduate Diploma In Teaching Profession	Consecutive	4 + 1	1–12	Specialist	BOTH 4 (Primary mathematics specialist) & 6 (Upper secondary – up to grade 11 and above)	Both

USA	Primary Concurrent	Concurrent	4	1–3/4/5	Generalist	2: Primary (grade 6 max)	Primary
	Primary Consecutive	Consecutive	4 + 1	1–3/4/5	Generalist	2: Primary (grade 6 max)	Primary
	Primary + Secondary Concurrent	Concurrent	4	4/5–8/9	Specialist	BOTH 4 (Primary mathematics specialist) & 5 (Lower secondary – grade 10 max)	Both
	Primary + Secondary Consecutive	Consecutive	4 + 1	4/5–8/9	Specialist	BOTH 4 (Primary mathematics specialist) & 5 (Lower secondary – grade 10 max)	Both
	Secondary Concurrent	Concurrent	4	6/7–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary
	Secondary Consecutive	Consecutive	4 + 1	6/7–12	Specialist	6: Upper secondary (up to grade 11 and above)	Secondary

Source: Tatto et al. (2012), pp. 29–32 teach at either the primary or the lower-secondary levels (Brese & Tatto, 2012; Tatto, 2013).

Before the data were collected, much collaboration went into developing workable definitions across the countries in the study. For instance, a teacher education institution was defined as a "secondary or post-secondary school, college, or university that offered a program or programs focusing in teacher preparation on a regular and frequent basis" (Tatto et al., 2012, p. 95). A program was defined as a "specific pathway within an institution that required students to undertake a set of courses and experiences that led to the award of a teaching credential or degree upon successful completion" (Tatto et al., 2012, p. 95).

After creating and arriving at agreements around definitions, the next task was to create the questionnaire. The IPQ included multiple sections to collect data about teacher preparation programs.[1] First, the questionnaire included questions about the organization and structure of the program(s) at the institution: How many different program types; concurrent vs. consecutive program(s); duration of program(s); grade level span(s) included; subject-matter specialization; number of future teachers in different program types; duration and nature of field experiences (both introductory and extended); and locus of control for teacher preparation programs. In most countries, guidelines and standards for teacher preparation were set at the state, provincial, or national level.

Program types were classified as either *concurrent* or *consecutive.* Concurrent program-types included studies in subject-matter content, pedagogy, and other courses in education, which were completed within the first phase of post-secondary education and resulted in one credential (e.g., bachelor of arts). Consecutive teacher education program-types consisted of two phases of post-secondary education: an initial university degree in a first phase followed by a second phase focused on pedagogy and the field experience, which result in a second credential (Tatto et al., 2012, p. 33). A large number of programs in the TEDS-M sample were concurrent programs.

Another important distinction in characterizing program types is the difference between preparing future teachers as generalists and as mathematics specialists. The data show that most future teachers planning to work in primary schools are prepared as generalists, although there is some variation in this across countries. In some countries, generalist teachers are expected to teach both primary and lower-secondary grades. In contrast, most future teachers of mathematics at the upper-secondary level are prepared as mathematics specialists and are expected to teach up to Grade 11 and above; some are expected to teach the earlier grades as well, up to Grade 10. The three most common program groupings in the participating counties were primary generalist (for teaching up to Grade 6 maximum), lower-secondary specialist (for teaching up to Grade 10 maximum), and secondary specialists (for teaching up to Grade 11 and above) (Tatto et al., 2012, p. 96).

The IPQ also included questions about program entry requirements and the possible outcomes of these requirements: to what extent future teachers' prior

[1] See Brese and Tatto (2012) for more details about the IPQ.

achievement in mathematics is a selection criterion; and how future teachers' primary and secondary achievement compared with achievement of peers in other college-level programs. Further, the TEDS-M researchers sought to capture the curriculum content of teacher preparation programs by asking about the number of hours allocated to field experiences and to required courses in areas such as the liberal arts, academic and school mathematics, mathematics pedagogy and pedagogy, and foundations.

Finally, programs reported quality assurance information. Entry requirements varied widely, with several countries requiring degrees in mathematics or acceptable scores on examinations as a condition for entry, and others countries allowing entry without a set mathematics standard. All programs participating in the study reported that student teachers were required to have passing grades and to have demonstrated competence in all courses and in their field experience to graduate. A comprehensive examination was another common requirement across institutions. The completion of a thesis was not a common requirement at the primary level, but was common at the secondary level (Tatto et al., 2012, p. 109).

Data on Teacher Educators

Although it could be argued that the teacher educators are the primary shapers of teacher education programs and are in a sense "street level bureaucrats" (Weatherley & Lipsky, 1977), research on teacher educators is rare. Recognizing the influence that teacher educators have over program design and outcomes, the TEDS-M study team decided to develop instruments to measure some dimensions of teacher educator backgrounds, and beliefs, and the kinds of opportunities to learn they provide to future teachers. This part of the study was primarily designed following Tatto's earlier work (1996, 1998, 1999).

Though limited, there are other examples of empirical studies of teacher educators, often discussing teacher educators' roles in modelling effective practices for future teachers (Lampert et al., 2013; Lunenberg, Korthagen, & Swennen, 2007). However, there is a dearth of research on the characteristics of teacher educators and their impact on future teachers, which demands further research.

For the TEDS-M study, teacher educators were defined as "persons with regular, repeated responsibility for teaching future teachers within a teacher-preparation program" (Tatto et al., 2012, p. 111). Teacher educators were classified into three groups: mathematics and mathematics pedagogy educators; general pedagogy educators; and educators belonging to both groups. Lack of universal agreement as to what constituted these different types of educators made it necessary for the research team to create common definitions (see Tatto et al., 2012, p. 84):

Educators of mathematics and mathematics pedagogy: Persons responsible for teaching one or more of the program's required courses in mathematics or mathematics pedagogy during the study's data-collection year at any stage of the institution's teacher preparation program.

General pedagogy educators: Persons responsible for teaching one or more of the program's required courses in foundations or general pedagogy (other than a mathematics or mathematics pedagogy course) during the study's data-collection year at any stage of the institution's teacher preparation program.

Educators belonging to both Groups 1 and 2 as described above: Persons responsible for teaching one or more of the program's required courses in mathematics and/or mathematics pedagogy and/or general pedagogy during the study's data-collection year at any stage of the institution's teacher preparation program.

We considered the teacher educators to be key "individuals through whom the intended teacher education curriculum becomes the implemented curriculum" (Tatto, 2013, p. 59). For that reason, the TEDS-M researchers developed a number of survey items to better understand educators' general background; their teaching, professional, and research experiences; and the opportunities to learn that they provide future teachers in their courses. The survey also included questions parallel to those asked of future teachers about their beliefs about teaching and learning mathematics, so that the responses of teacher educators and future teachers could be compared (see Brese & Tatto, 2012; Tatto et al., 2012, pp. 111–116).

The extent of teacher educators' backgrounds in mathematics, mathematics education, and education generally varied considerably across countries. In some countries, most teacher educators in mathematics content and pedagogy courses had doctoral degrees in mathematics, whereas in other countries this was rare. Common across all countries was the finding that few educators held doctorates in mathematics education. In the United States, a doctorate of some kind was customary, whereas in others a master's degree was far more common for teacher educators.

Overall, mathematics and mathematics pedagogy teacher educators considered mathematics to be their area of expertise, whereas teacher educators who were teaching general pedagogy reported mathematics not being their area of expertise (Tatto et al., 2012, p. 115).

Finally, teacher educators were asked whether they currently held or ever had held a certification or license to teach in K-12. Responses varied, with over 80% of teacher educators in some countries having teaching licenses, and less than 30% in other countries (Tatto et al., 2012, p. 116).

Chapters in Part I

Chapter 3, "Preparing High Quality Mathematics Primary Teachers: Exploring Program Strategies and Standards in the United States, Russia, Poland, and Chinese Taipei" by Peralta and Tatto, presents an investigation of the association between program outcomes, such as mathematics content knowledge and mathematics pedagogy knowledge, and program quality strategies as reported by program administrators. The authors use factor analyses to create indices of quality strategies and program standards. Linear regression models were estimated for four countries, including the United States, Russia, Poland, and Chinese Taipei. Findings show that program selectivity is the most important strategy related to program outcomes in the Russian and U.S. programs. In addition to entry and exit criteria regulations, standards regulating opportunities to learn while in the program and during field experiences were related to program outcomes in Chinese Taipei and Poland.

Chapter 4, "The Intended, Implemented, and Achieved Curriculum of Mathematics Teacher Education in the United States," by Tatto and Bankov, uses the United States as a case study. After describing the development of a systematic model for the comparative study of the curriculum of teacher education programs, the authors present the results of the curriculum analysis for programs preparing future primary and secondary teachers and the curriculum's association with future teachers' perception of their opportunities to learn and their knowledge at graduation in the United States. This study has important implications for the future study of the teacher education curriculum in an era of increased accountability.

Chapter 5, "Developing Diverse Teachers: Analyzing Primary Mathematics Teacher Education Programs Prioritizing Selection of Diverse Future Teachers," by Pippin, looks at issues of diversity related to teacher preparation programs. Education researchers indicate that diverse teachers can improve the achievement and school experiences of marginalized students, and that, in many countries, policy has been created to recruit and prepare more diverse teachers. Research that explores programs that prepare future primary mathematics teachers from underrepresented populations in less developed or non-Western countries, however, is lacking. In this chapter, Pippin describes important variation in the selection policies and goals of these programs, as well as future teachers' reported opportunities to learn to teach diverse students in the United States and three other countries.

Chapter 6, "A Comparative International Study of Differences in Beliefs between Future Teachers and Their Educators," by Rodriguez, Tatto, Palma, and Nickodem, investigates the extent to which differences exist in beliefs about teaching and learning mathematics between future teachers and their educators for five countries, including Chinese Taipei, Poland, the Russian Federation, Singapore, and the United States. The analysis is based on the methods of meta-analysis to estimate program effects within institutions and across institutions within a country. The findings show significant differences between future teachers and their educators on a number of beliefs about teaching and learning mathematics within and across programs. Variation in discrepancies between teacher educators' and future teachers' beliefs is explained by future teachers' opportunities to learn in the areas of mathematics pedagogy, general pedagogy, and field experiences provided by their teacher preparation programs. In addition, program coherence and mathematics content knowledge and mathematics pedagogical content knowledge moderated the magnitude of differences in beliefs between future teachers and their educators. The authors discuss implications for the design of the curriculum and experiences provided by teacher education programs.

References

Brese, F., & Tatto, M. T. (Eds.) (2012). *User guide for the TEDS-M international database.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Lampert, M., Franke, M. L., Kazemi, E., Gitomer, D., Turrou, A. C., Beasley, H., … Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education, 64*(3), 226–243.

Lunenberg, M., Korthagen, F., & Swennen, A. (2007). The teacher educator as a role model. *Teaching and Teacher Education, 23*(5), 586–601. https://doi.org/10.1016/j.tate.2006.11.001

Tatto, M. T. (1996). Examining values and beliefs about teaching diverse students: Understanding the challenges for teacher education. *Educational Evaluation and Policy Analysis, 18*, 155–180.

Tatto, M. T. (1998). The influence of teacher education on teachers' beliefs about purposes of education, roles and practice. *Journal of Teacher Education, 49*, 66–77.

Tatto, M. T. (1999). The socializing influence of normative cohesive teacher education on teachers' beliefs about instructional choice. *Teachers and Teaching, 5*, 111–134.

Tatto, M. T. (Ed.). (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries*, *Technical report*. Amsterdam, The Netherlands: IEA.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).* Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Weatherley, R., & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special-education reform. *Harvard Educational Review, 47*(2), 171–197.

Chapter 3
Preparing High Quality Mathematics Primary Teachers: Exploring Program Strategies and Standards in the United States, Russia, Poland, and Chinese Taipei

Yadira Peralta and Maria Teresa Tatto

Abstract Preparing high-quality mathematics primary teachers is a priority of teacher education programs as they seek to fulfill professional standards. Programs have used a number of strategies to achieve these goals including developing criteria for program admission and graduation, and regulating future teachers' opportunities to learn during their program and field experiences as they prepare to enter the classroom. Using data from the Teacher Education and Development Study in Mathematics (TEDS-M), this chapter presents an investigation of the relationship between program outcomes as indicated by an assessment of mathematics content knowledge and mathematics pedagogy knowledge, and program standards as reported by program administrators. Factor analyses were used to create indices of program standards in key areas of emphasis. Linear regression models were estimated for four countries, including the United States, Russia, Poland, and Chinese Taipei. Our findings show that program selectivity is the most important strategy related to program outcomes in the Russian and U.S. programs. In addition to entry and exit criteria regulations, standards regulating opportunities to learn while in the program and during field experiences were related to program outcomes in Chinese Taipei and Poland.

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Y. Peralta
University of Minnesota, Minneapolis, MN, USA
e-mail: peral026@umn.edu

M. T. Tatto (✉)
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: teresa.tatto@asu.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_3

Introduction

In the past decade, policies to regulate teacher education have intensified (Tatto, 2007). Whereas accreditation systems have existed in some countries since the mid-1950s—for example, in the United States, the National Council for Accreditation of Teacher Education (NCATE) was created as the mechanism to accredit teacher education institutions at the time—accreditation was seen more as a desirable rather than a necessary condition. This did not necessarily mean that programs were not monitoring their own quality all along, but engaging in accountability activities was seen as a low priority among the heavy demands of running a teacher education program.

In 1996, the publication of results from the IEA's TIMSS (Trends in International Mathematics and Science Study) created great concern across many nations, as comparisons indicated that a large number of children were not learning as expected; such results were later reinforced by the OECD's PISA (Programme for International Student Assessment) studies. The TIMSS studies, in particular, have focused on the school curriculum as the basis for their assessment frameworks. Whether pupils have had the opportunity to learn the content assessed has been a key question in efforts to use the IEA studies to inform education policy. Opportunity to learn has been seen primarily as dependent on how much and how well pupils and their teachers are able to cover the curriculum content. While a number of explanations exist regarding curriculum coverage, ranging from pupils' background characteristics to school resources, the notion that curriculum coverage is dependent on whether teachers have the knowledge needed to teach it has particularly captured the imagination of policy analysts and reformers alike. Consequently, great emphasis is placed on teacher education as a basis for teachers having the knowledge they need to teach effectively. Yet, analysts have found great variation in the content and outcomes of teacher education (Tatto & Hordern, 2017; Tatto et al., 2012).

The notion that teachers are exposed to highly variable opportunities to learn to teach, with highly variable results in their learning, has opened the door to the development of reforms directed at regulating teacher education. The Bologna agreement in the European Union gave rise to the first large-scale reform of higher education across the participating countries and was the first systemwide attempt at regulating teacher education through external means (see Council of the European Union, 2007, p. C300/7). As the "new accountability movement" began to dominate the global education discourse (Tatto, 2007), regulations and standard setting became the strategies of choice for improving educational quality—including teacher education—since the mid-2000s.

While there is a variety of standards across the world that seek to regulate the teaching profession and, as a consequence, teacher education programs, these seem to have developed around common themes (National Research Council, 2001). Teacher education program standards are typically directed at ensuring that future teachers acquire the abilities and develop the tools that will enable them to:

- understand and address their pupils' development and learning needs,
- know their subjects and how to teach them,

- manage their classrooms to create environments conducive to learning,
- be reflective about their practice and learn from their experiences and that of their pupils to improve their practice, and
- participate and become part of professional and local learning communities.

In the United States for instance, these standards align with those set by the National Board for Professional Teaching Standards (2016), with the Interstate New Teacher Assessment and Support Consortium (INTASC),[1] and with the Council for the Accreditation of Educator Preparation (CAEP),[2] all focused on current conceptions of teacher quality. These standards also resonated with the countries that participated in the Teacher Education and Development Study in Mathematics (TEDS-M).

Research Questions

Using data from the TEDS-M Institutional Program Questionnaire (IPQ) our intention in this chapter is to answer three questions directed at exploring the strategies that programs use to make sure that future primary teachers possess the knowledge and skills envisioned by their programs:

1. Do program standards directed at regulating admission and graduation requirements in teacher education programs result in more knowledgeable teachers?
2. Do other key standards that set priorities for the knowledge that is emphasized in the professional curriculum, and for the content and delivery of field experiences, result in more knowledgeable teachers?
3. Do these standards vary across countries?

The analysis in this chapter uses the information provided by administrators of primary programs in four of the 17 countries included in the TEDS-M study: the United States, Russia, Poland, and Chinese Taipei. These four countries were chosen because they were either actively developing and/or implementing standard-based reforms close to the time of the study. In the case of Russia and Poland, we were interested in knowing how these reforms played out in these two countries with a shared history in the development of their educational systems but also a departure after the dissolution of the Soviet Union. In the case of the United States and Chinese Taipei, there is some similarity in their systems, with Chinese Taipei adapting some reforms and curriculum models developed in the United States, but with higher levels of centralization, selectivity and rigor. In addition, future teachers in these four countries, had scores above the international mean of 500 (s.d. 100)[3]

[1] http://www.ccsso.org/Documents/2011/InTASC_Model_Core_Teaching_Standards_2011.pdf

[2] http://caepnet.org/standards/introduction

[3] With the exception of Poland for those teachers prepared to teach the early grades. The difference in scores between generalists and specialists in Poland provides important data concerning the relationship between standards and levels of knowledge attained by future primary teachers. Especially because those who are prepared as specialists in Poland attained very high scores com-

thus providing the opportunity to focus on what could be considered “best case” scenarios.

Insights from the Research Literature

A number of studies have suggested that the development of effective mathematics teachers depends on rigorous recruitment and selection strategies, and on teacher education programs’ ability to design coherently organized theoretical and practical opportunities to learn (Akiba, LeTendre, & Scribner, 2007; Feuer, Floden, Chudowsky, & Ahn, 2013; Grossman, Hammerness, & McDonald, 2009; Wang, Coleman, Coley, & Phelps, 2003). Increasingly, accreditation reviews of teacher education programs focus on the use of standards as a way to regulate admission and graduation requirements as well as program offerings (Eurydice, 2006; Tatto & Pippin, 2017). A comprehensive review of policies directed at improving teacher quality revealed that teacher education programs have established criteria for program admission to ensure that candidates have the required knowledge of the subject or subjects they are expected to teach and that the rigor of these policies seem to be characteristic of “top performing” countries (Barber & Mourshed, 2007). Indeed, because previous attainment is an important predictor of future attainment (Ingersoll, 2007), those programs in countries that emphasize a minimum standard of mathematics knowledge as a condition for program admission may be able to better prepare their future teachers than those that have no such requirement. Yet the few studies testing this notion in the United States have proven inconclusive (Casey & Childs, 2011; Levine, 2006; Mikitovics & Crehan, 2002), likely because of the variation in the criteria used or the degree to which this criterion is rigorously implemented.

The analysis of recruitment, selection, certification and accreditation policies carried out as part of the larger TEDS-M study is revealing. For instance, the analysis found that countries vary as to the level of mathematics required for program admission (Tatto et al., 2012). For instance, only in Chinese Taipei do regulations require that in addition to graduating from secondary school, candidates have one year of tertiary-level studies, plus a national examination in order to enter a teacher education program with mathematics as a required subject. Standards in Botswana, Poland (only for upper primary teachers), the Russian Federation, and Singapore require graduation from secondary school with a specific mathematics requirement. Graduation from secondary school, with no specific mathematics requirement, is

parable to future primary teachers in Chinese Taipei, the highest scoring country. The mathematical content knowledge assessment (MCK) and the mathematical pedagogical content assessment (MPCK) scores for future primary teachers are as follows: (a) lower primary up to grade 4 maximum in Poland (456.2/425), in the Russian Federation (535.5/511.9); (b) primary up to grade 6 maximum in Chinese Taipei (623.2/592.3), in the United States (517.5/543.6); (c) as primary specialists in Poland (614.2/574.8), in the United States (520/544.5) (Tatto et al., 2012).

the only requirement in Chile, Germany, Malaysia, the Philippines, Spain, Switzerland, Thailand, and the United States. The extent to which education policy can help improve the quality of those who will teach mathematics at the primary level, including programs' selection policies and monitoring, is an area that deserves much attention.

The TEDS-M study also found that standards for admission and graduation are typically complemented by standards that regulate teacher education curricular offerings and field experiences (typically via accreditation guidelines). The TEDS-M study supports findings from other studies that have found wide variation within and across countries, not only in terms of the standards used, but also in terms of the content and length of such experiences (Feuer et al., 2013; Grossman et al., 2009; Tatto & Hordern, 2017; Tatto & Pippin, 2017 ; Tatto et al., 2012; Youngs & Grogan, 2013). This is in spite of studies showing that coherence between a program's philosophy and a program's curricular offerings contributes to important cognitive gains (see, e.g., Tatto, 1996). Recent work on teacher education outcomes, including that done in TEDS-M, reveals that coherent programs that align their offerings with standards as part of accountability demands, such as accreditation exercises, may be better able to produce highly qualified teachers (Ingvarson et al., 2013; Tatto et al., 2012).

In this chapter, we examine the relationship between attained knowledge of mathematics and mathematics pedagogy at the end of teacher education and the role of standards regulating program admission, graduation, content, and practical experiences for future primary teachers.

Methods

Data Source and Instruments

The TEDS-M study Institutional Program Questionnaire (IPQ) asked program administrators questions related to the following:

- the extent to which the program had standards that regulated entry and exit requirements;
- the extent to which the program is required to comply with external regulations concerning who is admitted to the program;
- program standards that show how much emphasis is given to understanding how pupils learn and think in mathematics;
- program standards that show how much emphasis is given to the management of diversity in the classroom as a key ability needed to become an effective teacher; and
- program standards that show the degree to which the program emphasizes key field experiences such as learning to work with pupils, and doing research on one's own practice, and learning how to plan, reflect and instruct from mentors or supervisors, to equip them to deal with the realities of the classroom.

These questions served as the basis for the analyses presented in the second part of this chapter. The measures that are the foundation for the statistical analyses in the current study were created using these questions and are described below.

Program Selectivity and Monitoring Data from the IPQ provided two key indicators of programs' attempts at regulating program admission and graduation.

Entry Criteria Program administrators were asked whether the program "required a 'demonstrated high level of achievement in mathematics' as a characteristic or source of information that is used in selecting entering future teachers for the program" (Brese & Tatto, 2012b, p. 19). The questionnaire also asked "with reference to national norms, how do future teachers entering the program rate with respect to their prior academic achievement" with possible responses ranging from "generally very high achievers" (e.g. the top 10% of their age group) to "generally far below-average achievers" (for their age group) (Brese & Tatto, 2012b, p. 20).

Exit Criteria Administrators were asked whether future primary teachers were required to "write and defend a thesis" as part of the institutional requirements needed to successfully complete the program, and whether there was "a document that prescribed competencies or performance standards that the graduates of the program were expected to meet" (Brese & Tatto, 2012b, p. 28).

Although these questions seemed straightforward to answer, they were not in many cases. For instance, in Chinese Taipei, programs did not need to use admission criteria because the university and other external authority had done that already, as the international report explains: "In Chinese Taipei, students must be enrolled in their second or higher year of university (including masters and doctoral levels) before they can be admitted to a teacher education program. Although there is no specific secondary school mathematics requirement, students must pass the national university entrance examination, which has mathematics as a required test subject" (Tatto et al., 2012, pp. 44–45). Thus, while "no" would technically be the correct answer to our questions, there were de-facto entry requirements that assured high levels of mathematics competence among future primary teachers.

External/Internal Regulation for Program Admission Accountability policies are driven by the assumption that regulation is key to maintaining quality teacher-education programs. While this view has gained support increasingly in the policy area worldwide, others argue that teacher education programs should maintain autonomy. One of the most important decisions in this regard has to do with who sets the policies that govern which applicants are admitted to the program. The IPQ asked whether the policies that govern which applicants are admitted to the program are set by "each institution without reference to any outside requirements"; "regional or national authorities"; or "each institution, within guidelines set by regional or national authorities"; or, instead, "there is no selection and all applicants are admitted" (Brese & Tatto, 2012b, p. 19). In addition to indicating autonomy, this question also indicates the degree to which the institution is selective.

Program Content Standards There is extensive research on what makes a good teacher (e.g., Brophy & Good, 1986; Hammerness et al., 2005; Kennedy, 1999; Windschitl, Thompson, & Braaten, 2011). Some believe that good teachers are those who can successfully manage a classroom. Another view is that knowing how to tailor instruction according to the learning needs of students is more important than learning to manage students. Yet others view classroom management as essential in the creation of a productive learning environment. Others believe that deep knowledge of the subject—in this case mathematics—is the most important capacity characterizing effective teachers. Yet others argue that learning how to be a reflective practitioner is key to developing an effective practice. More recently, there has been a turn to practice as some argue that learning to teach is better done on the job (Zeichner, 2012).

As explained in the introduction, current standards bring together these views into a more holistic image of an effective teacher. Yet the data from the TEDS-M IPQ shows that there is variability within and across countries in the degree to which programs emphasize these standards.

Standards Emphasizing Mathematical Pedagogical Content Knowledge About Pupil Learning In the institutional questionnaire, in the section on pedagogical content knowledge, program administrators were asked,

> In the program requirements, guidelines and other documentation, how much weight is given to each of these goals?
>
> - Knowledge about pupil learning in mathematics,
> - knowing common pupil misunderstandings in mathematics,
> - knowing how to build on pupils' prior knowledge in mathematics (Brese & Tatto, 2012b, p. 26).

Standards Emphasizing General Pedagogical Knowledge About Managing, Teaching, and Assessing Learning, of Diverse Students In the institutional questionnaire, in the sections on general pedagogy/assessing learning/diversity, program administrators were asked,

> In the program requirements, guidelines and other documentation, how much weight is given to each of these goals?
>
> - Managing disruptive pupils;
> - specific strategies for teaching pupils with behavioral and emotional problems;
> - specific strategies and curriculum for teaching pupils with learning disabilities;
> - specific strategies and curriculum for teaching gifted pupils;
> - specific strategies and curriculum for teaching pupils from diverse cultural backgrounds; accommodating the needs of pupils with physical disabilities in your classroom;
> - working with children from poor or disadvantaged backgrounds;
> - conducting fair and valid summative assessments of pupil learning (Brese & Tatto, 2012b, pp. 26–27).

The data analyzed shows that, across the board, teacher education programs rarely provide opportunities to learn in these areas.

Program Field Experience Standards Program standards are also expressed in the extent to which they regulate opportunities to learn during field experiences. The field experience in pre-service teacher education programs is seen as a key part of learning to teach and is expected to bring future teachers in touch with pupils to help them understand how they learn, and to learn to research and reflect on their own practice.

Standards Emphasizing Working with Pupils The questionnaire asked

> How often were the following activities assigned as part of the introductory field experiences in this program?
>
> - tutoring individual pupils,
> - working with small groups of pupils.

Standards Emphasizing Doing Research on and in Practice The questionnaire asked

> How often were the following activities assigned as part of the introductory field experiences in this program?
>
> - collecting data for research projects (Brese & Tatto, 2012b, p. 31).

Standards Regulating the Role of Mentors/Supervisors During the Field Experience A key element of a well-designed field experience is the quality of the supervision future teachers receive. Teacher education programs typically rely on supervisors for a high-quality school experience.

Supervising/Modeling Instruction Of interest was whether and how supervisors provided support in mathematics instruction (content and pedagogy) as part of future teachers' extended practice. The following questions were part of the questionnaire:

> Please indicate whether supervisors/instructors/mentors in the extended teaching practice are likely to assume each of the following responsibilities.
>
> Responsibilities for helping future teachers to plan:
>
> - the mathematics content of a lesson;
> - the mathematics pedagogy of a lesson;
> - how to deal with pupils with learning problems;
> - how to deal with pupils with behavior problems.
>
> Responsibilities for giving oral feedback and fostering reflection:
>
> - giving future teachers oral feedback on the adequacy of the mathematics content in their teaching;
> - giving future teachers oral feedback on their pedagogical approach to teaching mathematics.

Supervising, Instructing, Modeling, Coaching When Working with Students Of interest was whether and how supervisors provided support to future teachers' when working with students as part of future teachers' extended practice. The following questions were part of the questionnaire:

> Please indicate whether supervisors/instructors/mentors in the extended teaching practice are likely to assume each of the following responsibilities.
>
> Responsibilities for instructing, modeling, coaching, etc.
>
> - teaching a lesson to primary or secondary school pupils that a future teacher is expected to observe;
> - taking charge of a class of primary or secondary school pupils to help a future teacher who has run into difficulties with the class (Brese & Tatto, 2012b, p. 33).

Together, these areas provide a comprehensive view of programs' standards, in the pursuit of quality. These will be used to investigate their relation with program outcomes for future primary teachers. In the next sections, we describe our methods and analysis.

Dependent Variables

Two program outcomes were considered for this analysis: mathematical content knowledge (MCK) and mathematical pedagogical content knowledge (MPCK). Both variables are IRT (item response theory) scales with mean of 500 and standard deviation of 100 available in the TEDS-M database (Brese & Tatto, 2012a).

Independent Variables

Variables derived from the IPQ described in the previous section were used for the analyses in this study. Factor analysis and principal components techniques were used to compute the derived variables. The final teacher-education programs unit weights were considered for the estimation in order to have design-unbiased derived variables as suggested by the TEDS-M User's Manual (Brese & Tatto, 2012a). Composite scores and factor variables were created as indicators of the degree to which standards (or norms) regulated program selectivity and monitoring, internal / external regulation for program admission, program content, and initial and extended field experiences. The derived measures are defined in Table 3.1. The source variables used to create the composite scores and factor analytic scales are presented in the Appendix.

Two factors were created to indicate the use of standards in program selectivity and monitoring: *Entry_criteria* and *Exit_criteria*. Both are factor analytic scales where zero is located at the central position, with higher values corresponding to more stringent requirements for program entry and exit/graduation.

SelfGov is a dichotomous variable regarding program's degree of internal or external regulation concerning admission standards. It is equal to one if the policies

Table 3.1 Definition of variables and description of their computation

Variable	Definition	Computation description
Program outcomes		
MCK	Mathematics content knowledge	IRT scores with mean of 500 and standard deviation of 100
MPCK	Mathematics pedagogy content knowledge	IRT scores with mean of 500 and standard deviation of 100
Program selectivity and monitoring		
Entry_criteria	Admission requirements to enter the program	Factor analytic scale, standardized scores
Exit_criteria	Graduation requirements to exit the program	Factor analytic scale, standardized scores
Governance		
SelfGov	Indicator of whether or not the admission norms are set by the program	Item MIC001 recoded (1 = self governance, 0 = external governance)
Content standards		
MPCK_ StudLearning	Emphasis given to knowledge about pupil learning in mathematics related aspects in the program requirements	Factor analytic scale, standardized scores
Managing_ StudDiversity	Emphasis given to knowledge about student diversity in the program requirements	Factor analytic scale, standardized scores
Program field experiences		
WorkPupils	Frequency with which tutoring individual pupils or working with small groups activities are assigned as part of the introductory field experiences in the program	Total score of items MIE003C and MIE003D
DataColl	Frequency with which data for research projects activities are assigned as part of the introductory field experiences in the program	Item MIE003I
Program supervisors		
SuppInstruction	Extent to which supervisors assume responsibilities related to mathematics instruction (content and pedagogy related)	Factor analytic scale, standardized scores
SuppWork_ Students	Extent to which supervisors assume responsibilities related to modeling when working with students	Factor analytic scale, standardized scores

that govern who is admitted to the program are set by each program without reference to any outside requirements, or simply all applicants are admitted. If *SelfGov* takes the value of zero, then regional or national guidelines are considered in the admission process. Because of the way we coded this variable, positive coefficients in our models (in our models below) represent program's self-governance and negative coefficients represent external regulations.

Two factor analytic scales where zero is located at the central position were created to measure program content standards. *MPCK_StudLearning* indicates how much emphasis is given by the program to future teachers acquiring knowledge about how pupils learn mathematics in the program requirements, guidelines, and other documentation. Similarly, *Managing_StudDiversity* represents the emphasis given by the program on future teachers learning how to manage students' diversity (which included classroom management, fair assessment, and cultural and economic diversity as well as diversity represented by students with special needs) in the program guidelines. Higher values in these two scales mean higher emphasis given to each aspect in the program standards.

Regarding program introductory field experiences standards, two aspects were salient. *WorkPupils* represents the weight given by the program to future teachers working with pupils such as tutoring individual pupils or working with small groups on specific school-related activities as part of the introductory field experiences. This variable represents the total score of two items in the IPQ. Higher values mean that activities related to working with pupils were assigned high priority in program standards. Program standards related to the weight given to experiences related to doing research projects such as data collection as part of the introductory field experiences are represented by the variable *DataColl*, and is generated solely by one item in the IPQ, with higher values meaning a high priority in program standards.

Regarding program extended teaching practice and standards for supervisors or mentor teachers, two aspects were key. These indicate standards regarding the responsibilities that program supervisors, instructors or mentor teachers are expected to assume in the extended teaching practice. *SuppInstruction* represents the extent to which supervisors are expected to help future teachers plan the content, the pedagogy of a lesson anticipating how to address learning and behavioral problems in pupils, and to provide feedback regarding the adequacy of the content and of the pedagogical approach to teaching mathematics. *SuppWork_Students* represents the extent to which the program expects supervisors, instructors or mentor teachers to assume responsibilities for modeling teaching, and helping future teachers when they encounter difficulties with the class. Higher values of the two factors mean higher expectations by the program that supervisors, instructors or mentors take responsibility for instructing, modeling or coaching future teachers as they observe them teaching lessons, and as they take charge of a class.

Models

Six linear regression models were examined. Models 1–6 defined below were estimated considering MCK and MPCK as outcome variables for the four different countries: the United States, Russia, Poland, and Chinese Taipei. Model 1 solely includes the program selectivity and monitoring measures (*Entry_criteria* and *Exit_criteria*); Models 2 through 5 factor in other key program strategies or (a) standards to ensure a high quality graduate including whether the program self-regulates or whether it is subject to external regulation as far as admission criteria is concerned, (b) standards that regulate program content and field experience, and (c) program standards for supervisors during field experiences. Model 6 considers all of these variables simultaneously.

Model estimation was performed using the *svyglm* function of the *survey* R package (Lumley, 2004) in order to reflect the complex sample design of TEDS-M.

Model 1

$$Y_i = \beta_0 + \beta_1 Entry\_criteria_i + \beta_2 Exit\_criteria_i + \varepsilon_i$$

Model 2

$$Y_i = \beta_0 + \beta_1 Entry\_criteria_i + \beta_2 Exit\_criteria_i + \beta_3 SelfGov_i + \varepsilon_i$$

Model 3

$$Y_i = \beta_0 + \beta_1 Entry\_criteria_i + \beta_2 Exit\_criteria_i + \beta_3 MPCK\_StudLearning_i + \beta_4 Managing\_StudDiversity_i + \varepsilon_i$$

Model 4

$$Y_i = \beta_0 + \beta_1 Entry\_criteria_i + \beta_2 Exit\_criteria_i + \beta_3 WorkPupils_i + \beta_4 DataColl_i + \varepsilon_i$$

Model 5

$$Y_i = \beta_0 + \beta_1 Entry\_criteria_i + \beta_2 Exit\_criteria_i + \beta_3 SuppInstruction_i + \beta_4 SuppWork\_Students_i + \varepsilon_i$$

Model 6

$$Y_i = \beta_0 + \beta_1 Entry\_criteria_i + \beta_2 Exit\_criteria_i + \beta_3 SelfGov_i + \beta_4 MPCK\_StudLearning_i + \beta_5 Managing\_StudDiversity_i + \beta_6 WorkPupils_i + \beta_7 DataColl_i + \beta_8 SuppInstruction_i + \beta_9 SuppWork\_Students_i + \varepsilon_i$$

Results

Descriptive statistics are shown in Tables 3.2, 3.3, 3.4, 3.5, and 3.6. Table 3.2 presents the weighted mean and standard deviation for the variables included in this study. Tables 3.3, 3.4, 3.5 and 3.6 are the unweighted variance-covariance matrices for each country showing the patterns of association for all the variables considered for the regression analyses.

Table 3.2 Weighted means and standard errors by country

Variable	U.S.	Russia	Poland	Chinese Taipei
MCK	517.150	528.420	490.700	622.770
	(4.146)	(11.072)	(3.113)	(3.152)
MPCK	543.740	506.640	473.080	592.440
	(2.519)	(10.455)	(2.100)	(2.221)
Entry_criteria	0.001	0.175	−0.637	−0.242
	(0.063)	(0.074)	(0.003)	(0.000)
Exit_criteria	−0.178	0.558	0.404	−0.728
	(0.045)	(0.000)	(0.002)	(0.000)
SelfGov	.334	.066	.363	.156
	(.061)	(.019)	(.003)	(.000)
MPCK_StudLearning	−0.172	0.623	−0.726	−1.085
	(0.168)	(0.071)	(0.008)	(0.000)
Managing_StudDiversity	−0.109	0.566	−0.067	−0.797
	(0.149)	(0.151)	(0.005)	(0.000)
WorkPupils	6.808	5.845	5.173	5.667
	(0.204)	(0.454)	(0.007)	(0.000)
DataColl	2.159	3.262	2.427	2.731
	(0.176)	(0.187)	(0.006)	(0.000)
SuppInstruction	−0.035	0.408	−0.431	−0.683
	(0.155)	(0.059)	(0.005)	(0.000)
SuppWork_Students	0.232	−0.265	−0.091	0.227
	(0.118)	(0.137)	(0.006)	(0.000)

Note. Standard errors are reported in parentheses

Estimation results of Models 1–6 for both MCK and MPCK as outcome variables are shown in Tables 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, 3.13 and 3.14. In order to make fair comparisons among Models 1–6, the analytic sample of each country contained only those observations that had information for all variables considered in the analyses. The bottom of Tables 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, 3.13 and 3.14 show the sample size (N) for each country. The analytic sample represents 60% of the total data for the United States, 79% for Russia, 53% for Poland, and 78% for Chinese Taipei. Thus, the data should be interpreted with caution; however, the complete data can be assumed to come from programs that have complete information, and these may be assumed to represent the "best case" scenarios (Patton, 1990).

We used the Akaike's Information Criteria (AIC) and deviance for model selection. AIC is a measure of the relative quality of a model, while deviance is a goodness of fit measure. Both statistics allow comparisons of non-nested models, and smaller values are preferred (Howell, 2013; McCullagh & Nelder, 1989). Notice that Model 6 is the best-fitting model according to both statistics, AIC and deviance, for both outcome variables (MCK and MPCK) in the four countries. Hence, we focus on describing the estimation results for Model 6.

Table 3.3 Unweighted correlation matrix for U.S.

	MCK	MPCK	Entry_criteria	Exit_criteria	SelfGov	StudLearn.	StudDiv.	WorkP.	DataColl	SuppInst.	SuppStud.
MPCK	.438	1.000									
Entry_criteria	.146	.103	1.000								
Exit_criteria	.006	.041	.325	1.000							
SelfGov	.053	.037	.149	.073	1.000						
MPCK_StudLearning	.000	−.026	.189	.347	.359	1.000					
Managing_StudDiversity	−.128	−.077	.252	.237	−.217	.239	1.000				
WorkPupils	.087	.080	.238	−.026	.000	−.074	.107	1.000			
DataColl	.101	.113	−.008	−.059	−.149	−.114	.100	.190	1.000		
SuppInstruction	.059	.040	.547	.433	.126	.447	.325	.184	.093	1.000	
SuppWork_Students	.062	.042	.414	.465	.001	.282	.121	.241	.035	.653	1.000

Table 3.4 Unweighted correlation matrix for Russia

	MCK	MPCK	Entry_criteria	SelfGov	StudLearn.	StudDiv.	WorkP.	DataColl	SuppInst.	SuppStud.
MPCK	.617	1.000								
Entry_criteria	.304	.303	1.000							
SelfGov	−.116	−.099	−.104	1.000						
MPCK_StudLearning	.194	.163	.185	.091	1.000					
Managing_StudDiversity	.137	.099	.410	−.142	.284	1.000				
WorkPupils	−.169	−.201	−.280	.175	−.024	.024	1.000			
DataColl	−.050	−.081	−.191	−.186	−.210	.106	.040	1.000		
SuppInstruction	.020	.009	−.122	.077	.157	.061	−.077	−.082	1.000	
SuppWork_Students	−.019	−.005	.104	−.042	−.015	.351	.053	−.013	.253	1.000

Table 3.5 Unweighted correlation matrix for Poland

	MCK	MPCK	Entry_criteria	Exit_criteria	SelfGov	StudLearn.	StudDiv.	WorkP.	DataColl	SuppInst.	SuppStud.
MPCK	.651	1.000									
Entry_criteria	.339	.217	1.000								
Exit_criteria	−.169	−.160	−.004	1.000							
SelfGov	−.030	.019	−.123	.124	1.000						
MPCK_StudLearning	−.041	−.027	.306	.220	.006	1.000					
Managing_StudDiversity	−.368	−.249	−.056	.251	.022	.420	1.000				
WorkPupils	−.053	−.057	.112	.348	.238	.225	.090	1.000			
DataColl	.077	.101	.131	−.115	−.054	.174	.164	.148	1.000		
SuppInstruction	−.022	−.054	.099	.223	.050	.449	.049	.292	−.135	1.000	
SuppWork_Students	−.154	−.137	−.190	.443	.350	.269	.281	.389	−.176	.503	1.000

Table 3.6 Unweighted correlation matrix for Chinese Taipei

	MCK	MPCK	Entry_criteria	Exit_criteria	SelfGov	StudLearn.	StudDiv.	WorkP.	DataColl	SuppInst.	SuppStud.
MPCK	.438	1.000									
Entry_criteria	−.081	−.093	1.000								
Exit_criteria	.038	.089	−.107	1.000							
SelfGov	.028	.090	−.639	.546	1.000						
MPCK_StudLearning	.043	−.015	.420	−.404	−.418	1.000					
Managing_StudDiversity	−.024	−.001	.305	.098	−.272	.551	1.000				
WorkPupils	−.121	−.066	−.403	.015	.131	−.672	−.221	1.000			
DataColl	−.049	−.009	−.661	−.218	.255	−.072	.070	.554	1.000		
SuppInstruction	.023	.062	−.336	.510	.732	−.114	.066	−.155	−.069	1.000	
SuppWork_Students	−.037	.040	.201	.687	.494	−.050	.197	−.196	−.127	.290	1.000

United States

For the United States, programs whose future primary teachers obtained high scores in the MCK assessment also reported having standards that set requirements for program admission (*Entry_criteria*) as indicated by a significant positive relationship. Conversely, programs that had standards emphasizing opportunities to learn to manage students' diversity (*Managing_StudDiversity*) show a significant and negative association with the MCK scores likely meaning that the most knowledgeable future primary teachers were not in programs that emphasized the need to learn how to address the needs of diverse students including those with special needs, and diverse cultural and social backgrounds (see Table 3.7).

Table 3.8 shows results for the MPCK assessment. Those programs whose future teachers had high MPCK scores reported having standards that regulated admission and graduation requirements (*Entry_criteria* and *Exit_criteria*). These programs also have standards that included initial field experiences emphasizing tutoring individual pupils or working with small groups (*WorkPupils*), and research projects(*DataColl*) as indicated by a significant and positive relationship. Programs whose future teachers had high scores in the MPCK assessment did not emphasize content standards about learning to manage and teach diverse students (*Managing_StudDiversity*) as indicated by a significant and negative parameter estimate.

Russian Federation

For the Russian Federation, programs whose future teachers obtained high scores in the MCK and MPCK assessments reported having standards that set requirements for admission to the program (*Entry_criteria*) with a positive and significant association (see Tables 3.9 and 3.10). In Russia admission policies are set by regional or national guidelines a regulation that shows a significant and positive association with high scores in the MCK assessment (note that our recoding assigns a value of "1" to program self-governance and a value of "zero" to programs' standards regulated by external guidelines for admission thus showing a negative estimation coefficient) (*SelfGov*) (see Table 3.9). For the Russian Federation, notice that the *Exit_criteria* variable was omitted for the analyses (see Tables 3.9 and 3.10). This was done because there was no variability in the data, i.e., all Russian institutions provided the same answers to the items that were considered for the creation of this factor scale. That is, all Russian institutions require students to write and defend a thesis in order to successfully complete the program, and they have a document that prescribes competences or performance standards that graduates are expected to meet.

While no significant, all the other standards show a positive relationship with higher scores in both assessments, with the exception of programs' lack of emphasis

Table 3.7 Estimation results for MCK as outcome variable in models 1–6 for U.S.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Program selectivity and monitoring						
Entry_criteria	22.312**	22.328**	27.075***	20.767**	23.299**	27.143***
	(7.984)	(8.090)	(8.524)	(7.787)	(8.627)	(8.126)
Exit_criteria	10.973	10.901	13.009	11.420	10.284	14.252
	(20.612)	(19.998)	(19.708)	(20.391)	(20.642)	(16.719)
Internal/External regulation						
SelfGov		−1.176				−12.396
		(8.074)				(8.447)
Program content standards						
MPCK_StudLearning			2.210			6.034
			(3.565)			(3.847)
Managing_StudDiversity			−12.515***			−17.862***
			(3.392)			(4.201)
Program field experience standards						
WorkPupils				2.038		3.565
				(2.684)		(2.197)
DataColl				1.064		3.251
				(4.540)		(4.193)
Program field experience's supervisors standards						
SuppInstruction					−2.399	−0.218
					(4.749)	(4.593)
SuppWork_Students					2.692	−3.236
					(6.514)	(5.638)

(continued)

Table 3.7 (continued)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
(Intercept)	519.096	519.476	518.478	503.004	518.263	492.375
	(6.752)	(7.863)	(6.798)	(21.346)	(6.745)	(18.209)
Log likelihood	−5120	−5120	−5111	−5118	−5119	−5105
Deviance	4371	4371	4286	4361	4368	4231
AIC	10246	10248	10232	10247	10249	10230
d.f.	20	19	18	18	18	13

Note. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; $N = 896$

Table 3.8 Estimation results for MPCK as outcome variable in models 1–6 for U.S.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Program selectivity and monitoring						
Entry_criteria	13.764*	13.769*	16.303**	11.772*	15.247*	15.443**
	(6.953)	(7.055)	(6.998)	(6.772)	(7.974)	(6.642)
Exit_criteria	14.817	14.791	17.901	14.244	16.329	20.329**
	(13.055)	(12.744)	(12.348)	(12.221)	(11.846)	(9.114)
Governance						
SelfGov		−0.427				−5.084
		(4.639)				(4.731)
Content standards						
MPCK_StudLearning			−1.771			−0.267
			(2.572)			(2.863)
Managing_StudDiversity			−6.747**			−10.970***
			(3.135)			(2.917)
Program field experience standards						
WorkPupils				1.650		2.578*
				(1.825)		(1.335)
DataColl				3.741		5.966*
				(3.664)		(3.128)
Program supervisors standards						
SuppInstruction					−2.113	0.833
					(3.800)	(3.883)
SuppWork_Students					−0.265	−4.987
					(3.791)	(3.976)
(Intercept)	546.379	546.517	545.890	526.966	546.635	518.577
	(3.641)	(4.217)	(3.957)	(15.898)	(3.443)	(11.956)
Log likelihood	−5127	−5127	−5124	−5125	−5127	−5118
Deviance	4448	4448	4414	4425	4446	4361
AIC	10261	10263	10258	10261	10265	10257
d.f.	20	19	18	18	18	13

Note. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; $N = 896$

on learning to teach diverse students and on working with individual or small groups of pupils in the initial field experience.

Poland

In the case of Poland, as indicated by a significant positive relationship (see Table 3.11), programs whose future teachers scored high in the MCK assessment reported having standards that set requirements for admission to the program

Table 3.9 Estimation results for MCK as outcome variable in models 1–6 for Russia

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Program selectivity and monitoring						
Entry_criteria	60.976***	55.757**	62.427***	53.403***	64.258***	57.747**
	(19.425)	(20.928)	(17.824)	(18.001)	(18.795)	(20.141)
Governance						
SelfGov		−44.739**				−37.595*
		(16.042)				(18.874)
Content standards						
MPCK_StudLearning			16.799			29.492
			(15.311)			(19.286)
Managing_StudDiversity			−5.996			−12.065
			(15.731)			(17.522)
Program field experience standards						
WorkPupils				−7.662		−6.451
				(5.524)		(5.655)
DataColl				11.876		15.468
				(16.573)		(17.644)

Program supervisors standards						
SuppInstruction					15.240	12.974
					(20.919)	(24.838)
SuppWork_Students					−10.731	−5.773
					(11.693)	(10.532)
(Intercept)	517.733	521.615	510.403	525.100	508.090	489.657
	(10.204)	(11.098)	(13.040)	(74.707)	(13.475)	(76.535)
Log likelihood	−10623	−10609	−10616	−10584	−10615	−10554
Deviance	7885	7767	7828	7549	7817	7303
AIC	21250	21225	21241	21176	21239	21127
d.f.	20	19	18	18	18	13

Note. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; $N = 1786$

Table 3.10 Estimation results for MPCK as outcome variable in models 1–6 for Russia

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Program selectivity and monitoring						
Entry_criteria	56.694***	52.959**	59.231***	48.701***	58.037***	53.643**
	(19.909)	(21.482)	(17.110)	(16.415)	(19.398)	(18.625)
Governance						
SelfGov		−32.019				−27.988
		(18.866)				(21.504)
Content standards						
MPCK_StudLearning			15.507			25.891
			(12.272)			(15.942)
Managing_StudDiversity			−7.272			−14.773
			(12.761)			(15.248)
Program field experience standards						
WorkPupils				−6.747		−6.011
				(4.598)		(4.866)
DataColl				6.460		10.332
				(14.872)		(15.462)

Program supervisors standards						
SuppInstruction					10.737	6.839
					(17.350)	(19.957)
SuppWork_Students					−2.766	3.622
					(10.359)	(10.221)
(Intercept)	496.706	499.484	490.714	516.47	491.355	490.919
	(9.865)	(10.685)	(13.495)	(64.876)	(12.124)	(65.699)
Log likelihood	−10517	−10509	−10510	−10490	−10515	−10469
Deviance	7004	6943	6950	6799	6986	6639
AIC	21039	21025	21029	20989	21038	20957
d.f.	20	19	18	18	18	13

Note. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; $N = 1786$

Table 3.11 Estimation results for MCK as outcome variable in models 1–6 for Poland

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Program selectivity and monitoring						
Entry_criteria	94.811***	95.926***	85.362***	94.651***	94.985***	90.651***
	(6.967)	(6.692)	(7.948)	(7.298)	(6.798)	(8.061)
Exit_criteria	−36.924***	−38.346***	−22.585***	−35.047***	−37.883***	−26.909***
	(4.851)	(4.771)	(5.378)	(5.505)	(6.850)	(7.045)
Governance						
SelfGov		8.966				7.719*
		(5.410)				(4.210)
Content standards						
MPCK_StudLearning			6.385**			4.347
			(2.735)			(3.243)
Managing_StudDiversity			−39.187***			−44.288***
			(2.663)			(2.761)
Program field experience standards						
WorkPupils				−1.255		−3.660*
				(2.056)		(1.984)
DataColl				1.017		9.035***
				(2.495)		(2.659)
Program supervisors standards						
SuppInstruction					0.122	−4.828*
					(2.785)	(2.684)
SuppWork_Students					0.894	16.140***
					(3.891)	(4.043)
(Intercept)	566.004	564.038	556.205	569.168	566.635	553.084
	(5.328)	(5.770)	(5.148)	(13.158)	(6.186)	(12.048)
Log likelihood	−6665	−6663	−6590	−6664	−6665	−6574
Deviance	7901	7883	6909	7897	7901	6713
AIC	13336	13335	13191	13339	13340	13169
d.f.	29	28	27	27	27	22

Note. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; $N = 1111$

(*Entry_criteria*), and these programs also exert self-governance concerning admission program policies (*SelfGov*). These programs also have standards emphasizing the need for future teachers to know how students learn and think about mathematics (*MPCK_StudLearning*), the need for introductory field experiences requiring research projects (*DataColl*), and the need for supervisors to allow future teachers to observe them when teaching a lesson, and to support them when taking charge of a class (*SuppWork_Students*). These high scoring programs, however, did not have requirements for graduation (*Exit_criteria*), and did not emphasize opportunities to learn about students' diversity (*Managing_StudDiversity*), working with individual

Table 3.12 Estimation results for MPCK as outcome variable in models 1–6 for Poland

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Program selectivity and monitoring						
Entry_criteria	74.767***	77.008***	62.237***	73.229***	74.618***	65.964***
	(6.215)	(6.489)	(6.397)	(5.997)	(6.881)	(6.681)
Exit_criteria	−35.704***	−38.558***	−30.658***	−31.974***	−33.586***	−28.351***
	(5.827)	(6.254)	(6.594)	(7.442)	(6.452)	(7.142)
Governance						
SelfGov		18.006**				19.453**
		(7.290)				(7.294)
Content standards						
MPCK_ StudLearning			11.780***			10.008**
			(3.774)			(4.149)
Managing_ StudDiversity			−27.182***			−30.129***
			(3.231)			(3.252)
Program field experience standards						
WorkPupils				−2.110		−4.099
				(2.406)		(2.694)
DataColl				5.199		7.911***
				(3.351)		(2.720)
Program supervisors standards						
SuppInstruction					−0.708	−2.552
					(3.596)	(3.169)
SuppWork_ Students					−1.725	5.539
					(2.725)	(3.304)
(Intercept)	535.128	531.180	531.839	530.938	533.717	526.151
	(4.666)	(4.588)	(5.201)	(11.571)	(5.048)	(12.546)
Log likelihood	−6737	−6733	−6710	−6735	−6737	−6698
Deviance	9004	8932	8565	8969	9001	8390
AIC	13481	13474	13430	13481	13485	13417
d.f.	29	28	27	27	27	22

Note. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; $N = 1111$

students in the initial field experience (*WorkPupils*), or expectations for supervisors' support with planning and mathematics instruction (*SuppInstruction*) as indicated by negative parameter estimates (see Table 3.11).

Table 3.12 shows that programs where future teachers scored high in the MPCK assessment had standards that regulated admission requirements to enter the program (*Entry_criteria*), and these programs had a system of self-governance in setting admission policies (*SelfGov*). Programs where future teachers were more knowledgeable in MPCK had standards that emphasized the need to acquire knowledge about student learning in mathematics (*MPCK_StudLearning*), and engagement in data collection for research projects in the introductory field experiences

(*DataColl*) as indicated by positive parameter estimates (see Table 3.12). These programs did not have standards that regulated requirements for program graduation (*Exit_criteria*), standards emphasizing opportunities to learn to manage students' diversity (*Managing_StudDiversity*), or program standards emphasizing field experiences that include working with pupils (*WorkPupils*) as indicated by a negative relationship with MPCK (see Table 3.12).

Chinese Taipei

For Chinese Taipei,[4] standards for admission and graduation are very high albeit implemented at the university level. Programs whose future teachers obtained high scores in MCK also had standards that placed emphasis in acquiring knowledge about pupils' learning in mathematics (*MPCK_StudLearning*) and on supervisors instructing, modeling and coaching future teachers when teaching a class (*SuppWork_Students*) as indicated by a positive and significant relationship. Because teacher education programs rely on university standards for program admission (as indicated by a negative relationship with *SelfGov*) that are stringent, the programs themselves do not have additional entry requirements. These programs also do not emphasize research projects as part of initial field experiences (*DataColl*) as indicated by a negative association with MCK (see Table 3.13). As shown in Table 3.14, programs whose future teachers had high scores in the MPCK assessment reported having no additional standards that regulated admission to the program (*Entry_criteria*), as indicated by a negative significant parameter estimate. However, as discussed before for MCK, this does not mean that there are no admission standards; rather, this indicates that the program itself does not set admission policies, and instead these are set by the institution (university) and in accordance with regional or national authorities (as expressed by the negative estimate in *SelfGov*). There was a positive and significant association between high scores in MPCK and standards that mandate that supervisors instruct, model and coach future teachers when teaching a class (*SuppWork_Students*). As was the case for MCK above, programs whose future teachers had high scores in MPCK do not emphasize research projects as part of initial field experiences (*DataColl*) as indicated by a negative association with MPCK scores.

The negative association of the requirements to enter the program with outcome variables must be taken with caution, and it is applicable to the situation within Chinese Taipei, which has very high standards to enter a program in the first place but set by authorities external to the programs themselves. As Tatto et al. (2012) noted, "students must be enrolled in their second or higher year of university (including master's and doctoral levels) before they can be admitted to a teacher

[4] Notice that the variable *SuppInstruction* was omitted from the estimation of Model 6 in both tables because when included in the model, some cell combinations for this variable and two other of variables in the model were empty, which was a problem for model convergence.

Table 3.13 Estimation results for MCK as outcome variable in models 1–6 for Chinese Taipei

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Program selectivity and monitoring						
Entry_criteria	-22.390^{***}	-43.283^{***}	-26.413^{***}	-44.513^{***}	-18.877^{*}	-139.531^{***}
	(7.179)	(7.523)	(7.801)	(6.277)	(9.506)	(24.376)
Exit_criteria	5.858	23.036^{***}	24.314^{***}	−6.584	28.628^{***}	−2.502
	(6.443)	(6.767)	(7.414)	(6.668)	(7.723)	(10.475)
Governance						
SelfGov		-41.731^{***}				-90.628^{***}
		(10.500)				(22.875)
Content standards						
MPCK_StudLearning			23.174^{***}			26.729^{***}
			(5.333)			(8.949)
Managing_StudDiversity			-17.208^{***}			−6.295
			(4.156)			(5.489)
Program field experience standards						
WorkPupils				-9.040^{***}		4.408
				(2.128)		(4.248)
DataColl				-22.430^{***}		-64.987^{***}
				(5.133)		(13.376)
Program supervisors standards						
SuppInstruction					-7.349^{*}	
					(4.211)	
SuppWork_Students					-18.169^{***}	39.394^{***}
					(5.354)	(13.111)

(continued)

Table 3.13 (continued)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
(Intercept)	621.628	635.618	645.520	719.689	638.174	768.893
	(6.868)	(7.464)	(8.813)	(11.451)	(9.129)	(20.051)
Log likelihood	−4253	−4249	−4245	−4234	−4249	−4228
Deviance	7395	7320	7229	7014	7314	6902
AIC	8513	8507	8500	8479	8509	8475
d.f.	29	28	27	27	27	23

Note. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; $N = 721$

Table 3.14 Estimation results for MPCK as outcome variable in models 1–6 for Chinese Taipei

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Program selectivity and monitoring						
Entry_criteria	−16.937***	−22.560***	−19.159***	−26.106***	−17.251**	−77.769***
	(5.450)	(5.289)	(6.048)	(4.762)	(7.136)	(19.626)
Exit_criteria	12.333**	16.956***	19.688**	7.655	17.180***	−6.090
	(4.934)	(5.236)	(7.122)	(5.808)	(5.752)	(7.999)
Governance						
SelfGov		−11.231				−43.790*
		(11.545)				(22.518)
Content standards						
MPCK_StudLearning			9.175*			4.952
			(5.140)			(9.323)
Managing_ StudDiversity			−5.570			2.532
			(4.209)			(5.747)
Program field experience standards						
WorkPupils				−5.118**		−0.128
				(1.948)		(3.884)
DataColl				−7.816*		−31.841**
				(4.538)		(11.666)
Program supervisors standards						
SuppInstruction					−2.519	
					(5.022)	
SuppWork_Students					−3.059	26.986**
					(5.236)	(10.296)
(Intercept)	597.331	601.096	607.664	642.049	599.761	664.985
	(5.031)	(5.808)	(8.720)	(9.051)	(7.106)	(19.670)
Log likelihood	−4094	−4094	−4092	−4087	−4094	−4086
Deviance	4757	4751	4732	4667	4753	4648
AIC	8195	8196	8195	8185	8198	8190
d.f.	29	28	27	27	27	23

Note. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; $N = 721$

education program. Although there is no specific secondary school mathematics requirement, students must pass the national university entrance examination, which has mathematics as a required test subject" (pp. 44–45). Therefore, the negative association between requirements to enter the program (*Entry_criteria*) and program outcomes (MCK and MPCK) comes from administrators correctly answering "no" to the question of teacher education entry requirement as these rules are set by the university, not by the program, but the readers must keep in mind that the standards to enter teacher education in Chinese Taipei are extremely high. Hence, program admission requirements are unique in comparison with the other three countries presented in this analysis.

Conclusion

This analysis allows us to better understand the role that admission and graduation requirements play in the outcomes of teacher education for different countries. While there is important variability, a general pattern shows that programs where future primary mathematics teachers obtained high scores in the knowledge assessments (MCK and MPCK) have standards that establish as a requirement for admission a demonstrated high level of mathematics knowledge and seem to be able to recruit individuals who are in general high achievers according to national norms. Even when programs do not set these admission standards, as in Chinese Taipei, the role of external standards as set by universities, or by regional or national agencies is important. In Chinese Taipei for instance, both admission and graduation standards are rigorous. Admission standards require a strong knowledge of the disciplines (e.g., candidates need to be enrolled in at least their second year of higher education or graduate studies to be accepted in a program), and for graduation future teachers are expected to write and defend a thesis in order to successfully complete the program. In Poland, future primary teachers who scored high in the knowledge assessments were in programs with rigorous admission requirements (e.g., all institutions require applicants to have upper-secondary school qualifications in mathematics) albeit with no graduation requirements. Having no graduation requirements may imply that the program sets formative standards and other strategies to regulate the content that is to be learned and the practical experiences that would ensure knowledgeable graduates. Similarly, program admission requirements were associated with high knowledge scores in the United States, and graduation requirements had a positive and significant correlation with mathematical pedagogical content knowledge. Russia also has rigorous admission and graduation requirements with the latter implemented uniformly across all programs requiring future teachers to write and defend a thesis in order to successfully complete the program. For most countries programs set and enforce admission and graduation standards, with the exception of Chinese Taipei, and in most countries, these standards are also regulated externally (at the university, local or national levels), with the exception of Poland.

Overall, programs whose future primary teachers scored high in MCK and MPCK do not emphasize learning how to manage student's diversity including students with special needs, but rather the emphasis in the programs is on providing future teachers with knowledge about how pupils learn mathematics and on building upon their prior knowledge.

With the exception of programs in Chinese Taipei, programs whose future primary teachers scored high in MCK and MPCK have standards that emphasize learning to do research and collect data in their initial/short field experiences over learning to work individually or in small groups with pupils.

With the exception of programs in Russia, where emphasis in the extended field experience is for supervisors' support with instructional planning and with feedback on teaching, programs whose future teachers scored high in the knowledge assessments had extended field experiences that required supervisors, instructors and mentor teachers to place a heavy emphasis on modeling how to teach lessons and on coaching future teachers when running into difficulties when taking charge of a class and teaching a lesson.

Thus, while programs emphasize standards regulating content and field experiences in different degrees, the most consistent factor related to highly knowledgeable teachers as they near the end of their programs are admission and graduation standards. Regardless of who regulates admission and graduation standards, the presence of such standards shows in every instance an important and significant relationship with higher mathematical and mathematical pedagogical content knowledge for future primary teachers as expressed by the scores obtained in the TEDS-M knowledge assessment. In fact, requiring a good level of mathematical knowledge prior to entering the program seems essential for future mathematics teachers' acquisition of the expected mathematical pedagogical content knowledge they will need to teach.

Appendix

Program Selectivity and Monitoring

Variable name:	Entry_criteria	
Description:	Admission criteria to regulate entry into the program	
Procedure:	Factor analytic scale, standardized scores	
Explained variance:	18.22%	
Source:	MIC002E	Demonstrated high level of achievement in mathematics as a characteristic or source of information that is used in selecting entering <future teachers> for this teacher preparation program
	MIC004	With reference to national norms, how do <future teachers> entering this program rate with respect to their prior academic achievement?

Variable name:	Exit_criteria	
Description:	Graduation criteria to exit the program	
Procedure:	Factor analytic scale, standardized scores	
Explained variance:	17.09%	
Source:	MID012H	Write a defend a thesis as part of the institutional requirements that <future teachers> have to meet to successfully complete this program
	MID013A	Is there a document that prescribes competencies or performance standards that the graduates of this program are expected to meet?

Governance

Original variable: MIC001	Created variable: SelfGov
Options:	Values:
The policies are set by each institution without reference to any outside requirements.	1
The policies are set by regional or national authorities.	0
The policies are set by each institution, within guidelines set by regional or national authorities.	0
There is no selection for this phase; all applicants are admitted.	1

Program Content Standards

Variable name:	MPCK_StudLearning	
Description:	Emphasis given to student learning as part of the goals of program requirements	
Procedure:	Factor analytic scale, standardized scores	
Explained variance:	78.09%	
Source:	In the program requirements, guidelines and other documentation, how much weight is given to each of the goals listed below?	
	MID011F	Knowledge about pupil learning in mathematics
	MID011G	Knowing common pupil misunderstandings in mathematics
	MID011H	Knowing how to build on pupils' prior knowledge in mathematics

Variable name:	Managing_StudDiversity	
Description:	Emphasis given to student diversity as part of the goals of program requirements	
Procedure:	Factor analytic scale, standardized scores	
Explained variance:	47.79%	
Source:	In the program requirements, guidelines and other documentation, how much weight is given to each of the goals listed below?	
	MID011J	Managing disruptive pupils
	MID011N	Conducting fair and valid summative assessments of pupil learning
	MID011Q	Specific strategies for teaching pupils with behavioral and emotional problems
	MID011R	Specific strategies and curriculum for teaching pupils with learning disabilities
	MID011S	Specific strategies and curriculum for teaching gifted pupils

	MID011T	Specific strategies and curriculum for teaching pupils from diverse cultural backgrounds
	MID011U	Accommodating the needs of pupils with physical disabilities in your classroom
	MID011V	Working with children from poor or disadvantaged backgrounds

Program Field Experiences

Variable name:	WorkPupils	
Description:	Tutoring individual pupils or working with small groups as part of the introductory field experience	
Procedure:	Total score	
Source:	How often are the following activities assigned as part of the introductory field experiences in this program?	
	MIE003C	Tutor individual pupils
	MIE003D	Work with small groups of pupils

Variable name:	DataColl	
Description:	Collect data for research projects as part of the introductory field experience	
Procedure:	–	
Source:	How often are the following activities assigned as part of the introductory field experiences in this program?	
	MIE003I	Collect data for research projects

Program Supervisors

Variable name:	SuppInstruction	
Description:	Supervisors support in mathematics instruction (content and pedagogy)	
Procedure:	Factor analytic scale, standardized scores	
Explained variance:	65.78%	
Source:	Please indicate whether in the extended teaching practice are likely to assume each of the following responsibilities	
	MIE009A	The mathematics content of a lesson
	MIE009B	The mathematics of a lesson
	MIE009J	Giving <future teachers> oral feedback on the adequacy of the mathematics content in their teaching
	MIE009K	Giving <future teachers> oral feedback on their approach to teaching mathematics

Variable name:	SuppWork_Students	
Description:	Supervisors support in working with students	
Procedure:	Factor analytic scale, standardized scores	
Explained variance:	53.46%	
Source:	Please indicate whether in the extended teaching practice are likely to assume each of the following responsibilities	
	MIE009C	How to deal with pupils with learning problems
	MIE009D	How to deal with pupils with behavior problems
	MIE009G	Teaching a lesson to or school pupils that a <future teacher> is expected to observe
	MIE009H	Taking charge of a class of or school pupils to help a <future teacher> who has run into difficulties with the class

References

Akiba, M., LeTendre, G., & Scribner, J. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher, 36*(7), 369–387.

Barber, M., & Mourshed, M. (2007). *How the best performing school systems come out on top*. London, UK: McKinsey & Co.

Brese, F., & Tatto, M. T. (Eds.). (2012a). *User guide for the TEDS-M international database*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Brese, F., & Tatto, M. T. (Eds.). (2012b). *User guide for the TEDS-M international database. Supplement 1*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Brophy, J. J., & Good, T. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York, NY: MacMillan.

Casey, C., & Childs, R. (2011). Teacher education admission criteria as measure of preparedness for teaching. *Canadian Journal of Education, 34*(2), 3–20.

Council of the European Union. (2007, 12 December). Conclusions of the council and of the representatives of the governments of the member states, on improving the quality of teacher education. *Official Journal of the European Union*, C300/6–9.

Eurydice Network. (2006). *Quality assurance in teacher education in Europe*. Brussels, Belgium: European Commission.

Feuer, M. J., Floden, R. E., Chudowsky, N., & Ahn, J. (2013). *Evaluation of teacher preparation programs: Purpose, methods, and policy options*. Washington, DC: National Academy of Education.

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching: Theory and Practice, 15*(2), 273–289.

Hammerness, K, Darling-Hammond, L, Bransford, J., with Berliner, D., Cochran-Smith, M. McDonald, M., & Zeichner K. (2005). How teachers learn and develop. In Darling-Hammond L., Bransford J. (Eds.), Preparing teachers for a changing world (pp. 358–389). San Francisco, CA: Jossey-Bass.

Howell, D. C. (2013). *Statistical methods for psychology* (8th ed.). Belmont, CA: Wadsworth/Cengage Learning.

Ingersoll, R. (2007). *A comparative study of teacher preparation and qualifications in six nations*. CPRE research reports. Retrieved from http://repository.upenn.edu/cpre_researchreports/4

Ingvarson, L., Schwille, J., Tatto, M. T., Rowley, G., Peck, R., & Senk, S. (2013). *An analysis of teacher education content, structure and quality assurance arrangements in TEDS-M countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Kennedy, M. (1999). The role of preservice teacher education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession* (pp. 54–85). San Francisco, CA: Jossey-Bass.

Levine, A. (2006). *Educating school teachers*. Washington, DC: The Education Schools Project.

Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software, 9*, 1–19.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). London, UK: Chapman and Hall.

Mikitovics, A., & Crehan, K. D. (2002). Pre-professional skills test scores as college of education admission criteria. *Journal of Educational Research, 95*(4), 215–223.

National Board for Professional Teaching Standards. (2016). *What teachers should know and be able to do*. Arlington, VA: Author.

National Research Council. (2001). *Testing teacher candidates: The role of licensure tests in improving teacher quality (Chapter 2)*. Washington, DC: The National Academies Press Retrieved from https://doi.org/10.17226/10090

Patton, M. (1990). *Qualitative evaluation and research methods* (pp. 169–186). Beverly Hills, CA: Sage.

Tatto, M. T. (1996). Examining values and beliefs about teaching diverse students: Understanding the challenges for teacher education. *Educational Evaluation and Policy Analysis, 18*, 155–180.

Tatto, M. T. (2007). *Reforming teaching globally*. Oxford, UK: Symposium Books (Reprinted in 2009 Information Age Publishers).

Tatto, M. T., & Hordern, J. (2017). The configuration of teacher education as a professional field of practice: a comparative study of mathematics education. In J. Furlong & G. Whitty (Eds.), *Knowledge and the study of education: an international exploration* (Oxford comparative education series, pp. 255–274). Oxford, UK: Symposium Books.

Tatto, M. T., & Pippin, J. (2017). The quest for quality and the rise of accountability systems in teacher education. In J. D. Clandinin & J. Husu (Eds.), *International handbook of research in teacher education*. Los Angeles, CA: SAGE.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M., & Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA teacher education and development study in mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Wang, A., Coleman, A., Coley, R., & Phelps, R. (2003). *Preparing teachers around the world*. Princeton, NJ: Educational Testing Service.

Windschitl, M., Thompson, J., & Braaten, M. (2011). Ambitious pedagogy by novice teachers. *Teachers College Record, 113*(7), 1311–1360.

Youngs, P., & Grogan, E. (2013). United States of America (USA). In I. J. Schwille, L. Ingvrason, & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopedia* (pp. 255–272). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Zeichner, K. (2012). The turn once again toward practice-based teacher education. *Journal of Teacher Education, 63*(5), 376–382.

Chapter 4
The Intended, Implemented, and Achieved Curriculum of Mathematics Teacher Education in the United States

Maria Teresa Tatto and Kiril Bankov

Abstract The first part of this chapter describes the development of a systematic model for the comparative study of the curriculum of teacher education programs used in the Teacher Education and Development Study in Mathematics (TEDS-M). We use the United States as a case study to describe in depth the intended, implemented, and achieved curriculum of teacher education. We present in detail the results of the curriculum analysis for programs preparing future primary and secondary teachers (or the intended curriculum). We contrast these results with the number of clock hours dedicated to different areas of teacher education, and with future teachers' perception of their opportunities to learn (as indicators of the implemented curriculum), and with future teachers' knowledge at graduation (as indicators of the achieved curriculum).

We discuss the implications of our findings for the study of the teacher education curriculum in an era of increased accountability.

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

M. T. Tatto (✉)
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: teresa.tatto@asu.edu

K. Bankov
University of Sofia, Sofia, Bulgaria
e-mail: kbankov@fmi.uni-sofia.bg

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_4

Introduction

While the notion of what should be taught to whom, when, and how has been a subject of concern, discussion, and contestation in schools and universities since at least the 1600s, it was not until the mid-1900s when formal school curriculum became a subject of systematic study within and across nations (Husen, 1967a, 1967b). The shift came after scholars affirmed that the curriculum can only have an impact on learning if students had the *opportunity* to learn it. While scholars had simply defined opportunity to learn (OTL) as "time allowed for learning" (e.g., Carroll, 1963), Husen was the first to define and use this term to explain the quality and depth of student learning. The concept has expanded to explore the links between the curriculum and OTL and teachers' perceptions of students' OTL, which have been found to correlate highly with student scores in mathematics test items (Mullis et al., 2007). The reframing of school learning as dependent upon access to opportunities to study assigned content of adequate quality and with sufficient depth not only revealed new paths to examining education access, equity, and social justice (McDonnell, 1995), but also brought to the forefront the importance of considering the connections between the intended, implemented, and achieved curriculum as a comprehensive measure of school efficiency.

Many studies have explored the influence of opportunities to learn the school curriculum on students' learning, and many of these studies have considered teaching quality an important mediating factor in school learning. Fewer studies exist that have explored the influence of the teacher education curriculum on teacher education students' OTL. Even fewer studies have examined the link between opportunities to learn and what is learned in teacher education, or how what teachers know is used when they interpret and convey curriculum intent in classroom situations. For a long time, educators assumed that future teachers learned the intended content and that passage through teacher education properly equipped them to be effective teachers. These assumptions have been brought into question (Levine, 2006). While critics of teacher education have always existed, future and current teachers seem to be the most severe critics (Kennedy, 2015), likely because they do not feel fully prepared upon graduation. This is an important issue as research has shown that the time a teacher dedicates to teaching different subjects in the primary curriculum—i.e., curriculum coverage—is correlated with how comfortable that teacher feels with the content with the implication that they sometimes neglect to teach what they haven't been effectively taught (Buchmann & Schmidt, 1981). While secondary teachers are presumed to be subject specialists and thus have a better knowledge of their subject, some seem to lack knowledge conveying conceptual understandings and abstract reasoning to pupils (Tatto et al., 2012). Thus, the study of teacher education OTL and their impact on future teachers' knowledge is important in helping clarify the factors that contribute to teaching quality and ultimately pupil learning.

This chapter presents the results of a study of the opportunities to learn contained in the teacher education curriculum in mathematics using the Teacher Education and Development Study in Mathematics (TEDS-M) data in the United States.

We seek to answer the following research questions for programs preparing future primary[1] and secondary teachers:

(a) What is the content of the curriculum as intended by teacher education programs as measured by the analysis of teacher education programs' syllabi?
(b) What is the implemented curriculum as measured by the number of clock hours dedicated to different domains as reported by program administrators, and by opportunities to learn as reported by future teachers?
(c) What is the achieved curriculum as measured by the scores attained by future teachers in mathematical and mathematical pedagogy knowledge assessments?

While cross-national analysis of the curriculum of teacher education using the TEDS-M data has been published elsewhere (Tatto & Hordern, 2017), in this chapter we chose to focus on the U.S. as a case to explore in more depth the connections between the intended, implemented, and achieved curriculum of mathematics teacher education than that allowed in previous publications.

The U.S. provides a unique best case scenario (Patton, 2002) to study curriculum alignment after decades of sustained efforts to develop standards for high-quality teaching and teacher education in mathematics. Since 1920 when the National Council of Teachers of Mathematics (NCTM) was founded, there has been sustained work on the development of standards to procure high-quality mathematics teaching and learning for each and every student. Later in 1991 the Association of Mathematics Teacher Educators (AMTE) was founded, to share ideas on effective ways of promoting the NCTM Standards, NCSM (the National Council of Supervisors of Mathematics), and MAA (Mathematical Association of America) recommendations on teaching school mathematics and developing programs to improve the mathematics education of practicing and future teachers. More recently in 2017, the AMTE released the *Standards for Preparing Teachers of Mathematics* whose designers see as:

> aspirational, advocating for practices that support candidates in becoming effective teachers of mathematics who guide student learning […] these standards will guide the improvement of individual teacher preparation programs and promote national dialogue and action related to the preparation of teachers of mathematics. (Association of Mathematics Teacher Educators, 2017, p. xi)

Thus, mathematics teachers and teacher educators have worked very hard in the creation and development of a coherent field.

Concerning the mathematics school curriculum, while standards existed since the early 1990s, individual and joint efforts to develop state curriculum standards intensified in the early 2000s when each state in the U.S. had developed their own learn-

[1] Note that in the U.S. 'primary' usually refers to grades K-3, while 'elementary' is used for grades K-5 or K-6. In this chapter we use the term elementary as used in TEDS-M (see Tatto et al., 2012, pp. 29–32 for specific definitions within countries as to what grades are included as primary or secondary).

ing standards. These efforts were followed in 2009 by the creation and implementation of the Common Core State Standards (CCSS) for Mathematics as an attempt at creating a more coherent school mathematics curriculum across the U.S. states.

Together these efforts signal a clear and sustained networked agenda that should result in strong curricular alignments. Our intention in this chapter is to explore this assumption in light of our research questions while providing an information-rich case intended to achieve depth of understanding (Patton, 2002). We hope that this work can serve as a template for future within-country analysis of the teacher education curriculum.

Literature Overview

Educational researchers have long pondered how to improve school learning, with most studies investigating the relationships between teaching and pupil learning outcomes. In a comprehensive review of the literature, Darling-Hammond, Holtzman, Gatlin, and Heilig (2005) showed that the research considered how a range of teachers' characteristics, such as degree, gender, and age, correlated with pupils' learning as measured by assessments. In most of these early studies, teacher preparation was understood to be represented by whether teachers had received a credential or whether they had higher levels of study, such as master's degrees, the assumption that teacher education provided the knowledge needed for teaching was rarely questioned.

The most significant research advances documenting how teachers learn to teach occurred in primary mathematics teaching. The work of Lampert, Ball, and their collaborators has moved the field forward in significant ways (Lampert & Ball, 1998; Lampert et al., 2013). Their work is distinctive in the in-depth exploration on how teachers learn key mathematics concepts and how they manage to teach them to diverse populations.

In 1991, Tatto et al. published the results of a comprehensive study of three approaches to teacher education in Sri Lanka. This was the first study that explored the links between programs' characteristics (such as the program's curriculum and OTL), and the outcomes of teacher preparation—for instance, links between graduates' knowledge of mathematics and mother tongue, teaching practices, and pupils' learning. The study showed that programs' selectivity, future teachers' previous knowledge of the subject, and the opportunities to learn designed by the programs were important influences on teachers' subsequent knowledge for teaching.

Calls for the careful study of curriculum variation cross-nationally intensified in the mid-1990s. A particularly important contribution was made by a literature review titled "Considerations of content and the circumstances of secondary school teaching," authored by Grossman and Stodolsky (1994a, 1994b). After providing a framework for understanding how a variety of contexts (schools, departments, professional communities, disciplines, and students) interact in forming the content and circumstances of secondary schooling, the authors call for research on teacher education, including systematic investigations of the formation of beliefs of prospective teachers, disciplinary socialization (both in liberal arts coursework and subject-specific methods courses), and differences in teaching across subjects both at the elementary and secondary levels. In addition, the authors suggest the need for further cross-national research to gain a better understanding of curriculum variation across similar contexts in different countries.

In the early 2000s, in an international review of the literature commissioned by the National Academy of Science, Tatto (2000) argued that the established method to evaluate the influence of teacher education on teacher quality and pupils' learning using credentials and years of study as independent variables in regression equations was problematic as these represent highly varied opportunities to learn to teach, likely as varied as those already understood from pupils, and that to fully understand what teachers knew as a result of teacher education, it was necessary to assess such knowledge against the curriculum studied. The result of this exploration was the development of the first cross-national large-scale study that assessed future teachers' knowledge as an outcome of teacher education looking at the OTL provided by programs (see Tatto et al., 2012). Researchers who participated in the TEDS-M study have used the TEDS-M database to investigate the relationships among opportunity to learn mathematics content and teaching methods among future teachers. Researchers also found that while there was a great deal of homogeneity in the topics that programs in different countries considered "basic" in the curriculum, there was significant heterogeneity within and among the 17 participating country-systems concerning more advanced topics (Blömeke, 2012). Other studies have explored the relationships between OTL and performance in the TEDS-M assessment as concerns Mathematics Content Knowledge (MCK) and Mathematical Pedagogical Content Knowledge (MPCK); see Tatto and Senk (2011).

Separately, researchers have examined how the kind of content courses taken (calculus, geometry, etc.) by teachers correlated with pupil learning. While many studies have used the so-called value-added approach, questions about the curriculum of teacher education continue to emerge. For instance, Boyd, Grossman, Lankford, Loeb, and Wyckoff (2009) used survey data on teachers' experiences across 31 programs in 18 institutions in New York, as well as multiple New York administrative data sets that included demographic data for students, teachers, and schools for each year from 2000–2001 through 2005–2006 to analyze the impact of teacher education on student achievement. The authors found significant variation across programs and institutions in the average effectiveness of their graduates, defined by value-added to student achievement in math and English Language Arts (ELA). Controlling for teacher background, two types of program characteristics were found to be significant predictors of effectiveness: program experiences that linked preparation to practice (e.g., student-teaching, capstone project, studying curricula, listening to a child read, planning a reading lesson, analyzing student math work) and subject content requirements. In a different but related study, Boyd et al. (2008) used multiple data sources (program documents, interviews, and surveys) to describe the pre-service training experiences and programs for elementary teachers of New York City public schools. The authors described the types of programs, the characteristics of teachers who entered the programs, the characteristics of teacher educators, and the curriculum. Although there were a large number of programs and program types in the area, the authors found programs to be more similar to one another than they were different. Boyd et al. did find some programs placing greater emphasis on mathematics, however, while others placed more emphasis on classroom management and working with families. They argued that these differences were minimal overall. The little between-program variation, at least in this narrow geographic setting, suggests that greater variation within programs may exist regarding individuals' opportunity to learn.

Researchers also have explored the relationship of the curriculum found in different routes into teaching and effectiveness with mixed results. For instance, Constantine et al. (2009) undertook an evaluation of teachers trained through different routes to certification. They used a purposive sample of 87 matched pairs of teachers who had certificates from traditional and alternative programs and randomly assigned them to pupils to analyze the impact of teacher education characteristics. Overall, neither the amount, the timing, nor the content of coursework in alternative or traditional teacher preparation were associated with teacher effectiveness as measured by pupil achievement of their respective students. Subgroup analysis revealed that traditionally certified teachers were more effective than alternatively certified with respect to mathematics outcomes. Traditionally certified teachers with three to four years of experience and reduced coursework in their preparation were more effective in mathematics than their alternatively trained counterparts.

While this does not represent a systematic review of the literature, it signals the direction of teacher education research in the United States, which has moved from the study of teachers' understandings of mathematical concepts (Clark & Peterson, 1986) to studies of the value-added of teacher education as indicated by pupils' learning (Boyd et al., 2009). The use of value-added methods illuminates some important aspects of the relationship between teaching and learning, but it does not help explain the how and why of the results obtained, and, specifically, does not assess the actual knowledge attained by teachers who engaged with the teacher preparation curriculum in a variety of programs.

This chapter's research questions are based on the notion that to understand the relationship between teacher education and teacher learning, it is important to explore in detail what the curriculum of teacher education is like within teacher education programs, paying particular attention to the intended, implemented, and achieved curriculum and its relationship with the knowledge to teach mathematics that future teachers attain before they begin to teach. The TEDS-M study provides the necessary data to explore these questions.

The TEDS-M Study of the Teacher Education Curriculum: Methods and Rationale

To collect data on the teacher education curriculum, the TEDS-M study used a fourfold strategy. To obtain indicators of the intended curriculum, country teams examined representative samples of syllabi in the programs selected for study and coded the topics as represented in these documents using a coding system similar to that used to analyze the school curriculum by McKnight and collaborators (1987) and by Valverde, Bianchi, Wolfe, Schmidt, and Houang (2002). A questionnaire asking future teachers what OTL they experienced as a result of moving through the program gave us indicators of the implemented/experienced curriculum. Another questionnaire, which asked administrators to report the number of contact hours dedicated to different domains during the teacher education program, gave us additional indicators of the implemented curriculum. A fourth source of data came from the results of the assessment of the mathematics content and mathematics pedagogical content knowledge of future teachers at the end of their program, providing us with indicators of the achieved curriculum. In TEDS-M, we added an external "relevance/alignment" dimension to the curriculum study by analyzing the degree to which the intended teacher education curriculum reflected the intended school curriculum as reflected in mathematics curricula at the country level.

An important step was to develop a codebook to analyze the syllabi. The study team developed a comprehensive typology of topics that emerged from a pilot study of the teacher education curriculum across the 17 countries that participated in TEDS-M. These topics were then used to code the syllabi and to map the intended, implemented, and achieved curriculum of teacher education, including such areas as (a) university- or tertiary-level mathematics; (b) school-level mathematics; (c) mathematics education/pedagogy; (d) education /pedagogy; (e) accommodating classroom diversity and reflection on practice; and (f) learning from school experience and the practicum.

The Intended Curriculum of Schools and Teacher Education Programs

Primary- and secondary-school curriculum standards are assumed to be an important influence on teacher education, as educators may decide to emphasize these standards, to varying degrees, in their program's own curriculum. Accordingly, TEDS-M explored the correspondence between mathematics school curricula and the teacher education mathematics curriculum.

The Mathematics Curriculum of Schools

TEDS-M analyzed the official school mathematics curricula from the participating countries using a revised curriculum analysis method from TIMSS 1996 (U.S. Department of Education, National Center for Education Statistics, 1997; Valverde et al., 2002). Curriculum analysis is a systematic method for obtaining comprehensive information from curricular documentation and materials. In the case of TEDS-M, the curriculum analysis examined the current primary- and secondary-school curriculum or standards in mathematics for each of the primary, middle school, and high school grades in each of the TEDS-M participating countries.[2] For countries with national or centralized education systems, the analysis of the school curriculum relied on national-level documents; for countries with a decentralized system, the analysis relied on the curriculum and /or standards used by the majority of the country's schools.

The analysis of the primary- and secondary-school mathematics curriculum and standards can be compared with the analysis of the teacher preparation curriculum to uncover the degree of alignment of teacher preparation programs with what teachers will eventually teach in schools.

The Teacher Education Curriculum

Building upon previous work analyzing primary and secondary curricula (Valverde et al., 2002), the TEDS-M research team developed a framework and a methodology for analyzing the curriculum of teacher education. This methodology allowed the exploration of the content associated with actual routes and programs, as well as the possible influence of this content on the professional knowledge of future teachers.

While there is usually ample documentation of school curricula, we found the same does not apply to curricula for teacher preparation programs. For instance, many countries that have national standards or syllabi for school subjects do not have national standards for teacher education. In teacher education, many courses either do not use detailed written syllabi, or the syllabus used differs from educator to educator and/or from institution to institution. In addition, if there are documents describing the expected or anticipated learning outcomes of the practicum, their analysis may prove challenging.

The initial phase of the study of the teacher preparation curricula therefore examined documents prescribing or describing the national curriculum of teacher education in countries where these existed, or an aggregated analysis of the local or institutional curricula when no documents at the national level were available.

The second phase in the analysis of the curricula of mathematics teacher education was at the institutional program level. For this, a protocol was developed to analyze curriculum documents from the teacher education mathematics curricula in the selected routes and programs in each participating country. The protocol examined the content covered by courses in the mathematics teacher education curriculum as an indicator of performance expectations for teacher certification or licensing.

[2] The TEDS-M Study Reports use the terms "primary generalist" to refer to future teachers prepared to teach all grades and subjects of primary schooling, "primary specialist" to refer to future teachers prepared to teach the primary grades with a mathematics specialization, "lower secondary" to refer to future teachers prepared to teach up to Grade 10 maximum, and "lower and upper secondary" high school to refer to programs that prepare teachers to teach up to Grades 11 and above. In this chapter, we retain the use of primary generalist and primary specialist, but use the United States' terms middle school and high school instead of lower and upper secondary, since we are using the U.S. data as our core example throughout this chapter.

Table 4.1 Guidelines for syllabi selection[a]

Academic courses syllabi (mathematics, mathematics, pedagogy and general pedagogy only)		Field experience/teaching practice syllabi
Type of course	Data-gathering methodology	
Required courses with one instructor	Obtain all syllabi	
Required courses with multiple sections and/or multiple instructors	All, but for courses with multiple sections taught by multiple instructors, randomly select syllabi from two instructors only	Get syllabi or practicum report or contract or any other institutional document related to the practice
Elective courses	Select those courses that are taken by most (at least 50%) of future teachers, if there are any[b]	

[a]*Source*: TEDS-M Manual for Syllabi Analysis at the Institutional Program Level
[b]Institutions may decide to use 2 steps: (1) compile a list of the elective courses across the program most frequently chosen by, for example, the last three cohorts of future teachers; (2) from that list select the 5–10 most highly ranked across the three cohorts

These analyses produced an initial profile of the intended curriculum in mathematics teacher education in terms of the knowledge, pedagogy, and dispositions future teachers are exposed to as they get ready to teach.

There were two options for the selection of the syllabi:

(a) Census of all relevant syllabi in all institutions of teacher education in TEDS-M sample. This option was comprehensive and required collection and coding of all relevant syllabi from all institutions.
(b) Census of all relevant syllabi in all institutions in a purposeful sample of institutions in the TEDS-M sample. The institutions could be selected because they train the most teachers, are the most influential, have the broadest subject matter coverage, or are seen as flagship institutions (e.g., those seen as providing exemplary practice).

Table 4.1 shows the guidelines for selection of syllabi from institutions.

All TEDS-M countries that completed the syllabi analysis tasks undertook a census of all relevant syllabi in all institutions of teacher education in the TEDS-M sample (aforementioned option "a"). Table 4.2 shows the average percentage of analyzed syllabi for the countries that completed the syllabi analysis tasks. Our data revealed some variability across institutions within and across countries. In some countries, it is possible to see that not all institutions had syllabi available for the courses they teach. This is often the case in general pedagogy and mathematics courses (often mathematics classes are not taught as part of the teacher education program, but as part of the mathematics department within the same institution, or in a different institution, as is the case in Singapore; thus, the syllabi may not be accessible to the teacher education institution). In some cases, the curriculum is assumed to be the same across institutions, especially in institutions in centralized countries and, in some cases, the mathematics pedagogy content and the mathematics content are integrated in the same syllabus.

Table 4.2 The number and average percentage of syllabi submitted and analyzed by countries

	Primary level syllabi				Secondary level syllabi			
Country	Math content	Math pedagogy	General pedagogy	Total	Math content	Math pedagogy	General pedagogy	Total
Botswana	1/100	1/100	1/100	3/100	46/100	7/100	1/50	54/98
Canada	9/36	24/86	51/71	84/86	20/15	31/65	32/45	83/65
Chile	30/100	63/100	555/100	648/100	24/100	41/100	207/100	272/100
Chinese Taipei	20/100	20/100	207/100	247/100	203/100	44/100	124/100	371/100
Germany	No information was provided				No information was provided			
Malaysia	Only one version (national level only)				43/100	9/100	19/100	71/100
Norway	31/97	0[a]	31/97	62/97	33/73	5/11[b]	37/82	75/82
Oman	No primary level				62/100	4/100	26/100	92/100
Philippines	164/64	16/53	186/48	366/53	390/62	18/31	248/43	656/51
Poland	4/100	124/96	1099/97	1227/97	782/95	110/98	146/81	1038/93
Singapore	40/85	20/100	40/85	100/85	0[d]	16/100	20/100	36/100
Spain	3/100	144/100	441/100	558/100	No secondary level			
Switzerland	42/75	108/100	675/100	825/100	71/91	36/91	177/90	284/91
Thailand[b]					273/100	60/100	237/99	570/99
US	177/73	92/87	634/100	903/100	657/87	73/78	524/96	1254/100
US[c]					258/82	74/91	312/100	644/100

[a]Mathematics Pedagogical content is included in the Mathematics Content
[b]For some institutions, Mathematics Pedagogical content is included in the Mathematics Content
[c]Both primary specialists and secondary levels
[d]There are not math courses in the institutions of this level

The third phase consisted of assigning codes to the selected syllabi. For this purpose, each syllabus was divided into smaller pieces called units and blocks. Units are large sections in the document describing the different types of knowledge that the program provides. In most cases, a typical unit is a course, which is represented by a syllabus. However, if a single course consists of two parts, e.g., classroom activities (lectures and discussions) and a field experience (future teachers' practice teaching in schools), then the course may be divided into two units. Another example of a course divided into separate units is when the course consists of some parts related to mathematics pedagogy and other parts that are mathematics content related to the school curriculum. In such cases, the course may be divided into two corresponding units: one for the parts that are mathematics-pedagogy related and the other for the rest of the course.

After identifying units, we engaged in the next task in the syllabi analysis process, that of identifying blocks. A block is a portion of a unit containing information about the specific topics that are covered by the particular class or course. Depending on the number of topics in the unit, the unit may be divided into many blocks, few blocks, or only one block. The aim of such partition of the syllabi is to analyze the units and the blocks in order to produce information about the content that is covered by each course.

Syllabi Analysis Framework The syllabi analysis framework describes the topics and the performance expectations that may be stated in each syllabus and provides corresponding numerical codes. The TEDS-M syllabi analysis framework includes three content topics areas: mathematics, mathematics pedagogy, and pedagogy. The framework also includes performance expectations to examine what future teachers are expected to do with the content. Performance expectations might include, for example, cognitive requirements in a syllabus or requirements for active participation during class, writing tests or papers, or field-based activities. In this section, we only focus on the content topic areas and do not consider performance expectations.

The structure of the coding framework is hierarchical, with unique codes for topics at each level: for example, 1.1.3 Geometry and 1.1.3.2 Euclidean Geometry. Topics with fewer numbers are in a higher level of the hierarchy (more general), while those with more numbers are in a lower level (more precise). Each block in a syllabus must be assigned one or more codes from the framework. The coders may use either the low-level topics (i.e., 1.1.3.2 Euclidean Geometry), or higher level (i.e., 1.1.3 Geometry). The goal, however, was to use the lowest possible code level to obtain the most concrete information about the content of the block.

Limitations

Analyzing documents is a complex task. While there were strict guidelines and detailed procedures to be followed, each country organized the coding of its own syllabi using local experts. The TEDS-M study organized several workshops to train each country's experts in this process. The purpose was not only to explain the procedure, but also to ensure a uniform understanding of partitioning the syllabi into units and blocks and assigning the codes. This was a way to diminish the subjectivity in the coding process. Local experts did not translate syllabi into other languages, but instead coded the syllabi in the language in which they were written. Translating all syllabi into English, besides being very resource-intensive, would also have created further opportunities for errors.

Some limitations still existed. For instance, in the selection of the curriculum documents, it became clear that syllabi falling within the same subject categories were highly variable (e.g., by format, structure, and content). In some institutions, the syllabi were short and general, giving just broad information about the courses. Other syllabi presented detailed descriptions of the topics that were covered by the programs. For some institutions, the syllabi were available on the program's website. In other cases, they had to be requested by contacting the appropriate person(s) at the institution. In the latter cases, it was sometimes not possible for all syllabi to be collected.

Even though the document analysis consists of low-inference methods, one should nevertheless consider that the coding process is subjective. Because of the variability in the syllabi, there is not a universal way to partition the documents into units and blocks and assign the codes. Since these steps were completed by a country's experts, it involved much of the experts' interpretation of the text in the document.

The Implemented Curriculum of Teacher Education Programs

Opportunities to Learn As Reported by Programs and Experienced by Future Teachers

The TEDS-M study included a survey of programs where questions about the clock hours dedicated to different domains were asked as an indicator of OTL provided to future teachers (see Table 4.18 for the definitions of these domains). The intent was to enable exploration of whether future primary teachers prepared as generalists had different times allocated to curriculum areas than those prepared as specialists. Similarly, these questions addressed an interest in the time allocation for future middle and high school teachers. These survey questions allowed us to have a measure of the OTL given to future teachers according to different approaches to teacher education.

In addition, the TEDS-M study included a number of questions addressed to future teachers to allow exploration of the OTL that future mathematics teachers report having across countries. These questions were used to develop five OTL scales to numerically represent aggregated future teachers' responses to the questions, and to make it possible to statistically analyze the data. We next explore these five OTL areas.

Opportunity to Learn University-Level Mathematics The scales in this area enabled exploration of whether future teachers have studied key mathematics topics (e.g., geometry-related topics, algebra, number theory, calculus, functions). Because

OTL in this area often occur before future teachers enter teacher preparation, these questions asked future teachers whether they had *ever* studied such topics.

Opportunity to Learn School-Level Mathematics These scales included questions that allow exploration of the interaction between mathematics topics studied by future teachers in their teacher preparation program and the school mathematics curriculum. In addition, this section included questions on the emphasis given to learning mathematics in more depth.

Opportunity to Learn Mathematics Education/Pedagogy These scales included questions that allow exploration of the interaction between mathematics content and pedagogy. Additional scales included questions about the use of learning strategies in mathematics. Future teachers were asked to indicate whether they had studied each topic as part of their teacher preparation program. Other questions asked how often future teachers engaged in a number of activities and learning strategies in mathematics in their teacher preparation program.

Opportunity to Learn General Knowledge for Teaching These scales included questions about topics considered relevant for all teachers to understand, such as educational theory, general principles of instruction, classroom management, and curriculum theory. As with the questions in the previous knowledge areas, these questions asked future teachers whether they studied such topics as part of their teacher preparation program.

Opportunity to Engage in School Experiences and in a Practicum These questions asked future teachers whether they spent time in the classroom in a primary or secondary school and, if so, how long; whether they had a school supervisor assigned to them; whether they engaged in particular activities and at what levels; and whether they found the school experience helpful. An additional set of questions asked about diverse characteristics of the practicum (e.g., the role of the mentor, feedback received, standards, methods used, and level of mathematics knowledge and pedagogy of the classroom teacher or mentor).

The Achieved Curriculum of Teacher Education Programs: Assessment of Mathematical Content and Mathematical Pedagogical Content Knowledge

The TEDS-M study included two assessments to measure future teachers' Mathematical Knowledge for Teaching (MKT), consisting of two constructs: mathematics content knowledge and mathematics pedagogical content knowledge. (For the theoretical origins of these constructs, consult Tatto et al., 2008.) The mathematics content knowledge framework included measures of knowledge for the domains of number, geometry, algebra, and data, and measures for cognitive domains, including application, knowing, and reasoning (Mullis et al., 2007).

The mathematics pedagogical content knowledge framework included measures of mathematics curricular knowledge, knowledge of planning for mathematics teaching and learning, and enactment of mathematics for teaching and learning. The frameworks are presented in detail in Tatto et al. (2012) and are included in Appendix 4.1, Tables 4.15, 4.16, and 4.17 of this chapter.

Findings: The Intended, Implemented, and Achieved Curriculum

In this section, we report the intended curriculum of schools and teacher education programs, and the implemented and achieved curriculum of teacher education at the primary and secondary levels in the United States. We begin with an examination of the school curriculum expectations for teachers.

The Intended Curriculum of Schools

The data below come from the analysis of the school curriculum guidelines or standards in the United States at the time of the study. The intended curriculum of schools was considered an important input for the TEDS-M study, as it greatly influences what is studied in teacher education programs.

Table 4.3 represents the information for the school mathematics curriculum in grades 1 to 8 in the United States at the time of the study. Each bullet shows the grade at which the corresponding topic is to be taught. For example, congruence and similarity is intended at Grades 4, 7, and 8; patterns, relations, and functions is intended for Grades 1 to 8; and vector geometry is not intended to be taught in Grades 1 to 8.

Table 4.3 shows that the mathematics school curriculum in the United States focuses on four main domains: number, geometry, algebra, and data. As could be expected, the study of numbers begins in Grade 1. Fractions and decimals are also introduced in Grade 1. In the United States, there is a relatively high emphasis on the estimation and number sense concepts topics. The geometry part of the U.S. curriculum gives a rich opportunity for students to study Euclidean, transformational, and 3D geometry. As expected, given the curriculum reforms implemented during the time of the analysis, algebraic and data ideas are included in the early grades curriculum (Grades 1 and 2). Vector and analytic geometry, trigonometry, and elementary analysis are not part of the intended mathematics curriculum for Grades 1 to 8 in the United States.

Table 4.3 The grades in which mathematics topics are studied according to the analysis of the school curriculum

Domains	Topics	Grade							
Number		1	2	3	4	5	6	7	8
	Whole numbers	●	●	●	●	●	●		
	Fractions and decimals	●	●	●	●	●	●	●	●
	Integer, rational and real numbers							●	●
	Other numbers and number concepts		●			●	●	●	●
	Estimation and number sense concepts		●	●	●	●	●	●	●
	Ratio and proportionality							●	●
Geometry	Measurement units	●	●	●	●	●	●	●	●
	Computations and properties of length, perimeter, area and volume			●	●	●	●	●	●
	Estimation and error		●	●	●	●	●	●	
	1-D and 2-D coordinate geometry			●	●	●	●	●	●
	Euclidean geometry	●	●	●	●	●	●	●	●
	Transformational geometry	●	●	●	●	●	●	●	●
	Congruence and similarity				●			●	●
	Constructions with straightedge and compass								
	3-D geometry	●	●	●	●	●	●	●	●
	Trigonometry and analytic geometry								
	Vector geometry								
	Simple topology								
Algebra	Patterns, relations, and functions	●	●	●	●	●	●	●	●
	Equations and formulas		●	●	●	●	●	●	●
	Elementary analysis								
Data	Data representation and analysis	●	●	●	●	●	●	●	●
	Uncertainty and probability		●	●	●	●	●	●	●

The Intended Curriculum of Teacher Education Programs

The data that emerged from the analysis of the syllabi of teacher education programs is included below and reflects the key content areas commonly covered: school mathematics (i.e., the mathematics covered in the school curriculum), academic or university mathematics, mathematics pedagogy, general pedagogy, and the activities that shape the field experience in schools.

School Mathematics Content

Table 4.4 displays the school curriculum topics covered across the different program types in the United States, which include programs preparing primary generalists (Grade 6 maximum), primary mathematics specialists, middle school (Grade 10 maximum), and high school (to Grade 11 and above) corresponding to the headings US2, US4, US5, and US6 respectively. The Table 4.4 shows that no topics were

Table 4.4 Percentage of teacher education institutions that taught each of the mathematics school curriculum topics according to the summary from the syllabi analysis

School mathematics topics	Percent of institutions/TPU cover the topic			
	US2[a]	US4[b]	US5[c]	US6[d]
	N = 63	N = 17	N = 19	N = 59
Whole numbers	35	29	42	10
Fractions and decimals	46	41	47	15
Integer, rational and real numbers	40	35	42	19
Other numbers and number concepts and number theory	37	29	32	20
Estimation and number sense concepts	33	24	32	12
Ratio and proportionality	30	29	37	5
Measurement units	14	12	16	3
Computations and properties of length, perimeter, area and volume	0	0	0	0
Estimation and error in measurement	10	18	16	8
1-D and 2-D coordinate geometry	21	18	16	8
Euclidean geometry	16	24	32	22
Transformational geometry	27	24	37	19
Congruence and similarity	25	24	32	19
Constructions with straightedge and compass	21	29	32	15
3-D geometry	14	18	32	8
Vector geometry	3	6	5	5
Simple topology	2	0	0	0
Patterns, relations and functions	29	24	26	17
Equations and formulas	11	18	21	10
Trigonometry and analytic geometry	6	12	16	20
Data representation and analysis	30	29	32	20
Uncertainty and probability	24	41	53	29
Elementary analysis	5	12	16	19
Validation and structure	17	18	11	20
Other school mathematics topics	13	24	16	14

Note 2: A topic is considered "taught in the program" if at least one of the subtopics is taught in the program

[a]Programs preparing primary generalists (Grade 6 maximum)

[b]Programs preparing primary mathematics specialists

[c]Programs preparing middle school (Grade 10 maximum) teachers

[d]Programs preparing high school (to Grade 11 and above) teachers

covered in more than 53% of institutions in any program type, and only one was covered by more than 50% of institutions (uncertainty & probability). Six topics were covered in the syllabi of between one quarter and less than half of the programs that prepare future primary and middle school teachers (first three columns of data or US2, US4 and US5). These are (1) whole numbers; (2) fractions and decimals; (3) integer, rational, and real numbers; (4) other numbers, number concepts, and number theory; (5) ratio and proportionality; and (6) data representation and analysis. Eight additional topics were uniformly covered in the syllabi of fewer than 25% of the institutions across primary-middle school program types: (1) measurement units; (2) estimation and error in measurement; (3) 1-D and 2-D coordinate geometry; (4) vector geometry; (5) equations and formulas; (6) trigonometry and analytic geometry; (7) elementary analysis; (8) and validation and structure. Estimation and number sense concepts are seen as relatively important topics to master for the primary generalists and middle school future teachers, but receive less emphasis in the syllabi of institutions that prepare mathematics specialists for primary and high school grades. The topic uncertainty and probability received more attention in programs preparing middle school teachers, but also in those preparing primary specialists and high school teachers. Computations and properties of length, perimeter, area and volume were marked as not covered in the syllabi of the institutions of any program. Most of the topics were covered in the syllabi of close to 25% of the programs preparing high school teachers.

While the topics in Tables 4.3 and 4.4 are not identical, they are close enough to make comparisons between the intended mathematics school curriculum and that of teacher preparation programs in the United States. The number domain is well covered in the intended curriculum of teacher preparation programs, which maps well to the number coverage in the school mathematics curriculum. This is not the case for Euclidean, transformational, and 3D geometry. Not many institutions that prepare primary teachers cover these topics, though they are included in the school curricula in Grades 1–8. The same is true for algebraic topics (patterns, relations and functions), except for programs that prepare primary generalists and middle school teachers. The data domain is emphasized in both school and teacher preparation curricula.

Academic/University Mathematics Content

Table 4.5 shows that few topics were covered in more than 75% of institutions in US2 program type, while a larger percentage of programs preparing future secondary teachers include university mathematics topics in their curriculum. Topics that were covered with the lowest frequency across program types were differential geometry, set theory, and topology (only in US1 and US4 program types). Topics covered with low frequency in programs preparing primary teachers, but with higher frequency in programs preparing high school teachers, were: axiomatic geometry, analytic/coordinate geometry, linear algebra, abstract algebra, number theory, beginning calculus, multivariate calculus, discrete mathematics, probability, theoretical and applied statistics, and other mathematics topics. The topics that appear

Table 4.5 Percentage of teacher education institutions that taught each of the university mathematics topics according to the summary from the syllabi analysis

University mathematics topics	Percent of institutions/TPU cover the topic			
	US2[a]	US4[b]	US5[c]	US6[d]
	N = 63	N = 17	N = 19	N = 59
Axiomatic geometry (including Euclidean axioms)	11	35	37	54
Analytic/coordinate geometry	13	35	47	54
Non-Euclidean geometry (e.g. geometry on a sphere)	6	29	37	44
Differential geometry	2	6	5	5
Topology	2	0	0	3
Linear algebra	8	35	37	78
Set theory	8	18	21	12
Abstract algebra (e.g., group theory, field theory, ring theory, ideals)	2	18	16	69
Number theory	24	18	37	51
Beginning calculus topics (e.g., limits, series, sequences)	10	29	42	76
Calculus (e.g., derivatives and integrals)	11	53	47	81
Multivariate calculus (e.g., partial derivatives, multiple integrals)	5	18	16	69
Advanced calculus or real analysis or measure theory	0	6	0	36
Differential equations	2	6	5	34
Functional analysis, theory of complex functions	3	18	16	37
Discrete mathematics, graph theory, game theory, combinatorics	11	47	42	58
Probability	11	47	58	75
Theoretical or applied statistics	22	47	53	69
Mathematical logic	11	18	32	47
Other mathematics topics	14	47	47	66

Note 2: A topic is considered "taught in the program" if at least one of the subtopics is taught in the program

[a]Programs preparing primary generalists (Grade 6 maximum)

[b]Programs preparing primary mathematics specialists

[c]Programs preparing middle school (Grade 10 maximum) teachers

[d]Programs preparing high school (to Grade 11 and above) teachers

with the greatest frequency in the syllabi are: axiomatic geometry, analytic geometry, non-Euclidean geometry, linear algebra, calculus, probability, and theoretical and applied statistics.

Mathematical Pedagogy Content

Table 4.6 shows the percentage of programs covering different topics in mathematics pedagogy, which is relatively high for all programs. It is not surprising that "mathematics instruction" is taught in all programs. A high percentage of programs

Table 4.6 Percentage of teacher education institutions that taught each of the mathematics pedagogy topics according to the summary from the syllabi analysis

Mathematical pedagogical topics	Percent of institutions/TPU cover the topic			
	US2[a]	US4[b]	US5[c]	US6[d]
	N = 63	N = 17	N = 19	N = 59
Theories/models of mathematics ability and thinking	27	18	21	8
Nature and development of mathematics ability and thinking	43	47	37	27
Aspects of mathematical ability and thinking	73	71	68	64
Mathematical problems and solutions	73	76	74	61
Mathematics instruction	89	76	84	78
Developing of mathematics teaching plans	70	47	58	56
Analyzing/observing/reflecting on mathematics teaching	59	59	53	54
Knowledge of mathematics standards and curriculum	81	65	63	69
Studying and selecting textbooks and instructional materials	30	35	42	19
Methods of presenting main mathematics concepts	57	71	79	53
Foundations of mathematics	25	24	32	61
Context of mathematics education	37	41	47	36
Affective Issues (beliefs, attitudes, anxiety, etc.)	21	6	5	8

Note 2: A topic is considered "taught in the program" if at least one of the subtopics is taught in the program
[a]Programs preparing primary generalists (Grade 6 maximum)
[b]Programs preparing primary mathematics specialists
[c]Programs preparing middle school (Grade 10 maximum) teachers
[d]Programs preparing high school (to Grade 11 and above) teachers

covered "aspects of mathematical ability and thinking," "mathematical problems and solutions," "developing of mathematics teaching plans," "analyzing/observing/reflecting on mathematics teaching," "knowledge of mathematics standards and curriculum," and "methods of presenting main mathematics concepts." Fewer programs reported including "affective issues" and "theories/models of mathematics ability and thinking" in their syllabi.

General Pedagogy Content

The percentage of programs covering General Pedagogy Topics (Table 4.7) is relatively high for all programs. The only topic that is covered by fewer than 20% of the programs is "counseling, advising students, and pastoral care." "Methods of educational research" is more frequently found in syllabi of programs preparing generalists than in programs that teach more specialized mathematics.

Table 4.7 Percentage of teacher education institutions that taught each of the general pedagogy topics according to the summary from the syllabi analysis

General pedagogy topics	Percent of institutions/TPU cover the topic			
	US2[a]	US4[b]	US5[c]	US6[d]
	N = 63	N = 17	N = 19	N = 59
History of education and educational systems	68	71	74	69
Educational psychology	98	100	100	90
Philosophy of education	87	71	74	78
Sociology of education	98	94	100	93
Introduction to education or theories of schools	98	94	95	93
Principles of instruction	97	88	95	86
Methods of educational research	52	35	32	42
Classroom management	95	88	95	85
Assessment and measurement theory	59	41	63	58
Counseling, advising students, and pastoral care	10	18	16	12
Instructional media and operation	87	71	74	81
Practical knowledge of teaching	98	94	100	92

Note 2: A topic is considered "taught in the program" if at least one of the subtopics is taught in the program
[a]Programs preparing primary generalists (Grade 6 maximum)
[b]Programs preparing primary mathematics specialists
[c]Programs preparing middle school (Grade 10 maximum) teachers
[d]Programs preparing high school (to Grade 11 and above) teachers

Field Experience in Schools

Syllabi were analyzed to explore what topics were covered in field experiences. The results are displayed in Table 4.8. The most common and uniform expectations across the syllabi were to observe a teacher, to write a report of the observation, and to meet with a supervisor to discuss practicum matters. The expectation that future teachers oversee a class as a teacher was found in fewer than 25% of the institutions. The following topics were covered in fewer than 25% of programs: supervise or organize social activities, participate in school level administration, design and carry out an action research project, and attend professional conferences. The topic "deliver instruction other than mathematics" was found frequently in the syllabi of institutions preparing primary generalists, presumably because these teachers are expected to teach other subjects in addition to mathematics.

Discussion of the Intended Curriculum of Teacher Education Programs

As seen in the data above, the degree to which school mathematics topics are covered in the syllabi of U.S. teacher education programs varys widely across programs. Number topics; transformation geometry; congruence and similarity; and data, uncertainty, and probability receive more attention in the programs. The

Table 4.8 Percentage of teacher education institutions that taught each of the field experience topics according to the summary from the syllabi analysis

Field experience topics	Percent of institutions/TPU cover the topic			
	US2[a]	US4[b]	US5[c]	US6[d]
	N = 63	N = 17	N = 19	N = 59
Observe a teacher	79	47	47	66
Serve as a teacher	13	12	11	14
Design instruction in mathematics	34	16	19	37
Design instruction other than mathematics	40	29	32	15
Deliver mathematics instruction	31	18	18	31
Deliver instruction other than mathematics	62	35	32	19
Supervise non-mathematics instruction	29	12	16	14
Assess students (full responsibility)	46	18	32	36
Supervise or organize social activities	13	0	0	19
Work with parents	38	18	26	37
Participate in formal school meetings of teachers	29	24	21	37
Participate in school level administration	5	0	0	2
Design and carry out an action research project	21	6	5	19
Discuss practicum experience with peers	46	12	26	44
Write report of observing teaching	63	35	47	56
Meet with supervisor to discuss practicum	63	47	47	56
Attend professional conferences	16	12	11	15
Other	6	6	5	3

Note 2: A topic is considered "taught in the program" if at least one of the subtopics is taught in the program

[a]Programs preparing primary generalists (Grade 6 maximum)

[b]Programs preparing primary mathematics specialists

[c]Programs preparing middle school (Grade 10 maximum) teachers

[d]Programs preparing high school (to Grade 11 and above) teachers

syllabi of the programs preparing high school teachers include fewer school mathematics topics and more university mathematics topics. Programs preparing primary generalists give less emphasis to including university mathematics in their curriculum. As could be expected, the university mathematics topics that appear with more frequency in the syllabi overall are axiomatic and analytic geometry, linear algebra, calculus, and probability and statistics. Relatively high percentages of institutions across all U.S. programs include most of the mathematics pedagogy topics. Less attention is paid to theories of mathematics ability and thinking, selecting textbooks, and affective issues. General pedagogy topics are present in the syllabi of almost all institutions. The only exceptions are topics having to do with counseling, advising students, and pastoral care. For field experiences, the more frequent topics are those connected with the observation of teachers, writing reports on the observations, and meeting with supervisors. Very few syllabi include topics related to participation in school-level administration. Based on the analysis of the data, future mathematics teachers in the United States, whether at the primary or secondary levels, spend a large proportion of their preparation time program exposed to OTL mathematics pedagogy and general pedagogy, and less time studying mathematics content, presumably under the assumption that such knowledge has been acquired elsewhere.

When comparing the expected curriculum of teacher education with the expected curriculum to be taught in schools (Tables 4.3 and 4.4), we found strong correspondence in the domains of number and data and weaker correspondence in the domains of geometry and algebra; teacher education programs, especially those preparing primary teachers, do not consistently cover topics in the geometry and algebra domains.

The Implemented Curriculum of Teacher Education As Reported by Programs

To create a robust indicator of the implemented curriculum, TEDS-M developed the institutional/program questionnaire (Tatto 2013, p. 59, 64–69), which, among other questions, asked programs officials to report how many contact hours were allocated to different domains (liberal arts, academic mathematics, school mathematics, mathematics pedagogy, professional foundations and theories, and general pedagogy) during the duration of the program.[3] The assumption in TEDS-M was that contact hours are important indicators of OTL and help to understand programs' approaches to learning to teach. While the number of contact hours is not a perfect indicator, by combining contact hours with other indicators, such as the OTL reported by future teachers and the knowledge levels attained by future teachers as indicated by our assessments, we can obtain a more holistic picture of the influence of teacher education on the processes and outcomes involved in learning to teach mathematics.

The program questionnaire also asked program officials about the frequency with which future teachers engaged in different activities during their field experiences. These included the opportunities to plan lessons, teach individual lessons to the whole class, tutor individual pupils, work with small groups of pupils, assist teachers in other ways, assist in school activities outside classroom, carry out case studies of selected pupils, carry out classroom observation, collect data for research projects, visit families in their homes, interview teachers and/or principals, and observe and/or participate in meetings.

[3] Teaching contact hours included lectures, class meetings, tutorial classes, and any other required meetings that bring future teachers together to meet as a group with staff of the teacher preparation program. If the courses were online, participants were asked to estimate the number of hours future teachers were required to interact with the instructor and the material. For simplicity, we call these "experiences," although participants were asked to insert the term that best fit the context (see Table 4.18 in Appendix 4.1 for a detailed definition of courses).

Because pre-service teachers in the United States are prepared in either concurrent or consecutive programs, we present our findings according to these categories.[4] We present the result in graphs, as it is easier to appreciate the contrasts between concurrent and consecutive routes, within the two key program modalities, i.e., for future primary teachers (generalists and specialists) and secondary teachers (middle and high school). The corresponding tables are in Appendix 4.2 and 4.3, Tables 4.19, 4.20, 4.21, and 4.22, and 4.23 through 4.42.

Contact Hours per Domain and Routes into Teaching

Future Primary Teachers Table 4.9 shows that institutions preparing primary generalist teachers (Grade 6 maximum) report on average about two times more contact hours in "liberal arts" than the institutions preparing primary mathematics specialists at the 25th and 75th percentiles. This difference is not as large at the median.

Based on reports of program officials, the contact hours dedicated to courses in "academic mathematics" is higher at each percentile for institutions preparing primary mathematics specialists in the consecutive routes than for primary specialists in concurrent routes. There are more hours dedicated to school mathematics in the concurrent routes for both generalists and specialists except at the 25th percentile for specialists; programs preparing specialists in consecutive routes reported slightly more contact hours in mathematics pedagogy. Institutions preparing primary generalists allocate on average about twice as many contact hours to the areas of professional foundations and general pedagogy than institutions preparing primary mathematics specialists.

Future Secondary Teachers Table 4.10 shows that there is wide variability in the amount of contact hours spent in liberal arts courses (see the 25th and 75th percentiles) among institutions preparing middle and high school teachers. This difference is not as large at the median. Consecutive route institutions reported, on average, very few contact hours in liberal arts. The percentiles of contact hours in academic mathematics is higher for institutions preparing future high school teachers; this is especially true for those in concurrent route programs. Programs preparing future high school teachers in concurrent routes allocate more time to school mathematics

[4] Concurrent Routes: The route is concurrent if its first phase consists of a single program that includes studies in the subject(s) future teachers will be teaching (academic studies), studies of pedagogy and education (professional studies), and practical experience in the classroom. Consecutive Routes: The route is consecutive if it consists of a first phase for academic studies (leading to a degree or diploma), followed by a second phase of professional studies and practical experience (leading to a separate credential/qualification). Thus, no route can be considered consecutive if the institution or government authorities do not award a degree, diploma, or official certificate at the end of the first phase. Moreover, it may be customary or required for future teachers to do the first and second phases in different institutions (Tatto et al., 2008, pp. 23–24).

Table 4.9 Weighted percentiles of hours per domain for the programs preparing primary generalist teachers and specialist teachers by routes

Generalists												
Domains	**Routes**											
	All programs				Concurrent route				Consecutive route			
	N	Percentiles			*N*	Percentiles			*N*	Percentiles		
		25	50	75		25	50	75		25	50	75
Liberal arts	366	240	495	768	336	240	518	816	30	0[a]	0[a]	0[a]
Academic mathematics	364	45	90	144	334	45	96	144	30	0[a]	0[a]	5
School mathematics	367	45	75	96	315	64	75	96	52	0[a]	0[a]	45
Mathematics pedagogy	447	45	48	75	364	45	48	75	83	15	37	45
Professional foundations and theories	461	90	144	240	378	90	144	210	83	105	262	330
General pedagogy	455	90	420	525	364	105	440	576	92	90	420	420
Specialists												
Domains	**Routes**											
	All programs				Concurrent route				Consecutive route			
	N	Percentiles			*N*	Percentiles			*N*	Percentiles		
		25	50	75		25	50	75		25	50	75
Liberal arts	79	45	450	480	79	45	450	480	0[a]			
Academic mathematics	94	45	45	270	79	38	45	120	15	45	45	345
School mathematics	60	0	38	90	55	0[a]	45	90	4	37	37	37
Mathematics pedagogy	98	45	45	45	79	38	45	45	20	45	45	105
Professional foundations and theories	98	45	45	150	79	45	45	150	20	45	45	45
General pedagogy	98	45	64	285	79	45	64	285	20	45	112	315

Source: TEDS-M Institutional Program Questionnaire, Questions 3, 4, 5, 6, 7, and 8, Part D
[a]A zero indicates that the program did not have available information

and mathematics pedagogy, but there were not important differences between the program types in consecutive routes in the percentiles of contact hours dedicated to these subjects (school mathematics and mathematics pedagogy). The same pattern can be found in general pedagogy across the programs and routes. Program officials reported that programs preparing future high school teachers allocate more hours to the study of professional foundations and theories than those preparing future middle school teachers.

Contact Hours Allocated to Field Experiences

Future Primary Teachers According to the data from the institutional questionnaire (Table 4.11) for both programs preparing primary teachers, the reported percentiles of hours for extended teaching practice are higher for the concurrent route. For the consecutive route, these percentiles are higher for the primary programs

Table 4.10 Weighted percentiles of hours per domain for the secondary programs preparing middle school teachers and high school teachers by routes

Middle school												
Domains	**Routes**											
	All programs				Concurrent route				Consecutive route			
	N	Percentiles			*N*	Percentiles			*N*	Percentiles		
		25	50	75		25	50	75		25	50	75
Liberal arts	79	45	450	480	79	45	450	480	0[a]			
Academic mathematics	94	45	45	270	79	38	45	120	15	45	45	345
School mathematics	60	0[a]	38	90	55	0[a]	45	90	4	37	37	37
Mathematics pedagogy	98	45	45	45	79	38	45	45	20	45	45	105
Professional foundations and theories	98	45	45	150	79	45	45	150	20	45	45	45
General pedagogy	98	45	64	285	79	45	64	285	20	45	112	315
High school												
Domains	**Routes**											
	All programs				Concurrent route				Consecutive route			
	N	Percentiles			*N*	Percentiles			*N*	Percentiles		
		25	50	75		25	50	75		25	50	75
Liberal arts	301	336	585	768	265	448	615	768	36	0[a]	0[a]	0[a]
Academic mathematics	333	270	528	644	293	330	544	675	40	0[a]	168	225
School mathematics	239	0	45	90	193	0	45	98	46	0[a]	0[a]	45
Mathematics pedagogy	368	45	48	90	288	45	48	90	81	45	48	96
Professional foundations and theories	368	48	135	210	305	48	96	200	81	45	225	285
General pedagogy	358	64	120	195	277	64	120	195	81	60	105	285

Source: TEDS-M Institutional Program Questionnaire, Questions 3, 4, 5, 6, 7, and 8, Part D
[a]A zero indicates that the program did not have available information

preparing mathematics specialists and lower for these that prepare primary generalist teachers. The reported percentiles of hours for the introductory field experiences are similar for the concurrent routes of both programs and the consecutive routes of the generalist programs, but they are much lower for the consecutive routes programs preparing mathematics specialists.

Field Experience Activities According to the data from the program questionnaire, for the programs preparing primary teachers, the least common activity in field experiences for both routes and programs is visiting families in their homes (Fig. 4.1). As could be expected, the activities connected with teaching are more common. Specialists in the consecutive routes have more opportunities to teach individual lessons to the whole class and opportunities to plan lessons. In contrast, program officials reported that future teachers from the programs preparing primary specialists have fewer opportunities than their generalists' counterparts to observe

Table 4.11 Weighted percentiles of hours of field experiences for the programs preparing primary generalist teachers and specialist teachers by routes

Generalists												
Field experience	**Routes**											
	All programs				Concurrent route				Consecutive route			
	N	Percentiles			*N*	Percentiles			*N*	Percentiles		
		25	50	75		25	50	75		25	50	75
Extended teaching practice	368	420	630	744	283	420	644	763	85	420	525	630
Introductory field experiences	387	70	110	150	295	70	120	170	92	90	90	120
Specialists												
Field experience	**Routes**											
	All programs				Concurrent route				Consecutive route			
	N	Percentiles			*N*	Percentiles			*N*	Percentiles		
		25	50	75		25	50	75		25	50	75
Extended teaching practice	67	620	640	800	47	640	800	830	20	620	620	640
Introductory field experiences	70	30	128	144	50	96	140	162	20	0[a]	30	30

Source: TEDS-M Program Institutional Questionnaire, Question 2, Part E
[a]A zero indicates that the program did not have available information

and participate in school meetings and to carry out case studies of selected pupils. In both modalities, future generalist and specialist teachers have limited opportunities to collect data for research projects, with the exception of specialists in the consecutive routes, who report occasional engagement. With few exceptions, the mean frequencies for consecutive routes are larger than those for the concurrent routes.

Future Secondary Teachers The data from the reports of the program officials for the programs preparing secondary teachers are presented in Table 4.12. For both programs, the reported percentiles of hours for extended teaching practice are higher for the concurrent route. For the consecutive routes, these percentiles are higher in the programs preparing future middle school teachers. The reported percentiles of hours for the introductory field experiences are higher for teachers prepared to teach in the middle school grades in the concurrent routes, and the opposite is true for those prepared to teach the high school grades.

Field Experience Activities According to the program officials' reports, the least common activity in both routes for programs preparing middle and high school teachers is visiting families in their homes, followed by assisting in school activities outside of the classroom (shown in Fig. 4.2). Relatively few programs for future high school teachers reported engaging future teachers in collecting data for research projects—a finding that is consistent with the results from the syllabi analysis (Table 4.8, "Design and carry out an action research project…").

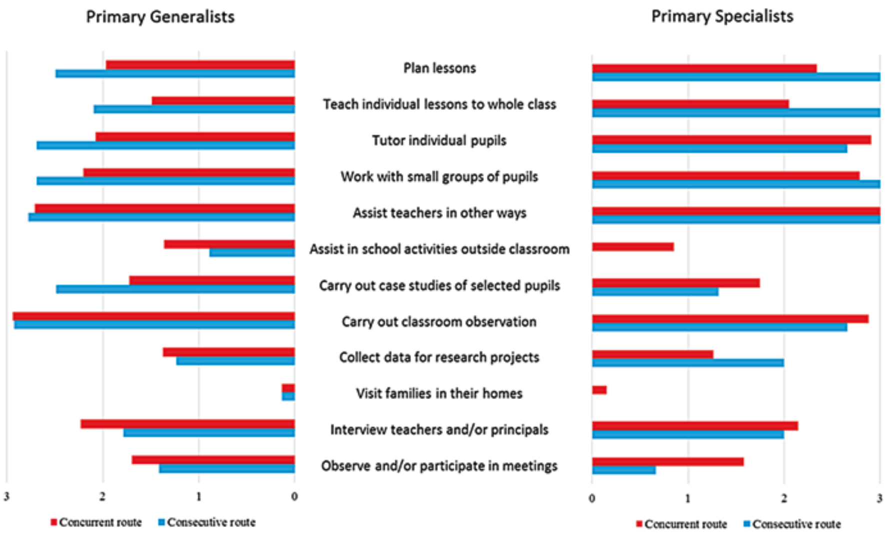


Fig. 4.1 Weighted mean frequency of the activities in the field experience for the programs preparing primary generalist and specialist teachers by routes. (Source: TEDS-M Institutional Program Questionnaire, Question 3, Part E. The scale is 0 = Not at all, 1 = Rarely, 2 = Sometimes, 3 = Usually. Activities: Ability to plan lessons, teach individual lessons to whole class, tutor individual pupils, work with small groups of pupils, assist teachers in other ways, assist in school activities outside classroom, carry out case studies of selected pupils, carry out classroom observation, collect data for research projects, visit families in their homes, interview teachers and/or principals, and observe and/or participate in meetings)

According to the program questionnaire, future middle school teachers in consecutive routes rarely observe and/or participate in meetings. As could be expected, the activities connected with teaching were more common.

Discussion of the Implemented Curriculum of Teacher Education As Reported by Programs

The program questionnaire data indicates that concurrent routes in all programs have the most contact hours dedicated to liberal arts curriculum, in accordance with a traditional higher education curriculum. Teacher education programs following a consecutive model reported low to no coverage of subjects in the liberal arts domain; this is not surprising since, in consecutive programs, teacher education occurs separately from the typical undergraduate university curriculum. The second domain with high mean contact hours across all programs and both routes is general pedagogy. This is an indication that general pedagogy is an important domain in the intended curriculum for future mathematics teachers in the United States. Consistent with the understanding that high school teachers should know mathematics well, the

Table 4.12 Weighted percentiles of hours of field experiences for the secondary programs preparing middle and high school teachers by routes

Middle school												
Field experience	**Routes**											
	All programs				Concurrent route				Consecutive route			
	N	Percentiles			*N*	Percentiles			*N*	Percentiles		
		25	50	75		25	50	75		25	50	75
Extended teaching practice	67	620	640	800	47	640	800	830	20	620	620	640
Introductory field experiences	70	30	128	144	50	96	140	162	20	0[a]	30	30
High school												
Field experience	**Routes**											
	All programs				Concurrent route				Consecutive route			
	N	Percentiles			*N*	Percentiles			*N*	Percentiles		
		25	50	75		25	50	75		25	50	75
Extended teaching practice	321	490	550	640	249	490	560	640	72	420	525	608
Introductory field experiences	311	40	75	100	239	40	60	100	72	40	75	104

Source: TEDS-M Program Institutional Questionnaire, Question 2, Part E
[a]A zero indicates that the program did not have available information

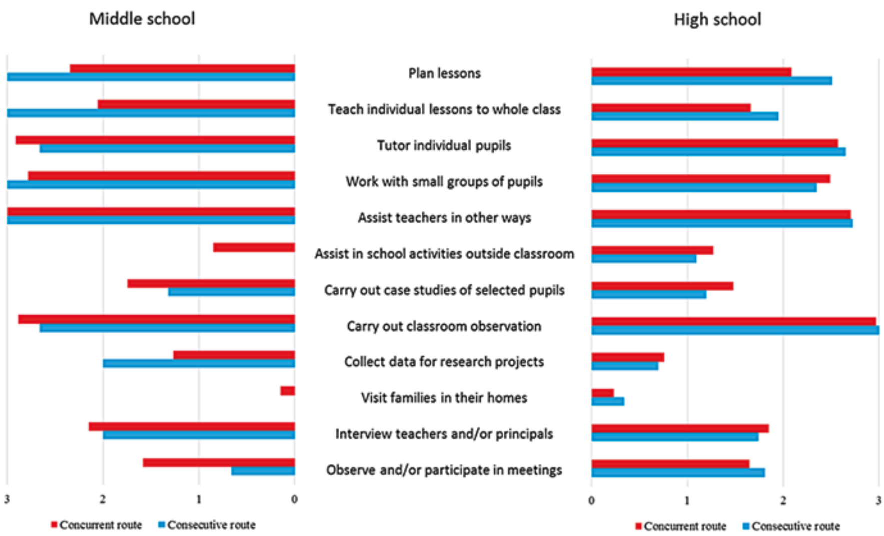


Fig. 4.2 Weighted mean frequency of the activities in the field experience for the secondary programs preparing middle and high school teachers by routes. (Source: TEDS-M Institutional Program Questionnaire, Question 3, Part E. The scale is: 0 = Not at all, 1 = Rarely, 2 = Sometimes, 3 = Usually. Activities: Ability to – plan lessons, Teach individual lessons to whole class, Tutor individual pupils, Work with small groups of pupils, Assist teachers in other ways, Assist in school activities outside classroom, Carry out case studies of selected pupils, Carry out classroom observation, Collect data for research projects, Visit families in their homes, Interview teachers and/or principals, and Observe and/or participate in meetings)

programs preparing these teachers allocate more contact hours for academic mathematics. This is especially true for concurrent route programs, since consecutive programs typically assume—and, for future high school teaching, require—previous coursework in mathematics. Contrary to our assumption, the mean value for the contact hours dedicated to the study of school mathematics and mathematics pedagogy domains is small in both routes and all programs. These two domains are directly connected with the professional work of mathematics teachers; thus, one might assume that a large portion of contact hours should be dedicated to them, this finding is in contrast with what is indicated about the expected teacher preparation curriculum in the United States. Based on the analysis of our data, the emphasis in preparing future mathematics teachers in the United States is placed on liberal arts and general pedagogy and less on mathematics pedagogy and academic mathematics, with the exception of future high school teachers studying in concurrent routes, who are expected to have good preparation in academic or university level mathematics.

The Implemented Curriculum of Teacher Education As Reported by Future Teachers

In TEDS-M, future teachers' reports of their OTL were used to arrive at a more complete understanding of the implemented curriculum; we were interested in learning more about the experienced curriculum from the viewpoint of future teachers.

Opportunities to Learn School Mathematics

Among the programs preparing primary teachers, a high percentage of future teachers, both generalists and specialists, reported studying numbers; measurement; geometry; functions, relations, and equations; and data representation, probability, and statistics (Fig. 4.3). In contrast, the mean percentage of future teachers who reported studying calculus and validation, structuring, and abstracting was low. There are only small differences between the two routes: (a) among programs preparing primary generalists (Grade 6 maximum), future teachers in the consecutive routes study on average less geometry; functions, relations, and equations; and data representation, probability, and statistics; (b) among programs preparing primary mathematics specialists, future teachers in the consecutive routes report higher OTL than those in concurrent routes.

In both types of secondary programs (Fig. 4.4) a high percentage of future teachers reported studying numbers; measurement; geometry; functions, relations, and equations; and data representation, probability, and statistics (with the exception of the consecutive route for the programs preparing future middle school teachers). The mean percentage of future middle school teachers who reported studying calculus and validation, structuring, and abstracting was generally low. In contrast, high percentages of

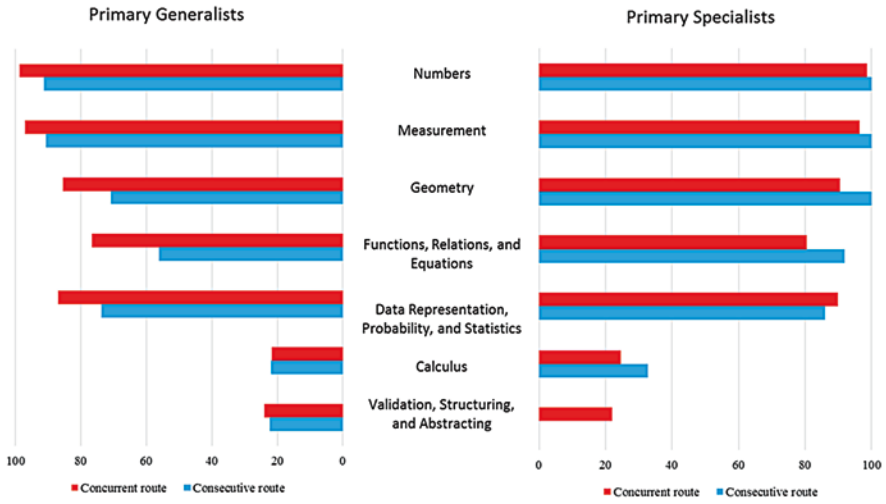


Fig. 4.3 Weighted mean percent of future teachers reporting studying school mathematics topics in the programs preparing primary generalists (Grade 6 maximum) and specialists by routes. (Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 2, Part B)

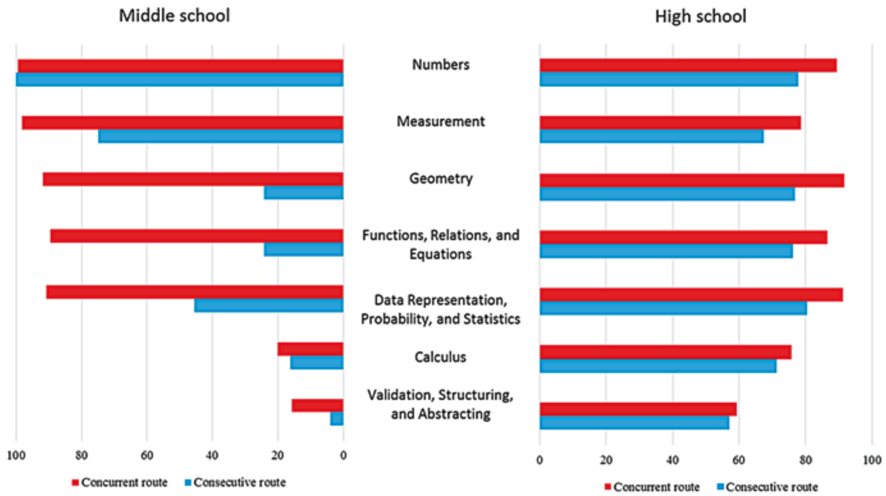


Fig. 4.4 Weighted mean percent of future teachers reporting to study school mathematics topics in the programs preparing middle (Grade 10 maximum) and high (to Grade 11 and above) school teachers by routes. (Source: TEDS-M Future Teachers of Secondary Mathematics Questionnaire, Question 2, Part B)

those in the programs preparing future high school teachers report having studied those topics. Overall, future high school teachers from concurrent routes study more school mathematics topics than their counterparts from the consecutive routes. This difference is larger for the programs preparing middle school teachers.

Opportunities to Learn University Mathematics

In the programs preparing primary teachers, the highest mean percentages of future teachers reporting that they study university mathematics topics are in programs preparing mathematics specialists in consecutive routes, except for those who are prepared in concurrent routes who reported studying more topics on probability and number theory (Fig. 4.5). The future teachers in programs preparing primary generalists reported studying less topology and multivariate and advanced calculus. The future teachers in programs preparing primary specialists from the consecutive routes reported studying more university mathematics topics than their counterparts from the concurrent routes.

In the programs preparing secondary teachers, the future teachers who reported studying more topics in university mathematics were those preparing to teach the high school grades (Fig. 4.6). Overall, future teachers in the concurrent routes reported receiving less OTL university mathematics with low means in topology, advanced calculus, and theory of functions. In the programs preparing middle school teachers, future teachers from the concurrent routes reported higher means than their counterparts from the consecutive routes (with some exceptions for instance in multivariate

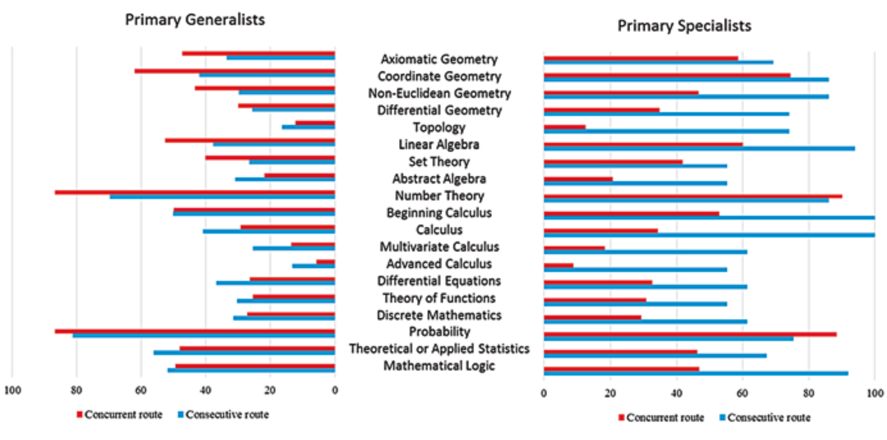


Fig. 4.5 Weighted mean percent of future teachers reporting to study university mathematics topics in the programs preparing primary generalists (Grade 6 maximum) and primary specialists by routes. (Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 1, Part B)

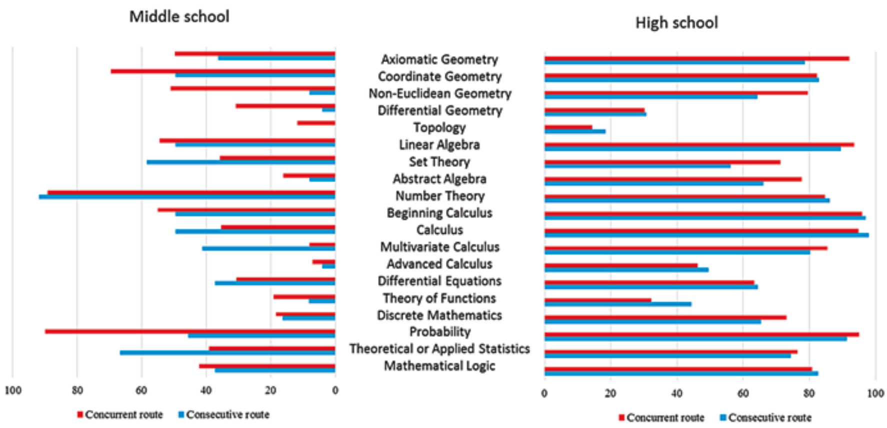


Fig. 4.6 Weighted mean percent of future teachers reporting studying university mathematics topics in the programs preparing middle (Grade 10 maximum) and high (to Grade 11 and above) school teachers by routes. (Source: TEDS-M Future Teachers of Secondary Mathematics Questionnaire, Question 1, Part B)

calculus). The relatively high mean percentages are on topics: analytic/coordinate geometry, axiomatic geometry, number theory, calculus, probability and statistics.

Opportunities to Learn Mathematics Pedagogy

Among future teachers in primary programs, those who were prepared as specialists reported equal or more OTL in mathematics pedagogy domains except in two areas—foundations of mathematics and context of mathematics education—while those in the programs preparing primary generalists reported higher OTL (Fig. 4.7).

In the secondary programs, future teachers who were prepared to teach middle school grades reported more OTL than future teachers who were prepared to teach high school in all domains but foundations of mathematics and context of mathematics education. For these topics, future middle school teachers in consecutive routes programs reported lower OTL than their counterparts (see Fig. 4.8).

Opportunities to Learn General Pedagogy

Among future primary teachers, we found some important differences in the mean percentage of general pedagogy topics reported across programs and routes. Future primary teachers prepared as specialists in consecutive route programs reported very different opportunities to learn theories of schooling and history of

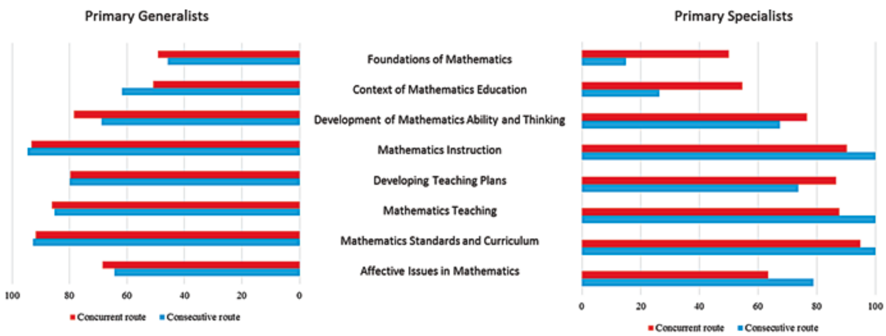


Fig. 4.7 Weighted mean percent of future teachers reporting studying mathematics pedagogy topics in the programs preparing primary generalists (Grade 6 maximum) and primary specialists by routes. (Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 4, Part B)

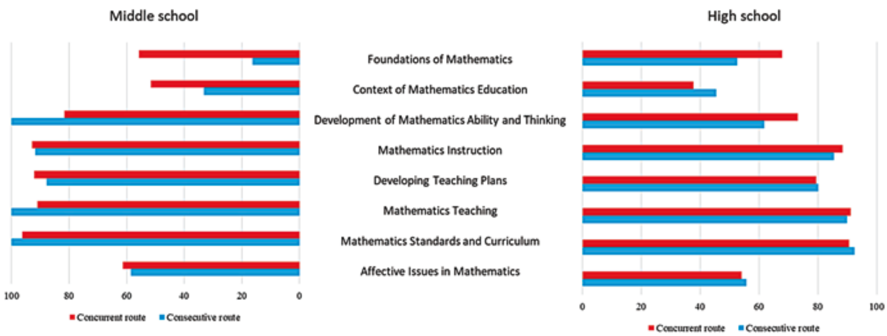


Fig. 4.8 Weighted mean percent of future teachers reporting studying mathematics pedagogy topics in the programs preparing middle (Grade 10 maximum) and high (to Grade 11 and above) school teachers by routes. (Source: TEDS-M Future Teachers of Secondary Mathematics Questionnaire, Question 4, Part B)

education compared with their counterparts in concurrent programs or those prepared as generalists. Opportunities to learn methods of educational research were also reported in lower percentages than other topics such as knowledge of teaching, educational psychology, and philosophy of education (see Fig. 4.9).

Among future middle school teachers, higher percentages reported having OTL in all the domains than those prepared for the high school grades, regardless of route. In general, future teachers from the concurrent routes reported more opportunities to learn than their counterparts from the consecutive routes in all areas, with the exception of opportunities to learn assessment and measurement, which was higher for future middle school teachers on the consecutive route, and methods of educational research for both program types, which was higher for both future middle school and high school teachers on the consecutive routes (Fig. 4.10).

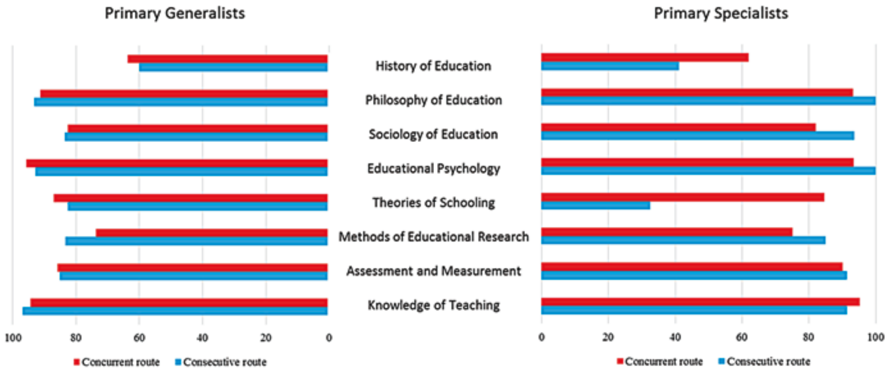


Fig. 4.9 Weighted mean percent of future teachers reporting to study general pedagogy topics in the programs preparing primary generalists (Grade 6 maximum) and specialists by routes. (Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 7, Part B)

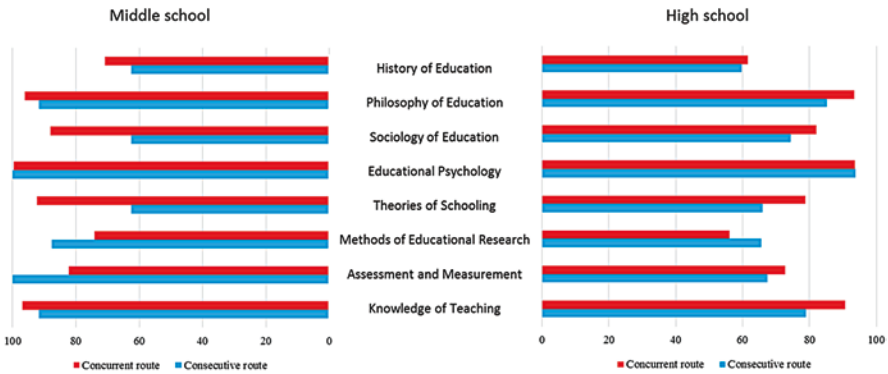


Fig. 4.10 Weighted mean percent of future teachers reporting studying general pedagogy topics in the programs preparing middle (Grade 10 maximum) and high (to Grade 11 and above) school teachers by routes. (Source: TEDS-M Future Teachers of Secondary Mathematics Questionnaire, Question 7, Part B)

Opportunities to Learn from School Experiences

Among future primary teachers, the pattern of responses concerning opportunities to engage in school experience activities is similar for both routes across both types of programs (generalists and specialists). The only exception is in the topic "test out findings from educational research about pupils' difficulties." For the programs preparing primary generalists, the mean for reported opportunities to learn this topic is higher for future teachers in the concurrent routes. For those preparing as primary mathematics specialists, the pattern is the opposite (Fig. 4.11).

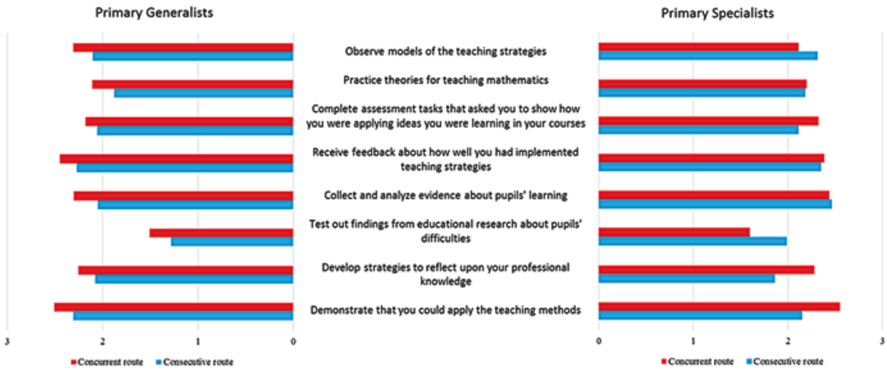


Fig. 4.11 Weighted mean frequency of school experience activities required in the programs preparing primary generalists (Grade 6 maximum) and specialists by routes. (Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 13, Part B. The scale is 0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Often)

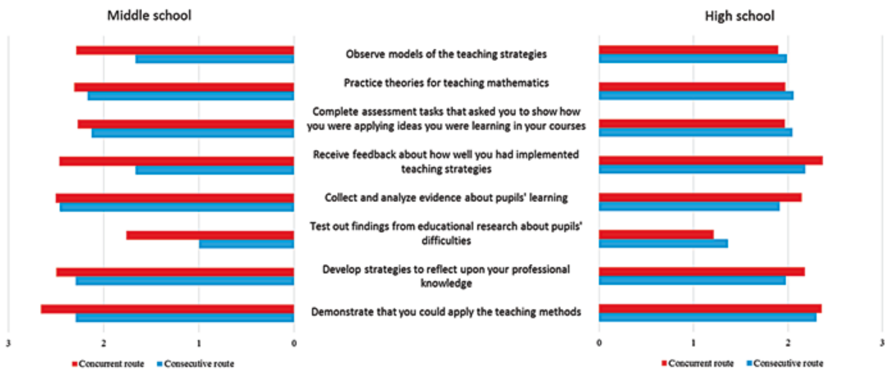


Fig. 4.12 Weighted mean frequency of school experience activities required in the programs preparing middle (Grade 10 maximum) and high school (to Grade 11 and above) teachers by routes. (Source: TEDS-M Future Teachers of Secondary Mathematics Questionnaire, Question 13, Part B. The scale is 0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Often)

Among secondary future teachers, the pattern of responses concerning opportunities to engage in school experiences activities is similar for both routes across both program types. The only exception is "test out findings from educational research about pupils' difficulties." Future middle school teachers in concurrent routes programs report a higher mean percentage of this OTL, than their counterparts in the consecutive routes programs. For the programs preparing high school teachers, the pattern is the opposite (Fig. 4.12).

Discussion of the Implemented Curriculum of Teacher Education as Reported by Future Teachers

The mean percentage of future teachers who reported studying school mathematics topics is high, with the exceptions of calculus and validation, structuring, and abstracting. As could be expected, a high percentage of high school future teachers reported studying university or academic mathematics. Differences between the routes are more noticeable in the programs preparing primary specialists, as future teachers in consecutive routes reported studying more university mathematics than their counterparts in the concurrent routes. In the mathematics pedagogy area, fewer OTL overall are reported in the domains of foundations of mathematics and context of mathematics education. It is not surprising that the mean percentage of future teachers who reported studying general pedagogy is high for all topics. In the school experience activities, the least frequent activity was related to doing research about pupils' difficulties.

The Achieved Curriculum of Teacher Education

Based on the data from the assessments, TEDS-M produced two scales to describe and analyze future teachers' mathematical content knowledge and mathematical pedagogical content knowledge as the outcomes (or the achieved curriculum) of teacher education. As described in the methods section of this chapter, the Mathematical Content Knowledge (MCK) and the Mathematical Pedagogical Content Knowledge (MPCK) were measured by assessments using frameworks that are presented in Appendix 4.1 of this chapter. The knowledge results were scaled using Item Response Theory (IRT) to have a mean of 500 and standard deviation of 100. Future teachers also answered a questionnaire that included questions about OTL, among others (Tatto, 2013, pp. 49–55). The OTL results were scaled using IRT to have a mean of 10 (indicating a neutral position), and standard deviation of 1. Tables 4.13 and 4.14 below show these scales for primary and secondary future teachers. The inclusion of the outcome scales and the OTL scales in the tables help us to explore the possible associations between future teachers' reported experiences and the achieved curriculum.[5]

Primary Future Teachers

The MCK and the MPCK assessment results for future primary teachers are above the international mean (500) in both routes and all programs. This is especially clear for the MPCK scale. These results indicate a good level of knowledge among the U.S. future primary teachers in mathematics pedagogy. Although the sample size of the consecutive programs preparing primary mathematics specialists is relatively small, the mean scores of these specialist future teachers on both MCK

[5] For a multilevel analysis of the associations between the knowledge assessment results and the OTL for future primary and secondary teachers consult chapters 8 and 14 of this book.

Table 4.13 Comparison of the weighted statistics on OTL and knowledge scales for future teachers in primary programs

Scales	Programs and routes											
	Primary generalists programs						Primary specialists programs					
	Concurrent route			Consecutive route			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Mathematical content knowledge[a]	14,926	518.22	68.76	1233	509.06	83.24	2356	519.46	61.48	77	536.01	98.44
Mathematical pedagogical content knowledge[a]	14,926	544.06	66.59	1233	537.68	78.51	2356	543.08	70.01	77	587.79	121.64
OTL – university level math – geometry[b]	20,276	1.82	1.39	2002	1.31	1.49	3432	2.15	1.30	172	3.15	1.15
OTL – university level math – discrete structures and logic[c]	20,276	2.77	1.64	2002	2.47	1.95	3432	2.88	1.67	172	4.44	1.75
OTL – university level math –continuity and functions[d]	20,261	1.24	1.40	2002	1.66	1.70	3432	1.47	1.54	172	3.78	1.39
OTL – university level math – probability and statistics[e]	20,261	1.35	0.67	2001	1.37	0.72	3432	1.35	0.66	172	1.43	0.86
OTL – school level math – numbers measurement geometry[f]	20,261	2.81	0.51	2006	2.53	0.86	3432	2.85	0.46	172	3.00	0.00
OTL – school level math – functions probability calculus[b]	20,261	2.09	1.06	2006	1.74	1.35	3432	2.17	0.98	172	2.11	0.61
OTL – math education pedagogy – instruction[d]	15,989	4.19	1.13	1407	4.17	1.12	2655	4.23	1.07	161	4.52	0.50
OTL – math ed pedagogy – class participation[g]	15,958	10.95	1.51	1407	10.78	1.50	2655	11.02	1.58	161	11.85	2.21
OTL – math ed pedagogy – solving problems[g]	25,927	10.30	1.43	1407	9.70	1.33	2655	9.89	1.64	161	11.45	2.15
OTL – math ed pedagogy – instructional practice[g]	15,896	11.83	1.89	1407	11.30	1.78	2644	11.41	1.71	161	12.76	2.66
OTL – math ed pedagogy – assessment uses[g]	15,888	11.58	0.08	1407	10.74	0.09	2644	11.27	0.28	161	11.80	0.63
OTL – math ed pedagogy – assessment practice[g]	15,927	12.26	1.94	1407	11.85	1.98	1644	12.07	2.03	161	12.56	2.91

[a]The international mean for MCK and MPCK scales is 500; the standard deviation is 100

[b]The scale is based on the future teachers' responses on 4 topics, asking whether they have ever studied each topic. The values of 0, 1, 2, 3, or 4 equal the number of topics studied

[c]The scale is based on the future teachers' responses on 6 topics, asking whether they have ever studied each topic. The values of 0, 1, 2, 3, 4, 5 or 6 equal the number of topics studied

[d]The scale is based on the future teachers' responses on 5 topics, asking whether they have ever studied each topic. The values of 0, 1, 2, 3, 4, or 5 equal the number of topics studied

[e]The scale is based on the future teachers' responses on 2 topics, asking whether they have ever studied each topic. The values of 0, 1, or 2 equal the number of topics studied

[f]The scale is based on the future teachers' responses on 3 topic, askings whether they have ever studied each topic. The values of 0, 1, 2, or 3 equal the number of topics studied

[g]Rasch score scale where 10 is located at the neutral position. Numbers greater than 10 indicate greater than neutral OTL; numbers smaller than 10 indicate lower than neutral OTL. Smaller numbers indicate less OTL, greater numbers indicate more OTL

Table 4.14 Comparison of the weighted statistics on OTL and knowledge scales for future teachers in secondary programs

Scales	Programs and routes											
	Middle school (grade 10 maximum) programs						High school (to grade 11 and above) programs					
	Concurrent route			Consecutive route			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Mathematical content knowledge[a]	2702	468.39	45.97	147	455.95	52.73	1864	552.46	56.17	392	554.85	61.39
Mathematical pedagogical content knowledge[a]	2702	470.29	53.04	147	478.71	54.39	1864	543.87	81.60	392	534.62	67.63
OTL – university level math – geometry[b]	4037	2.01	0.05	196	0.95	0.12	2242	2.84	0.07	603	2.57	0.04
OTL – university level math – discrete structures and logic[c]	4037	2.56	1.62	196	2.62	1.24	2242	4.80	1.27	603	4.46	1.45
OTL – university level math –continuity and functions[d]	4037	1.36	0.09	196	1.82	0.20	2237	3.85	0.15	603	3.87	0.09
OTL – university level math – probability and statistics[e]	4037	1.29	0.04	196	1.12	0.01	2237	1.71	0.03	603	1.66	0.05
OTL – school level math – numbers Measurement geometry[f]	4037	2.89	0.07	196	1.99	0.10	2237	2.60	0.08	603	2.22	0.13
OTL – school level math – Functions Probability calculus[b]	4037	2.16	0.05	196	0.90	0.27	2237	3.13	0.05	603	2.85	0.14
OTL – Math education pedagogy – instruction[d]	2937	4.34	0.12	196	4.38	0.32	1931	4.03	0.09	439	4.04	0.28
OTL – Math Ed pedagogy – Class participation[g]	2992	11.45	0.12	196	11.84	0.20	1931	11.06	0.05	439	11.54	0.09
OTL – Math Ed pedagogy – solving problems[g]	2992	9.92	0.06	196	9.57	0.32	1931	11.26	0.09	439	10.27	0.20
OTL – Math Ed pedagogy – Instructional practice[g]	2992	11.64	0.15	196	11.56	0.00	1931	11.57	0.07	439	10.82	0.15
OTL – Math Ed pedagogy – Assessment uses[g]	2992	11.43	0.29	196	10.07	0.25	1931	11.22	0.12	439	10.72	0.14
OTL – Math Ed pedagogy – Assessment practice[g]	2992	11.96	0.27	196	12.25	0.26	1931	11.81	0.05	439	11.65	0.21

[a]The international mean for MCK and MPCK scales is 500; the standard deviation is 100

[b]The scale is based on the future teachers' responses on 4 topics, asking whether they have ever studied each topic. The values of 0, 1, 2, 3, or 4 equal the number of topics studied

[c]The scale is based on the future teachers' responses on 6 topics, asking whether they have ever studied each topic. The values of 0, 1, 2, 3, 4, 5 or 6 equal the number of topics studied

[d]The scale is based on the future teachers' responses on 5 topics, asking whether they have ever studied each topic. The values of 0, 1, 2, 3, 4, or 5 equal the number of topics studied

[e]The scale is based on the future teachers' responses on 2 topics, asking whether they have ever studied each topic. The values of 0, 1, or 2 equal the number of topics studied

[f]The scale is based on the future teachers' responses on 3 topics, asking whether they have ever studied each topic. The values of 0, 1, 2, or 3 equal the number of topics studied

[g]Rasch score scale where 10 is located at the neutral position. Numbers smaller than 10 indicate less OTL, numbers greater than 10 indicate more OTL

and MPCK scales is higher than for those in the concurrent routes. This result indicates that consecutive route programs preparing primary specialist teachers may be able to recruit individuals with a high level of mathematics knowledge, while also providing OTL that support future teachers in learning not only mathematics content, but also pedagogical content. Thus, a question worth exploring is whether consecutive route programs do a better job of preparing primary mathematics specialist teachers than those in the concurrent routes, and whether these results are due to the ability to control the quality of their recruits. No such differences were found between the two routes for the programs preparing primary generalists.

A closer examination of the OTL results is suggestive. The higher means in OTL (except for OTL – school level math – functions, probability, and calculus) are in the consecutive route of the programs preparing primary mathematics specialists (Table 4.13). At the same time, looking at the concurrent route programs for primary mathematics specialists, we can see that the means on the OTL scales are generally higher than those in the programs preparing primary generalists. This is consistent with the expectation that programs preparing mathematics specialists provide more OTL mathematics than the programs preparing generalist teachers.

The means of the last three OTL scales (instructional practice, assessment issues, and assessment practice) are above neutral, indicating that future primary teachers in the United States have high levels of OTL on these topics. A finding that may indicate increased concern with accountability demands introduced in 2001 as a result of the federal law known as the No Child Left Behind Act.

Surprisingly, given teacher education reforms that have followed an inquiry approach to teaching in the United States, the means of the OTL 'math education pedagogy – solving problems' scale for consecutive programs preparing primary generalists and for concurrent programs preparing primary mathematics specialists are below the neutral point. This finding indicates that the future teachers prepared in these routes have less OTL to learn to solve problems and to learn to reason around mathematics concepts (this is a finding evident in other chapters in this book). This finding may indicate a new orientation of the teacher education curriculum to accommodate competing demands as noted in the previous paragraph.

In sum, the data in Table 4.13 indicate that the MPCK of future primary teachers in the United States is above the international average in both routes and all programs. For the programs preparing primary generalists in both routes, a number of OTL scales show a mean above the midpoint (neutral point). These are in mathematics at the university level: probability & statistics; and in mathematics at the school level: numbers measurement geometry. In the area of mathematics education pedagogy these included: class participation; instruction and instructional practice; assessment uses; and assessment practice. For the programs preparing primary specialists via consecutive routes, all OTL scales received an above average score. For the concurrent routes, the picture is similar to that for the programs preparing primary generalists, with the exception of mathematics at the school level: functions, probability, and calculus where more topics were reported as studied. Concerning school level mathematics, all future teachers reported having the opportunity to learn the three topics included in the scale (e.g., numbers, measurement, and geometry).

Secondary Future Teachers

For future secondary teachers, the means for MCK and MPCK are above the international means in both routes for programs that prepare teachers to teach high school mathematics (Table 4.14). This result may indicate a rigorous selection strategy for these individuals. There are no differences in the achievement results for the different routes. The situation is not the same for teachers prepared to teach middle school mathematics, in both routes with means for both MCK and MPCK below the international mean.

Looking at the OTL scales, the means for university- and school-level mathematics are, in general, higher for the programs preparing future teachers to teach high school than for those preparing them to teach middle school. The larger means in the OTL scales (with few exceptions) are in the programs preparing high school teachers. This is consistent with the intended curriculum in these programs, where the institutions, on average, dedicated more contact hours to academic mathematics than other programs and routes (see Table 4.5).

The means for OTL mathematics education pedagogy (e.g., solving problems) are, in general, higher for teachers in programs preparing high school teachers than for those preparing middle school teachers. On the other hand, the means for math education pedagogy OTL are, in general, higher for future teachers in programs preparing middle school teachers.

Of interest are the relatively lower means for OTL school-level mathematics in programs preparing middle school teachers in consecutive routes. A possible explanation is that consecutive programs pay more attention to professional studies and practical experience and less to school mathematics, assuming that the future teachers have already studied school mathematics during the first phase of their education, or even during their school years.

The means of the last five scales indicating OTL mathematics education pedagogy (e.g., class participation, instructional practice, assessment uses, and assessment practice), are above the neutral point. These results indicate that the future secondary teachers in the United States have greater OTL on these topics.

In sum, the data in Table 4.14 indicate that the MPCK of future teachers in the programs preparing high school teachers in the United States is above the international average. In contrast, it is below the international average for programs preparing middle school teachers. A number of OTL scales have a mean above the midpoint (neutral point) indicating higher than average opportunities to learn among future middle school teachers. This is the case for OTL mathematics at the university level: probability & statistics; and at the school level: numbers, measurement, and geometry, functions, probability, and calculus (concurrent route only). In regards to mathematics education pedagogy future middle school teachers reported higher than average OTL on class participation, instruction and instructional practice, assessment uses, and assessment practice.

Discussion of the Achieved Curriculum in the Context of Programs' Opportunities to Learn

Future teachers in the United States reported high OTL when asked if they had studied school mathematics topics, while the data from the intended curriculum of the programs do not show particularly high expectations in this domain. This finding may indicate a disconnect between the *intended* and the *implemented* curriculum, as teacher educators may plan to dedicate more time to mathematics pedagogy assuming that future teachers know already the mathematics of the school curriculum. Yet as they encounter their students and understand their knowledge limitations, educators may find themselves struggling to teach the mathematics of the school curriculum at the same time that they teach mathematics pedagogy.

The future middle school and primary teachers' reports are consistent with the information from the syllabi analysis concerning *university* mathematics topics. Future teachers' responses on OTL topics are also consistent with the findings from the analysis of the intended curriculum in the domain of general pedagogy; that is, future teachers reported high OTL when asked if they had studied general pedagogy topics. Future teachers' responses about school experience activities are in agreement with the syllabi analysis information as well. Future teachers reported studying most of the topics in the mathematics pedagogy domain (except for foundations of mathematics and context of mathematics education), a finding consistent with program expectations in this domain (see Table 4.6).

Overall, the finding, based on the syllabi analysis information, that U.S. future teachers are expected to be well prepared in mathematics pedagogy and general pedagogy was consistent with future teachers' responses to the OTL questions.

It can be expected that the mean percent of OTL for university mathematics varies among programs and routes. Not surprisingly, future teachers prepared to teach as primary specialists (especially in consecutive routes) and those prepared to teach high school grades reported more OTL in this area than their counterparts prepared as primary generalists and middle school teachers. Similarly, in the area of school mathematics, future teachers prepared to teach high school grades also reported a high mean percentage of OTL about almost all topics. The future teachers in the other three program types reported high OTL on number; measurement; geometry; functions, relations, and equations; and data representation, probability and statistics (except in middle school consecutive route programs). The future teachers' responses about OTL from school experiences are relatively uniform, with medium to high mean frequencies in the list of these particular OTLs.

Throughout this chapter, we have explored the intended, implemented and achieved curriculum of teacher education in the United States seeking evidence of alignment. Our findings indicate that opportunities to learn as measured through the intended and the implemented curriculum differ across program types reflecting some alignment with the knowledge that is expected at these different levels. That is, generally (i) programs preparing primary specialists offer more OTL on mathematics and mathematics pedagogy domains than programs preparing primary generalists; and (ii) programs preparing high school teachers offer more OTL in the mathematics and mathematics pedagogy domains than programs preparing middle school teachers. Table 4.13 above gives information about the validity of statement (i). For the concurrent routes, the means of the OTL scales in the programs preparing primary specialists are higher than for those preparing primary generalists. For the consecutive route, this difference is even greater. Table 4.14 above provides information about the validity of statement (ii). For most of the OTL scales, the means in the programs preparing high school teachers are higher than for those preparing middle school teachers.

Our findings indicate that opportunities to learn are correlated with the achieved curriculum (MCK and MPCK), which we see as another indication of alignment. The data in Table 4.13 supports the hypothesis that OTL has a positive correlation with MCK and MPCK: the highest MCK and MPCK means are for the future teachers in primary specialists programs of the consecutive route, where the means of most of the OTL scales are also the highest. Table 4.14 shows that the MCK and MPCK means for the programs and routes preparing high school teachers are higher than for those preparing middle school teachers. This supports the hypothesis that OTL has a positive correlation with MPCK. These hypotheses are further explored using multilevel analysis in chapters 8 (for future primary teachers), and 14 (for future secondary teachers) of this book.

Conclusion

The main goal of teacher preparation programs and institutions is to prepare highly qualified teachers who can face the challenges of school teaching. To meet this goal, institutions organize their courses by taking into account the general requirements for teacher preparation, the needs of their students, and the needs of the school system. In this chapter, we used the data from the TEDS-M study to explore the curriculum of mathematics teacher preparation, focusing on the United States. We studied the intended curriculum by examining teacher education syllabi, the implemented curriculum as it is described by future teachers' OTL, and by program administrators, and the achieved curriculum as indicated by future teachers' responses to our assessments. We focused on five aspects of the curriculum, namely school mathematics, university mathematics, mathematics pedagogy, general pedagogy, and school experience.

Teacher preparation institutions follow different routes into teaching and, accordingly, design courses to prepare mathematics teachers for different school levels and contexts. The different opportunities to learn also reflect differences in performance. What we have found is that mathematics teacher preparation programs offer all five aspects of the curriculum mentioned above, following what seems to be a universal norm (Tatto & Hordern, 2017). But we also found variability in the amount and character of the OTL as implemented across programs. This variation in the breadth and depth of OTL can be seen as important evidence for the need for program alignment across domains as programs seek to improve the quality of the preparation of future mathematics teachers.

The data shows that there is overall consistency across the intended, implemented, and achieved curriculum for teacher preparation. We found, however, that the correspondence between the mathematics curriculum of schools and the intended curriculum of the teacher preparation institutions is mixed. The correspondence is more evident in the domains of number and data; it is less so in the algebra and geometry domains.

The analysis of the intended teacher education curriculum shows relatively strong emphasis in the areas of mathematics pedagogy and general pedagogy and less in mathematics knowledge, particularly among primary generalists and middle school future teachers. Yet there is much variability within and across institutions, as shown in the analysis of the time allocated to the different domains, and which corresponds to variation in TEDS-M assessment results.

According to the data, there is important variation between different programs and routes and different institutions in the attention that is paid to different mathematics topics. International comparisons in the knowledge assessments (Tatto et al., 2012)[6] show that the performance of U.S. future mathematics teachers on the knowledge scales is above the international mean, except for future teachers in the programs that prepare middle school mathematics teachers. This indicates that, in general, middle school teachers are not as well prepared to teach the middle school mathematics curriculum as their high school counterparts.

The findings from the U.S. curriculum study have implications for the future of mathematics teacher education in an era of increased accountability and increased diversity in the student population. There has been legislation proposed at the federal level concerning the introduction of regulations into teacher education closely linked to the Council for the Accreditation of Educator Preparation (CAEP) standards.[7] While this legislation was rescinded in early 2017 under a new administration, the accreditation of teacher preparation is still subject to the CAEP standards, which are likely to promote greater uniformity in the organization and curriculum of

[6] For the primary level both generalists and specialists are below yet very close to the international mean, in the mathematics and mathematics pedagogy scales (see Tatto et al., 2012, pp. 140 and 144). For the secondary level, middle school future teachers scores are below the international mean, whereas for high school future teachers the scores are above the international mean, in the mathematics and in the mathematics pedagogy scales (see Tatto et al., 2012, pp. 148 and 150).

[7] http://caepnet.org/standards/

teacher education. Thus, although controversial, many in the United States may continue to argue for more regulation of teacher education programs in light of the great variability in OTL that exists across teacher education programs and the relatively modest levels of knowledge of mathematics that some teachers, especially primary and middle school teachers, possess.

Calls for more rigorous graduation criteria, including knowledge assessments for program graduates, may lead to changes in the teacher education curriculum within states or across states. However, it will not be an easy task to change the teacher education curriculum to be more rigorous, as this change may have important implications, including requiring changes in recruitment and selection procedures across institutions. The findings presented in this chapter about the United States may help inform those searching to design a more effective and consistent curriculum for mathematics teacher education for primary and secondary teachers.

A. Appendices

Appendix 4.1

Table 4.15 Mathematics framework: content knowledge domains

Number	Whole numbers[ps]
	Fractions and decimals[ps]
	Number sentences[ps]
	Patterns and relationships[ps]
	Integers[ps]
	Ratios, proportions, and percent[ps]
	Irrational numbers[ps]
	Number theory[ps]
Geometry	Geometric shapes[ps]
	Geometric measurement[ps]
	Location and movement[ps]
Algebra	Patterns[ps]
	Algebraic expressions[ps]
	Equations/formulas and functions[ps]
	Calculus and analysis[s]
	Linear algebra and abstract algebra[s]
Data	Data organization and representation[ps]
	Data reading and interpretation[ps]
	Chance[ps]

Source: TEDS-M Framework 2008; TIMSS 2007 content domain assessment framework (Mullis et al., 2007); TIMSS 2008 advanced assessment frameworks (Garden et al., 2006)
Note: [p]primary level; [s]secondary level

Table 4.16 Mathematics framework: cognitive domains

Knowing	
Recall	Recall definitions; terminology; number properties; geometric properties; notation.
Recognize	Recognize mathematical objects, shapes, numbers and expressions; recognize mathematical entities that are mathematically equivalent.
Compute	Carry out algorithmic procedures for addition, multiplication, division, subtraction with whole numbers, fractions, decimals, and integers; approximate numbers to estimate computations; carry out routine algebraic procedures.
Retrieve	Retrieve information from graphs, tables, or other sources; read simple scales.
Measure	Use measuring instruments; use units of measurement appropriately; estimate measures.
Classify/order	Classify/group objects, shapes, numbers, and expressions according to common properties; make correct decisions about class membership; order numbers and objects by attributes.
Applying	
Select	Select an efficient/appropriate operation, method, or strategy for solving problems where there is a known algorithm or method of solution.
Represent	Display mathematical information and data in diagrams, tables, charts, or graphs; generate equivalent representations for a given mathematical entity or relationship.
Model	Generate an appropriate model, such as an equation or diagram, for solving a routine problem.
Implement	Follow and execute a set of mathematical instructions; draw figures and shapes according to given specifications.
Solve routine problems	Solve routine or familiar types of problems (e.g., use geometric properties to solve problems); compare and match different representations of data; use data from charts, tables, graphs, and maps to solve routine problems.
Reasoning	
Analyze	Determine and describe or use relationships between variables or objects in mathematical situations; use proportional reasoning; decompose geometric figures to simplify solving a problem; draw the net of a given unfamiliar solid; visualize transformations of three-dimensional figures; compare and match different representations of the same data; make valid inferences from given information.
Generalize	Extend the domain to which the result of mathematical thinking and problem solving is applicable by restating results in more general and more widely applicable terms.
Synthesize/integrate	Combine (various) mathematical procedures to establish results, and combine results to produce a further result; make connections between different elements of knowledge and related representations, and make linkages between related mathematical ideas.
Justify	Provide a justification for the truth or falsity of a statement by reference to mathematical results or properties.
Solve non-routine problems	Solve problems set in mathematical or real-life contexts where future teachers are unlikely to have encountered closely similar items, and apply mathematical procedures in unfamiliar or complex contexts; use geometric properties to solve non-routine problems.

Source: TIMSS 2007 cognitive domain assessment framework (Mullis et al., 2007)

Table 4.17 Mathematical Pedagogical Content Knowledge (MPCK) framework[a]

Mathematical curricular knowledge	Establishing appropriate learning goals
	Knowing different assessment formats
	Selecting possible pathways and seeing connections within the curriculum
	Identifying the key ideas in learning programs
	Knowledge of mathematics curriculum
Knowledge of planning for mathematics teaching and learning [pre-active]	Planning or selecting appropriate activities
	Choosing assessment formats
	Predicting[b] typical students' responses, including misconceptions
	Planning appropriate methods for representing mathematical ideas
	Linking the didactical methods and the instructional designs
	Identifying different approaches for solving mathematical problems
	Planning mathematical lessons
Enacting mathematics for teaching and learning [interactive]	Analyzing or evaluating students' mathematical solutions or arguments
	Analyzing the content of students' questions
	Diagnosing typical students' responses, including misconceptions
	Explaining or representing mathematical concepts or procedures
	Generating fruitful questions
	Responding to unexpected mathematical issues
	Providing appropriate feedback

[a]This framework paid attention to the temporal dimension of teacher knowledge as well as the way in which mathematical categories refer to different types of knowledge

[b]Attention to choice of verbs may prove useful in distinguishing between pre-active and interactive dimensions of the categories

Table 4.18 Teaching contact hours: definition of courses

Courses	Definition
Liberal arts (except mathematics)	These are experiences of a general or theoretical nature designed to develop judgment and understanding about human beings' relationship to the social, cultural, and natural environment (e.g., natural and social sciences, languages, drama, music, art, philosophy, and religion).
Academic mathematics	These experiences aim to provide mathematics knowledge to a more general population of university students, that may or may not include future teachers, and are designed to treat content beyond the mathematics learned at the secondary school level, that is, mathematics at the university level (e.g., "Abstract Algebra", "Functional Analysis", "Differential Equations", etc.).
Mathematics content related to the school mathematics curriculum	These are experiences dealing mainly with the structure, sequence, content, and level of competence required from pupils to successfully learn from the school mathematics curriculum (primary or secondary levels). Examples of such courses are "Structure and Content of the Lower Secondary Mathematics Curriculum", "Development and Understanding of the School Mathematics Curriculum", etc.
Mathematics pedagogy	Courses dealing with the methods of teaching and learning mathematics (e.g., mathematics pedagogy, didactics of mathematics). These courses could include treatment of pupils' cognition (e.g., how one learns mathematics) or pupils' thinking in relation to mathematics concepts. Examples of such types of units are courses like "Learner Diversity and the Teaching of Subject Matter: Mathematics," "Primary and Middle School Mathematics: Teaching Developmentally", etc.
Professional foundations and theories	Courses on the study of education utilizing such disciplines as history, philosophy, sociology, psychology, social psychology, anthropology, economics, and political science, or such interdisciplinary fields as comparative and international education, multicultural education, community and adult education, and many others. All such study stresses diverse perspectives in understanding, analyzing, and implementing educational theory and practice.
General pedagogy (not mathematics)	Courses on the "art or science of teaching" providing instruction on the correct use of teaching strategies. In addition, these courses include the study of the correlation of those teaching strategies with the instructor's own philosophical beliefs of teaching and pupils' background knowledge and experiences, personal situations, social and classroom environment, as well as setting learning goals.

Appendix 4.2: Institutional/Program Questionnaire: Analysis of Field Experience Activities

Table 4.19 Weighted mean frequency for field experience activities in the programs preparing future primary generalists (Grade 6 maximum)

Activities	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Plan lessons	461	2.06	1.19	378	1.96	1.23	83	2.49	0.89
Teach individual lessons to whole class	461	1.60	1.18	378	1.49	1.19	83	2.09	0.98
Tutor individual pupils	461	2.19	0.89	378	2.08	0.93	83	2.69	0.47
Work with small groups of pupils	461	2.29	0.88	378	2.20	0.93	83	2.69	0.47
Assist teachers in other ways	461	2.72	0.47	378	2.71	0.48	83	2.78	0.42
Assist in school activities outside classroom	458	1.27	0.81	375	1.36	0.84	83	0.89	0.53
Carry out case studies of selected pupils	457	1.85	0.96	378	1.72	0.96	79	2.49	0.71
Carry out classroom observation	461	2.93	0.25	378	2.94	0.25	83	2.92	0.27
Collect data for research projects	458	1.35	1.03	375	1.37	1.10	83	1.23	0.59
Visit families in their homes	458	0.13	0.41	375	0.13	0.42	83	0.13	0.34
Interview teachers and/or principals	457	2.15	0.83	378	2.23	0.83	79	1.79	0.74
Observe and/or participate in meetings	461	1.64	0.79	378	1.69	0.80	83	1.42	0.71

Source: TEDS-M Institutional Program Questionnaire, Question 3, Part E. The scale is: 0 = Not at all, 1 = Rarely, 2 = Sometimes, 3 = Usually

Table 4.20 Weighted mean frequency for field experience activities in the programs preparing future primary mathematics specialists

Activities	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Plan lessons	92	2.44	0.82	79	2.34	0.85	13	3.00	0.00
Teach individual lessons to whole class	92	2.19	1.06	79	2.05	1.09	13	3.00	0.00
Tutor individual pupils	92	2.87	0.33	79	2.91	0.29	13	2.66	0.49
Work with small groups of pupils	92	2.81	0.39	79	2.78	0.41	13	3.00	0.00
Assist teachers in other ways	92	3.00	0.00	79	3.00	0.00	13	3.00	0.00
Assist in school activities outside classroom	92	0.73	1.08	79	0.85	1.13	13	0.00	0.00
Carry out case studies of selected pupils	92	1.69	0.67	79	1.75	0.59	13	1.32	0.99
Carry out classroom observation	92	2.85	0.36	79	2.88	0.32	13	2.66	0.49
Collect data for research projects	92	1.37	0.87	79	1.26	0.89	13	2.00	0.00
Visit families in their homes	92	0.13	0.33	79	0.15	0.36	13	0.00	0.00
Interview teachers and/or principals	92	2.13	0.33	79	2.15	0.36	13	2.00	0.00
Observe and/or participate in meetings	92	1.45	0.98	79	1.58	0.98	13	0.66	0.49

Source: TEDS-M Institutional Program Questionnaire, Question 3, Part E. The scale is: 0 = Not at all, 1 = Rarely, 2 = Sometimes, 3 = Usually

Table 4.21 Weighted mean frequency for field experience activities in the programs preparing future middle school (Grade 10 maximum) teachers

Activities	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Plan lessons	92	2.44	0.82	79	2.34	0.86	13	3.00	0.00
Teach individual lessons to whole class	92	2.19	1.06	79	2.05	1.09	13	3.00	0.00
Tutor individual pupils	92	2.87	0.33	79	2.91	0.29	13	2.66	0.49
Work with small groups of pupils	92	2.81	0.39	79	2.78	0.41	13	3.00	0.00
Assist teachers in other ways	92	3.00	0.00	79	3.00	0.00	13	3.00	0.00
Assist in school activities outside classroom	92	0.73	1.08	79	0.85	1.13	13	0.00	0.00
Carry out case studies of selected pupils	92	1.69	0.67	79	1.75	0.59	13	1.32	0.99
Carry out classroom observation	92	2.85	0.36	79	2.88	0.32	13	2.66	0.49
Collect data for research projects	92	1.37	0.87	79	1.26	0.89	13	2.00	0.00
Visit families in their homes	92	0.13	0.34	79	0.15	0.36	13	0.00	0.00
Interview teachers and/or principals	92	2.13	0.34	79	2.15	0.36	13	2.00	0.00
Observe and/or participate in meetings	92	1.45	0.98	79	1.58	0.98	13	0.66	0.49

Source: TEDS-M Institutional Program Questionnaire, Question 3, Part E. The scale is: 0 = Not at all, 1 = Rarely, 2 = Sometimes, 3 = Usually

Table 4.22 Weighted mean frequency for field experience activities in the programs preparing future high school (to Grade 11 and above) teachers

Activities	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Plan lessons	346	2.19	1.00	265	2.09	1.06	81	2.51	0.71
Teach individual lessons to whole class	346	1.73	1.03	265	1.66	1.02	81	1.95	1.02
Tutor individual pupils	346	2.59	0.56	265	2.57	0.59	81	2.65	0.48
Work with small groups of pupils	346	2.46	0.64	265	2.49	0.68	81	2.35	0.48
Assist teachers in other ways	346	2.71	0.57	265	2.70	0.60	81	2.72	0.45
Assist in school activities outside classroom	346	1.23	0.86	265	1.27	0.85	81	1.09	0.90
Carry out case studies of selected pupils	339	1.41	0.90	259	1.48	0.94	81	1.19	1.04
Carry out classroom observation	346	2.97	0.16	265	2.97	0.18	81	3.00	0.00
Collect data for research projects	346	0.74	0.82	265	0.76	0.75	81	0.70	1.01
Visit families in their homes	346	0.25	0.48	265	0.23	0.47	81	0.34	0.48
Interview teachers and/or principals	346	1.82	0.86	265	1.85	0.84	81	1.74	0.95
Observe and/or participate in meetings	346	1.68	0.79	265	1.64	0.84	81	1.81	0.62

Source: TEDS-M Institutional Program Questionnaire, Question 3, Part E. The scale is: 0 = Not at all, 1 = Rarely, 2 = Sometimes, 3 = Usually

Appendix 4.3: Analysis of the TEDS-M Future Teachers' Questionnaire Responses of OTL Topics

School Mathematics

Table 4.23 Weighted mean percent of future teachers reporting to study school mathematics topics in the programs preparing future primary generalists (Grade 6 maximum)

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Numbers	22,268	98	14.0	20,261	99	11.4	2006	91	28.2
Measurement	22,268	96	18.7	20,261	97	17.2	2006	91	29.1
Geometry	22,268	84	36.4	20,261	86	35.2	2006	71	45.5
Functions, relations, and equations	22,249	75	43.4	20,250	77	42.3	1999	56	49.6
Data representation, probability, and statistics	22,250	86	35.0	20,244	87	33.8	2006	74	44.0
Calculus	22,237	22	41.2	20,230	22	41.2	2006	22	41.3
Validation, structuring, and abstracting	22,268	24	42.6	20,261	24	42.7	2006	22	41.7

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 2, Part B

Table 4.24 Weighted mean percent of future teachers reporting to study school mathematics topics in the programs preparing future primary mathematics specialists

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Numbers	3604	99	11.0	3432	99	11.3	172	100	0.0
Measurement	3597	97	18.1	3426	96	18.6	172	100	0.0
Geometry	3604	91	28.7	3432	90	29.4	172	100	0.0
Functions, relations, and equations	3604	81	39.2	3432	80	39.7	172	92	27.3
Data representation, probability, and statistics	3604	90	30.3	3432	90	30.1	172	86	34.8
Calculus	3604	25	43.3	3432	25	43.1	172	33	47.1
Validation, Structuring, and abstracting	3604	21	40.6	3432	22	41.3	172	0	0.0

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 2, Part B

Table 4.25 Weighted mean percent of future teachers reporting to study school mathematics topics in the programs preparing future middle school (Grade 10 maximum) teachers

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Numbers	4233	99	7.7	4037	99	7.9	196	100	0.0
Measurement	4233	97	17.0	4037	98	13.7	196	75	43.5
Geometry	4233	89	31.7	4037	92	27.5	196	24	43.0
Functions, relations, and equations	4233	87	34.2	4037	90	30.6	196	24	43.0
Data representation, probability, and statistics	4233	89	31.6	4037	91	28.9	196	46	49.9
Calculus	4233	20	40.0	4037	20	40.1	196	16	37.0
Validation, structuring, and abstracting	4233	15	35.9	4037	16	36.4	196	4	19.7

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 2, Part B

Table 4.26 Weighted mean percent of future teachers reporting to study school mathematics topics in the programs preparing future high school (to Grade 11 and above) teachers

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Numbers	2840	87	33.8	2237	89	30.9	603	78	41.6
Measurement	2840	76	42.6	2237	79	41.0	603	67	46.9
Geometry	2840	89	31.9	2237	92	27.6	603	77	42.2
Functions, relations, and equations	2840	84	36.4	2237	86	34.2	603	76	42.6
Data representation, probability, and statistics	2840	89	31.5	2237	91	28.4	603	80	39.8
Calculus	2840	75	43.4	2237	76	42.9	603	71	45.3
Validation, structuring, and abstracting	2840	59	49.2	2237	59	49.2	603	57	49.6

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 2, Part B

University Mathematics

Table 4.27 Weighted mean percent of future teachers reporting to study university mathematics topics in the programs preparing future primary generalists (Grade 6 maximum)

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Axiomatic geometry	22,267	46	49.8	20,265	47	49.9	2002	34	47.2
Coordinate geometry	22,278	60	49.0	20,276	62	48.5	2002	42	49.4
Non-Euclidean geometry	22,261	42	49.4	20,259	43	49.6	2002	30	45.7
Differential geometry	22,233	29	45.6	20,231	30	45.7	2002	26	43.6
Topology	22,171	13	33.1	20,197	12	32.7	1974	16	36.9
Linear algebra	22,200	51	50.0	20,226	53	49.9	1974	38	48.4
Set theory	22,219	39	48.7	20,217	40	49.0	2002	27	44.2
Abstract algebra	22,258	22	41.7	20,261	22	41.2	1997	31	46.2
Number theory	22,262	85	35.5	20,260	87	34.0	2002	70	46.0
Beginning calculus	22,263	50	50.0	20,261	50	50.0	2002	50	50.0
Calculus	22,263	30	45.9	20,261	29	45.4	2002	41	49.2
Multivariate calculus	22,257	15	35.2	20,255	13	34.1	2002	25	43.5
Advanced calculus	22,263	6	24.5	20,261	6	23.2	2002	13	33.8
Differential equations	22,201	27	44.5	20,199	26	44.1	2002	37	48.2
Theory of functions	22,216	26	43.7	20,214	25	43.5	2002	30	45.9
Discrete mathematics	22,178	27	44.6	20,181	27	44.4	1997	31	46.5
Probability	22,263	86	34.4	20,261	87	33.8	2002	81	39.1
Theoretical or applied statistics	22,247	49	50.0	20,245	48	50.0	2002	56	49.6
Mathematical logic	22,263	50	50.0	20,261	49	50.0	2002	52	50.0

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 1, Part B

Table 4.28 Weighted mean percent of future teachers reporting to study university mathematics topics in the programs preparing future primary mathematics specialists

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Axiomatic geometry	3604	59	49.2	3432	59	49.2	172	69	46.2
Coordinate geometry	3604	75	43.3	3432	74	43.6	172	86	34.8
Non-Euclidean geometry	3604	49	50.0	3432	47	49.9	172	86	34.8
Differential Geometry	3604	37	48.2	3432	35	47.6	172	74	44.0
Topology	3604	15	36.1	3432	12	33.1	172	74	44.0
Linear algebra	3604	62	48.6	3432	60	49.0	172	94	23.7
Set theory	3557	42	49.4	3385	42	49.3	172	55	49.9
Abstract algebra	3604	22	41.6	3432	21	40.4	172	55	49.9
Number theory	3604	90	30.1	3432	90	29.8	172	86	34.8
Beginning calculus	3604	55	49.7	3432	53	49.9	172	100	0.0
Calculus	3604	37	48.4	3432	34	47.5	172	100	0.0
Multivariate calculus	3604	20	40.2	3432	18	38.6	172	61	48.9
Advanced calculus	3604	11	31.4	3432	9	28.5	172	55	49.9
Differential equations	3604	34	47.4	3432	33	46.9	172	61	48.9
Theory of functions	3604	32	46.6	3432	31	46.1	172	55	49.9
Discrete mathematics	3604	31	46.2	3432	29	45.5	172	61	48.9
Probability	3604	88	32.6	3432	89	31.9	172	75	43.2
Theoretical or applied statistics	3604	47	49.9	3432	46	49.9	172	67	47.1
Mathematical logic	3604	49	50.0	3432	47	49.9	172	92	27.3

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 1, Part B

Table 4.29 Weighted mean percent of future teachers reporting to study university mathematics topics in the programs preparing future middle school (Grade 10 maximum) teachers

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Axiomatic geometry	4216	49	50.0	4037	50	50.0	180	36	48.2
Coordinate geometry	4233	69	46.4	4037	70	46.0	196	50	50.1
Non-Euclidean geometry	4233	49	50.0	4037	51	50.0	196	8	27.3
Differential geometry	4233	30	45.7	4037	31	46.2	196	4	19.7
Topology	4233	11	31.7	4037	12	32.3	196	0	0.0
Linear algebra	4233	54	49.8	4037	55	49.8	196	50	50.1
Set theory	4233	37	48.3	4037	36	48.0	196	58	49.4
Abstract algebra	4233	16	36.4	4037	16	36.7	196	8	27.3
Number theory	4233	89	30.9	4037	89	31.1	196	92	27.3
Beginning calculus	4233	55	49.8	4037	55	49.8	196	50	50.1
Calculus	4233	36	48.0	4037	35	47.8	196	50	50.1
Multivariate calculus	4177	10	29.4	3981	8	27.1	196	41	49.4
Advanced calculus	4233	7	25.3	4037	7	25.5	196	4	19.7
Differential equations	4233	31	46.3	4037	31	46.1	196	37	48.5
Theory of functions	4233	19	38.9	4037	19	39.3	196	8	27.6
Discrete mathematics	4233	18	38.6	4037	18	38.7	196	16	37.0
Probability	4233	88	32.6	4037	90	30.1	196	46	49.9
Theoretical or applied statistics	4233	41	49.1	4037	39	48.8	196	67	47.2
Mathematical logic	4233	42	49.3	4037	42	49.4	196	37	48.5

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 1, Part B

Table 4.30 Weighted mean percent of future teachers reporting to study university mathematics topics in the programs preparing future high school (to Grade 11 and above) teachers

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Axiomatic geometry	2836	89	31.1	2233	92	27.1	603	79	41.0
Coordinate geometry	2845	82	38.0	2242	82	38.1	603	83	37.7
Non-Euclidean geometry	2845	76	42.5	2242	80	40.3	603	64	48.0
Differential geometry	2837	30	46.0	2234	30	45.9	603	31	46.2
Topology	2840	15	36.0	2237	14	35.2	603	18	38.8
Linear algebra	2845	93	26.1	2242	93	24.7	603	89	30.7
Set theory	2840	68	46.6	2237	71	45.2	603	56	49.7
Abstract algebra	2845	75	43.2	2242	78	41.7	603	66	47.4
Number theory	2845	85	35.8	2242	85	36.1	603	86	34.6
Beginning calculus	2840	96	19.2	2237	96	19.8	603	97	16.9
Calculus	2832	95	20.8	2237	95	22.2	596	98	14.0
Multivariate calculus	2831	84	36.3	2228	85	35.2	603	80	39.8
Advanced calculus	2823	47	49.9	2232	46	49.9	591	50	50.0
Differential equations	2840	64	48.1	2237	63	48.2	603	65	47.9
Theory of functions	2837	35	47.6	2234	32	46.7	603	44	49.7
Discrete mathematics	2834	72	45.1	2231	73	44.3	603	65	47.6
Probability	2834	94	23.2	2231	95	21.6	603	91	28.1
Theoretical or applied statistics	2840	76	42.7	2237	76	42.4	603	74	43.7
Mathematical logic	2840	81	39.0	2237	81	39.3	603	83	37.9

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 1, Part B

Mathematics Pedagogy

Table 4.31 Weighted mean percent of future teachers reporting to study mathematics pedagogy topics in the programs preparing future primary generalists (Grade 6 maximum)

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Foundations of mathematics	22,628	38	48.4	20,597	38	48.6	2031	32	46.5
Context of mathematics education	22,628	40	48.9	20,597	39	48.9	2031	43	49.5
Development of mathematics ability and thinking	22,628	60	49.0	20,597	61	48.8	2031	48	50.0
Mathematics instruction	22,628	72	45.0	20,597	72	44.7	2031	66	47.6
Developing teaching plans	22,628	61	48.7	20,597	62	48.6	2031	55	49.7
Mathematics teaching	22,628	66	47.3	20,597	67	47.0	2031	59	49.2
Mathematics standards and curriculum	22,628	71	45.6	20,597	71	45.3	2031	64	47.9
Affective issues in mathematics	22,628	52	49.9	20,597	53	49.9	2031	45	49.7

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 4, Part B

Table 4.32 Weighted mean percent of future teachers reporting to study mathematics pedagogy topics in the programs preparing future primary mathematics specialists

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Foundations of mathematics	2816	48	50.0	2655	50	50.0	161	15	35.7
Context of mathematics education	2816	53	49.9	2655	54	49.8	161	26	44.1
Development of mathematics ability and thinking	2816	76	42.7	2655	77	42.3	161	67	47.0
Mathematics instruction	2816	91	28.9	2655	90	29.6	161	100	0.0
Developing teaching plans	2816	86	34.9	2655	87	34.1	161	74	44.1
Mathematics teaching	2816	88	32.2	2655	88	33.0	161	100	0.0
Mathematics standards and curriculum	2816	95	21.4	2655	95	22.0	161	100	0.0
Affective issues in mathematics	2816	64	47.9	2655	63	48.2	161	79	41.0

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 4, Part B

Table 4.33 Weighted mean percent of future teachers reporting to study mathematics pedagogy topics in the programs preparing future middle school (Grade 10 maximum) teachers

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Foundations of mathematics	3133	53	49.9	2937	56	49.7	196	16	37.0
Context of mathematics education	3133	50	50.0	2937	52	50.0	196	33	47.2
Development of mathematics ability and thinking	3133	83	37.8	2937	82	38.7	196	100	0.0
Mathematics instruction	3133	93	25.8	2937	93	25.6	196	92	27.6
Developing teaching plans	3133	92	27.4	2937	92	26.9	196	88	32.9
Mathematics teaching	3133	92	27.7	2937	91	28.5	196	100	0.0
Mathematics standards and curriculum	3133	97	18.3	2937	96	18.9	196	100	0.0
Affective issues in mathematics	3133	61	48.7	2937	61	48.7	196	58	49.4

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 4, Part B

Table 4.34 Weighted mean percent of future teachers reporting to study mathematics pedagogy topics in the programs preparing future high school (to Grade 11 and above) teachers

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Foundations of mathematics	2367	65	47.7	1928	68	46.7	439	53	50.0
Context of mathematics education	2370	39	48.8	1931	38	48.5	439	46	49.9
Development of mathematics ability and thinking	2370	71	45.4	1931	73	44.4	439	62	48.7
Mathematics instruction	2370	88	32.6	1931	88	32.0	439	85	35.3
Developing teaching plans	2370	79	40.4	1931	79	40.5	439	80	40.0
Mathematics teaching	2355	91	28.7	1915	91	28.4	439	90	30.2
Mathematics standards and curriculum	2370	91	28.7	1931	91	29.2	439	93	26.3
Affective issues in mathematics	2370	54	49.8	1931	54	49.8	439	56	49.7

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 4, Part B

General Pedagogy

Table 4.35 Weighted mean percent of future teachers reporting to study general pedagogy topics in the programs preparing future primary generalists (Grade 6 maximum)

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
History of education	17,254	63	48.2	15,846	64	48.1	1407	60	49.0
Philosophy of education	17,234	91	27.9	15,826	91	28.1	1407	93	25.3
Sociology of education	17,234	83	37.8	15,826	83	37.9	1407	84	37.1
Educational psychology	17,254	95	21.0	15,846	96	20.5	1407	93	25.9
Theories of schooling	17,241	87	34.1	15,833	87	33.7	1407	83	37.9
Methods of educational research	17,254	74	43.6	15,846	74	44.0	1407	83	37.2
Assessment and measurement	17,228	86	34.8	15,820	86	34.8	1407	85	35.7
Knowledge of teaching	17,234	95	22.6	15,826	94	22.9	1407	97	17.6

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 7, Part B

Table 4.36 Weighted mean of percent future teachers reporting to study general pedagogy topics in the programs preparing future primary mathematics specialists

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
History of education	2816	61	48.8	2655	62	48.6	161	41	49.4
Philosophy of education	2797	94	24.6	2636	93	25.3	161	100	0.0
Sociology of education	2812	83	37.8	2651	82	38.3	161	94	24.4
Educational psychology	2812	94	24.1	2651	93	24.8	161	100	0.0
Theories of schooling	2816	82	38.7	2655	85	36.1	161	33	47.0
Methods of educational research	2816	76	42.9	2655	75	43.3	161	85	35.7
Assessment and measurement	2812	90	29.7	2651	90	29.8	161	91	28.1
Knowledge of teaching	2812	95	21.7	2651	95	21.2	161	91	28.1

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 7, Part B

Table 4.37 Weighted mean percent of future teachers reporting to study general pedagogy topics in the programs preparing future middle school (Grade 10 maximum) teachers

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
History of education	3151	70	45.7	2955	71	45.4	196	63	48.5
Philosophy of education	3151	96	20.0	2955	96	19.4	196	92	27.6
Sociology of education	3151	86	34.2	2955	88	32.4	196	63	48.5
Educational psychology	3151	100	6.4	2955	100	6.6	196	100	0.0
Theories of schooling	3151	90	29.5	2955	92	26.8	196	63	48.5
Methods of educational research	3151	75	43.3	2955	74	43.8	196	88	32.9
Assessment and measurement	3151	83	37.2	2955	82	38.2	196	100	0.0
Knowledge of teaching	3151	97	18.0	2955	97	17.1	196	92	27.6

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 7, Part B

Table 4.38 Weighted mean percent of future teachers reporting to study general pedagogy topics in the programs preparing future high school (to Grade 11 and above) teachers

Topics	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
History of education	2366	61	48.7	1927	62	48.7	439	60	49.1
Philosophy of education	2366	92	27.4	1927	93	24.9	439	85	35.6
Sociology of education	2366	81	39.5	1927	82	38.4	439	74	43.8
Educational psychology	2366	94	24.6	1927	93	24.7	439	94	24.2
Theories of schooling	2366	76	42.5	1927	79	40.9	439	66	47.5
Methods of educational research	2366	58	49.4	1927	56	49.7	439	66	47.6
Assessment and measurement	2345	72	45.1	1906	73	44.6	439	67	46.9
Knowledge of teaching	2362	88	32.0	1927	91	29.2	435	79	40.8

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 7, Part B

School Experience Activities

Table 4.39 Weighted mean frequency future teachers reported for school experience activities required in the programs preparing future primary generalists (Grade 6 maximum)

Activities	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Observe models of the teaching strategies	16,906	2.29	0.77	15,539	2.31	0.77	1367	2.11	0.82
Practice theories for teaching mathematics	16,906	2.09	0.88	15,539	2.11	0.87	1367	1.88	0.96
Complete assessment tasks that asked you to show how you were applying ideas you were learning in your courses	16,871	2.17	0.84	15,512	2.18	0.84	1359	2.06	0.93
Receive feedback about your teaching strategies	16,835	2.44	0.75	15,481	2.45	0.74	1355	2.27	0.85
Collect and analyze evidence about pupils' learning	16,843	2.28	0.80	15,488	2.30	0.79	1355	2.05	0.89
Test out findings from educational research about pupils' difficulties	16,786	1.49	0.96	15,464	1.51	0.95	1322	1.28	0.99
Develop strategies to reflect upon your professional knowledge	16,776	2.24	0.82	15,449	2.26	0.82	1327	2.08	0.86
Demonstrate that you could apply the teaching methods	16,721	2.49	0.71	15,394	2.51	0.69	1327	2.31	0.80

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 13, Part B. The scale is 0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Often

Table 4.40 Weighted mean frequency future teachers reported for school experience activities required in the programs preparing future primary mathematics specialists

Activities	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Observe models of the teaching strategies	2735	2.12	0.87	2573	2.11	0.87	161	2.31	0.88
Practice theories for teaching mathematics	2691	2.20	0.90	2572	2.20	0.90	119	2.18	0.82
Complete assessment tasks that asked you to show how you were applying ideas you were learning in your courses	2714	2.31	0.82	2553	2.32	0.81	161	2.11	0.96
Receive feedback about your teaching strategies	2714	2.38	0.77	2553	2.38	0.77	161	2.35	0.87
Collect and analyze evidence about pupils' learning	2714	2.44	0.73	2553	2.44	0.74	161	2.46	0.61
Test out findings from educational research about pupils' difficulties	2714	1.62	0.97	2553	1.60	0.95	161	1.99	1.09
Develop strategies to reflect upon your professional knowledge	2701	2.26	0.85	2553	2.28	0.80	148	1.86	1.41
Demonstrate that you could apply the teaching methods	2701	2.53	0.71	2553	2.55	0.68	148	2.15	1.06

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 13, Part B. The scale is 0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Often

Table 4.41 Weighted mean frequency future teachers reported for school experience activities required in the programs preparing future middle school (Grade 10 maximum) teachers

Activities	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Observe models of the teaching strategies	3144	1.25	0.70	2947	2.29	0.69	196	1.66	1.03
Practice theories for teaching mathematics	3137	1.30	0.70	2941	2.31	0.73	196	2.17	0.80
Complete assessment tasks that asked you to show how you were applying ideas you were learning in your courses	3137	2.25	0.73	2941	2.27	0.86	196	2.13	0.83
Receive feedback about your teaching strategies	3137	2.30	0.73	2941	2.46	0.70	196	1.66	0.74
Collect and analyze evidence about pupils' learning	3137	2.26	0.86	2941	2.51	0.69	196	2.46	0.70
Test out findings from educational research about pupils' difficulties	3137	2.41	0.73	2941	1.76	0.97	196	1.00	0.82
Develop strategies to reflect upon your professional knowledge	3152	2.50	0.69	2956	2.50	0.66	196	2.29	0.94
Demonstrate that you could apply the teaching methods	3152	1.71	0.98	2956	2.66	0.61	196	2.29	0.94

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 13, Part B. The scale is 0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Often

Table 4.42 Weighted mean frequency future teachers reported for school experience activities required in the programs preparing future high school (to Grade 11 and above) teachers

Activities	Routes								
	All programs			Concurrent route			Consecutive route		
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Observe models of the teaching strategies	2241	1.91	0.89	1831	1.89	0.90	410	1.98	0.83
Practice theories for teaching mathematics	2241	1.99	0.93	1831	1.97	0.93	410	2.06	0.95
Complete assessment tasks that asked you to show how you were applying ideas you were learning in your courses	2235	1.98	0.96	1824	1.96	0.964	410	2.05	0.94
Receive feedback about your teaching strategies	2221	2.33	0.81	1810	2.36	0.81	410	2.18	0.82
Collect and analyze evidence about pupils' learning	2235	2.10	0.88	1824	2.14	0.86	410	1.91	0.92
Test out findings from educational research about pupils' difficulties	2233	1.24	0.88	1822	1.21	0.85	410	1.36	1.00
Develop strategies to reflect upon your professional knowledge	2228	2.14	0.87	1822	2.17	0.84	405	1.97	0.98
Demonstrate that you could apply the teaching methods	2228	2.35	0.80	1822	2.36	0.80	405	2.30	0.84

Source: TEDS-M Future Teachers of Primary Mathematics Questionnaire, Question 13, Part B. The scale is 0 = Never, 1 = Rarely, 2 = Occasionally, 3 = Often

References

Association of Mathematics Teacher Educators. (2017). *Standards for preparing teachers of mathematics*. Retrieved from http://amte.net/standards

Blömeke, S. (2012). Content, professional preparation, and teaching methods: How diverse is teacher education across countries? *Comparative Education Review, 56*(4), 684–714.

Boyd, D., Grossman, P., Hammerness, K. M., Lankford, H., Loeb, S., McDonald, M., … Wyckoff, J. (2008). Surveying the landscape of teacher education in New York City: Constrained variation and the challenge of innovation. *Educational Evaluation and Policy Analysis, 30*, 319–343.

Boyd, D., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 416–440.

Buchmann, M., & Schmidt, W. (1981). *The school day and content commitments*. East Lansing, MI: Institute for Research on Teaching.

Carroll, J. (1963). A model for school learning. *Teachers College Record, 64*, 723–733.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255–296). New York, NY: Macmillan.

Constantine, J., Player D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification, final report* (NCEE 2009-4043). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Heilig, J. V. (2005). Does teacher preparation matter? Evidence about teacher certification, teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 13*(42), 1–48.

Garden, R., Lie, S., Robitaille, D. F., Angell, C., Martin, M. O., Mullis, I. V. S., … Arora, A. (2006). *TIMSS Advanced 2008 assessment frameworks*. Chestnut Hill, MA: Boston College.

Grossman, P. L., & Stodolsky, S. S. (1994a). Considerations of content and the circumstances of secondary school teaching. In L. Darling-Hammond (Ed.), *Review of Research in Education* (pp. 179–221). Washington, DC: American Educational Research Association.

Grossman, P., & Stodolsky, S. (1994b). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher, 24*(8), 5–11.

Husen, T. (Ed.). (1967a). *International study of achievement in mathematics: A comparison of twelve countries* (Vol. 1). New York, NY: Wiley.

Husen, T. (Ed.). (1967b). *International study of achievement in mathematics: A comparison of twelve countries* (Vol. II). New York, NY: Wiley.

Kennedy, M. (2015). Parsing the practice of teaching. *Journal of Teacher Education, 67*(1), 6–17.

Lampert, M., & Ball, D. (1998). *Teaching, multimedia and mathematics: Investigations of real practice*. New York, NY: Teacher's College Press.

Lampert, M., Franke, M. L., Kazemi, E., Gitomer, D., Turrou, C., Beasley, H., … Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education, 64*(3), 226–243.

Levine, A. (2006). *Educating school teachers*. Washington, DC: The Education Schools Project.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis, 17*(3), 305–322.

McKnight, C. C., Crosswhite, F. J., Dossey, J. A., Kifer, E., Swafford, J. O., Travers, K. J., et al. (1987). *The underachieving curriculum: Assessing US school mathematics from an international perspective*. Champaign, IL: Stipes Publishing Company.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2007). *TIMSS 2007 Assessment frameworks*. Chestnut Hill, MA: Boston College.

Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.

Tatto, M. T. (2000). *Assessing what we know about teacher quality and development: Empirical indicators and methodological issues in comparative perspective*. Report commissioned by the Board on Comparative and International Studies in Education (BICSE) National Academy of Sciences/National Research Council, Washington, DC.

Tatto, M. T. (Ed.). (2013). *The teacher education and development study in mathematics (TEDS-M). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report*. Amsterdam: International Association for the Evaluation of Student Achievement.

Tatto, M. T., & Hordern, J. (2017). The configuration of teacher education as a professional field of practice: a comparative study of mathematics education. In G. Whitty & J. Furlong (Eds.), *Knowledge and the study of education: An international exploration*, *Oxford Comparative Education Series* (pp. 255–274). Oxford, UK: Symposium Books.

Tatto, M. T., Nielsen, H. D., Cummings, W. C., Kularatna, N. G., & Dharmadasa, D. H. (1991). *Comparing the effects and costs of different approaches for educating primary school teachers: The case of Sri Lanka*, *BRIDGES Research Report Series* (Vol. 10). Cambridge, MA: Harvard Institute for International Development.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, Michigan State University, and IEA.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Tatto, M. T., & Senk, S. (2011). The mathematics education of future primary and secondary teachers: Methods and findings from the teacher education and development study in mathematics. *Journal of Teacher Education, 62*, 121–137.

U.S. Department of Education, National Center for Education Statistics. (1997). *The condition of education 1997*. Washington, DC: U.S. Government Printing Office Retrieved from https://nces.ed.gov/pubs97/97388.pdf

Valverde, G., Bianchi, L., Wolfe, R., Schmidt, W., & Houang, R. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Norwell, MA: Kluwer Academic Publishers.

Tatto. M.T. (ed.) (2013). The Teacher Education and Development Study in Mathematics (TEDS-M). Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries: Technical Report. Amsterdam: International Association for the Evaluation of Student Achievement.

Chapter 5
Developing Diverse Teachers: Analyzing Primary Mathematics Teacher Education Programs Prioritizing Selection of Diverse Future Teachers

James Pippin

Abstract Education research indicates that the academic achievement and school experiences of marginalized students improve when they have diverse teachers. Therefore, many national and state governments have implemented policies to recruit and prepare a more diverse group of teachers. However, little research has explored programs that prepare future primary mathematics teachers from marginalized populations in less developed or non-Western countries. This chapter describes variation in the selection policies and goals of these programs, as well as future teachers' reported opportunities to learn to teach diverse students in the United States, Chinese Taipei, Philippines, and Thailand. Results indicate some diversity among future teachers and alignment between program goals and teacher-reported opportunities to learn to teach diverse students.

Introduction

Due to greater access to education, international migration, and demographic shifts, teachers in many countries are working in schools with increasingly diverse populations of students (OECD, 2010; UNESCO, 2014). In many of these contexts, achievement gaps persist between traditionally marginalized students and their more advantaged classmates (Clark, 2014). Facing challenges to educational access,

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

J. Pippin (✉)
Michigan State University, East Lansing, MI, USA
e-mail: pippinja@msu.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_5

marginalized children often are unable to enroll in or complete primary school (Sabates, Akyeampong, Westbrook, & Hunt, 2010).

While any approach to minimizing or eliminating student achievement gaps and drop-out rates must be multipronged, one strategy is to develop a diverse teacher labor force that includes more teachers from marginalized populations (Villegas & Davis, 2007, 2008; Villegas & Irvine, 2010). Given the different regions and countries included in this study, it is important to note here that the term "diverse" or "marginalized" can be understood differently within and across cultures and countries. For example, scholars from the United States often highlight differences or diversity in teachers' racial, ethnic, or linguistic backgrounds or characteristics while those in countries with more homogeneous populations might focus more on differences in socioeconomic status. Clearly there are other important indicators of diversity as well, including age, gender, sexual orientation, disability, location, religion, and refugee status, to name a few. In this chapter, I generally use the terms "diverse" or "diversity" to refer to these characteristics and "marginalized" to refer to individuals or groups who traditionally experience disadvantages related to these characteristics.

A growing body of literature suggests that teachers from marginalized populations can serve as role models for marginalized students (Dee, 2005; Faez, 2012; Klopfenstein, 2005; Miller & Endo, 2005), are best positioned to understand and teach these students (Achinstein & Ogawa, 2011; Milner, 2006; Santoro & Reid, 2006) and often agree to teach and remain in "hard-to-staff" districts or schools (Achinstein, Ogawa, Sexton, & Freitas, 2010; Santoro & Reid, 2006). Perhaps most importantly, their presence can improve the achievement and school experiences of marginalized students (Dee 2004; Egalite, Kisida, & Winters, 2015).

Despite the advantages just cited, teachers from marginalized populations are underrepresented in schools in the United States (Boser, 2014; Ingersoll & May, 2011; Villegas, Strom, & Lucas, 2012) and abroad (OECD, 2010; Yen, 2009). In the United States, teachers from marginalized populations are also underrepresented in content areas (e.g. bilingual, mathematics, and science education) that are considered critical to the academic success of marginalized students (Flores, Clark, Claeys, & Villarreal, 2007). Given the benefits to these students of learning from qualified teachers from diverse populations and the importance of addressing achievement gaps in key content areas at the primary level, recruiting and preparing knowledgeable and diverse primary teachers is an important task for policymakers and educators globally.

In an effort to increase the number, quality, and diversity of teachers in America, the U.S. Department of Education (2010) recently launched a national teacher recruitment campaign. At least 31 states in the United States have also implemented a range of policies and programs to recruit and prepare teacher candidates from marginalized populations (Villegas et al., 2012). Governments in other countries also acknowledge the need to implement policies and programs that prepare greater numbers of teachers from marginalized populations. For example, Australia (Commonwealth of Australia, 2003), Canada (Ryan, Pollack, & Antonelli, 2009), the Netherlands (Wolff, Severiens, & Meeuwisse, 2010), and New Zealand (Howard, 2010) have implemented such policies. Yet countries with these policies and programs are typically highly-developed and/or Western countries. Little is known about programs which select and prepare future teachers from marginalized populations in less-developed countries or in other regions of the world.

One region that has been largely unexamined is Asia. Although numerous news reports inform readers of the superior test scores demonstrated by students in China (Shanghai), South Korea, Japan, Singapore, and Chinese Taipei on international standardized tests (e.g. Coughlin, 2015), there is considerable variation in the education systems, student populations, and preparation of teachers throughout the region. Therefore, in addition to the United States, which has a history of implementing diversity oriented policies, this study analyzes data from—Chinese Taipei, Philippines, and Thailand—which represent this variation across the region.

In addition to a lack of literature regarding programs that prepare teachers from marginalized populations in Asia, it is unclear whether and how these programs meet their goals in terms of preparing future primary teachers who have opportunities to learn about, and are prepared to teach, marginalized students.

This chapter draws upon the literature regarding teacher diversity and teacher education for marginalized populations to consider the following research questions:

1. What are the selection criteria and program goals of teacher preparation programs that emphasize the selection of future primary mathematics teachers from marginalized populations in the United States, Chinese Taipei, Philippines, and Thailand?
2. Do these programs meet their selection and program goals regarding future primary mathematics teachers' opportunities to learn about, and preparedness to teach, marginalized children?

To answer these questions, I analyze nationally representative data from the Teacher Education Development Study in Mathematics (TEDS-M), a large-scale cross-national study of teacher preparation programs and future mathematics teachers. I hypothesize that future primary teachers in programs that emphasize the selection of candidates from marginalized populations will report more opportunities to teach students from disadvantaged backgrounds, as well as those with behavioral or emotional problems and learning disabilities. These future teachers will likely have fewer opportunities to learn to teach gifted students and those with physical disabilities.

Literature Review

Education researchers identify two primary reasons for increasing the diversity of the teacher workforce: *demographic* and *democratic* imperatives (Achinstein & Ogawa, 2011). The *demographic* imperative arises from the persistent gap between the backgrounds of students and their teachers. For example, in the United States, students of color make up 45% of the student population, while teachers of color only make up 17% of the teacher work force (Boser, 2011). In fact, the same report notes that every state in the United States has a teacher diversity gap (i.e., difference between the percentage of students of color and teachers of color). Internationally, similar teacher diversity gaps are increasing throughout countries in the Organization for Economic Cooperation and Development (OECD). In Australia, for instance, 23% of school children speak a language other than English and 4% of children are from Indigenous populations; however, only 13% of teachers are from non-English speaking backgrounds and less than 1% are from Indigenous populations (Cruickshank, 2004; Santoro, 2015). Policymakers in these countries recognize a need for coherent policies and plans to attract, develop, and retain teachers from marginalized populations (Keane & Heinz, 2016; OECD 2010), highlighting the *democratic* imperative to improve educational experiences and outcomes (e.g., achievement and drop-out rates) of marginalized students (Achinstein & Ogawa, 2011).

Within the past 15 years, several studies have found evidence that teachers from marginalized populations are instrumental in addressing this democratic imperative. First, teachers from marginalized populations often serve as role models for marginalized students. In fact, studies of pre-service and in-service teachers of color indicate that these teachers entered the profession with the intent to serve as role models for students of color (Villegas et al., 2012). The intentions of teachers of color to be role models for their students seem to have the intended outcomes. For example, Miller and Endo (2005) conducted a qualitative study of eight students of color (three of the participants identified as Asian, two as African-American, and three as Latin American) enrolled in a teacher education program at a large Midwestern university in the United States. The authors found that participants were strongly influenced to become teachers by their own family members who were teachers and by the teachers of color they had in school. Teachers in these studies support the idea that increasing the proportion of teachers from marginalized populations in schools can have a positive influence on students by more accurately representing the population, demonstrating a more equitable distribution of power in society, and presenting possibilities of a different future. Scholars in other countries have found similar results. For example, Santoro and Reid (2006) conducted 3-year longitudinal case studies of 25 newly graduated teachers from Indigenous populations and found that these teachers were effective cultural experts (e.g., facilitating students' understanding of Indigenous content), cultural bridges between Indigenous and non-Indigenous students and colleagues, and role models for students from Indigenous populations.

Teachers from marginalized populations may also be best positioned to understand and teach marginalized students (Achinstein & Ogawa, 2011). Drawing on the work of scholars like Ladson-Billings (1994) and Gay and Kirkland (2003) in a qualitative study of educational researchers, Milner (2006) highlights the culturally informed relationships often cultivated between students and teachers of color in the United States. These relationships develop because of teachers' insider awareness and understanding of students' lived experiences in and outside the classroom. Relationships with students enable teachers to engage more meaningfully with parents as well, which in turn reinforces success with students. Faez (2012) found similar results in Canada; in a mixed-methods study of 25 culturally and linguistically diverse teaching candidates from Ontario University, the author found evidence that compared to their classmates, these candidates tended to be more empathetic towards English language learners. As a result of their own personal and professional experiences learning the language, they were able to better understand and meet the needs of their English language learners.

The intention to serve as role models and engage in culturally informed relationships seems to encourage teachers from marginalized populations to agree to teach and remain in hard-to-staff districts or schools (Achinstein et al., 2010). In fact, a recent analysis of data from the Schools and Staffing Survey found that U.S. teachers from marginalized populations are two to three times more likely than other teachers to work in schools in urban communities and schools characterized by high poverty and large numbers of minority students (Ingersoll, Merrill, & Stuckey, 2014). Similarly, Santoro and Reid (2006) noted that due to cultural norms of community, Australian teachers from Indigenous populations may be more likely than non-Indigenous teachers to remain in teaching positions located in schools near their ancestral lands.

Perhaps most importantly, recent studies indicate that teachers from marginalized populations can improve the achievement and school experiences of marginalized students. For example, Dee (2004) analyzed data from Tennessee's Project STAR large-scale randomized experiment, which randomly assigned both students and teachers to small class sizes, regular class sizes, and regular class sizes where the teacher had the help of an aide. The random assignment of teachers and students offered a unique opportunity to examine educational outcomes of students paired with a teacher from their own race. Findings indicated that exposure to a teacher from the same race resulted in significant achievement gains for Black and White students (Dee, 2004).

Klopfenstein (2005) analyzed the Texas Schools Microdata Panel (TSMP) data, which includes a sample of 20,091 Black high school geometry students (9,120 males, 10,971 females) in grades 9–11 who attended Texas high schools between 1997 and 1998, to determine the likelihood that a Black math student would choose to enroll in a rigorous math class the following year if he/she had a Black teacher. The author found that as the proportion of female Black math teachers increased by one standard deviation above the mean, the probability of male Black students enrolling in subsequent rigorous math classes increased by 9.6%. Additionally, increasing the proportion of male Black math teachers by one standard deviation above the mean increased the probability that female Black students enrolled in rigorous math classes by 6.2%.

Finally, Egalite et al. (2015) analyzed student data from the universe of public school students taking the Florida Comprehensive Assessment Test (FCAT) in grades 3–10 between the years of 2001 and 2009. More than 2.9 million students linked to 92,000 teachers are included in the dataset. The authors found significant positive impacts of a race match between students and teachers in reading (0.005 SD) and math (0.013 SD) in the primary grades. The impact was higher for Black students (0.019 SD in math and 0.004 SD in reading). The authors noted that while these numbers may seem small, they represent achievement from just one year.

It is important to recognize that these studies do not imply that teachers from majority populations cannot be effective teachers of diverse students or that teachers of color are necessarily effective with students of color (Achinstein et al., 2010). Many have argued that all teachers should experience preparation programs that provide opportunities to learn knowledge and practices that enable them to effectively teach marginalized students (e.g., Gay, 2002; Ladson-Billings, 2000). International scholars have drawn similar conclusions; for example, in Canada (Faez, 2012), results indicated that high levels of empathy for English language learners found among linguistically and culturally diverse teachers were not necessarily associated with high levels of inclusive pedagogies or ideology. Similarly, in Australia (Santoro & Reid, 2006), teachers from Indigenous populations expressed frustration with their perpetual status as cultural representatives in schools and argued that all teachers should be prepared to teach marginalized students. Scholars and policymakers alike have worked to set standards for teacher preparation programs that would accomplish this task.

In the past two decades, numerous education organizations in the United States have sought to define the standards and characteristics of exemplary teacher preparation programs (e.g., Council for the Accreditation of Educator Preparation, 2015; National Commission on Teaching and America's Future, 1996; National Council on Teacher Quality, 2015; Teacher Education Accreditation Council, 2009). During this time, education scholars have also conducted a host of studies on teacher preparation programs (see Cochran-Smith & Villegas, 2015 and Cochran-Smith et al., 2015 for extensive reviews). Through these efforts, several characteristics of exemplary teacher preparation programs—those that produce teachers who are prepared to teach in ways that are learner-centered and learning-centered (Darling-Hammond, 2000)—have emerged.

Teachers educated in exemplary programs are prepared to teach students from a range of backgrounds, needs, and interests in ways that support deep thinking and learning that is demonstrated in content proficiency. Research indicates that programs producing teachers with such skills tend to have some features in common: They (a) carefully oversee student teaching experiences, (b) match the student teaching experiences with their likely teaching assignments, (c) require substantial coursework in reading and mathematics, (d) provide opportunities to learn specific practices that are then used in student teaching, (e) provide opportunities to learn about the local curriculum, (f) require a capstone project, and (g) have a significant number of tenured faculty (Darling-Hammond, 2010).

Villegas and Irvine (2010) helpfully identified five practices of effective teachers of students of color: having high expectations of students, using culturally relevant teaching, developing caring and trusting relationships with students, confronting issues of racism through teaching, and serving as advocates and cultural brokers. Other studies found that engaging preservice teachers in sustained reflection about diversity and social justice and providing field experiences in which they learn and build on cultural and linguistic strengths of students can prove helpful in limiting discrimination (Sleeter & Owuor, 2011).

Akiba (2011), analyzing pre and post surveys of 243 pre-service teachers enrolled in diversity courses and accompanying field experience, found three characteristics of teacher preparation programs that were related to positive changes in future teachers' beliefs about diversity in personal and professional contexts: a focus on the classroom as a learning community, instructors who modeled constructivist and culturally-responsive teaching, and field experiences with opportunities for learning about and understanding diverse students.

Finally, focusing on a new teacher education project in the highly diverse city of Auckland, New Zealand, Cochran-Smith et al. (2016) argued that there are six practices teacher education programs can adopt to put equity at the center. These practices include (a) selecting content that is worthwhile and then organizing opportunities to learn that are aligned with content and outcomes; (b) connecting content to students' lives and backgrounds; (c) facilitating supportive learning-centered environments; (d) supporting teaching and learning with evidence; (e) providing professional engagement opportunities and an inquiry stance; (f) and recognizing that various classroom, school, and societal challenges contribute to inequality.

These studies beg the question of whether equity-focused teacher preparation programs have similar goals across regions and if future teachers in these programs recognize and reach these goals.

Methods

To answer the research questions, I use descriptive statistics and tests of mean differences to analyze data from the International Association for the Evaluation of Educational Achievement's (IEA) Teacher Education and Development Study in Mathematics (TEDS-M). TEDS-M is a 17-country international comparative study of how primary and secondary future teachers are prepared to teach mathematics, the mathematical and pedagogical knowledge gained in their preparation programs, the programs in which they learn to teach, and the policies, practices, and contexts that influence their development (Tatto et al., 2009). TEDS-M includes nationally representative cross-sectional data from nearly 25,000 future teachers of mathematics (more than 15,000 future primary teachers and over 9,000 future secondary teachers) and nearly 5,000 teacher educators working in 500 teacher preparation institutions (with 451 preparing primary teachers and 339 preparing secondary teachers) (Tatto, 2013).

Sampling Procedures

TEDS-M used a stratified multi-stage probability sampling design for surveys of the two groups studied here: (a) future primary school teachers of mathematics in their last year of teacher preparation and (b) the institutions in which these teachers received their preparation (Dumais, Meinck, Tatto, Schwille, & Ingvarson, 2013). TEDS-M researchers first randomly selected teacher preparation institutions in each country

Table 5.1 Final sample sizes for preparation programs and future primary mathematics teachers (weighted)

Country/Region	Institutions	Future Primary Teachers
Chinese Taipei	8	1,619
Philippines	31	1,177
Thailand	30	753
United States	173	6,481
Total	242	10,030

and then randomly selected future teachers from a list of appropriate future teachers (e.g., primary level, last year of preparation program, etc.) within each institution. The target population of teacher preparation institutions was defined as the "set of secondary or post-secondary schools, colleges, or universities offering structured opportunities to learn (i.e., a program or programs) on a regular and frequent basis to future teachers of mathematics within a teacher preparation route" (Dumais et al., 2013, 84). The target population of future teachers included those future teachers in their last year of preparation who were "enrolled in an institution offering formal opportunities to learn to teach mathematics, and explicitly intended to prepare individuals qualified to teach mathematics in any of Grades 1–8" (Dumais et al., 2013, 84).

Sample

TEDS-M researchers set the minimum required sample sizes for each participating country at 50 teacher preparation institutions per level and 400 future teachers per level (Dumais & Meinck, 2013). To address the research questions for this study, I focus on the United States, Chinese Taipei, Philippines, and Thailand. To serve the purposes of this study, I narrowed the sample to the institutions that had an equity focus, identifying institutions that reported that selecting future primary teacher candidates from groups underrepresented in the teaching profession was *somewhat important* or *very important* (item MIC002G on the Institutional Program Questionnaire). Table 5.1 describes the final sample sizes for each country in this study.

Because of the complexity of the sampling design, estimation weights were developed for the surveys of teacher preparation institutions and future primary teachers of mathematics. All analyses in this study used the appropriate institutional- and teacher-level weights.

Measures

Institutional and future primary teacher data in TEDS-M were collected through the Institutional Program Questionnaire (IPQ) and the Future Teachers' Questionnaire (FTQ). These instruments were developed through a collaboration among the International Study Center at Michigan State University (MSU), the Australian

Council for Educational Research (ACER), and the national research coordinators (NRCs) in each country. The instruments were extensively tested in both a pilot study in 2005 and a field trial in 2006 and were designed to use parallel questioning where appropriate to facilitate comparison of constructs (Tatto, Rodriguez, Ingvarson, et al., 2013). While all TEDS-M instruments were developed in English, researchers in the 17 national centers followed rigorous and extensive translation and translation verification processes to ensure the highest quality translation and cultural adaptations, while maintaining international comparability (Malak-Minkiewicz & Berzina-Pitcher, 2013).

The Institutional Program Questionnaire is comprised of questions in nine categories: program description, future teacher background, selection policies, program content, field experience, program accountability and standards, staffing, program resources, and reflections on the program (Tatto, Rodriguez, Ingvarson, et al., 2013). Of particular interest to this study were the selection policies and the program goals. More specifically, this study is concerned with the consideration that institutions gave to various future teacher characteristics when selecting program candidates. In addition, the study examines the program goals regarding preparing future teachers with knowledge of pupils and diversity. Table 5.2 outlines the variables of interest and how they were measured. Because no program-level scales were developed, I analyze individual items.

The Future Teacher Questionnaire (FTQ) includes questions on future teachers' general background, program learning opportunities, and beliefs about teaching and learning mathematics. With respect to teachers' background, the FTQ gathers information on age, gender, socioeconomic status, diversity, educational attainment, academic achievement, reasons for becoming a teacher, and future in teaching. Program learning opportunities included future teachers' reporting on opportunities to learn a range of mathematics topics, mathematics education, education pedagogy, teaching for diversity, school and practicum experience, and coherence of teacher preparation program. Finally, the FTQ asked future teachers to report their beliefs about the nature of mathematics, learning mathematics, mathematics achievement, their own preparedness for teaching mathematics, and program effectiveness.

This study is concerned with future teachers' demographic characteristics and opportunities to learn how to teach diverse students. Table 5.3 outlines the FTQ items and measures for examined constructs.

Survey Administration

In each participating country/region, a national research coordinator (NRC) directed the implementation of TEDS-M. The NRCs were aided by instructions from a nine-unit series of administration manuals explaining procedures for (a) conducting the field test, (b) contacting institutions, (c) verifying instrument translation, (d) producing, assembling, and laying out the instrument, (e) listing and sampling within the institution, (f) administering the survey, (g) scoring constructed –response items, (h) analyzing syllabi, and (i) creating data files (Brese, Becker, Berzina-Pitcher, Tatto, & Carstens, 2013).

Table 5.2 Institutional program questionnaire: constructs and measures of interest

Content	Construct	Measure
Selection policies[a]	Educational attainment	Overall level of attainment (grades) in final year of secondary school
	Educational performance	Performance on national or state exam at end of final year of secondary school
		Performance on institutional exam for admission
	Suitability for teaching	Including personal qualities, experience, and motivation (assessed through interviews or application)
	Math achievement	Demonstrated math achievement
	Gender	Male/Female
	Under-represented groups	Belonging to groups under-represented in the teaching profession
	Order of application	Order in which candidates apply
	Location	Region of residence
	Age	In years
Program goals: knowledge of pupils and diversity[b]	Child development	Studying child development
	Behavioral or emotional problems	Specific strategies for teaching pupils with behavior or emotional problems
	Learning disabilities	Specific strategies and curriculum for teaching pupils with learning disabilities
	Gifted pupils	Specific strategies and curriculum for teaching gifted pupils
	Diverse backgrounds	Specific strategies and curriculum for teaching pupils from diverse cultural backgrounds
	Physical disabilities	Accommodating the needs of pupils with physical disabilities
	Disadvantaged pupils	Working with children from poor or disadvantaged backgrounds

[a]Response options: Not considered, Not very important, Somewhat important, or Very important
[b]Response options: Little or no weight, Some weight, Moderate weight, or Major weight

Analysis

To describe and compare the selection criteria of these equity-focused preparation programs, I tested mean differences between countries regarding the emphasis placed on the selection of future teachers from underrepresented populations and the emphasis placed on other selection criteria, including gender, age, region, overall achievement, demonstrated achievement, national exam scores, institutional exam scores, and interviews. To describe program goals, I generated descriptive statistics of the variables MID011A-AB. These 28 items (i.e., variables MID011A-AB) asked respondents to characterize the weight (e.g., little or no weight, some weight, moderate weight, or major weight) given to a series of program goals as identified in the teacher preparation program requirements,

Table 5.3 Future teacher questionnaire: constructs and measures of interest

Content	Construct	Measure
General background	Age	In years
	Gender	Male/female
	Socioeconomic status	Number of books in future teachers' home, ranging from "0-10" to "more than 200 books"
		Possession of specific items, including eight common items (e.g., computer)
		Highest educational attainment of mother
		Highest educational attainment of father
	Diversity	Frequency of speaking the language of the test at home, ranging from "always" to "never"
	Academic achievement	Highest year or grade of mathematics studied
		Most advanced mathematics course studied
		Grades received in secondary school
	Reason for becoming a teacher	Included nine possible choices rated by significance
	Future in teaching	Included five choices
Opportunities to learn[a]	Teaching diverse students	Develop strategies for teaching students with behavior and emotional problems
		Develop specific strategies and curriculum for teaching pupils with learning disabilities
		Develop specific strategies and curriculum for teaching gifted pupils
		Develop specific strategies and curriculum for teaching pupils from diverse backgrounds
		Accommodate the needs of pupils with physical disabilities
		Work with children from poor or disadvantaged backgrounds

[a]Response options: Never, Rarely, Occasionally, or Often

guidelines, or other documentation. I paid special attention to seven items (MID011P-V) of the 28 which asked specifically about "Knowledge of Pupils and Diversity." For example, one item asked respondents to characterize the weight that the program gave to teaching future teachers to develop specific strategies and curriculum for teaching students from diverse cultural backgrounds.

Next, to determine if these equity-focused programs are meeting their selection and program goals, I used descriptive statistics and tests of mean differences to analyze the reports of future primary mathematics teachers in these programs regarding their demographics and opportunities to learn about, and preparedness to teach, marginalized children. Demographic characteristics used included language, gender, age, socioeconomic status, parents' education, secondary school grades, reason for becoming a teacher, future in teaching, and program effectiveness. I also present descriptive statistics on future primary teachers' reports on the extent to which they had opportunities to learn about teaching students with behavioral or emotional problems, learning disabilities, physical disabilities, and those who are gifted, disadvantaged, or have diverse cultural backgrounds.

In all analyses, I used the International Association for the Evaluation of Educational Achievement's (IEA) IDB Analyzer, a program designed by the IEA Data Processing and Research Center that allows researchers to use SPSS software to analyze IEA datasets without writing computer code. Importantly, the IDB Analyzer automatically accounts for the nested structure of the TEDS-M dataset when computing statistics and standard errors. For program-level analyses, I used institutional weights and the TARGETP (i.e., target primary) variable to limit analysis to programs preparing *primary* teachers (see Table 5.1 for final institution and future teacher sample sizes). For teacher-level analyses, I used teacher weights and the TARGETP variable. The mean differences were tested for significance based on t-tests using the pooled standard errors estimated with the IDB Analyzer, and using an Excel macro developed by the IEA to compute the standard error of the difference (Gonzalez, 2010) using formulas commonly found in standard statistics textbooks.

Country/Region Backgrounds

Before describing the results of the analysis, it is important to provide brief background summaries of each country/region's population and economy, as well as statistics regarding their education systems and teachers working within those systems.

Chinese Taipei

From Table 5.4, it is clear that Chinese Taipei has a small and relatively homogeneous population that enjoys a high-income level and a fairly equitable distribution of that income (a Gini coefficient of 34.2, ranking 94th out of 141 countries with respect to income inequality; CIA World Factbook, 2014). In fact, very few people in Chinese Taipei live below the poverty line (1.5%). While educational statistics for Chinese Taipei are few, it is clear that people enjoy a fairly high level of literacy (96.1%).

In spite of its wealth and relative homogeneity, there is still a need for a more diverse teacher labor force. For example, Liu and Tsung (2007) focus on the aboriginal population and cite a number of challenges for providing equitable educational opportunities to marginalized children in Indigenous communities: low household income, few educational programs in mountainous regions, and barriers for teachers gaining access to professional development. The authors argue that teacher education programs should recruit more aborigine teachers, encourage all teacher candidates to work with aboriginal communities, and provide future teachers with the development opportunities necessary to effectively teach the students in these

Table 5.4 Country/Region background summaries, including population, economy, education, and teachers

	Chinese Taipei	Philippines	Thailand	United States
Population[a] size (in millions)	23.3	107.7	67.7	318.9
Ethnic groups	Taiwanese (including Hakka) 84%	Tagalog 28.1%	Thai 95.9%	White 72.4%
		Cebuano 13.1%	Burmese 2%	Black 12.6%
	Mainland Chinese 14%	Ilocano 9%	Other 1.3%	Asian 4.8%
	Indigenous 2%	Bisaya/Binisaya 7.6%	Unspecified 0.9%	Amerindian and Alaska native 0.9%
		Hiligaynon Ilonggo 7.5%		Native Hawaiian and other Pacific islander 0.2%
		Bikol 6%		Two or more races 2.9%
		Waray 3.4%		
		Other 25.3%		
% Below poverty line	1.5	26.5	13.2	15.1
% Urban	n/a	48.8	34.1	82.4
Economy[a, b]				
Gross domestic product (PPP)	$39,600	$4,700	$9,900	$52,800
Income level[2]	High	Lower middle	Upper middle	High
Gini coefficient	34.2	44.8	39.4	45.0
Education[c]				
Expenditures, % GDP	n/a	2.7	5.8	5.4
% Literate	96.1[1]	95.4	93.5	99.0
% Net primary enrollment, male	n/a	87.9	90.0	95.4
% Net primary enrollment, female	n/a	89.5	89.4	96.1
% Net secondary enrollment, male	n/a	56.4	69.9	88.8
% Net secondary enrollment, female	n/a	66.9	78.4	90.2
Teachers[d]				
% Female, primary	n/a	89.7	59.7	87.2
% Female, secondary	n/a	76.4	51.4	62.0

(continued)

Table 5.4 (continued)

	Chinese Taipei	Philippines	Thailand	United States
Pupil teacher ratio, primary	n/a	31.4	16.3	14.4
Pupil teacher ratio, secondary	n/a	34.8	19.9	14.7

[a]CIA World Factbook – https://www.cia.gov/library/publications/the-world-factbook/
[b]World Bank – http://data.worldbank.org/country/
[c]UNICEF – http://www.unicef.org/infobycountry/
[d]UNESCO Institute for Statistics (UIS) – http://data.uis.unesco.org/

communities. It appears that the Ministry of Education would agree with this assessment; in its Objectives for 2015, the ministry states that it seeks to "intensify the education and training of highly-skilled Indigenous people and increase the number of qualified Indigenous teachers" (Republic of China Ministry of Education, 2015, paragraph 8).

The Teacher Education Act of 1994 outlines the basic admissions criteria set by the Ministry of Education for teacher education programs in Chinese Taipei. Hsieh, Lin, Chao, and Wang (2013) note, for instance, that students must be in their second (or higher) year of university before they can apply to a teacher preparation program. However, the authors note that individual universities can set their own selection processes, which may include test scores of knowledge, language, and attitudes, or evaluations of character and extracurricular activities.

Philippines

In contrast to Chinese Taipei, the Philippines has a large, mostly rural population that is highly diverse in terms of race/ethnicity. Filipinos have the lowest income among the four countries in this study, and a relatively unequal distribution of that income (Gini coefficient of 44.8, ranking 42nd out of 141 countries; CIA World Factbook, 2014). Indeed, more than one-fourth of Filipinos live below the poverty line. Although Filipinos place a high value on educational attainment, the government spends just 2.7% of Gross Domestic Product (GDP) on education. Literacy rates are high, but net primary and secondary enrollments are the lowest in the sample. Nearly 90% of primary teachers and more than 75% of secondary teachers are female, and class sizes (based on headcount) are the largest in the countries for which data is available.

Scant literature exists on teacher diversity in the Philippines, but existing reports indicate large disparities in the distribution of educational opportunities for students within and across regions of the country (Mesa, 2007). The clear diversity of the country's population also suggests that a more diverse teacher labor force may benefit marginalized students across the country.

In recent years, teacher education students in the Philippines have come from economically disadvantaged families. Ogena, Brawner, and Ibe (2013) note that this phenomenon is due to the historically low entrance standards and low tuition that have made teaching a popular field of choice among many Filipinos. Efforts to screen candidates through examinations and increase standards have largely failed, and requirements to enter education programs remain lower than the requirements to enter most other university programs (Ogena et al., 2013).

Thailand

Thailand has a smaller and more homogeneous population than the Philippines. As of 2008, roughly 142,000 Burmese refugees lived in nine camps along the Thailand-Burma border (Oh & van der Stouwe, 2008), but most of the country's population are Thai. Like the Philippines, Thailand is largely a rural society with low levels of income with average distribution compared to other countries (Gini coefficient of 39.4, ranking 64th out of 141 countries; CIA World Factbook, 2014); nevertheless, more than 13% of the population lives below the poverty line. Of the countries in this sample, Thailand reports the largest proportion of Gross Domestic Product (GDP) spending on education, and net primary and secondary enrollment rates are higher than in the Philippines. However, literacy rates in Thailand are the lowest of the four countries sampled. Just over half of primary and secondary teachers are female, and class sizes (based on headcounts) appear to be quite small.

Because of the relatively homogeneous population of Thailand, policymakers are more concerned with the changing nature of the economy from an agriculture-based one to a service-oriented one; clearly, this change requires changes to the education of the country's workforce (Dechsri & Pativisan, 2013). Dechsri and Pativisian note that to achieve this goal, primary education in Thailand focuses on giving children the knowledge and skills to negotiate social change and live a peaceful life, while secondary education seeks to help students explore their abilities, aptitudes, and interests and prepares students for new jobs and further study.

Preparing students for Thailand's social and economic future presents teachers with numerous challenges, including transitioning to child-centered instruction, helping students develop critical thinking skills, and increasing time spent on math and science curricula (Dechsri & Pativisan, 2013). Thailand's increasingly market-based economy is also associated with growing income inequality.

United States

The United States has the largest population in the sample, and it is moderately diverse in terms of race/ethnicity compared to the other three countries. Most Americans live in urban areas and enjoy the highest level of income of the countries in this study. However, compared to the other countries, income in the United States is the most unequally distributed (Gini coefficient of 45.0, ranking 41st out of 141 countries; CIA World Factbook, 2014), and more than 15% of the population lives below the poverty line. Literacy rates and net primary and secondary enrollment rates are the highest in this sample, and large proportions of teachers at all levels are female. Finally, class sizes in the United States (based on headcounts) are the lowest in the sample.

Policymakers in the United States have expressed a need for a more diverse teacher labor force. For example, Education Secretary Arne Duncan developed an initiative in 2010 to recruit and develop teachers from marginalized populations (Bireda & Chait, 2013). Furthermore, American scholars have conducted extensive research on the diversification of the teacher labor force (e.g., Achinstein et al., 2010).

Results

The first research question of this study seeks to investigate the selection criteria and program goals of teacher preparation programs that emphasize the selection of future primary mathematics teachers from marginalized populations. Table 5.5 presents descriptive statistics for the importance that equity-focused teacher preparation programs placed on each source of information used to select future primary teachers.

Table 5.5 Importance of information used by equity-focused programs to select future primary teacher

	Chinese Taipei		Philippines		Thailand		United States	
	M	*SE*(*M*)	*M*	*SE*(*M*)	*M*	*SE*(*M*)	*M*	*SE*(*M*)
Gender	1.60	(0.27)	2.16	(0.49)	1.45^{TU}	(0.12)	1.07	(0.04)
Age	1.27^{CP}	(0.32)	2.28^{PU}	(0.20)	2.00^{TU}	(0.18)	1.00	(0.00)
Region	1.13^{CP}	(0.16)	$2.16^{PT,PU}$	(0.18)	1.55	(0.14)	1.37	(0.17)
Overall achievement	$1.40^{CP,CT}$	(0.49)	3.70^{PU}	(0.20)	3.14^{TU}	(0.12)	2.36	(0.24)
Demonstrated achievement	$1.40^{CP,CT,CU}$	(0.35)	$3.85^{PT,PU}$	(0.13)	2.97	(0.16)	2.70	(0.14)
National exam	1.40^{CP}	(0.49)	$3.79^{PT,PU}$	(0.17)	2.67	(0.20)	2.45	(0.23)
Institutional exam	$1.40^{CP,CT}$	(0.49)	3.72	(0.20)	3.19	(0.17)	2.81	(0.41)
Interviews	$4.00^{CT,CU}$	(0.00)	3.93^{PT}	(0.09)	3.48	(0.09)	3.49	(0.17)

Note. Response options for each selection criteria: 1 = not considered, 2 = not very important, 3 = somewhat important, and 4 = very important. Statistically significant differences ($p < .01$) between countries/regions identified by superscript, where CP indicates a significant difference between Chinese Taipei and Philippines, CT indicates a significant difference between Chinese Taipei and Thailand, CU indicates a significant difference between Chinese Taipei and the United States, PT indicates a significant difference between Philippines and Thailand, PU indicates a significant difference between the Philippines and United States, and TU indicates a significant difference between Thailand and the United States

In Chinese Taipei, equity-focused teacher preparation programs placed relatively little importance on any of these sources of information except for the interviews. Here, interviews refer to information gathered about the candidate's suitability for teaching and may include items such as the person's experience and personal qualities. Notwithstanding the low emphasis placed on other sources of information, Chinese Taipei showed greater variation than other countries in the importance placed on candidates' scores on exams and their overall achievement. In the Philippines, programs placed the greatest importance on interviews, but also considered achievement and exam scores to be quite important in the selection process. Programs in Thailand also placed the greatest importance on interviews, with institutional exams and overall achievement second. These programs appeared to be less interested in candidate scores on the national exam than the institutional exam. Programs in the United States also focused on the interview as the most important source of information, with achievement and exam scores trailing. The United States is the only country in which age and gender were not considered.

Looking across the four countries, the general pattern is that interviews of candidates were, by far, the most important source of information for choosing future primary teacher candidates in equity-focused programs. Compared with other countries, interviews were especially important in Chinese Taipei, where results indicate statistically significant differences from Thailand and the United States. Gender, age, and region appear to matter little to these programs, except in the Philippines where they were significantly more important than the other countries/regions. A focus on gender, age, and region in the Philippines could indicate that teacher preparation programs are trying to recruit more male future teachers (89.7% of primary teachers are female) and future teachers from regions populated by minority groups. From these data, the rationale for selecting future teachers based on age is unclear. There was considerable variation across countries/regions in terms of the importance placed on achievement and examinations. Perhaps most notable was the little importance placed on these sources of information in Chinese Taipei.

Providing some insight into the information that is most important to equity-focused programs when selecting primary future teacher candidates, Table 5.6 highlights the weight that programs gave to seven aspects of knowledge of pupils and diversity: child development, children with behavioral or emotional challenges, children with learning disabilities, gifted pupils, pupils from different cultural backgrounds, children with physical disabilities, and children from disadvantaged backgrounds.

In Chinese Taipei, equity-focused programs gave the most weight to ensuring that future teachers study child development in general and moderate weight to teaching children with behavioral and emotional problems and those from disadvantaged backgrounds. Programs in the Philippines also gave the most weight to child development, but appear to give considerable weight to nearly all aspects of preparing teachers to teach diverse students. There also appears to be minimal variation across programs in terms of the weight given to these aspects of knowledge of diverse students in the Philippines. Programs in Thailand gave the most weight to child development, with behavioral and emotional challenges and learning disabilities

Table 5.6 Program weight given to each item within program goal: knowledge of pupils and diversity

	Chinese Taipei		Philippines		Thailand		United States	
	M	*SE*(*M*)	*M*	*SE*(*M*)	*M*	*SE*(*M*)	*M*	*SE*(*M*)
Child development	3.47	(0.40)	3.59	(0.25)	3.53^{TU}	(0.10)	2.96	(0.17)
Behavioral/emotional	2.87	(0.23)	3.39	(0.28)	3.18	(0.14)	2.75	(0.16)
Learning disabilities	2.27	(0.38)	3.39	(0.28)	2.82	(0.16)	3.01	(0.24)
Gifted pupils	2.13	(0.37)	2.94^{PU}	(0.07)	2.68	(0.16)	2.38	(0.12)
Diff. cultural backgrounds	2.27	(0.32)	2.94	(0.07)	2.68	(0.18)	2.92	(0.18)
Physical disabilities	2.13	(0.37)	2.94^{PT}	(0.07)	2.46	(0.16)	2.57	(0.15)
Disadvantaged backgrounds	2.60	(0.42)	3.38^{PT}	(0.25)	2.55	(0.16)	2.78	(0.14)

Note. Response options for each item: 1 = little or no weight, 2 = some weight, 3 = moderate weight, and 4 = major weight. Statistically significant differences ($p < .01$) between countries identified by superscript, where CP indicates a significant difference between Chinese Taipei and Philippines, CT indicates a significant difference between Chinese Taipei and Thailand, CU indicates a significant difference between Chinese Taipei and the United States, PT indicates a significant difference between Philippines and Thailand, PU indicates a significant difference between the Philippines and United States, and TU indicates a significant difference between Thailand and the United States

following. In the United States, child development, learning disabilities, and different cultural backgrounds were quite similar in terms of their weight within programs. Learning to teach gifted pupils received the least weight in the United States.

Across all four countries, programs gave the least weight to teaching students with physical disabilities and gifted pupils. However, among countries/regions there were few statistically significant differences in program weight given to each item. Compared to the United States, equity-focused programs in Thailand placed greater emphasis on child development and programs in the Philippines concentrated more on gifted pupils. Programs in the Philippines also differed from those in Thailand in that they placed more weight on preparing future teachers to teach students with physical disabilities and disadvantaged backgrounds.

Demographics of Future Primary Mathematics Teachers in Equity-Focused Preparation Programs

The second research question in this study investigates whether equity-focused teacher preparation programs met their selection and program goals regarding future primary mathematics teachers' opportunities to learn about, and preparedness to teach, marginalized children.

Table 5.7 presents demographic data of the future primary teachers in equity-focused programs in the four countries in this study. TEDS-M did not gather information on future teacher ethnicity, but did ask respondents to indicate the frequency with which they speak the language of the test at home. Since the instruments were

Table 5.7 Demographic characteristics of future primary mathematics teachers in equity-focused preparation programs

	Chinese Taipei		Philippines		Thailand		United States	
	M	*SE*(*M*)	*M*	*SE*(*M*)	*M*	*SE*(*M*)	*M*	*SE*(*M*)
Speak test language at home[a]	2.13[CP,CT,CU]	(0.03)	2.94[PT,PU]	(0.02)	1.95[TU]	(0.04)	1.06	(0.02)
Gender[b]	.26[CP,CU]	(.02)	.14[PT]	(.01)	.27[TU]	(.01)	.10	(.02)
Age in years	23.29[CP,CT]	(0.20)	20.69[PT,PU]	(0.17)	22.30[TU]	(0.03)	24.47	(0.67)
Number of books at home[c]	3.54[CP,CT,CU]	(0.07)	2.05[PT,PU]	(0.04)	2.74[TU]	(0.06)	4.02	(0.07)

Note. Statistically significant differences ($p < .01$) between countries/regions identified by superscript, where CP indicates a significant difference between Chinese Taipei and Philippines, CT indicates a significant difference between Chinese Taipei and Thailand, CU indicates a significant difference between Chinese Taipei and the United States, PT indicates a significant difference between Philippines and Thailand, PU indicates a significant difference between the Philippines and United States, and TU indicates a significant difference between Thailand and the United States

[a]Response options: 1 = Always, 2 = Almost Always, 3 = Sometimes, and 4 = Never

[b]0 = Female and 1 = Male, so mean can be interpreted as proportion male

[c]Response options: 1 = none or few, 2 = one shelf, 3 = one bookcase, 4 = two bookcases, 5 = three or more bookcases

translated into the official or primary language of each country/region, one can assume that future teachers who speak a different language are more likely to be members of a marginalized population. Other variables indicating diversity are gender, age, and number of books at home (an indicator of socioeconomic status).

In Chinese Taipei, future primary mathematics teachers in equity-focused preparation programs almost always spoke the language of the test at home, although respondents' responses indicate that at least some future teachers sometimes spoke a different language at home. Twenty-six percent of future primary teachers in Chinese Taipei were male, the average age was 23 years, and they had on average at least one bookcase at home. More future primary teachers in the Philippines spoke a language different from the test language. Fourteen percent of these future teachers were male, average age was 21 years, and they had an average of one shelf of books at home. Future teachers in Thailand tended to speak the language of the test most of the time. Twenty-seven percent of Thai primary future teachers were male, averaged 22 years, and had an average of one shelf of books. Finally, future teachers in the United States overwhelmingly reported always speaking the language of the test at home. Just 10 % were male, the average age was 24 years, and they averaged two bookcases at home.

Across the four countries/regions, most future primary teachers spoke the language of the test at home. However, future teachers in equity-focused preparation programs in the United States were the least diverse in terms of language; nearly all reported speaking the language of the test at all times. Future teachers in the Philippines were the most diverse in terms of language. In Chinese Taipei and Thailand, more than one-fourth of teachers were male, proportions that are significantly different from 14% in the Philippines and 10% in the United States. Future teachers across the four countries were in their early twenties on average; yet future teachers in the United States and Chinese Taipei were significantly older than future teachers in other countries. Finally, in terms of socioeconomic status (as measured by number of books at home), there were significant differences between future teachers in Chinese Taipei and the United States, who reported having significantly more books at home (at least one bookcase full) than future teachers in the Philippines and Thailand.

Opportunities to Learn about Pupil Diversity as Reported by Future Primary Mathematics Teachers in Equity-Focused Preparation Programs

Addressing programs' ability to ensure that all teachers learn strategies to effectively teach diverse students, the second research question investigates future primary teachers' reports of their opportunities to learn content related to program goals for teaching diverse students. Here, aspects of student diversity considered include opportunities to learn to teach children with behavioral or emotional behavioral challenges, students with disabilities, gifted pupils, those from different cultural backgrounds, those who have physical disabilities, and those who come from disadvantaged backgrounds. Table 5.8 presents this data from future teachers in

Table 5.8 Opportunities to learn to teach for diversity reported by future primary mathematics teachers in equity-focused preparation programs

	Chinese Taipei		Philippines		Thailand		United States	
	M	*SE*(*M*)	*M*	*SE*(*M*)	*M*	*SE*(*M*)	*M*	*SE*(*M*)
Composite[a]	9.52[CP,CT,CU]	(0.09)	11.85[PT]	(0.10)	10.11[TU]	(0.10)	11.35	(0.22)
Behavioral/emotional[b]	2.60[CP,CU]	(0.05)	3.39[PT,PU]	(0.02)	2.67[TU]	(0.05)	3.07	(0.10)
Learning disabilities[b]	2.45[CP,CU]	(0.05)	3.17[PT]	(0.07)	2.61[TU]	(0.05)	3.15	(0.10)
Gifted pupils[b]	1.81[CP,CT,CU]	(0.05)	3.01[PT]	(0.05)	2.45[TU]	(0.05)	2.77	(0.11)
Diff. cultural backgrounds[b]	2.33[CP,CU]	(0.04)	3.03[PT]	(0.06)	2.27[TU]	(0.05)	3.14	(0.07)
Physical disabilities[b]	1.93[CP,CT,CU]	(0.06)	3.30[PT,PU]	(0.05)	2.35[TU]	(0.06)	2.73	(0.10)
Disadvantaged backgrounds[b]	2.71[CP,CT,CU]	(0.06)	3.42[PT]	(0.04)	2.90[TU]	(0.04)	3.27	(0.10)

Note. Statistically significant differences ($p < .01$) between countries/regions identified by superscript, where CP indicates a significant difference between Chinese Taipei and Philippines, CT indicates a significant difference between Chinese Taipei and Thailand, CU indicates a significant difference between Chinese Taipei and the United States, PT indicates a significant difference between Philippines and Thailand, PU indicates a significant difference between the Philippines and United States, and TU indicates a significant difference between Thailand and the United States

[a]Rasch score scale where 10 is located at the neutral position

[b]Response options: 1 = Never, 2 = Rarely, 3 = Occasionally, 4 = Often

equity-focused programs in the United States, Chinese Taipei, Philippines, and Thailand. This data includes teacher scores on a composite Rasch model scale (the row labeled "Composite" in Table 5.8) of these aspects of diversity, where 10 is located in the neutral position (Tatto, Rodriguez, Reckase, Rowley, & Lu, 2013).

In Chinese Taipei, the composite score for future teachers' opportunities to learn to teach diverse students fell near the neutral location of 10. Future teachers in Chinese Taipei reported that they had the most opportunities to learn to teach students from disadvantaged backgrounds, those who have behavioral or emotional problems, and those with learning disabilities. Yet, on average, they experienced these opportunities only occasionally. Future teachers in Chinese Taipei enjoyed fewer opportunities to learn to teach gifted pupils and student with physical disabilities. In the Philippines, the composite score for learning to teach diverse students was nearly 12, indicating that future primary teachers in equity-focused programs had greater than average opportunities learn to teach diverse students. Filipino future teachers reported opportunities to learn about each type of diverse student at least occasionally, with the greatest frequencies for students from disadvantaged backgrounds, those who have behavioral or emotional problems, and those with physical disabilities. Thailand's future teachers' composite score was very close to neutral. Thai future teachers reported the most opportunities to learn to teach students from disadvantaged backgrounds, those who have behavioral or emotional problems, and those with learning disabilities. Future teachers in Thailand had fewer opportunities to learn to teach students with different cultural backgrounds. In the United States, future primary teachers' composite score of 11.35 indicated greater than average opportunities to learn to teach diverse students overall. American future teachers in equity-focused programs had the most frequent opportunities to teach students from disadvantaged backgrounds, those with learning disabilities, and those with different cultural backgrounds. They had less frequent opportunities to teach students with physical disabilities.

Across the United States, Chinese Taipei, Philippines, and Thailand, future teachers reported the most opportunities to learn to teach students from disadvantaged backgrounds, those who have behavioral or emotional problems, and those with learning disabilities. Teaching gifted students and students with physical disabilities were generally addressed less often in equity-focused programs. Future teachers in Chinese Taipei and Thailand reported the fewest opportunities to learn to teach diverse students. Looking at the composite Rasch score and individual items it is clear that future teachers in equity-focused programs in Chinese Taipei had the fewest opportunities to learn to teach diverse students.

Discussion

The purpose of this study was to first describe the selection criteria and program goals of teacher preparation programs that emphasize the selection of future primary mathematics teachers from marginalized populations in the United States, Chinese

Taipei, Philippines, and Thailand. Furthermore, I sought to determine whether these programs met their selection and program goals in providing future teachers opportunities to learn about teaching diverse students. As noted earlier in the chapter, given the different regions and countries included in this study, it is important to note that the term "diversity" can be understood differently among and within cultures. The questionnaire items analyzed in this study include several indicators of the diversity of future teachers (language, age, gender, and socioeconomic status) and the students they are prepared to teach (behavioral and emotional challenges, learning disabilities, physical disabilities, giftedness, cultural background, and socioeconomic status). Yet these indicators clearly do not cover all forms of diversity; others might include factors like parent education levels, distance to school, religion, and refugee status. Therefore, in the following sections I strive to pay careful attention to individual contexts when interpreting results across countries/regions.

Chinese Taipei

Chinese Taipei is a developed society that does not have a very diverse population. It seems reasonable, then, that the demographic profile of future primary mathematics teachers indicates that on average candidates have relatively high socioeconomic status and almost always speak the language of the test at home. Yet the presence of an Indigenous population warrants attention to the educational opportunities that Indigenous students have to succeed. In fact, the Ministry of Education has declared its intent to intensify the education of Indigenous peoples and increase the number of qualified Indigenous teachers (Republic of China Ministry of Education, 2015). In addition, the government is moving to use the education system to "revitalize" the language and traditions of the Hakka, a "sub-group" of Han Chinese that comprises 20% of the population (Taiwan Today, 2015). The children of increasing numbers of migrants and foreign spouses, though still relatively few, also challenge future and in-service teachers to learn to teach a diverse student population. As a result, more teacher education programs are including multicultural education coursework, and the government encourages in-service teachers to also update their skills through such courses (Liu & Lin, 2011).

Given the efforts of policymakers and education scholars identified above, it seems appropriate that teacher preparation programs would attempt to teach all future teacher candidates how to teach diverse students. Even if proportions of diverse populations remain relatively small, educational research from the United States suggests that exposure to teachers from marginalized populations and multicultural education practices supports the learning and social development of *all* students (Achinstein & Ogawa, 2011). The goals of equity-focused teacher preparation programs in Chinese Taipei appear to align with this research. As in the Philippines and Thailand, programs in Chinese Taipei gave the most weight to the goal of overall child development for diverse students, as seen in Table 5.6. Additional goals included teaching students with behavioral or emotional problems, those from disadvantaged backgrounds, and those with different cultural backgrounds.

When comparing future teachers' reported opportunities to learn how to teach diverse students in Chinese Taipei, there was fairly close alignment with the program goals reported in Table 5.6. Future teachers reported the most opportunities to learn how to teach disadvantaged students, those with behavioral or emotional problems, those with learning disabilities, and those with different cultural backgrounds.

Philippines

The Philippines is the poorest and most diverse of the countries/regions studied. Future teachers in the Philippines also had the lowest socioeconomic status of the countries/regions studied and spoke a language other than that used in the test more often. Given the diversity of the population and disparities in educational opportunities across the country (Mesa, 2007), one would expect that equity-focused teacher preparation programs would seek to help future teachers learn to teach all types of students.

The results indicate that this was, indeed, the case in the Philippines at the time of the TEDS-M data collection. The top goals of the sampled teacher preparation programs were to prepare future teachers to teach students with behavioral or emotional problems, learning disabilities, and students from disadvantaged backgrounds. Future teachers in these programs reported fairly frequent opportunities to learn to teach each type of student, with the most frequent being students from disadvantaged backgrounds, those who have behavioral or emotional problems, and students with physical disabilities.

While these results indicate a bit of a mismatch between program goals and teachers' reported opportunities to learn (e.g. future teachers reported opportunities to learn about disadvantaged students most frequently, when this appears to be somewhat lower in the ranking of program goals), this may be mitigated by the fact that future teachers reported relatively high frequencies of opportunities to learn about all types of students.

Thailand

Like the Philippines, Thailand is a relatively poor country, but with a more homogeneous population. Future teachers in equity-focused preparation programs reported almost always speaking the language of the test at home and also reported having relatively few books in their homes. Literature from Thailand (e.g. Dechsri & Pativisan, 2013) suggests that learning to teach students from different cultures may not be high on the list of goals for teacher preparation programs. This may be because teachers need to be prepared to help their students gain skills that will support them in a society transitioning to a market orientation.

In some ways, a lack of emphasis on learning to teach students from other cultures seems evident in the goals of the programs sampled in Thailand. For example, equity-focused programs in Thailand placed the most weight on teaching students with behavioral or emotional problems, those with learning disabilities, and those from different cultures (as well as gifted pupils). However, future teachers reported the most frequent opportunities to learn to teach disadvantaged students, those with behavioral or emotional problems, and those with learning disabilities. Interestingly, although it appears to be given some weight within program goals, future teachers reported the fewest opportunities to learn about students with different cultural backgrounds. In fact, overall, the average of future teachers' reports on opportunities to learn about diverse students was closest to *rarely*.

United States

Although the United States has a fairly diverse population, its future teachers were the least diverse (in terms of language) of the studied countries/regions and had the highest socioeconomic status. It is important to note here that in the United States, alternative teacher programs such as Teach for America tend to attract and train more diverse teachers than traditional preparation programs (Boser, 2011, 2014). This study did not capture such programs, but rather traditional teacher preparation programs that emphasize selecting future teachers from marginalized populations.

In terms of the goals of equity-focused preparation programs, in the United States programs gave the most weight to learning to teach children with learning disabilities, overall child development, students with different cultural backgrounds, and those from disadvantaged backgrounds. Yet, as in the Philippines, future teachers in these programs reported the most frequent opportunities to learn to teach students from disadvantaged backgrounds, even though this was given less weight in the program goals than other goals. Following disadvantaged backgrounds, future teachers reported learning to teach students with learning disabilities, those from different cultural backgrounds, and those with behavioral or emotional problems.

Across Countries/Regions

When looking at the results across the countries/regions studied, it appears that, for the most part, the goals of equity-focused programs and future teachers' reported opportunities to learn to teach diverse students were similar. Furthermore, these goals appeared to align with the demographic and economic conditions in each country. There were also considerable differences across countries/regions in terms of the frequency of future teachers' opportunities to learn. For example, as noted above, future teachers in Thailand and Chinese Taipei rarely reported having relatively limited opportunities to learn about teaching diverse students.

These findings in Thailand and Chinese Taipei may be due to the prevalence of the concept of *horizontal equity* or the equal treatment of equals in these countries (Brown, 2006). Many countries and states strive to provide all children with equal access to schooling with similar levels of funding, teachers with similar credentials, and so on. An assumption often underlying these efforts is that if students have equal educational opportunities, then their educational outcomes can be attributed to their own effort or merit. This assumption would explain the fact that preparation programs in Chinese Taipei place major weight on the goal of overall child development and far less weight on aspects of student diversity.

Despite the prevailing concept of horizontal equity, gaps in the educational outcomes of various student populations persist. Brown (2006) argued that in societies in which populations have been systematically disadvantaged, the equal treatment of equals is unlikely to result in educational equity. Instead, *vertical equity*, or the "unequal, but equitable, treatment of unequals" may hold more promise. This approach to equity acknowledges that different groups of students have different starting locations, which necessitates different kinds of teaching or support to reach similar educational outcomes.

Conclusion

Education research indicates that teachers from marginalized populations can improve the educational experiences and outcomes of marginalized students. Given the lack of diversity in the teacher labor force in many countries, scholars argue that there are demographic and democratic imperatives for increasing the number of teachers from marginalized populations. Many teacher preparation programs across the United States strive to do just that. Yet little is known about similar efforts of programs in other countries. This study analyzed TEDS-M data on programs preparing future primary mathematics teachers and emphasizing the selection of candidates from underrepresented populations in the United States, Chinese Taipei, Philippines, and Thailand. I found some evidence of diversity among future teachers in equity-focused preparation programs and general alignment between program goals for teaching about diverse students and future teachers' reported opportunities to learn to teach diverse students.

Future research should investigate whether or not the selection criteria, program goals, and future teachers' opportunities to learn to teach diverse students in equity-focused preparation programs are significantly different from traditional programs in the same countries. Future research should also determine differences in the quality of future teachers prepared through equity-focused programs and traditional programs in terms of their mathematical knowledge, mathematical pedagogical knowledge, and beliefs about mathematics.

References

Achinstein, B., & Ogawa, R. T. (2011). *Change(d) agents: New teachers of color in urban schools*. New York, NY: Teachers College Press.

Achinstein, B., Ogawa, R. T., Sexton, D., & Freitas, C. (2010). Retaining teachers of color: A pressing problem and a potential strategy for "hard-to-staff" schools. *Review of Educational Research, 80*(1), 71–107.

Akiba, M. (2011). Identifying program characteristics for preparing pre-service teachers for diversity. *Teachers College Record, 113*(3), 658–697.

Bireda, S., & Chait, R. (2013). *Increasing teacher diversity: Strategies to improve the teacher workforce*. Washington, DC: Center for American Progress.

Boser, U. (2011). *Teacher diversity matters: A state-by-state analysis of teachers of color*. Washington, DC: Center for American Progress.

Boser, U. (2014). *Teacher diversity revisited: A new state-by-state analysis*. Washington, DC: Center for American Progress.

Brese, F., Becker, A., Berzina-Pitcher, I., Tatto, M. T., & Carstens, R. (2013). Survey operations procedures. In M. T. Tatto (Ed.), *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report* (pp. 93–101). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Brown, K. (2006). "New" educational injustices in the "new" South Africa: A call for justice in the form of vertical equity. *Journal of Educational Administration, 44*(5), 509–519.

CIA World Factbook. (2014). Retrieved from https://www.cia.gov/library/publications/the-world-factbook

Clark, J. V. (Ed.). (2014). *Closing the achievement gap from an international perspective*. London, UK: Springer.

Cochran-Smith, M., Ell, F., Grudnoff, L., Haigh, M., Hill, M., & Ludlow, L. (2016). Initial teacher education: What does it take to put equity at the center? *Teaching and Teacher Education, 57*, 67–78.

Cochran-Smith, M., & Villegas, A. M. (2015). Framing teacher preparation research: An overview of the field, part 1. *Journal of Teacher Education, 66*(1), 7–20.

Cochran-Smith, M., Villegas, A. M., Abrams, L., Chavez-Moreno, L., Mills, T., & Stern, R. (2015). Framing teacher preparation research: An overview of the field, Part 2. *Journal of Teacher Education, 66*(2), 109–121.

Commonwealth of Australia. (2003). *Australia's teachers: Australia's future. Advancing innovation, science, technology, and mathematics*. Canberra, Australia: Commonwealth of Australia.

Coughlin, S. (2015). Asia tops biggest global school rankings. *BBC News*. Retrieved from http://www.bbc.com/news/business-32608772

Council for the Accreditation of Educator Preparation. (2015). *CAEP accreditation standards*. Retrieved from http://edsource.org/wpcontent/uploads/commrpt.pdf

Cruickshank, K. (2004). Towards diversity in teacher education: Teacher preparation of immigrant teachers. *European Journal of Teacher Education, 27*(2), 125–138.

Darling-Hammond, L. (2000). Forward. In L. Darling-Hammond (Ed.), *Studies of excellence in teacher education: Preparation in the undergraduate years* (pp. v–xi). Washington, DC: American Association of Colleges for Teacher Education.

Darling-Hammond, L. (2010). Teacher education and the American future. *Journal of Teacher Education, 61*(1–2), 35–47.

Dechsri, P., & Pativisan, S. (2013). Thailand. In J. Schwille, L. Ingvarson, & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in 17 countries* (pp. 243–255). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *The Review of Economics and Statistics, 86*(1), 195–210.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review, 95*(2), 158–165.

Dumais, J., & Meinck, S. (2013). Estimation of weights, participation rates, and sampling error. In M. T. Tatto (Ed.), *The Teacher Education and Development Study in Mathematics: Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report* (pp. 129–160). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Dumais, J., Meinck, S., Tatto, M. T., Schwille, J., & Ingvarson, L. (2013). Sampling design. In M. T. Tatto (Ed.), *The teacher education and development study in mathematics: Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report* (pp. 79–92). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review, 45*, 44–52.

Faez, F. (2012). Diverse teachers for diverse students: Internationally educated and Canadian-born teachers' preparedness to teach English language learners. *Canadian Journal of Education, 35*(3), 64–84.

Flores, B. B., Clark, E. R., Claeys, L., & Villarreal, A. (2007). Academy for teacher excellence: Recruiting, preparing, and retaining Latino teachers through learning communities. *Teacher Education Quarterly, 34*(4), 53–69.

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education, 53*, 106–116.

Gay, G., & Kirkland, K. (2003). Developing cultural critical consciousness and self-reflection in preservice teacher education. *Theory into Practice, 42*(3), 181–187.

Gonzalez, E. (2010). *Calculating standard errors in TIMSS, PIRLS, ICCS, TEDS-M and PISA* (unpublished training material). Hamburg, Germany: IEA Data Processing and Research Center, International Association for the Evaluation of Educational Achievement (IEA).

Howard, J. (2010). The value of ethnic diversity in the teaching profession: A New Zealand case study. *International Journal of Education, 2*(1), 1–22.

Hsieh, F. J., Lin, P. J., Chao, G., & Wang, T. Y. (2013). Chinese Taipei. In J. Schwille, L. Ingvarson, & J. Holdgreve-Resendez (Eds.), *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in 17 countries* (pp. 69–87). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Ingersoll, R., Merrill, L., & Stuckey, D. (2014). *Seven trends: The transformation of the teaching force, updated April 2014* (Report No. 80). Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania.

Ingersoll, R. M., & May, H. (2011). *Recruitment, retention and the minority teacher shortage* (Report No. 69). Philadelphia, PA: The Consortium for Policy Research in Education.

Keane, E., & Heinz, M. (2016). Excavating an injustice? Nationality/ies, ethnicity/ies and experiences with diversity of initial teacher education applicants and entrants in Ireland in 2014. *European Journal of Teacher Education, 39*(4), 507–527.

Klopfenstein, K. (2005). Beyond test scores: The impact of black teacher role models on rigorous math taking. *Contemporary Economic Policy, 23*(3), 416–428.

Ladson-Billings, G. (1994). *The dreamkeepers: Successful teachers of African American children*. San Francisco, CA: Jossey-Bass.

Ladson-Billings, G. (2000). Fighting for our lives. *Journal of Teacher Education, 51*, 206–214.

Liu, K., & Tsung, W. K. L. (2007). Cultivating aboriginal cultures and educating aboriginal children in Taiwan. *Childhood Education, 83*(5), 282–287.

Liu, M., & Lin, T. B. (2011). The development of multicultural education in Taiwan: Overview and reflection. In C. A. Grant & A. Portera (Eds.), *Intercultural and multicultural education: Enhancing global connectedness* (pp. 157–176). New York, NY: Routledge.

Malak-Minkiewicz, B., & Berzina-Pitcher, I. (2013). Translation and translation verification of the TEDS-M research instruments. In M. T. Tatto (Ed.), *The Teacher Education and Development*

Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Mesa, E. P. (2007). *Measuring educational inequality in the Philippines*. Discussion paper, School of Economics, University of the Philippines. Retrieved from: http://www.econstor.eu/bitstream/10419/46638/1/539787426.pdf

Miller, P. C., & Endo, H. (2005). Journey to becoming a teacher: The experiences of students of color. *Multicultural Education, 13*(1), 2–9.

Milner, H. R. (2006). The promise of Black teachers' success with Black students. *Educational Foundations, 20*(34), 89–104.

National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future*. New York, NY: Author.

National Council on Teacher Quality. (2015). *2014 teacher prep review: A review of the nation's teacher preparation programs*. Washington, DC: Author.

Ogena, E. B., Brawner, F. G., & Ibe, M. D. (2013). Philippines. In J. Schwille, L. Ingvarson, & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in 17 countries* (pp. 159–175). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Oh, S. A., & van der Stouwe, M. (2008). Education, diversity, and inclusion in Burmese refugee camps in Thailand. *Comparative Education Review, 52*(4), 589–617.

Organization for Economic Cooperation and Development. (2010). *Educating teachers for diversity: Meeting the challenge*. Paris, France: Author.

Republic of China Ministry of Education. (2015). *Ministry of Education objectives for 2015 (January–December)*. Retrieved from http://english.moe.gov.tw/ct.asp?xItem=16072&ctNode=11410&mp=1

Ryan, J., Pollack, K., & Antonelli, F. (2009). Teacher diversity in Canada: Leaky pipelines, bottlenecks, and glass ceilings. *Canadian Journal of Education, 32*(3), 591–617.

Sabates, R., Akyeampong, K., Westbrook, J., & Hunt, F. (2010). *School drop out: Patterns, causes, changes, and policies*. Background paper prepared for the Education for all global monitoring report 2011: The hidden crisis: Armed conflict and education. Brighton, United Kingdom: Centre for International Education School of Education and Social Work University of Sussex. Retrieved from http://unesdoc.unesco.org/images/0019/001907/190771e.pdf

Santoro, N. (2015). The drive to diversify the teaching profession: Narrow assumptions, hidden complexities. *Race, Ethnicity, and Education, 18*(6), 858–876.

Santoro, N., & Reid, J. (2006). All things to all people': Indigenous teachers in the Australian teaching profession. *European Journal of Teacher Education, 29*(3), 287–303.

Sleeter, C. E., & Owuor, J. (2011). Research on the impact of teacher preparation to teach diverse students: The research we have and the research we need. *Action in Teacher Education, 33*, 524–536.

Taiwan Today. (2015). *Preserving cultural diversity*. Retrieved from http://taiwantoday.tw/ct.asp?xItem=226114&ctNode=2198&mp=9

Tatto, M. T. (2013). Chapter 1: Overview of the Teacher Education and Development Study in Mathematics. In M. T. Tatto (Ed.), *Teacher Education and Development Study in Mathematics (TEDS-M): Technical report: Policy, practice, and readiness to teach primary and secondary mathematics*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Tatto, M. T., Rodriguez, M., Ingvarson, L., Rowley, G., Maeda, Y., & Byun, S. Y. (2013). Chapter 4: Development of the TEDS-M survey questionnaires. In M. T. Tatto (Ed.), *Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics: Technical Report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Tatto, M. T., Rodriguez, M., Reckase, M., Rowley, G., & Lu, Y. (2013). Scale development and reporting: Opportunities to learn, beliefs, and mathematics knowledge for teaching. In M. T.

Tatto (Ed.), *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report* (pp. 164–173). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Tatto, M. T., Schwille, J. S., Senk, S., Ingvarson, L. C., Peck, R., & Rowley, G. L. (2009). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework: Policy, practice, and readiness to teach primary and secondary mathematics*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Teacher Education Accreditation Council. (2009). *TEAC principles and standards for teacher education programs*. Retrieved from http://www.teac.org/wp-content/uploads/2009/03/quality-principles-for-teacher-education-programs.pdf

United Nations Educational, Scientific, and Cultural Organization. (2014). *Education for all global monitoring report: Teaching and learning, achieving quality for all*. Oxford, United Kingdom: Author.

United States Department of Education. (2010). *U.S. Department of Education launches national teacher recruitment campaign*. Washington, DC: Author Retrieved from http://www.ed.gov/news/press-releases/us-department-education-launches-national-teacher-recruitment-campaign

Villegas, A. M., & Davis, D. (2008). Preparing teachers of color to confront racial/ethnic disparities in educational outcomes. In M. Cochran-Smith, S. Feiman-Nemser, & J. McIntyre (Eds.), *Handbook of research in teacher education: Enduring issues in changing contexts* (pp. 583–605). Mahwah, NJ: Lawrence Earlbaum.

Villegas, A. M., & Davis, D. E. (2007). Approaches to diversifying the teacher force: Attending to issues of recruitment, preparation, and retention. *Teacher Education Quarterly, 24*(4), 137–147.

Villegas, A. M., & Irvine, J. J. (2010). Diversifying the teaching force: An examination of major arguments. *The Urban Review, 42*(3), 175–192.

Villegas, A. M., Strom, K., & Lucas, T. (2012). Closing the racial/ethnic gap between students of color and their teachers: An elusive goal. *Equity and Excellence in Education, 45*(2), 283–301.

Wolff, R., Severiens, S., & Meeuwisse, M. (2010). Chapter 6 Attracting and retaining diverse student teachers. In *Educating teachers for diversity: Meeting the challenge*. Paris, France: OECD.

Yen, S. H. (2009). Effective strategies for teaching Taiwanese minority students with low achievement and low socioeconomic backgrounds. *Asia Pacific Education Review, 10*, 455–463.

Chapter 6
A Comparative International Study of Differences in Beliefs Between Future Teachers and Their Educators

Michael C. Rodriguez, Maria Teresa Tatto, José Palma, and Kyle Nickodem

Abstract In this study we investigate the extent to which differences exist in beliefs about teaching and learning mathematics between future teachers and their teacher educators/professors across five countries, including Chinese Taipei, Poland, Russian Federation, Singapore, and the United States. The analyses include teacher educators and future primary and secondary teachers grouped within institutions within countries and is based on meta-analytic methods to estimate effects within institutions and synthesize them across institutions within a country. We found significant differences in a number of beliefs about teaching and learning mathematics, and for some countries, the differences between future teachers and their educators vary across institutions. In a number of important and somewhat consistent ways, variation in differences in beliefs is explained by future teachers' opportunities to learn in the areas of mathematics pedagogy, general pedagogy, and field experiences provided by their teacher preparation programs. In addition, program coherence, mathematical content knowledge, and mathematical pedagogical knowledge also moderate the magnitude of differences in beliefs between future teachers and their educators. We discuss the implications of our findings for the design of the curriculum and experiences provided by teacher education programs.

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

M. C. Rodriguez (✉) · J. Palma · K. Nickodem
Department of Educational Psychology, University of Minnesota, Minneapolis, MN, USA
e-mail: mcrdz@umn.edu; palm0304@umn.edu; nicko013@umn.edu

M. T. Tatto
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: maria.tatto@asu.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_6

Introduction

Although there is wide agreement that beliefs are an important influence on teaching, there is no compelling evidence that beliefs can be effectively influenced by teacher preparation or that holding a particular belief is an intrinsic characteristic of those individuals who become good teachers (Tatto & Coupland, 2003; Tatto et al., 2012). Yet teacher education programs do allocate a portion of their curriculum to addressing beliefs as part of future teachers' cognitive processes (Calderhead, 1996; Clark & Peterson, 1986; Shavelson, 1981). The key assumption is that there are close connections among knowledge, beliefs, and action (American Mathematical Society, 2012). In spite of the attention this assumption has received among teacher educators, the field suffers from lack of agreement concerning definitions and best practices (Kane, Sandretto, & Heath, 2002).

In a review of the literature, Tatto and Coupland (2003) discussed the wide range of definitions for *belief* and recognized the lack of clarity about what Pajares (1992) once referred to as a messy construct. They found that beliefs about subject matter content and pedagogy were the focus of a significant number of studies, with less attention given to other areas, such as beliefs about technology and about diverse students, among others.

Some researchers have investigated the interaction between teacher educators' pedagogy and future teachers' learning with mixed results. Lunenberg, Korthagen, and Swennen (2007) focused on a set of prospective teachers and teacher educators in the Netherlands to investigate the prevalence of modelling by teacher educators. They found that few of the ten teacher educators in the study recognized modelling of effective practice as a part of the training they offered their prospective teachers and lamented the lack of a body of knowledge about teacher educators' roles.

Using TEDS-M data, Hsieh, Law, and Shy (2011) compared prospective teachers' and teacher educators' beliefs about the effectiveness and quality of their teacher education programs. Through descriptive statistics, they showed notable differences between prospective teachers' and teacher educators' reports. Most importantly, they found that teacher educators tended to rate their programs' effectiveness far higher than the prospective teachers themselves. They speculated that this could lower teacher educators' motivation to evaluate and improve their programs.

Although some researchers have investigated the connection of teacher beliefs with program activities, fewer have explored the association between teacher educators' and future teachers' beliefs. In three studies, Tatto explored this question in samples of U.S. teacher education programs using entry and exit level data in three areas: beliefs about teaching diverse students; beliefs about instructional choice; and beliefs about purposes of education, teachers' roles, and practices (Tatto, 1996, 1998, 1999). In these studies, a consistent finding was that programs with teacher educators who more consistently conceive of learning mathematics as a process of inquiry saw important and positive changes in their student teachers' views at the end of their programs.

In sum, future teachers' beliefs are considered an important teacher education outcome, and teacher educators are key to understanding the kinds of beliefs teacher education programs value and whether the opportunities to learn (OTL) provided by these programs influence these beliefs, including mathematics content knowledge and the pedagogy for teaching mathematics. Here, we are able to investigate the differences in beliefs between future teachers and their educators, given future teacher performance on the assessment of mathematical content knowledge (MCK) and mathematical pedagogical content knowledge (MPCK).

Research Questions

We address the following questions:

(a) How are the beliefs about teaching and learning mathematics held by future teachers similar to or different from those held by educators in their institutions by the end of the teacher preparation program?
(b) To what extent are OTL, MCK, and MPCK associated with differences between beliefs of future teachers and beliefs of their educators across institutions?
(c) Do differences between teacher educator and future teacher beliefs and the effects of OTL vary across countries?

We investigated these questions using data collected in the TEDS-M study from teacher educators and their future teachers in primary and secondary programs in five countries: Chinese Taipei, Poland, Russia, Singapore, and the United States of America. These countries obtained high or very high scores in our assessments of MCK and MPCK (see Tatto et al., 2012), thus representing systems producing future teachers with firm knowledge of the subject. Studying the U.S. is of particular interest in light of the efforts that have reformed mathematics teacher education since at least the 1990s and the strong research base that has in many cases influenced reforms in several other countries as well, including Chinese Taipei and Singapore.

To measure beliefs, we relied on the work of scholars engaged in international comparative research (Deng, 1995; Grigutsch, Raatz, & Törner, 1998; Ingvarson, Beavis, & Kleinhenz, 2007; Muis, 2004; Tatto, 1996, 1998, 1999; Tatto & Coupland, 2003). Because the central goal of this study was to investigate the outcomes of teacher education, we focused on those areas of beliefs judged to influence (or be influenced by) the process of learning to teach mathematics. This includes beliefs about the nature of mathematics (e.g., Can mathematics be considered as the application of a set of rules and procedures to be memorized and mastered, or is it more a process of inquiry or problem solving?), beliefs about learning mathematics (e.g., Is mathematics better learned via teacher direction or through active learning?), beliefs about mathematics achievement (e.g., Is one naturally good at mathematics or can one become good at mathematics?), and beliefs about the effectiveness of the teacher preparation program as a whole.

In addition, because our intention was to investigate the association between beliefs and future teachers' OTL in their programs, we included measures of OTL that future teachers had in their mathematics pedagogy courses (foundations, instruction, participation and reading in classes; solving problems; instructional planning and practice; and learning about standards and assessments); in their pedagogy courses (e.g., in the areas of the social sciences, in learning to teach diverse students and to reflect on and improve practice); and from school experience (e.g., learning how to connect classroom learning in the university to school practice, and how to learn from supervisors). We also asked about the degree to which the program experienced by future teachers seemed coherent. We used the assessment results indicating future teachers' MCK and MPCK both as control variables and to note associations of interest where they appeared. We develop these concepts below (for more detail regarding each measure, see Tatto et al., 2008 and Tatto, 2013).

Defining and Measuring Beliefs

A belief is a cognitive act or condition where a proposition (in measurement terms, a statement indicating a belief) is taken to be true (Egan, 1986). In psychological theory, there are a number of types of beliefs, including propositional attitudes, subjective probabilities, inferences, and associations (Egan, 1986). The specific nature or structure of the belief is of less interest to us here; rather, we are interested in beliefs as cognitive states regarding the teaching and learning of mathematics that future teachers and their educators hold to be true. Beliefs are a cognitive basis for values and behaviors (Ajzen & Fishbein, 1977), a connection with evidence in teaching practice (Aguirre & Speer, 1999) (more on this below). We investigated five teacher-relevant beliefs as measured in TEDS-M.

Beliefs About the Nature of Mathematics Information about this area was collected from questions that explored how future teachers perceive mathematics as a subject—for example, as formal, structural, procedural, or applied (Grigutsch et al., 1998; Ingvarson et al., 2007). We examined two related beliefs in this area: (a) mathematics as the application of a set of rules and procedures and (b) mathematics as a process of inquiry.

Beliefs About Learning Mathematics This area included questions about the appropriateness of particular instructional activities, questions about students' cognition processes, and questions about the purposes of mathematics as a school subject. We included two related beliefs: (c) learning mathematics through teacher direction and (d) learning mathematics through active learning.

Beliefs About Mathematics Achievement This area comprised future teachers' beliefs about various teaching strategies used to facilitate the learning of mathematics and about how mathematics learning may take place. It also included questions exploring the application of attribution theory to teaching and learning interactions (e.g., innate ability for learning mathematics). We included one specific belief about mathematics achievement: (e) mathematics achievement is a fixed ability.

Beliefs About Preparedness for Teaching Mathematics The fourth area of beliefs concerned the extent to which future teachers perceived their teacher preparation as having given them the capacity to carry out the central tasks of teaching and to meet the demands of their first year of practice. The questions asked about (f) preparedness to teach mathematics (in areas such as assessment, management of learning environments, and practices for engaging students in effective learning) and the extent to which teachers become active members of their professional community (Ingvarson et al., 2007).

Defining and Measuring OTL

TEDS-M investigators used the concept of OTL as central to explaining the impact of teacher preparation on teacher learning. Torstén Husén first defined and used this term to explain student learning in IEA's First International Study of Achievement in Mathematics (FIMS). Husén (1967) defined OTL as the extent to which students had the opportunity to study particular topics or learn problem solving techniques relevant to a test. If students have not had such opportunities, they might generalize skills and abilities from related topics or problem solving techniques, but the chance of answering unique test items correctly is diminished (as described in Burstein, 1993).

TEDS-M investigators explored OTL in teacher preparation. Inclusion of OTL in the study served several purposes: as an explanation of differences in levels of knowledge; as an indicator of curricular variation among countries; as an aspect of fairness (e.g., appropriateness of language of test items); and as a representation of the diversity of content, both overall and for distinct groups of teachers (Floden, 2002).

The TEDS-M survey included multiple scales to allow exploration of the various OTL that future mathematics teachers have across countries. These included OTL in the following areas: (a) university- or tertiary-level mathematics, (b) school-level mathematics, (c) mathematics education/pedagogy, and (d) general education/pedagogy. They also included (e) how to accommodate classroom diversity and to reflect on practice, (f) how to learn from school experience and the practicum, and (g) the extent to which a teacher education program had OTL that were considered as coherent.

In this chapter we focus on the OTL involving mathematics education pedagogy, general education pedagogy, teaching for diversity, reflecting on practice in schools and through practicum experiences, and OTL in a coherent teacher education program.

OTL Mathematics Education/Pedagogy This area included (a) the foundations of mathematics education, (b) mathematics instruction, (c) participation in mathematics education courses, (d) doing readings in mathematics education courses,

(e) solving problems in class, (f) studying instructional practice, (g) studying instructional planning, and (h) studying the uses of assessment and (i) assessment practices.

OTL General Knowledge for Teaching This area included (j) the social science of education, (k) general educational applications, (l) teaching for diversity, (m) teaching for reflection on practice, and (n) teaching for improving practice.

OTL in Schools and Through the Practicum A far more extensive section asked more in-depth questions regarding in-school experience. This area included (o) connecting classroom learning to practice, (p) the reinforcement of teacher preparation program goals in the school setting, and (q) the quality of supervising- teacher feedback.

OTL in a Coherent Teacher Education Program The (r) coherence area included items exploring program consistency across courses and experiences offered to future teachers, and whether there are explicit standards with expectations for what future teachers should learn from their respective programs.

Mathematical Content and Mathematical Pedagogical Content Knowledge

MCK and MPCK as measured in TEDS-M for future primary and secondary teachers is described in detail in Tatto et al. (2008), Tatto et al. (2012), and Tatto (2013). MCK was defined and assessed in four content domains to be consistent with the school curriculum across the globe and with TIMSS frameworks (Mullis et al., 2007). It included number, algebra, geometry, and data, across the three cognitive domains of application, knowledge, and reasoning. The MPCK assessment was developed by the TEDS-M researchers and measured two sub-domains: mathematical curricular knowledge and knowledge of planning for mathematics teaching and learning (Tatto et al., 2008, pp. 37–38). In this study, we used scales developed from these measures as control variables (for more detail on how TEDS-M scales were constructed, see Tatto, 2013).

Methods

Here we investigate similarities and differences when comparing beliefs of future teachers and their educators about teaching and learning mathematics. Educators' views are considered program expectations conveyed to future teachers. Thus, for programs to be effective, the expectation would be that, at the end of the program, future teachers' views would be closely aligned to those of their educators. Yet these expectations are mediated by other factors in complex ways. The diverse OTL made available to future teachers are important mediators, as is the content knowledge they possess. It is possible that those who are more mathematically knowledgeable have fixed ideas about how they themselves learned mathematics and that these views may go untouched by program efforts.

To explore these propositions, our analyses grouped educators, future primary teachers (FPT), and future secondary teachers (FST) within institutions within each of five countries, including Chinese Taipei, Poland, the Russian Federation, Singapore, and the U.S. The approach taken is one based on the methods of meta-analysis (see Cooper, Hedges, & Valentine, 2009), estimating effects within institutions and synthesizing them across institutions within country. We carefully describe the steps followed for conducting the analyses.

Data Source

Data were obtained from the TEDS-M study, including future teachers and their teacher educators from Chinese Taipei, Poland, the Russian Federation, Singapore, and the U.S. The models used to address the research questions included two beliefs about the nature of mathematics: (a) mathematics is a set of rules and procedures and (b) mathematics is a process of inquiry; two beliefs about learning mathematics: (c) mathematics is learned through teacher direction and (d) mathematics is learned through active learning; beliefs about math achievement: (e) mathematics is a fixed ability; and beliefs about the teacher preparation program as a whole: (f) perceptions of preparedness for teaching mathematics.

Measures of OTL in mathematics education pedagogy included the degree to which future teachers learned about (a) the foundations of education pedagogy and (b) mathematics instruction; experienced opportunities for (c) class participation, (d) class readings, and (e) solving problems during their mathematics education pedagogy courses; and had opportunities to explore and experience (f) instructional practice, (g) instructional planning, (h) assessment uses, and (i) assessment practice.

Measures of OTL in general education pedagogy included the degree to which future teachers learned about (j) the social sciences relevant to education pedagogy, (k) the application of education pedagogy, (l) teaching for diversity, (m) teaching for reflection on practice, and (n) teaching for improving practice.

Measures of OTL in school experience or practicum included the degree to which future teachers had the opportunities to experience (o) connecting classroom learning to practice and (p) supervising teacher reinforcement of university goals for practicum; and future teacher perceptions of (q) supervising teacher feedback quality. Finally we include (r) a measure of overall program coherence.

Measures of both MCK and MPCK were used as control variables, primarily to control for differences in mathematics knowledge, skills, and abilities across institutions. We report the effects of these control variables when they were notable.

The mean of each belief variable was computed for future teachers and educators separately within each institution, and the mean of each OTL variable was computed for future teachers within each institution. TEDS-M sampling weights for future teachers and educators were used when computing each institution's mean in all countries, with one exception. The U.S. educator sample was not a representative sample, so educator weights were not used to compute the belief means within each institution.

Analysis

The analysis procedures are presented in two sections below. The first section describes the data preparation process and computation of standardized mean differences in beliefs between educators and future teachers. The second section specifies the regressions used to model standardized mean differences in beliefs as a function of OTL variables.

Standardized Mean Differences Analysis The approach to evaluate the differences between future teachers and educators across institutions was based on meta-analytic methods. Future teachers cannot be associated with educators in a one-to-one manner, since future teachers have multiple educators, and not all educators have contact with each future teacher. However, future teachers and educators do exist within a common institution. Each institution is considered a study, within which we can estimate the difference in beliefs between future teachers and educators. The difference is estimated as Cohen's *d*, the standardized mean difference (Borenstein, 2009). Because each difference (effect) is estimated with different levels of precision, largely based on sample size, a weighted-least-squares (WLS) analysis is required. Weights are estimated based on the sampling variance of *d* (essentially, the sampling error). The weighted average difference is estimated and a test of homogeneity (Q) of these effects is conducted, answering the question of whether there is a common effect (difference in beliefs) across institutions. The Q-test is reported to indicate whether effects vary significantly within a country or, instead, the mean effect (mean difference in beliefs) can adequately represent all institutions in the country. In the case where these differences vary significantly, that variation is modeled, and such models are explored in the second section. The IDB Analyzer (IEA, 2015) was used to compute the means for future teachers and educators, and the R statistical package (R Core Team, 2015) was used to complete the remaining analyses. In each of the following steps, statistics were computed within country.

Step 1. Compute effect sizes $d_{jk} = \frac{\left(\overline{x}_{\mathrm{FT}} - \overline{x}_{\mathrm{E}}\right)_{jk}}{s_k}$, where d_{jk} is the standardized mean difference (i.e., Cohen's *d*) for belief *k* in institution *j*; $\left(\overline{x}_{\mathrm{FT}} - \overline{x}_{\mathrm{E}}\right)_{jk}$ is the

Table 6.1 International standard deviations for beliefs of future primary and secondary level teachers

Belief	Primary *SD*	Secondary *SD*
Process of inquiry	1.57	1.57
Rules and procedures	1.24	1.36
Teacher-directed	0.86	0.95
Active learning	1.33	1.43
Fixed ability	1.04	1.05
Preparedness	1.93	1.79

Table 6.2 Number of primary and secondary level future teachers, educators, and institutions

	Primary			Secondary		
	FT	Ed	Inst	FT	Ed	Inst
Chinese Taipei	923	115	11	273	80	8
Poland	1,833	676	70	268	362	21
Russian Federation	2,219	1,086	49	2,093	1,158	48
Singapore	378	74	1	393	74	1
United States	1,081	726	47	463	681	43

Note: *FT* future teachers, *Ed* educators, *Inst* Institutions

difference between the mean of future teachers (FT) and the mean of educators (E) within institution j for belief k; and s_k is the future teacher population weighted standard deviation for belief k from the international sample (see Table 6.1). Using s_k provides for the same reference standard deviation (scaling factor) for all countries.

Step 2. Compute variance components $v_{jk} = \frac{n_{\mathrm{FT}jk} + n_{\mathrm{E}jk}}{n_{\mathrm{FT}jk} n_{\mathrm{E}jk}} + \frac{d_{jk}^2}{2\left(n_{\mathrm{FT}jk} + n_{\mathrm{E}jk}\right)}$ (Borenstein, 2009), where $n_{\mathrm{FT}jk}$ is the sample size for future teachers within institution j for belief k and $n_{\mathrm{E}jk}$ is the sample size for educators within institution j for belief k. These may vary slightly across beliefs based on the number of participants receiving scores on a given belief k. The maximum $n_{\mathrm{FT}jk}$, $n_{\mathrm{E}jk}$, and j for each country are listed in Table 6.2. These variance components are also referred to as sampling error variances, which serve as the basis for weights in WLS analyses.

Step 3. Compute effect size weights $w_{jk} = \frac{1}{v_{jk}}$ for belief k in institution j. This is used to compute the weighted average difference across institutions and in the WLS regression analyses below.

Step 4. Compute weighted average effect size $d_{\bullet k} = \frac{\sum w_{jk} d_{jk}}{\sum w_{jk}}$ for belief k across institutions. This average $d_{\bullet}$ is computed for each country.

Step 5. Compute the standard error of the average effect size $SE\left(d_{\bullet k}\right)=\sqrt{\frac{1}{\sum w_{jk}}}$ for belief *k*. This is computed for each country.

Step 6. Compute $z_k = \frac{d_{\bullet k}}{SE\left(d_{\bullet k}\right)}$, which allows for statistical testing of whether the country mean difference for belief *k* is different from zero.

Step 7. Compute the Q-statistic $Q_k = \sum \frac{\left(d_{jk} - d_{\bullet k}\right)^2}{v_{jk}}$. The *Q*-statistic follows a Chi-Square distribution with *j*-1 degrees of freedom, where *j* is the number of effects (institutions). This is a homogeneity test statistic where the null hypothesis (H_0) states that for belief *k*, the d_{jk} share a common population effect size; it is a test of the homogeneity of d_{jk} across institutions.

Regression Analysis The regression model for the effect size as a function of *p* OTL predictor variables was specified as follows. This regression was estimated as a WLS regression, employing the weights w_{jk}. When the typical ordinary-least squares (OLS) regression is used, we assume homoscedasticity, an assumption of constant variance across the regression variables. When effect-size statistics are used in regression, we know that they are estimated with different levels of precision, thus reflecting heteroscedasticity. This can be controlled in the regression by modeling the variation in the precision of each estimated effect size, by weighting each effect size by a function of its precision. Thus, the following WLS regression is used: $d_{jk} = \beta_0 + \beta_1(OTL_1)_{jk} + \beta_2(OTL_2)_{jk} + \ldots + \beta_p(OTL_p)_{jk} + e_{jk}$.

For each belief *k*, initial models were fitted by regressing differences (d_{jk}) on all *p* OTL variables. The backward elimination procedure using the *step* function in the R software was used to simplify each model. The *step* function looks to minimize the Akaike Information Criterion (AIC), an information-based model fit criterion that identified variables as contributing to the model. Because the AIC is a measure of relative quality of a given model compared to others, it was used here for evaluating model fit and comparisons across models. The tables of regression results include those variables that remained in each model based on the AIC criterion, and indicate each coefficient statistical significance. Separate models are reported for each country.

To support interpretation across OTL measures, the scores were standardized within each country. The regression coefficients are interpreted such that one standard-deviation change in the OTL predictor is associated with a change of magnitude, given the associated coefficient, in the average standardized mean difference ($d_\bullet$) between future teachers and their educators.

Results

Exploring Differences in Beliefs

As described in the methods section, our primary outcome for this study is the standardized mean difference in beliefs between future teachers and their educators, which is a common effect size that expresses the difference in means in terms of standard deviations. For each institution, we estimated the standardized mean difference between future teachers and their educators for a given belief (future teacher belief mean score minus educator belief mean score). These differences are summarized in an effect size for each institution, *d*, the standardized mean difference.

The *d* for a particular institution and the average ***d.*** for a given country indicate the magnitude of differences in beliefs between future teachers and their educators. If future teachers held the same beliefs as their educators, these effects would be zero.

As the *d*s were estimated based on the mean score of future teachers minus the mean score of educators, $\left(\bar{x}_{\mathrm{FT}} - \bar{x}_{\mathrm{E}}\right)_{jk}$, when *d* is positive, future teachers have higher scores on the particular belief than do educators (that is, they are more likely to agree with specific items comprising the belief than are their educators). In addition, the standard deviation of effects within a country provides an indication of whether a common effect or common difference in beliefs is held across institutions. A large standard deviation indicates that the average ***d.*** for a country might not be a meaningful value because of the variability across institutions within a country—that is, in some institutions, the differences in beliefs may be very small and in others may be very large. We explore inter-institutional variability through a series of regression models, using OTL indices as explanatory variables in the next section. The average ***d.*** and associated standard errors for each belief per country are represented visually in Fig. 6.1.

In general, across beliefs and across countries, the differences observed were larger for FPTs than for FSTs. One notable exception is in Singapore, where the observed differences were larger for FSTs across most beliefs. Additionally, for most beliefs and countries there tended to be more variability in the effect sizes between institutions for FPTs than between institutions for FSTs. That is, institution-level differences in beliefs seemed to vary more for FPTs than FSTs. This may be because secondary teachers are prepared as mathematics specialists; thus, mathematics' stronger grammar (Bernstein, 1999) may result in stronger norms within these programs across institutions and more coherence in views among future teaches and their educators (Tatto 1996, 1998, 1999).

For each country, the average standardized mean difference (***d.***) between future teachers' and educators' beliefs, the standard error of ***d.***, and the standard deviation of institution-based *d*s across institutions are reported in Tables 6.3 to 6.8.

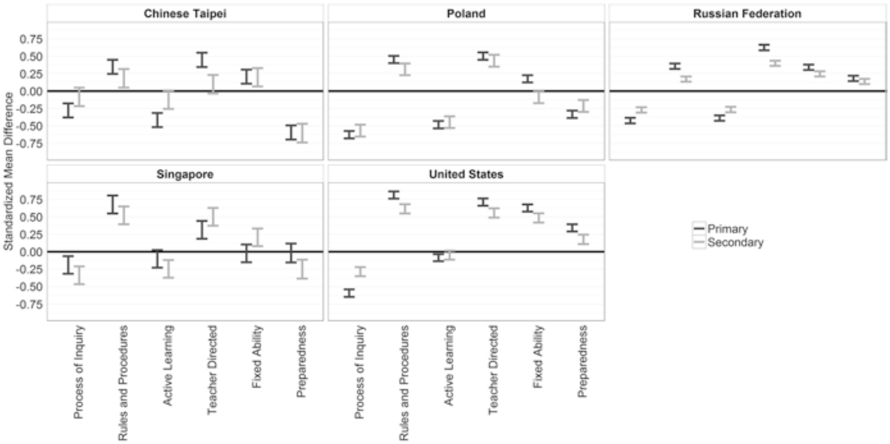


Fig. 6.1 Standardized mean differences between educators and future teachers on beliefs about teaching and learning mathematics with 95% confidence intervals

Table 6.3 Mean differences, standard errors, and standard deviations for process of inquiry

	Primary					Secondary				
	d.	*SE*(***d.***)	*SD*(d_{jk})	*Min*	*Max*	***d.***	*SE*(***d.***)	*SD*(d_{jk})	*Min*	*Max*
Chinese Taipei	−0.28*	0.10	0.32	−1.14	0.48	−0.08	0.13	0.35	−1.06	0.41
Poland	−0.63*	0.05	0.50†	−2.17	0.38	−0.57*	0.09	0.44	−1.65	0.10
Russian Federation	−0.42*	0.04	0.29	−1.24	0.26	−0.27*	0.04	0.30	−1.25	0.72
Singapore	−0.19	0.13				−0.34*	0.13			
United States	−0.59*	0.05	0.45†	−1.91	0.75	−0.29*	0.07	0.36	−1.57	0.87

Note. The d_{jk} statistic is the standardized mean difference between future teachers and their educators for institution *j* for belief *k*; ***d.*** is the average *d* within country
*$d.$ is significantly different than zero at $p < .05$; †*Q*-test of homogeneity is significant at $p < .05$

Beliefs About the Nature of Mathematics With the exception of FPTs in Singapore and FSTs in Chinese Taipei, future teachers were less likely to believe that mathematics is a process of inquiry than their educators (as indicated by the statistically significant negative ***d.*** in Table 6.3); that is, educators generally view mathematics as a process of inquiry to a higher degree than do future teachers. Conversely, future teachers were more likely than their educators to believe that mathematics involves the remembering and application of definitions, formulas, mathematical facts and procedures and applying a set of rules and procedures that prescribe how to solve a problem (Table 6.4). In most countries, the standardized mean differences between future teachers and educators were stable across institutions. However, for FPTs in Poland, there was significant heterogeneity in effect size between institutions on the view that mathematics is a process of inquiry, where $Q(69) = 89.40$ ($p = .05$), and of mathematics as applying rules and procedures, where $Q(69) = 96.46$ ($p = .02$). Recall that the Q-statistic conveys information about

Table 6.4 Mean differences, standard errors, and standard deviations for rules and procedures

	Primary					Secondary				
	d.	*SE*(*d.*)	*SD*(d_{jk})	*Min*	*Max*	*d.*	*SE*(*d.*)	*SD*(d_{jk})	*Min*	*Max*
Chinese Taipei	0.35*	0.10	0.21	0.08	1.04	0.18	0.13	0.30	−0.50	0.55
Poland	0.45*	0.05	0.52†	−0.97	3.05	0.31*	0.08	0.33	−0.18	0.91
Russian Federation	0.36*	0.04	0.38†	−0.74	1.49	0.17*	0.04	0.23	−0.45	0.65
Singapore	0.68*	0.13				0.52*	0.13			
United States	0.81*	0.05	0.38	0.11	1.52	0.62*	0.07	0.39	−0.15	1.89

Note. The d_{jk} statistic is the standardized mean difference between future teachers and their educators for institution *j* for belief *k*; ***d.*** is the average *d* within country
*$\boldsymbol{d.}$ is significantly different than zero at $p < .05$; †*Q*-test of homogeneity is significant at $p < .05$

the variability of effects (differences in beliefs) within country—that is, the extent to which the mean effect adequately represents a common difference across institutions within the country. Here we find significant variability in differences, meaning that there is significant variability in the magnitude of differences across institutions. For example, because the *Q*-statistic is significant in Poland (see the flag associated with the SD in Table 6.3), we know that the mean effect does not adequately describe the differences in beliefs in Poland. That is, there is much variation in the mean (average effect) across institutions likely due to the fact that some institutions prepare specialist and others generalist teachers, so the mean in this case does not describe the differences in beliefs for many institutions—simply because there is significant variation. This significant variation in differences in beliefs across institutions is investigated through regression analyses in the next section of this chapter.

Beliefs About Learning Mathematics FPTs and FSTs in the U.S. (FPT: $d. = -0.09$, $z = -1.71$; FST: $d. = -0.05$, $z = -0.79$) and Singapore (FPT: $d. = -0.10$, $z = -0.81$; FST: $d. = -0.25$, $z = -1.95$) did not differ on average from their educators on the belief that mathematics should be learned through active and independent learning, investigations and discussion; whereas in Chinese Taipei (FPTs only), Poland, and the Russian Federation, FPTs and FSTs belief scores on active learning were significantly lower (i.e., less likely to believe that mathematics should be learned through active learning) than their educators, likely signaling strong cultural beliefs and a reaction against inquiry-based learning (see Table 6.5). Significant variability in effect sizes across institutions was seen for FPTs in Chinese Taipei ($Q[10] = 21.57$, $p = .017$) and Poland ($Q[69] = 124.64$, $p < .001$) and FSTs in the Russian Federation ($Q[47] = 107.84$, $p < .001$). Again, this variation in future teacher and educator differences across institutions is investigated through regression analyses in the next section of this chapter.

With the exception of FSTs in Chinese Taipei ($d. = 0.10$, $z = 0.74$), FPTs and FSTs had significantly higher scores than their educators on the belief that learning mathematics should be teacher-directed and based on attending to teachers' explanations and learning how to solve problems quickly and correctly (Table 6.6). Whereas the effect size estimates were homogeneous for FSTs in most countries,

Table 6.5 Mean differences, standard errors, and standard deviations for active learning

	Primary					Secondary				
	d.	*SE*(*d.*)	*SD*(d_{jk})	*Min*	*Max*	*d.*	*SE*(*d.*)	*SD*(d_{jk})	*Min*	*Max*
Chinese Taipei	−0.42*	0.10	0.47†	−1.25	0.52	−0.12	0.13	0.22	−0.63	0.32
Poland	−0.48*	0.05	0.59†	−3.01	0.92	−0.45*	0.08	0.43	−1.91	0.25
Russian Federation	−0.39*	0.04	0.31	−0.90	0.55	−0.27*	0.04	0.39†	−1.02	0.87
Singapore	−0.10	0.13				−0.25	0.13			
United States	−0.09	0.05	0.38	−0.89	0.67	−0.05	0.07	0.38	−1.03	2.39

Note. The d_{jk} statistic is the standardized mean difference between future teachers and their educators for institution *j* for belief *k*; ***d.*** is the average *d* within country
****d.*** is significantly different than zero at $p < .05$; †*Q*-test of homogeneity is significant at $p < .05$

Table 6.6 Mean differences, standard errors, and standard deviations for teacher directed learning

	Primary					Secondary				
	d.	*SE*(*d.*)	*SD*(d_{jk})	*Min*	*Max*	*d.*	*SE*(*d.*)	*SD*(d_{jk})	*Min*	*Max*
Chinese Taipei	0.45*	0.10	0.60†	−0.18	2.77	0.10	0.13	0.31	−0.54	0.81
Poland	0.50*	0.05	0.62†	−0.84	3.70	0.44*	0.09	0.48†	−0.51	2.81
Russian Federation	0.63*	0.04	0.39†	−0.65	1.36	0.40*	0.04	0.23	−0.16	1.11
Singapore	0.31*	0.13				0.50*	0.13			
United States	0.71*	0.05	0.45†	−0.32	1.84	0.56*	0.07	0.39	−0.53	1.48

Note. The d_{jk} statistic is the standardized mean difference between future teachers and their educators for institution *j* for belief *k*; ***d.*** is the average *d* within country
****d.*** is significantly different than zero at $p < .05$; †*Q*-test of homogeneity is significant at $p < .05$

there was significant variability for FPTs in Chinese Taipei ($Q[10] = 34.62$, $p < .001$), Poland ($Q[69] = 136.70$, $p < .001$), the Russian Federation ($Q[48] = 95.85$, $p < .001$), and the U.S. ($Q[46] = 76.20$, $p = .003$).

Beliefs About the Ability to Learn Mathematics With the exception of Singapore, FPTs were more likely than their educators to believe that mathematics achievement is due to a natural ability to do well at mathematics and that this ability is fixed (i.e., some people are good at mathematics and some are not) (see Table 6.7). For FSTs, this trend was only found in the Russian Federation (***d.*** = 0.25, $z = 6.57$) and the U.S. (***d.*** = 0.48, $z = 7.24$). Furthermore, the U.S. was the only country to show variability in effect size between institutions, both for FPTs ($Q[46] = 105.21$, $p < .001$) and for FSTs ($Q[42] = 75.85$, $p = .001$) and their educators—again, indicating significant variation in the differences in beliefs across institutions. This signals a great degree of heterogeneity in differences in beliefs across teacher education institutions in the U.S.

Perceptions of Preparedness for Teaching Mathematics The most country-to-country variation was seen in the views held by future teachers and their educators concerning preparedness for teaching mathematics (see Table 6.8). Whereas FPTs and FSTs in Chinese Taipei (FPT: ***d.*** = −0.59, $z = -5.84$; FST: ***d.*** = −0.60, $z = -4.49$) and Poland (FPT: ***d.*** = −0.33, $z = -6.24$.; FST: ***d.*** = −0.21, $z = -2.53$) report that

Table 6.7 Mean differences, standard errors, and standard deviations for mathematics as a fixed ability

	Primary					Secondary				
	d.	*SE(d.)*	*SD*(d_{jk})	*Min*	*Max*	*d.*	*SE(d.)*	*SD*(d_{jk})	*Min*	*Max*
Chinese Taipei	0.21*	0.10	0.27	−0.29	0.93	0.20	0.13	0.32	−0.47	0.88
Poland	0.18*	0.05	0.45	−0.85	1.47	−0.09	0.08	0.26	−0.72	0.31
Russian Federation	0.34*	0.04	0.25	−0.63	0.98	0.25*	0.04	0.19	−0.56	0.62
Singapore	−0.03	0.13				0.20	0.13			
United States	0.63*	0.05	0.53†	−0.57	1.67	0.48*	0.07	0.58†	−0.77	1.66

Note. The d_{jk} statistic is the standardized mean difference between future teachers and their educators for institution *j* for belief *k*; ***d.*** is the average *d* within country
*$d.$ is significantly different than zero at $p < .05$; †Q-test of homogeneity is significant at $p < .05$

Table 6.8 Mean differences, standard errors, and standard deviations for preparedness for teaching

	Primary					Secondary				
	d.	*SE(d.)*	*SD*(d_{jk})	*Min*	*Max*	*d.*	*SE(d.)*	*SD*(d_{jk})	*Min*	*Max*
Chinese Taipei	−0.59*	0.10	0.40	−1.57	0.08	−0.60*	0.13	0.32	−1.65	−0.13
Poland	−0.33*	0.05	0.47	−1.37	1.69	−0.21*	0.08	0.45	−1.03	0.69
Russian Federation	0.18*	0.04	0.32†	−1.28	0.88	0.14*	0.04	0.33†	−0.68	1.03
Singapore	−0.02	0.14				−0.25	0.14			
United States	0.34*	0.05	0.44†	−0.49	1.40	0.18*	0.07	0.59†	−1.22	1.42

Note. The d_{jk} statistic is the standardized mean difference between future teachers and their educators for institution *j* for belief *k*; ***d.*** is the average *d* within country
*$d.$ is significantly different than zero at $p < .05$; †Q-test of homogeneity is significant at $p < .05$

they are prepared for teaching by their institutions at a lower level than their educators report, FPTs and FSTs in the Russian Federation (FPT: $d. = 0.18$, $z = 4.66$; FST: $d. = 0.14$, $z = 3.64$) and the U.S. (FPT: $d. = 0.34$, $z = 3.64$; FST: $d. = 0.18$, $z = 2.60$) report being much more prepared for teaching than did their educators. Furthermore, for FPTs and FSTs, there is significant variability in effect sizes between institutions in the Russian Federation (FPT: $Q[48] = 68.23$, $p = .029$; FST: $Q[47] = 77.00$, $p = .004$) and the U.S. (FPT: $Q[46] = 70.37$, $p = .012$; FST: $Q[42] = 74.98$, $p = .001$), whereas Chinese Taipei and Poland had similar effect sizes across institutions. This denotes a larger degree of heterogeneity across these institutions in the U.S. and the Russian Federation.

Modeling Variability Through Regression

A series of regressions were completed to model the variability in effect sizes across institutions—testing whether differences in beliefs across institutions could be explained as a function of OTL. Similar regressions were run for each country. The

WLS regressions revealed few consistencies in OTLs predicting the difference between future teacher and educator beliefs across countries (see Tables 6.9 to 6.14). Note that for all of the regressions, a positive intercept indicates that, on average, when all other variables are held constant, future teachers had higher scores on the given belief (that is, they were more likely to agree with the statements comprising the belief) than their educators. Therefore, an OTL with a significant positive coefficient suggests that a one standard-deviation increase in the OTL will increase the difference in belief positively, potentially widening the gap in the level of belief between future teachers and their educators. Conversely, if an OTL has a negative coefficient, every one standard-deviation increase in the OTL will shift the difference in belief between future teachers and their educators negatively.

For instance, in Table 6.12 the intercept for FPTs in the U.S. regarding teaching as active learning is −0.12, meaning that, for institutions with average OTL, educators are more likely to believe in teaching as active learning than their FPTs do. The coefficient for OTL to apply mathematics to real-world problems and to distinguish between procedural and conceptual mathematics when teaching (instructional practice) is 0.40, indicating that a 1.0 standard deviation (SD) increase in the OTL from instructional practices is associated with a shift in that difference of 0.40 ($-0.12 + 0.40 = 0.28$, essentially reversing the difference). Because of the complexity in interpretation of coefficients, we focus on the extent to which OTL predicts variability in differences in beliefs, finding overall that future teachers and their educators hold different beliefs, and this appears to vary, in some cases, as a function of OTL. To support interpretations of the regression coefficients, we mention the statistically significant effects and note that when the coefficient is the same sign as the intercept, that effectively increases the difference in beliefs estimated by the intercept; when the coefficient is the opposite sign as the intercept, that effectively reverses the difference, or at least reduces the difference to a point where additional OTL may actually reverse the difference.

We completed regressions for three countries with sufficient numbers of institutions: Poland, the Russian Federation, and the U.S. We found that some OTL variables, MCK, and MPCK were often significant predictors of belief differences between educators and future teachers. Conversely, other types of OTL were rarely, if ever, significant predictors of the difference in beliefs between future teachers and their educators. We examine these below.

Mathematics as a Process of Inquiry Across each country (the three with sufficient numbers of institutions to support regression), for primary and secondary programs, educators were more likely to agree that mathematics learning is a process of inquiry than future teachers (see the negative intercepts in Table 6.9 and negative averages in Fig. 6.1), with the WLS regression models explaining 43% (Poland FPTs and FSTs) to 55% (Russian Federation FSTs) of the variance across institutions in differences between future teachers and educators. OTL about national or state standards and assessments as related to pupils' learning (assessment practice), was shown to reduce and reverse the difference in this belief for FSTs in Poland as well as for FPTs and FSTs in the Russian Federation (with more OTL assessment

Table 6.9 WLS regression results for future primary and secondary teachers on mathematics learning as a process of inquiry

	Primary			Secondary		
Predictors	POL	RUS	U.S.	POL	RUS	U.S.
R^2	.43	.46	.53	.43	.55	.44
Intercept	−0.68	−0.43	−0.65	−0.59	−0.29	−0.31
OTL variables						
Foundations		0.05				−0.13
Instruction						
Class participation					0.18*	
Class reading	0.08	−0.08				
Solving problems	−0.15	−0.10		−0.15	−0.15*	0.09
Instructional practice		0.12	0.23*		0.24*	
Instructional planning		−0.15	−0.19		−0.21*	0.40*
Assessment uses	−0.15		−0.15		−0.31*	
Assessment practice		0.20*		0.29*	0.41*	
Social Science					0.06	
Application	−0.29*		−0.20*	−0.24*		
Teach for diversity	0.34*	−0.10			−0.09	
Teach for reflection						
Teach for improving						−0.51*
Classroom learning to practice	0.16*	0.23*	0.13			0.25*
Reinforcement of goals		−0.11*				0.19*
Feedback quality					−0.09	−0.13
Program coherence			0.15		0.11	−0.32*
Knowledge variables						
Mathematical content			0.28*		−0.07	−0.43*
Mathematical pedagogical content	0.38*		−0.18			0.34*

Note: *POL* Poland, *RUS* Russian Federation, *U.S.* United States of America
*$p < .05$

practice, future teachers approach agreeing with their educators, reversing the differences). Similarly, OTL from school experiences, such as observing and practicing teaching, and collecting and analyzing evidence about pupil learning as a result of their teaching methods (classroom learning to practice), also reduced the difference for FPTs in Poland and the Russian Federation, and for FSTs in the U.S. The associations between the measures of knowledge and beliefs in this case were not consistent across countries and levels of teacher preparation programs. Overall, OTL explained about half of the variation in differences in beliefs.

Mathematics as a Set of Rules and Procedures In all cases, the future teachers who participated in our study were more likely to agree that mathematics consists of the application of a set of rules and procedures (consistent with their lower scores on beliefs regarding mathematics as a process of inquiry) relative to their educators (see the positive intercepts in Table 6.10 and positive averages in Table 6.4). The

Table 6.10 WLS regression results for future primary and secondary teachers on mathematics learning as a set of rules and procedures

	Primary			Secondary		
Predictors	POL	RUS	U.S.	POL	RUS	U.S.
R^2	.23	.54	.42	.78	.50	.47
Intercept	0.50	0.35	0.81	0.29	0.18	0.62
OTL variables						
Foundations		0.12*	0.09			
Instruction				−0.45*		
Class participation		−0.23*	0.09	0.38*	0.09*	0.14*
Class reading				0.13*		
Solving problems						
Instructional practice		0.31*				
Instructional planning	0.18*	−0.23		0.31*	0.16*	
Assessment uses			0.18*			
Assessment practice			−0.17*		−0.17	
Social Science		−0.10	−0.09		0.07	
Application						
Teach for diversity	−0.24*	−0.15				−0.37*
Teach for reflection		0.18			−0.11	
Teach for improving						0.26*
Classroom learning to practice				0.38*		
Reinforcement of goals	−0.13		−0.14		0.12*	
Feedback quality		−0.10	0.15*	−0.29*		
Program coherence		0.20*				
Knowledge variables						
Mathematical content	−0.35*	−0.34*	−0.13*	−0.36*	−0.17*	0.21
Mathematical pedagogical content		0.23		0.40*	0.12	−0.44*

Note: *POL* Poland, *RUS* Russian Federation, *U.S.* United States of America
*$p < .05$

WLS regression models explained 23% (Poland primary) to 78% (Poland secondary) of the variance in differences between future teachers and educators. Future teachers' opportunity to ask questions, participate in discussion, and to teach a class during their program (class participation) tends to reverse the difference in beliefs in the Russian Federation primary programs, but increase the differences for FSTs in all three countries. Similarly, having opportunities to develop instructional plans to accommodate pupils' diverse learning needs significantly increased the differences in views between teacher educators and their FPTs and FSTs in Poland and FSTs in the Russian Federation, an outcome that may occur because instructional plans are more attuned to school norms than to program norms, and, in some cases, these norms differ. In most cases, institutions where future teachers had higher levels of MCK also had views that were more closely aligned with those of their educators (that is, they were more likely to reject the view that learning mathematics for the most part means memorizing and applying a set of rules and procedures).

Table 6.11 WLS regression results for future primary and secondary teachers on mathematics teaching as teacher directed

	Primary			Secondary		
Predictors	POL	RUS	U.S.	POL	RUS	U.S.
R^2	.54	.55	.41	.43	.39	.50
Intercept	0.55	0.63	0.73	0.44	0.40	0.53
OTL variables						
Foundations	−0.21*	0.18*			0.07	−0.13
Instruction		−0.12		−0.35	0.08	
Class participation		−0.16*			0.06	
Class reading					0.09	
Solving problems	0.24*					−0.15
Instructional practice		0.28*			0.12	0.21
Instructional planning	0.38*	−0.27*	−0.15	0.27		−0.21
Assessment uses			0.34*		−0.21*	
Assessment practice				−0.27		
Social science						0.12
Application	0.27*					−0.16
Teach for diversity	−0.53*				0.21*	−0.16
Teach for reflection	0.23*					
Teach for improving	−0.30*			0.36	−0.28*	0.33*
Classroom learning to practice	−0.23*			0.19		
Reinforcement of goals			−0.16			
Feedback quality	0.16*	−0.10			0.07	
Program coherence		0.22*				
Knowledge variables						
Mathematical content		−0.16*			−0.15*	0.16
Mathematical pedagogical content	−0.61*		−0.21*			−0.30*

Note: *POL* Poland, *RUS* Russian Federation, *U.S.* United States of America
$^*p < .05$

Teacher-Directed Mathematics Learning In all cases, future teachers were more likely to believe that mathematics instruction should be teacher-directed relative to their educators (see the positive intercepts in Table 6.11 and positive averages in Table 6.6). Regarding variation in these differences, the models resulted in R^2 values ranging from .39 (Russian Federation secondary) to .55 (Russian Federation primary), with very few consistent predictors across countries and teacher level. None of the OTL measures were significant predictors for more than two of the six teacher types by country combinations. Even then, the direction of the impact varied across country and teacher level. In two cases for MCK and in three cases for MPCK, institutions with future teachers with higher knowledge scores were less likely to believe that mathematics instruction should be teacher-directed, toward a view generally more common among their educators. Overall, OTL appears to explain about half of the variation in differences in this belief, although there is variation across countries.

Table 6.12 WLS regression results for future primary and secondary teachers on mathematics teaching as active learning

	Primary			Secondary		
Predictors	POL	RUS	U.S.	POL	RUS	U.S.
R^2	.23	.10	.48	.83	.40	.21
Intercept	−0.52	−0.39	−0.12	−0.58*	−0.26	−0.08
OTL variables						
Foundations				−0.20	−0.18*	−0.13
Instruction						0.25*
Class participation			−0.10	0.88*		
Class reading				0.17*		
Solving problems						
Instructional practice			0.40*			
Instructional planning						
Assessment uses						
Assessment practice	−0.25*		−0.30*			
Social Science						
Application			−0.08			
Teach for diversity				−0.74*	−0.23*	
Teach for reflection		0.11*	0.12	0.48*		
Teach for improving						
Classroom learning to practice	0.28*			0.63*	0.22	
Reinforcement of goals						
Feedback quality				−0.23*	−0.22*	−0.17*
Program coherence			−0.14*	−0.36*	0.25*	
Knowledge variables						
Mathematical content			0.15*	−1.47*		
Mathematical pedagogical content	0.33*			1.74*	0.13*	

Note: *POL* Poland, *RUS* Russian Federation, *U.S.* United States of America
* $p < .05$

Active Learning of Mathematics In all cases, and relative to their educators, future teachers were less likely to believe that mathematics instruction should allow time for investigations and discussion so that pupils can figure out their own solutions to mathematical problems and to understand why an answer is correct (mathematics as active learning; see the negative intercepts in Table 6.12). The model only explained 10% of the variation in the gap in this belief between educators and their FPTs in the Russian Federation, but explained 83% of the variation for FSTs in Poland. There was little consistency in the predictors for FPTs; however, in all three countries, higher quality of feedback provided by the supervising teacher (supervising teacher feedback quality) uniformly increased the differences between FSTs and their educators regarding the notion that mathematics learning can be inquiry-based. In addition, having more OTL about standards and assessments, including the analysis of assessment results in relation to pupils' learning (assessment practice), increased the difference in views among educators and their FPTs in

Table 6.13 WLS regression results for future primary and secondary teachers on mathematics learning as a fixed ability

	Primary			Secondary		
Predictors	POL	RUS	U.S.	POL	RUS	U.S.
R^2	.58	.48	.38	.94	.21	.55
Intercept	0.23	0.34	0.64	−0.17	0.25	0.43
OTL variables						
Foundations		0.10*		−0.24*		
Instruction	−0.06			−0.67*		
Class participation		−0.10*		0.30*		−0.28*
Class reading				0.16*		
Solving problems				0.18*		−0.20
Instructional practice						
Instructional planning		−0.08	−0.30*	0.66*		−0.37*
Assessment uses			0.42*			
Assessment practice				−0.37*		
Social Science	0.14					0.31*
Application	0.14*					−0.53*
Teach for diversity	−0.13		0.19			
Teach for reflection		0.10		0.25*		
Teach for improving			−0.16			0.54*
Classroom learning to practice	−0.12			0.56*	0.09	
Reinforcement of goals	0.14*					
Feedback quality		−0.07		−0.46*	−0.09	0.40*
Program coherence		0.06		−0.18*		0.41*
Knowledge variables						
Mathematical content		−0.11		−0.80*	−0.12*	0.19
Mathematical pedagogical content	−0.17*		−0.12	0.65*	0.13*	

Note: *POL* Poland, *RUS* Russian Federation, *U.S.* United States of America
* $p < .05$

Poland and the U.S. For FSTs, learning how to address the learning needs of diverse students (teaching for diversity) increased the difference in views among educators and future teachers in Poland and the Russian Federation. Finally, higher levels of program coherence were related to larger differences between future teachers and their educators concerning the notion that mathematics learning can be inquiry-based; this was true for FPTs in the U.S. and for FSTs in Poland, but this difference was reversed for the Russian Federation's FSTs (e.g., 1 SD increase in program coherence reversed the difference in beliefs about active learning, $-0.26 + 0.25 = -0.01$). MCK did not have a consistent effect on differences in this belief, but MPCK had significant effects in Poland on primary and secondary programs: Greater MPCK tended to reverse differences.

Mathematics as Fixed Ability In all cases except one (Poland's secondary programs), future teachers were more likely to believe that mathematics learning is a function of fixed ability than their teacher educators (see the positive intercepts in Table 6.13). The WLS models explained 21% (Russian Federation secondary) to 94% (Poland secondary) of the variation in the differences of belief that doing well in mathematics can be explained by a natural and fixed ability, with little consistency in the significance of OTL predictors in the models between country and teacher level. Class participation and instructional planning significantly reversed the difference in this belief between FPTs and their educators in the Russian Federation and FSTs and their educators in the U.S. Although the opposite effect was found for FSTs in Poland, opportunities to participate in discussions and teach a class (class participation) and to develop instructional plans (instructional planning) also reversed the difference, so that future teachers were relatively more likely than their educators to believe that mathematics learning is a function of fixed ability, a finding that is consistent with the average differences (given intercepts) of other countries and teacher levels.

To help clarify the overall effects in these models, consider the R^2 value of .94 in Table 6.13 for Poland's FSTs, regarding the belief that doing well in mathematics can be explained by a natural ability or a mathematical mind (mathematics as fixed ability). An R^2 equal to .94 tells us that the OTL variables explained nearly all of the variation in differences across institutions. On average, Poland's FSTs believe mathematics learning is a fixed ability just slightly less than their educators, whereas future teachers in most other countries had strong beliefs *supporting* the fixed ability notion relative to their educators (with essentially no difference in Singapore elementary programs). However, these differences in future teacher and educator beliefs about the extent to which mathematics learning can be the result of natural ability toward mathematics vary across institutions (in some institutions, the differences are very small, in others they are large). For Poland's secondary programs, the OTL variables explain much of this variation in differences. Another way of saying this, in regression-based language, is that differences in beliefs between future teachers and educators begin to look more similar (less variable) across institutions by accounting for variation in OTL and conditioning on MCK and MPCK as well.

Preparedness for Teaching In Poland, in both primary and secondary programs, future teachers were less likely to believe than their educators that they were well prepared for teaching; in contrast, in both program levels in the Russian Federation and the U.S., future teachers were more likely to believe than their educators that they were well prepared (see negative intercepts for Poland in Table 6.14). The WLS models produced R^2 values ranging from .22 (Poland primary) to .99 (Poland secondary) for differences in this belief. Programs offering OTL to teach diverse students, however, significantly increased differences in the perceptions of preparedness to teach among FPTs and FSTs in Poland. Learning in a coherent program significantly increased differences in the perceptions of preparedness among educators and their FSTs in Poland and the U.S.—that is, in more coherent programs, future teachers felt even less prepared in Poland and more prepared in the U.S. to teach relative to the views of their educators. In institutions where future teachers demonstrated higher levels of MCK in Poland primary and secondary programs, differ-

Table 6.14 WLS regression results for future primary and secondary teachers on preparedness for teaching

	Primary			Secondary		
Predictors	POL	RUS	U.S.	POL	RUS	U.S.
R^2	.22	.43	.62	.99	.61	.43
Intercept	−0.35	0.20	0.33	−0.18	0.17	0.11
OTL variables						
Foundations		0.11*		0.25*		
Instruction				0.19*	0.19*	0.21
Class participation			0.12*		0.10*	
Class reading	−0.08		−0.09			
Solving problems		−0.11			−0.11	
Instructional practice				−0.20*		
Instructional planning	0.20*			−0.18	0.17*	
Assessment uses		0.18*		0.95*		
Assessment practice			0.28*	−0.41*		
Social Science					−0.09	
Application						
Teach for diversity	−0.29*	−0.13*		−0.85*	−0.09	
Teach for reflection	0.23*					
Teach for improving	−0.20		0.10	0.77*		
Classroom learning to practice				−0.63*	0.15	
Reinforcement of goals		−0.16*		0.52*		
Feedback quality		0.11		0.29*	−0.12	0.23
Program coherence		0.14*		−0.34*		0.31*
Knowledge variables						
Mathematical content	−0.19*			−0.37*	−0.14*	0.24
Mathematical pedagogical content				0.46*		−0.23

Note: *POL* Poland, *RUS* Russian Federation, *U.S.* United States of America
*$p < .05$

ences in perceptions about preparedness were increased; whereas greater MPCK reversed the difference in Poland secondary programs. In the U.S., regarding differences in perceptions of preparedness (future teachers having more positive perceptions of preparedness than their educators), greater MCK increased these differences; however, greater MPCK reversed these differences.

Discussion

The TEDS-M conceptual framework included beliefs as an important outcome of teacher education and preparation. It is assumed that teachers' beliefs about teaching and learning mathematics influence their attitudes, dispositions, and practices, which in turn influence pupils' beliefs and behaviors. We investigated the extent to which differences in beliefs exist between future teachers and their educators in their teacher preparation programs in five countries. Because there is not a one-to-one correspondence between future teachers and educators (not all students have the same instructors, and not all instructors teach all students), we used the group of future teachers (separately for primary and secondary education programs) and the group of educators involved in teacher preparation at the institution as the relevant levels of comparison. We then looked across institutions within a country to see if the differences in beliefs varied. We found significant variation in future teacher and teacher educator differences for most beliefs in most countries. We also found that this variation can be explained significantly by OTL in various components of their education programs, including mathematics pedagogy. That is, the differences in beliefs between future teachers and educators are a function of the teacher education curriculum and experiences, and, in some cases, of future teachers' MCK and MPCK.

Interestingly, differences between future teachers and educators averaged across institutions within a country are remarkably similar across countries. That is, when, on average, future teachers are less likely to agree with their educators on a given belief in one country, we find similar patterns in the other countries, possibly indicating a universal norm. For example, regarding the belief that learning mathematics is a process of inquiry, FPTs and FSTs were less likely to agree with this belief (i.e., they had lower scores on this belief) than their educators, on average, across all five countries. This was also true for believing that mathematics teaching should involve active learning (future teachers were less likely to agree with this belief than their educators). In contrast, there were beliefs future teachers were more likely to agree with than their educators, including the beliefs that mathematics is a set of rules and procedures, that learning mathematics should be teacher-directed, and that mathematics ability is fixed (though with no real differences evident in Singapore's elementary and Poland's secondary programs). These are the trends we find in differences in beliefs on average across institutions; however, there is variation in these differences across institutions.

Finally, as expected, differences between future teachers and their educators on perceptions of being prepared for teaching varied across countries, but not between primary and secondary programs within a country. FPTs and FSTs perceived themselves to be less prepared for teaching than did their educators in Chinese Taipei and Poland (all significant differences). FPTs and FSTs perceived themselves to be more prepared for teaching than did their educators in the Russian Federation and the U.S. (all significant differences). These differences were not significant in Singapore, where, on average, future teachers and their educators perceived teacher candidates to be equally well prepared for teaching.

We also noted that in some cases, there was significant variation in these differences across institutions within a country. We observed significant variation across institutions preparing FPTs in Poland and the U.S. on the belief about mathematics learning as a process of inquiry; the same was true in institutions preparing FPTs in Poland and the Russian Federation with respect to the belief that mathematics learning consists of memorizing a set of rules and procedures. Regarding the belief that mathematics learning should be teacher-directed, significant variation in institutional differences from educators was found for FPTs in all five countries and for FSTs in Poland. Regarding mathematics as a fixed ability, only in the U.S. did we see significant variation across institutions. Finally, regarding being prepared for teaching, significant variation across institutions was seen in both the Russian Federation and the U.S.

Even though the variation across institutions may not have been at a statistically significant level, some variation was observed in all cases. We chose to use regression to evaluate the extent to which variation in differences in beliefs was a function of the OTL in education and mathematics pedagogy curriculum and experiences across institutions. The primary message from these regressions is that, for many countries, as much as one-half or more of the variation across institutions was accounted for by OTL. This indicates that the extent to which future teachers hold levels of beliefs different from their educators is a function of teacher education curriculum and experiences, controlling for MCK and MPCK. This suggests that future teacher beliefs may be malleable, but such an inference can only be suggested preliminarily. The design of the TEDS-M study was correlational and not causal. Unfortunately, a national longitudinal or experimental study is logistically and financially prohibitive to investigate this particular question in a direct and rigorous way. Nevertheless, the evidence presented here, with replication across multiple institutions and countries, provides relevant information regarding the association of common elements of teacher education curriculum and experiences with differences between future teacher and educator beliefs.

Looking globally across the 19 OTL measures, there are a few notable effects. First, OTL had a relatively consistent level of effects overall on institutional variation in beliefs differences. For instance, the OTL measures resulted in 19 (out of 108 estimated coefficients) significant effects explaining variation in differences for beliefs that mathematics learning should be teacher-directed and 23 significant effects explaining variation in differences for beliefs that mathematics skill is a function of a natural and fixed ability. We also see OTL resulting in 25 significant effects for differences in perceptions about being prepared for teaching.

Looking at individual OTL measures also indicates a few common effects. The measure of program coherence had the most number of significant effects on differences in beliefs between future teachers and their educators (11 significant effects), followed by OTL for mathematics education foundations, participating in class, instructional practice and assessment practice, mathematics instructional planning, teaching for diversity, and in-school experiences that promote connecting classroom learning to practice and the quality of the in-school supervising teacher (all with nine or ten significant effects). OTL areas that had less effect included OTL in the social sciences of education (one significant effect) and to solve problems in mathematics education courses (three significant effects). The other OTL areas had five to seven significant effects across belief areas, teacher levels, and countries (out of 36 possible coefficients).

Also note that we included MCK and MPCK in the regression equations, primarily to control for differences in mathematics knowledge, skills, and abilities. We found that these control variables were significant in 17 cases for MCK and 14 cases for MPCK—more often than any of the OTL variables. The most consistent effect in this regard was the effect of MCK on the differences in beliefs regarding mathematics learning as memorizing a set of rules and procedures (five significant effects out of six possible combinations of countries and teacher levels). When looking at differences in beliefs regarding mathematics as a set of rules and procedures, future teachers are more likely to hold this belief than their educators in all three countries (positive mean differences in Table 6.4). In each case, with the exception of FSTs in the U.S., in institutions with higher average MCK, these differences were significantly smaller—that is, in institutions where future teachers had higher levels of MCK, they were less likely to view mathematics learning as memorizing a set of rules and procedures relative to their educators (see the negative coefficients for MCK in Table 6.10).

There is one limitation worth mentioning at this point. Here we investigated differences in beliefs between future teachers and their educators. We did not attend to the overall level of beliefs (strong agreement versus weak agreement with beliefs). If program faculty are clear and consistent about their beliefs regarding the teaching and learning of mathematics, they may hope to impart those beliefs through the curriculum, instruction, and experiences provided to their future teachers. It is also likely that most programs have not articulated a consistent or coherent set of beliefs or that teacher preparation educators are intentional in imparting beliefs regarding mathematics teaching and learning. That was not directly evaluated here, and we note that there is substantial variability in beliefs within institutions. Our focus was on variation in the differences at the institutional level and across these institutions within countries. This allowed us to evaluate the extent to which these differences in beliefs were significant, whether they varied across institutions, and the extent to which that institutional variability was a function of curriculum and experiences. Thus, OTL, as expressed in the teacher education curriculum, the experiences in the program, and the experiences in the field, and teacher educators' own views about teaching and learning to teach, are the potential levers that teacher preparation programs can control.

References

Aguirre, J., & Speer, N. M. (1999). Examining the relationship between beliefs and goals in teacher practice. *The Journal of Mathematical Behavior, 18*(3), 327–356.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84*(5), 888–918.

American Mathematical Society. (2012). *The mathematical education of teachers II*. Washington, DC: Author. Retrieved from https://www.cbmsweb.org/archive/MET2/met2.pdf

Bernstein, B. (1999). Vertical and horizontal discourse: An essay. *British Journal of Sociology of Education, 20*(2), 157–173.

Borenstein, M. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 221–235). New York, NY: Russell Sage Foundation.

Burstein, L. (1993). Prologue: Studying learning, growth and instruction cross-nationally: Lessons learned about why and why not to engage in cross-national studies. In L. Burstein (Ed.), *The IEA study of mathematics III: Student growth and classroom processes*. New York, NY: Pergamon Press.

Calderhead, J. (1996). Teachers: Beliefs and knowledge. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 709–725). New York, NY: Macmillan.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255–296). New York, NY: Macmillan.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.

Deng, Z. (1995). *Estimating the reliability of the teacher questionnaire used in the Teacher Education and Learning to Teach (TELT) study* (National center for research on teacher learning technical series, 95-1). East Lansing, MI: Michigan State University. Retrieved from http://education.msu.edu/ncrtl/PDFs/NCRTL/TechnicalSeries/TS-95-1.pdf

Egan, O. (1986). The concept of belief in cognitive theory. In L. P. Moss (Ed.), *Annals of theoretical psychology* (Vol. 4, pp. 315–350). New York, NY: Plenum Press.

Floden, R. (2002). The measurement of opportunity to learn. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 231–266). Washington, DC: National Academy Press.

Grigutsch, S., Raatz, U., & Törner, G. (1998). Einstellungen gegenüber Mathematik bei Mathematiklehrern [Mathematics teachers' epistemological beliefs about the nature of mathematics]. *Journal für Mathematik-Didaktik (Journal of Mathematics Education), 19*, 3–45.

Hsieh, F. J., Law, C. K., & Shy, H. Y. (2011). Mathematics teacher education quality in TEDS-M: Globalizing the views of future teachers and teacher educators. *Journal of Teacher Education, 62*(2), 172–187.

Husen, T. (Ed.). (1967). *International study of achievement in mathematics: A comparison of twelve countries* (Vol. 1). New York: Wiley

IEA. (2015). IDB analyzer 3.2. [software]. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement. Retrieved from http://www.iea.nl/eula.html

Ingvarson, L., Beavis, A., & Kleinhenz, E. (2007). Factors affecting the impact of teacher education courses on teacher preparedness: Implications for accreditation policy. *European Journal of Teacher Education, 30*(4), 351–381.

Kane, R., Sandretto, S., & Heath, C. (2002). Telling half the story: A critical review of research on the teaching beliefs and practices of university academics. *Review of Educational Research, 72*(2), 177–228.

Lunenberg, M., Korthagen, F., & Swennen, A. (2007). The teacher educator as a role model. *Teaching and Teacher Education, 23*(5), 586–601.

Muis, K. (2004). Personal epistemology and mathematics: A critical review and synthesis of research. *Review of Educational Research, 74*(3), 317–377.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2007). *TIMSS 2007 assessment frameworks*. Chestnut Hill, MA: Boston College.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*, 307–332.

R Core Team. (2015). *R: A language and environment for statistical computing software*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Shavelson, R. J., & Stem, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research, 51*, 455–498.

Tatto, M. T. (1996). Examining values and beliefs about teaching diverse students: Understanding the challenges for teacher education. *Educational Evaluation and Policy Analysis, 18*, 155–180.

Tatto, M. T. (1998). The influence of teacher education on teachers' beliefs about purposes of education, roles and practice. *Journal of Teacher Education, 49*, 66–77.

Tatto, M. T. (1999). The socializing influence of normative cohesive teacher education on teachers' beliefs about instructional choice. *Teachers and Teaching, 5*, 111–134.

Tatto, M.T. (Ed.). (2013). The Teacher Education and Development Study in Mathematics (TEDS-M). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement. Retrieved from http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/TEDS-M_technical_report.pdf

Tatto, M. T., & Coupland, D. (2003). Teaching and measuring attitudes in teacher education. In J. Raths & A. McAninch (Eds.), *Teacher beliefs and classroom performance: The impact of teacher education (Advances in teacher education)* (Vol. 6, pp. 123–181). Greenwich, CT: Information Age Publishing.

Tatto, M.T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, Michigan State University. Retrieved from https://msu.edu/user/mttatto/documents/TEDS_FrameworkFinal.pdf

Tatto, M.T., Schwille, J., Senk, S.L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M. & Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).* Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement. Retrieved from http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/TEDS-M_International_Report.pdf

Part II
Exploring Future Teacher Characteristics, Knowledge, Beliefs and Opportunities to Learn

Chapter 7
Introduction: Exploring Future Teacher Characteristics, Knowledge, Beliefs and Opportunities to Learn in the TEDS-M Study

Maria Teresa Tatto and Wendy M. Smith

Abstract This chapter provides an introduction to Part II of the book, which focuses on the future teachers that participated in the Teacher Education and Development Study in Mathematics (TEDS-M), a cross-national study of teacher education programs that prepare future primary and secondary mathematics teachers. The study collected data from 13,871 future primary teachers and 8,207 future secondary teachers through a novice teacher questionnaire (NTQ), and from assessments of mathematics knowledge (MCK) and mathematics pedagogical content knowledge (MPCK). The chapter includes a brief overview of the characteristics of the future teachers, the methods of data collection, and the challenges encountered when administering the questionnaire and the assessments. The data on background

The text in this Part II Introduction contains shortened and slightly edited versions of text that has appeared in the following publications: Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M., & Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).* Amsterdam: the Netherlands. International Association for the Evaluation of Student Achievement (IEA). Brese, F., & Tatto, M.T. (Eds.) (2012). *User guide for the TEDS-M international database.* Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA). Tatto, M.T. (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report.* Amsterdam, the Netherlands: International Association for the Evaluation of Student Achievement (IEA). Text cited directly or indirectly from those sources will not be made recognizable. An extensive report on the descriptive findings on the characteristics of future teachers in the study can be found in Tatto et al. (2012) and Tatto (2013). We summarize the key findings here and reproduce some figures and tables to orient the reader; the chapters in Part II however are original contributions written exclusively for this book.

M. T. Tatto (✉)
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: teresa.tatto@asu.edu

W. M. Smith
University of Nebraska, Lincoln, NE, USA
e-mail: wsmith5@unl.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_7

characteristics includes age, gender, previous career, highest level of education reached before teacher education, and highest level of mathematics reached before teacher education. The assessments measured future teachers' abilities to demonstrate key learning outcomes of teacher education programs, including their knowledge (mathematical content knowledge and mathematical pedagogical content knowledge) and beliefs (about the nature of mathematics, the nature of teaching mathematics, the nature of learning mathematics, self-efficacy, and preparedness to teach). Standing between the characteristics of future teachers and their learning outcomes are the characteristics of teacher education programs, which include the opportunities to learn that they afford future teachers. The chapters explore how key opportunities to learn offered by teacher education programs influence future teachers' knowledge and beliefs. The first six chapters focus on future primary teachers, whereas the last two chapters include the study of future secondary teachers.

Future Teachers

In Part II of the book, authors focus on the future teachers themselves. The study of future teachers included the exploration of their background characteristics such as age, gender, previous career, highest level of education reached before teacher education, and highest level of mathematics reached before teacher education. In addition, the TEDS-M team studied the future teachers' abilities to demonstrate the key learning outcomes of teacher education programs, including their knowledge (mathematical content knowledge and mathematical pedagogical content knowledge) and beliefs (about the nature of mathematics, the nature of teaching mathematics, the nature of learning mathematics, self-efficacy, and preparedness to teach). Standing between the characteristics of future teachers and their learning outcomes are the characteristics of teacher education programs, which include the opportunities to learn that they afford future teachers.

Through the chapters in Part II of the book, various authors explore how teacher education influences future teachers' knowledge and beliefs. In the first six chapters, authors focus on future primary teachers, whereas in the last two chapters, authors address future secondary teachers.

In this introduction, we provide a brief overview of the characteristics of the future teachers, the methods of data collection, and the challenges encountered when collecting data.

Summary of Future Teacher Characteristics

Future teachers were defined as "students enrolled in teacher education programs designed to prepare them to teach mathematics at the primary or secondary levels" (Tatto et al., 2012, p. 116). In total, data were obtained from 13,871 future primary teachers and 8,207 future secondary teachers. Data about the future teachers were collected through both a novice teacher questionnaire (NTQ), which covered background characteristics, beliefs, and opportunities to learn (OTL), and assessments of mathematics knowledge (MCK) and mathematics pedagogical content knowledge (MPCK).

The study team found that, overall, future teachers tended to be young. Across programs, the mean age of the future teachers in their last year of teacher preparation ranged from about 21–29 years. Future secondary teachers tended to be slightly older, on average, than primary future teachers. Teachers at both levels tended to be female across countries (Tatto et al., 2012).

As a proxy for prior achievement, the NTQ included questions for future teachers to rate the usual level of marks or grades received in secondary school. It was not possible for TEDS-M to collect pre/post data on future teachers due to the large variance in when and how future teachers enter their teacher preparation programs. Because programs varied in duration from 1 to 5 years, and teachers sometimes changed courses of study in college; no common baseline data could be collected on an international scale. Overall, most future teachers reported being "usually near the top of my year level," or "generally above average for my year level" before entering the teacher education program. However, in some countries and more often at the primary levels, the more common answer was a step below. Across countries, it was consistently found that the teachers trained for teaching at higher levels had higher prior achievement levels (Tatto et al., 2012). Though these are self-reports, it is important to note that they are a strong predictor of both MCK and MPCK (e.g., Qian & Youngs, 2015).

The NTQ included questions for future teachers about their socio-economic status, using the following indicators: number of books in the homes of the students' parents or guardians, the availability of a variety of educational resources in those homes, and the highest level of education completed by their male and female parents or guardians, all commonly used proxies for socio-economic status in the literature.

Researchers show that family education level varied considerably by country, with some countries (typically those less wealthy) having many future teachers report primary school as their parents' highest level of education attainment, and other countries having more than 20% of future teachers report that their parents attained advanced degrees (Tatto et al., 2012).

The NTQ also included questions for future teachers on whether the language they spoke at home was the same as the language of the questionnaire, and whether the future teachers were natural citizens or immigrants in the country in which they were becoming a teacher. Although most future teachers in most countries reported speaking the language of the assessment at home, significant proportions in some countries[1] reported that they sometimes or never spoke the language of the assessment at home (Tatto et al., 2012, p. 122).

[1] These included including Botswana (90%), Chinese Taipei (about 30%), Malaysia (about 87%), Oman (about 28%), the Philippines (about 95%), Singapore (about 43%), and Thailand (about 39%).

Future teachers were also asked about previous careers and their commitment to being a teacher in the future. The proportion of future teachers reporting a prior career varied across countries, but it tended to be less than one-third of respondents. Regarding the reasons for choosing teaching, the most common response across program-types that teachers gave was "like working with young people." Other reasons varied across program-types; for example, more secondary than primary teachers selected "I love mathematics." The least frequently chosen responses were "good student in school," "availability of teaching positions," and "I am attracted by teacher salaries" (Tatto et al., 2012, pp. 122–126).

Summary of Future Teacher Beliefs and Opportunities to Learn

Beliefs held by teachers and students are an important influence on teaching and learning (Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Peterson, Fennema, Carpenter, & Loef 1989; Wilkins, 2008). However, there is little conclusive evidence that beliefs can be effectively influenced by teacher preparation (Tatto, 1999; Tatto & Coupland, 2003).

For this reason, the TEDS-M study team collected data about three aspects of future teachers' mathematics-related beliefs to allow researchers to investigate the assocations between teacher education and the beliefs of future teachers:

1. beliefs about the nature of mathematics;
2. beliefs about learning mathematics; and
3. beliefs about mathematics achievement.

The development of the TEDS-M questionnaire scales was informed in part by work done in the participating countries and by the Teaching and Learning to Teach Study at Michigan State University (Deng, 1995; Ingvarson, Beavis, Danielson, Ellis, & Elliott, 2005; Ingvarson, Beavis, & Kleinhenz 2007; Tatto 1996, 1998, 2003); the questionnaire ultimately contained five belief scales covering the three areas above.

Among the future teachers who answered the questionnaires, items expressing beliefs most consistent with cognitive-constructivist views of mathematics learning (e.g., mathematics is a process of inquiry; learning mathematics requires active involvement) attracted much greater support than the items expressing beliefs most consistent with the procedural-rules-guided views of mathematics learning (e.g., mathematics is a set of rules and procedures; learning mathematics requires following teacher direction). This pattern was common across countries, but not universal. The latter two beliefs were more prevalent than the former two in Georgia (the country where the range of beliefs across programs was also greatest), the Philippines, Malaysia, and, to some extent, in Botswana and Thailand. Differences between patterns of response for the future primary teachers and for the future lower-secondary teachers were relatively small.

Summary of Future Teachers' Knowledge Outcomes: MCK and MPCK

Studying the knowledge that future teachers have at the end of their formal teacher education is important for two main reasons. First, teachers' knowledge influences the mathematics achievement of their students (Baumert et al., 2010; Hill, Rowan, & Ball, 2005). Second, studying knowledge as a key outcome of teacher education in combination with rich information about the teacher education experience and important control variables can shed light on the influence that teacher education may have had on future teachers' knowledge.

Knowledge for teaching requires both *content knowledge* and *pedagogical content knowledge* (National Research Council, 2010; Shulman, 1987). The TEDS-M research team drew on prior work in this area to design the instruments used to measure the mathematics content knowledge (MCK) and the mathematics pedagogical content knowledge (MPCK) of future teachers at the end of their programs. The subdomains of each are presented below in Tables 7.1, 7.2, and 7.3.

For a more detailed description of the development of the MCK and MPCK assessments, see Chapter 3 of the conceptual framework (Tatto, Schwille, Senk, Ingvarson, Peck, & Rowley, 2008), the TEDS-M international report (Tatto et al., 2012), and the TEDS-M technical report (Tatto, 2013).

The tables containing the results of the MCK and MPCK assessment and the answers to beliefs and opportunities to learn questions are extensive. We advise the reader to consult the *TEDS-M International Report* (Tatto et al., 2012), and

Table 7.1 Mathematics content knowledge framework, by content subdomain

Subdomain	Sample topics
Number and operations	Whole numbers, fractions, and decimals Number sentences Patterns and relationships Integers Ratios, proportions, and percentages Irrational numbers Number theory
Geometry and measurement	Geometric shapes Geometric measurement Location and movement
Algebra and functions	Patterns Algebraic expressions Equations/formulas and functions Calculus and analysis[a] Linear algebra and abstract algebra[a]
Data and chance	Data organization and representation Data reading and interpretation Chance

Source. Tatto et al., (2012, p. 130)
[a]Lower-secondary level only

Table 7.2 Mathematics content knowledge framework, by cognitive domain

Subdomain	Sample behaviors
Knowing	Recall, recognize, compute, retrieve, measure, classify/order
Applying	Select, represent, model, implement, solve routine problems
Reasoning	Analyze, generalize, synthesize/integrate, justify, solve non-routine problems

Source. Tatto et al., (2012, p. 130)

Table 7.3 Mathematics pedagogical content knowledge (MPCK) framework

Subdomain	Sample topics
Mathematics curricular knowledge	Knowing the school mathematics curriculum Establishing appropriate learning goals Identifying key ideas in learning programs Selecting possible pathways and seeing connections within the curriculum Knowing different assessment formats and purposes
Knowledge of planning for mathematics teaching and learning	Selecting appropriate activities Predicting typical students' responses, including misconceptions Planning appropriate methods for representing mathematical ideas Linking didactical methods and instructional designs Identifying different approaches for solving mathematical problems Choosing assessment formats and items
Enacting mathematics for teaching and learning	Explaining or representing mathematical concepts or procedures Generating fruitful questions Diagnosing students' responses, including misconceptions Analyzing or evaluating students' mathematical solutions or arguments Analyzing the content of students' questions Responding to unexpected mathematical issues Providing appropriate feedback

Source. Tatto et al., (2012, p. 131)

specifically Chapter 5, "The Mathematics Content Knowledge and Mathematics Pedagogical Content Knowledge of Future Primary and Lower Secondary Teachers" (pp. 129–151), Chapter 6, "Beliefs about Mathematics and Mathematics Learning" (pp. 153–173), and Chapter 7, "Opportunity to Learn" (pp. 175–197).

A final observation should be made about the quality of the data collected and used for this book. The data collection and the study as a whole were done following the strict guidelines that the IEA (International Association for the Evaluation of Educational Achievement) sets for all its studies. The field trial, which was intended to be a test of all procedures and instruments with a reduced sample, revealed the challenging nature of the study, with issues ranging from the inaccuracy of future teachers' data as kept by programs (with implications for sampling and response rates) to refusals to answer, to problems of measurement fit for some items. The knowledge acquired in the field trial supported a strong and rigorous study as the TEDS-M team developed methods and procedures to find the target population, obtained the IEA response rates required for reporting (85–100%), and fine-tuned study instruments to secure measures with validity evidence and reliable scores. Thus, the surveys were completed with high response rates and coverage of the target populations, in most cases meeting the very high IEA standards for sampling and response rates. There were, however, limited instances in which the IEA standards were not met, yet the response rates still compared favorably with general experience in higher education surveys, especially in those cases in which the targeted participants are all volunteers. TEDS-M thus lays the foundation for future rigorous cross-national research in teacher education, and allows for the exploration of the TEDS-M database to conduct secondary analysis.

Overview of Chapters in This Section

This section features eight chapters exploring the association between individual and program characteristics and teacher education outcomes. As the main TEDS-M findings are reported elsewhere (Tatto et al., 2012; Ingvarson et al., 2013), these chapters take a closer look at the associations among variables of interest.

One of the groundbreaking aspects of TEDS-M was the information collected about future teachers' opportunities to learn a variety of mathematical and pedagogical skills. The chapters of this section focus on subsets of these measures of OTL, and the associations with other future teacher outcomes. We recognize that future teachers' self-reported OTL may not align completely with the OTL teacher educators believe their programs offer, but they have been shown to be fairly well aligned with future teachers' knowledge and skills. The OTL items on the NTQ encompass both mathematical content and pedagogical strategies. Authors explore these questions across the next eight chapters.

First, in Chap. 8, "The Mathematical Education of Primary Teachers," and later in Chap. 14, "The Mathematical Education of Secondary Teachers," Tatto uses multivariate analyses to explore the association between future teachers' individual characteristics (e.g., levels of achievement in previous schooling), program characteristics (e.g., programs' selection policies and opportunities to learn the content and the pedagogy of the mathematics school curriculum), future teachers' beliefs, and future teachers' knowledge of mathematics and mathematics pedagogy. Results support the use of teacher education policies directed at raising the level of subject knowledge required for program selection/graduation and increasing the level of cognitive demand of the mathematics and mathematics pedagogy opportunities to learn offered to future primary and secondary mathematics teachers.

In Chap. 9, “How Primary Future Teachers’ Knowledge Is Shaped by Teacher Preparation,” Qian and Youngs take a closer look at the associations between elementary future teachers’ opportunities to learn in mathematics courses and mathematics methods courses and their mathematics content knowledge and mathematics pedagogical content knowledge in Chinese Taipei, Singapore, and the United States. The authors found evidence that elementary future teachers’ knowledge is affected by the content of mathematics courses taken and by the number of topics addressed in mathematics methods courses.

In Chap. 10, “Opportunities to Learn Mathematics Pedagogy and Connect Classroom Learning to Practice: A Study of Future Teachers in the United States and Singapore,” Kutaka, Smith, and Males explore future mathematics specialists’ opportunities to learn how to teach mathematics using latent class analysis to differentiate among groups of prospective mathematics specialists with potentially different opportunities to learn mathematics pedagogy within the United States and Singapore. After identifying different learning and practicum experiences in university and in the field experience for each country, the authors discuss implications for teacher-preparation programs.

In Chap. 11, “Preparing Primary Mathematics Teachers to Learn to Work with Students from Diverse Backgrounds,” and Chap. 15, “An International Study of the Relationship between Learning to Teach Students from Diverse Backgrounds and Mathematical Knowledge for Teaching in Future Secondary Mathematics Teachers,” Dyer analyzes the associations between future primary and secondary teachers’ opportunities to learn to teach students from diverse backgrounds, and their mathematical knowledge for teaching. Using multilevel modeling, Dyer creates separate models for primary and secondary future teachers. In Chap. 11, Dyer found that primary mathematics specialist teachers with more opportunities to learn to teach students from diverse backgrounds had lower levels of mathematical knowledge for teaching (as measured by MCK and MPCK). Primary generalist teachers do not consistently show the same results across all countries, with some showing higher and others showing lower levels of mathematical knowledge for teaching associated with greater opportunities to learn to teach students from diverse backgrounds. The results suggest teachers who are better prepared for the mathematical aspects of teaching tend to be less prepared for addressing the needs of diverse learners. In the Chapter focusing on secondary teachers, Dyer found a negative association within teacher preparation programs; teachers with more opportunities to address the learning needs of students from diverse backgrounds have lower levels of mathematical knowledge for teaching. These results suggest that teachers with tools for addressing the learning needs of students from diverse backgrounds may lack adequate mathematical preparation.

Chapter 12, “Differences in Beliefs and Knowledge for Teaching Mathematics: An International Study of Future Teachers,” by Kutaka, Smith, and Albano, uses multi-level modeling to examine the associations between mathematics content knowledge (MCK) and prospective primary teachers’ beliefs about the nature of mathematics and about learning mathematics, across 15 countries. A series of multilevel models were fit to four program groups (lower-primary, primary, primary/

secondary, and primary mathematics specialists) with future teachers nested within institutions. Future teachers in the same country preparing to teach at different grade levels do not endorse the same kinds of beliefs across programs, nor did they observe consistent patterns of associations between belief types and MCK within programs. Across programs, endorsing the belief mathematics is a process of inquiry was associated with higher MCK, whereas endorsing the belief mathematics is a set of rules and procedures was associated with lower MCK. There was a less consistent pattern of associations between belief in mathematics as a fixed ability and MCK. The chapter concludes with possible explanations for programmatic differences between and within countries, grounded in a discussion of program features such as entry requirements, program types and credentials, as well as school curriculum organization and content.

Chapter 13, "Future Teachers' and Teacher Educators' Perceptions of Learning Mathematics Instruction and Relationships to Knowledge," by Ayieko, examines associations between opportunities to learn mathematics instruction for conceptual understanding and elementary future teachers' knowledge for teaching mathematics in three countries: Poland, Russia, and the United States. The perception of the opportunities to learn mathematics instruction among future teachers did not match those of their teacher educators. A comparison of the teacher educators' and future teachers' responses suggest that the future teachers in the three countries had fewer opportunities to learn mathematics instruction for conceptual understanding than was intended by the teacher educators. The pattern of associations from a multilevel regression analysis in each of the selected countries show variations across contexts and categories of knowledge. In particular, opportunities to learn how to make distinctions between procedural and conceptual knowledge, and how to show why a procedure works, were significantly related to future teachers' knowledge for teaching mathematics between programs in the different countries.

References

Baumert, J. R., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C.-P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal, 26*(4), 499–531. https://doi.org/10.3102/00028312026004499

Deng, Z. (1995). Estimating the reliability of the teacher questionnaire used in the Teacher Education and Learning to Teach (TELT). *National Center for Research on Teacher Learning Technical Series, 95*(1). Retrieved from http://ncrtl.msu.edu/

Hill, H., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*, 371–406.

Ingvarson, L., Beavis, A., Danielson, C., Ellis, L., & Elliott, A. (2005). *An evaluation of the bachelor of learning management at central queensland university*. Canberra, Australia: Australian Government Department of Education, Science and Technology.

Ingvarson, L., Beavis, A., & Kleinhenz, E. (2007). Factors affecting the impact of teacher education courses on teacher preparedness: Implications for accreditation policy. *European Journal of Teacher Education, 30*(4), 351–381.

Ingvarson, L., Schwille, J., Tatto, M.T., Senk, S.L., Bankov, K., Rodriguez, M., Reckase, M., Rowley, G., & Peck, R. (2013). An analysis of teacher education context, structure, and quality-assurance arrangements in TEDS-M countries. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement.

National Research Council. (2010). *Preparing teachers: Building evidence for sound policy*. Committee on the study of teacher preparation programs in the United States, Center for Education. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. Retrieved from https://www.nap.edu/read/12882/chapter/1

Peterson, P. L., Fennema, E., Carpenter, T. P., & Loef, M. (1989). Teacher's pedagogical content beliefs in mathematics. *Cognition and Instruction, 6*(1), 1–40. https://doi.org/10.1207/s1532690xci0601_1

Qian, H., & Youngs, P. (2015). The effect of teacher education programs on future elementary mathematics teachers' knowledge: A five-country analysis using TEDS-M data. *Journal of Mathematics Teacher Education., 19*, 371. https://doi.org/10.1007/s10857-014-9297-0

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Tatto, M. T. (1996). Examining values and beliefs about teaching diverse students: Understanding the challenges for teacher education. *Educational Evaluation and Policy Analysis, 18*(2), 155–180.

Tatto, M. T. (1998). The influence of teacher education on teachers' beliefs about purposes of education, roles, and practice. *Journal of Teacher Education, 49*(1), 66–77.

Tatto, M. T. (1999). The socializing influence of normative cohesive teacher education on teachers' beliefs about instructional choice. *Teachers and Teaching, 5*(1), 111–134.

Tatto, M. T. (2003, April). *Evaluating the effectiveness of the teacher preparation program at Michigan State University: Analyzing survey and ethnographic evidence*. Presentation to the annual meeting of the American Educational Research Association, Chicago, IL.

Tatto, M. T. (Ed.). (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M). Policy, Practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report.* Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Tatto, M. T., & Coupland, D. (2003). Teaching and measuring attitudes in teacher education. In J. Raths & A. McAninch (Eds.), *Teacher beliefs and classroom performance: The impact of teacher education (Advances in Teacher Education)* (Vol. 6, pp. 123–181). Greenwich: Information Age Publishing.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, Michigan State University, and IEA.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).* Amsterdam, The Netherlands: IEA.

Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education, 11*(2), 139–164.

Chapter 8
The Mathematical Education of Primary Teachers

Maria Teresa Tatto

Abstract This chapter reports the results of a cross-national study designed to examine the mathematics knowledge and the mathematical pedagogical content knowledge attained by prospective primary teachers at the end of their formal preparation and before they begin to teach. The study used survey methods to collect data from nationally representative samples of pre-service university-based teacher education programs and their future teachers in Botswana, Chile, Chinese Taipei, Germany, Malaysia, the Philippines, Poland, Russia, Singapore, Spain, Switzerland, Thailand, and the United States. Descriptive and multivariate analyses show that future teachers' individual characteristics, such as levels of achievement in previous schooling, programs' selection policies, and opportunities to learn the content and the pedagogy of the mathematics school curriculum, were associated with higher levels of knowledge and dispositions toward teaching and learning mathematics. Results support teacher education policies directed at (a) raising the level of subject knowledge required for program selection and graduation and (b) increasing the level of complexity and cognitive demand of the opportunities to learn mathematics and mathematics pedagogy offered to future primary mathematics teachers.

Introduction

Teachers' knowledge of mathematics alone does not guarantee high-quality teaching. However, this knowledge is seen as a prerequisite to teach mathematics effectively (Akiba, LeTendre, & Scribner, 2007; National Science Board, 2004; Staub &

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

M. T. Tatto (✉)
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: teresa.tatto@asu.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_8

Stern, 2002; Van Dooren, Verschaffel, & Onghena, 2002). For years, scholars have argued that other kinds of knowledge—of teaching, of learning, of students and their context, of the curriculum—are just as important (see Ball, 1991; Ball & Bass, 2000; Hill & Ball, 2009; Shulman, 1987). While there is disagreement concerning the balance of what and how much future teachers need to know to be successful, there is agreement that teachers are a key element in improving the overall level of mathematical proficiency among growingly diverse societies, and that good teacher preparation is essential (Conference Board of the Mathematical Sciences [CBMS], 2012; NRC, 2010). Accordingly, much emphasis has been placed on understanding what knowledge counts for effective teacher preparation, a question that has been the focus of sustained research across the globe for the last 17 years (see Ball, 2003; CBMS, 2012; Kilpatrick & Swafford, 2002; National Commission on Mathematics and Science Teaching for the 21st Century, 2000).

As curricula become more complex and demanding, and as nations across the world call for access to high-quality teachers for all children (UNESCO, 2014, 2016), it is important to have a better understanding of how this complex balance of what teachers need to know is managed in settings where teachers are known to be knowledgeable and effective. Cross-national studies are particularly useful in providing a comparative empirical basis to assess the effectiveness of different approaches to teacher education.

This chapter contributes to the literature on the effectiveness of mathematics teacher education by exploring the degree to which future teachers' individual characteristics and program policies and opportunities to learn (OTLs) contribute to the development of knowledgeable mathematics primary teachers.[1] This study uses data from the Teacher Education and Development Study in Mathematics (TEDS-M), the first large-scale cross-national study collaboratively designed and implemented by mathematicians, mathematics teacher educators, and local researchers in the participating countries.

Studying the Effectiveness of Pre-service Teacher Education Programs

The late 1990s saw an increased urgency in universal calls to improve learning in schools and to focus on teachers as part of both the problem and the solution (see the Delors Report, 1996, and the World Education Report on Teachers and Teaching, 1998, both published by UNESCO). Increased interest in teachers resulted in studies examining teaching and learning in classrooms, especially in countries where pupils had high scores in international mathematics and science tests (Hiebert et al.,

[1] 'Note that in the U.S. 'primary' usually refers to grade K-3, while 'elementary' is used for grades K-5 or K-6. In this chapter the term elementary is used as was used in TEDS-M (see Tatto et al., 2012 pp. 29–32 for specific definitions within countries as to what grades are included as primary or secondary).'

2003; Stigler & Hiebert, 1997; Stigler, Gallimore, & Hiebert, 2000). These studies helped the field to move beyond the examination of teacher characteristics, to searching for evidence related to teachers' cognition as an explanation for good teaching, concluding that effective teachers seem to possess a special kind of *teaching knowledge,* or what education experts in Europe call *didactique* (Develay, 1998) and elsewhere *pedagogical content knowledge* (Shulman, 1987). Such knowledge can theoretically be learned through deliberately organized experiences such as pre-service teacher education (for calls to expand this research to reflect the practice of teaching, see Grossman & McDonald, 2008). More recent developments, supported by alternative theories characterized by an emphasis on school-based learning, have resulted in the creation of alternative routes into teaching.

Policy-relevant research on the most effective approaches to becoming a teacher has developed slowly, in contrast with the rapidity of policy-driven change. Yet a number of studies have contributed important insights calling for the need to consider teacher education as one of the most important policy levers for increasing teaching quality (in the United States, see Cochran-Smith & Zeichner, 2015; Darling-Hammond, 2006; Mewborn & Stinson, 2007; in Singapore, Ginsburg, Anstrom, & Pollock, 2005; in Sri Lanka, Tatto, Nielsen, Cummings, Kularatna & Dharmadasa, 1993; Tatto & Kularatna 1993; and in the field of comparative studies, Comiti & Ball, 1996; Tatto, 2008, 2017).

Research from the mathematics education community has provided strong evidence that teacher education has an important influence on teachers' knowledge and effective practice (in the United States, see Ball, 2003; Even & Ball, 2009; Even & Tirosh, 2002; Hill, 2007; in Mexico, Luschei, 2011; Santibañez, 2002; Tatto 1999a, 1999b; in other countries, Baumert et al., 2010; Schmidt, Blomeke, & Tatto, 2011).

In spite of these advances, the field lacks examples of authoritative self-evaluation studies of teacher education programs. The TEDS-M study is the first large-scale cross-national study of the outcomes of teacher education, and it provides the most rigorous publicly available database to date to explore the outcomes of mathematics pre-service teacher education.

Research Questions

Using the TEDS-M database, this chapter explores the following questions about future primary teachers at the end of their pre-service university-based teacher education:

1. What is the level and depth of the mathematical and mathematical pedagogical content knowledge future teachers attain? Is this knowledge similar or different across countries?
2. How are specific characteristics of future teachers (such as socioeconomic status (SES), age, gender, prior attainment, beliefs) associated with their attained levels of knowledge?

3. What are some of the key learning opportunities available to future teachers in their teacher education programs, and how are these associated with their attained levels of knowledge?

Previous Research

This section includes an abridged review of studies that inform and support the research questions that guide this study. These studies have to do with research on program characteristics such as structure, selectivity, and content, including OTLs; research on future teachers' characteristics such as background (SES, age, gender, and prior attainment) and beliefs; and research on mathematical and mathematical pedagogical knowledge needed for teaching.

Program Characteristics

Teacher education programs are usually described by their general or structural characteristics such as duration; route, which can be either traditional (typically linked with higher education institutions such as universities) or alternative (ranging widely and having a strong practical school-based component); and credentials granted. Less common are program descriptions based on their specific or content characteristics such as their curriculum, the opportunities to learn they make available to future teachers (type and content of courses and school experiences they offer), and whether these curricular offerings and experiences are internally coherent (see Chapter 4 of this book and Tatto & Hordern, 2017 for examples using the TEDS-M syllabus analysis data).

Education scholars such as Darling-Hammond (2006), Kennedy (2016), and Zeichner and Conklin (2005) have pointed out that it is precisely the latter program features that tend to have more impact on future teachers' knowledge. However, structural and content factors interact in complex ways.

A review of research by Coggshall, Bivona, and Reschly (2012, p. 6) found that research studies (such as those by Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Boyd et al., 2008; and National Center for Analysis of Longitudinal Data in Education Research, 2012) suggest that indeed program content (i.e., what is taught) is important for program effectiveness, but so are other aspects of programs, such as program selectivity and program structure, which may be seen as determining the types of OTLs provided (i.e., five-year programs make it possible for candidates to have access to one full year of clinical experiences and obtain a graduate credential).

The importance of program selectivity and OTLs is supported in the United States by studies such as the NYC Pathways (Boyd et al., 2009, 2008), and the Mathematica study (Constantine et al., 2009). Both studies found that the most meaningful features of program effectiveness are whether the program is selective and the depth and breadth of the OTLs they provide.

The research of Schmidt, Blomeke and Tatto (2011), and Blomeke, Suhl, Kaiser, and Dohrmann, (2012) on programs' OTL for future teachers supports the relationship of certain key characteristics of program structure with both OTL (such as subject and pedagogy content and professional preparation) and teachers' knowledge. These two studies conclude that key program structure features that influence OTL have to do with a program's selectivity, the target grades for which teachers are being prepared, whether teachers are prepared as generalists or specialists, and whether the program has a field experience component.

Opportunities to Learn

An OTL is an experience with an anticipated or intended learning outcome. OTLs have been seen as central to explaining mathematics learning since the concept was introduced in the late 1960s (Husen, 1967) and explained as that pupils who were not exposed to mathematical concepts and experiences in school would find it difficult to demonstrate such knowledge in achievement tests. According to McDonnell (1995), the concept of OTL "has changed how researchers, educators, and policymakers think about the determinants of student learning" (p. 317). More recently, OTL has been used to explore determinants of teacher learning (Blomeke et al., 2012; Tatto et al., 2012).

Indeed, teacher education programs' OTLs can be considered one of the most important program characteristics that can contribute to future primary teachers' knowledge and skills. Based on an extensive review of the literature, Floden (2002) explained that inclusion of OTL in education studies serves a number of purposes: as an explanation of differences in levels of knowledge; as an indicator of curricular variation among entities; as an aspect of fairness (e.g., appropriateness of language of test items); and as a representation of the diversity of content, both overall and for distinct groups. In spite of the recognized need for more research on the influence of the teacher education curriculum as expressed, for instance in the variety of courses offered by teacher education programs (Wayne & Youngs, 2003; Tatto & Hordern, 2017), these studies are rare.

The degree to which these OTLs are consistent across courses and experiences offered to future teachers, and whether there are explicit standards with expectations for what future teachers should learn from their respective programs, served to define a program's *coherence*. Thus, program effectiveness depends on the degree of curriculum coherence, as more coherent programs are better able to influence future teachers' cognition about teaching and learning mathematics (Tatto, 1996, 1998, 1999a, 1999b).

Future Teacher Characteristics

Background Much research in the past has focused on the characteristics of teachers as a way to explain good teaching (Greenwald, Hedges, & Laine, 1996); yet teachers' characteristics, while important moderating factors, may not be the best predictors of teaching effectiveness. Wayne and Youngs' (2003) examination of the literature spanning more than 25 years (from 1975–2002) yielded 21 studies that failed to establish a definitive link between teacher characteristics (defined as ratings of teachers' colleges, teachers' test scores, teachers' degrees and coursework, and teachers' certification status) and effectiveness. While some studies find correlations between mathematics content knowledge and attributes such as SES, gender, and age, more informative studies found significant relationships between measures of ability (such as previous school performance or test scores) and effective teaching (Darling-Hammond, 2000; Monk, 1994; Mullens, Murnane, & Willett, 1996).

Importantly, Henry, Bastian, and Smith (2012) found that individual academic attainment prior to entering a teacher education program was correlated with teaching effectiveness; thus, while teacher background characteristics are important indicators of factors that may moderate program effectiveness, measures of knowledge and ability seem to better account for program outcomes (Ball, 1990a, 1990b).

Beliefs Teachers' beliefs have been considered an essential component of what makes a teacher effective and thus have been an important feature of teacher education. Researchers such as Fennema and Franke (1992) have argued since the early 1990s that teacher beliefs cannot be separated from teacher knowledge, a conclusion also supported by Grouws (1992) and by the work of DeCorte, Op't Eynde, and Verschaffel (2002). While this might lead one to conclude that teachers' beliefs can be altered as they develop knowledge of teaching and learning, there is no conclusive evidence that beliefs can be effectively influenced by teacher preparation (Tatto & Coupland, 2003). Yet there are a number of small course evaluation studies that provide evidence of the influence of primary teacher preparation courses on teachers' beliefs (see for instance Wilkins & Brand, 2004). Previous work done in this area on a larger scale by the Teaching and Learning to Teach Study (National Center for Research on Teacher Learning) helped distinguish among sets of beliefs relevant to mathematics teaching and the role of program norms in influencing views about teaching and learning (Deng, 1995; Tatto, 1996, 1998, 1999a, 1999b). In these studies, a questionnaire was developed to measure program influence by asking teacher educators and future teachers about their beliefs in aspects considered central to teaching (such as views about learning to teach diverse students) at the beginning and end of their programs. The analysis examined the aggregated views of future teachers and their educators and concluded that in programs with strong norms about inquiry-based teaching and learning, graduates' views had changed to fit program's norms (as expressed by teacher educators' views); while in programs with weak norms, graduates' views remained unchanged.

Of special importance for future mathematics teachers, of course, are beliefs about the nature of mathematics and beliefs about learning mathematics. How future teachers perceive the nature of mathematics as a subject (e.g., mathematics as formal, structural, procedural, or applied) can reflect how they have experienced learning mathematics themselves and how they are likely to approach teaching mathematics (Grigutsch, Raatz, & Törner, 1998; Ingvarson, Beavis, Danielson, Ellis, & Elliott, 2005; Ingvarson, Beavis, & Kleinhenz, 2007). Beliefs about mathematics learning have consequences for how teachers ultimately plan and deliver instruction and such beliefs include notions about the appropriateness of particular instructional activities, about students' cognition processes, and about the purposes of mathematics as a school subject. If it is indeed a goal of teacher education to challenge naïve views acquired via the apprenticeship of observation (Lortie, 1975; Mewborn & Tyminski, 2006), then exploring graduates' views across nations could provide important insights for, and about the nature of, teacher education.

Mathematics and Mathematics Knowledge Needed for Teaching

Teachers' professional knowledge has been conceptualized in a variety of ways throughout the years, but probably the most significant work came in the 1980s with the re-conceptualization of the complex kinds of knowledge that teachers need to be able to teach well, and the focus on what it means to know it. In 1987, Shulman argued that there was a knowledge base for teaching that could be understood as a "combination of content knowledge, pedagogical content knowledge, general pedagogical knowledge, curriculum knowledge, knowledge of learners, knowledge of educational contexts, and knowledge of educational ends, purposes, and values" (p. 8). Two of these concepts find global resonance in the international literature (Delaney, Ball, Hill, Schilling, & Zopf, 2008; Tatto et al., 2012) and have provided a common framework to study teacher education: (a) *content knowledge* (CK), is defined as the set of accumulated "knowledge, skills and dispositions that are to be learned by school children" (pp. 8–9), but which also has a base in the disciplines and ideas about what it means to know in those content areas; and (b) *pedagogical content knowledge* (PCK), which includes a "blending of content and pedagogy into an understanding of how particular topics, problems, or issues are organized, represented and adapted to the diverse interests and abilities of learners, and presented for instruction" (p. 8).

In subsequent decades, mathematics education scholars have produced a large body of work on prospective elementary teachers' knowledge of both different aspects of mathematics—such as division (Ball, 1990a, 1990b), proportional reasoning (Simon & Blume, 1994), fundamental knowledge of elementary mathematics operations (Ma, 1999)—and mathematical knowledge for teaching (An, Kulm, & Wu, 2004; Even & Ball, 2009; Hill, Rowan, & Ball, 2005; Hill, Sleep, Lewis, & Ball, 2007). These and other studies have been developing consensus as to the knowledge base for mathematics teaching in a field once seen by some as highly incoherent (see Begle, 1979; but also Ball, 1991), and have developed the theoretical basis to understand how content knowledge and pedagogical content knowledge contribute to the knowledge that is needed for teaching mathematics (Hill et al., 2005) and the extent to which teacher education provides future teachers with the opportunities to learn such knowledge (Boero, Dapueto, & Parenti 1996; Hill, Sleep, Lewis, & Ball, 2007).

The greatest challenge, however, has been in operationalizing what it means to possess this knowledge and in the development of valid and reliable measures of knowledge for teaching (Delaney et al., 2008). Early in the 2000s, education researchers (Ball, Lubienski, & Mewborn, 2001; Hammer & Elby, 2002) questioned whether knowledge for teaching can be measured only at a theoretical level and without consideration of "situated knowledge" (e.g., whether and how teachers are able to use mathematical knowledge in the course of their work). Indeed, a major issue evident in the research on teacher education is how to develop valid and reliable measures of teacher knowledge, and, in this era of accountability, the extent to which programs must demonstrate future teachers' levels of attained theoretical and practical knowledge.

Framework

The question in this study is the extent to which the ability of teacher education programs to produce mathematically knowledgeable future primary teachers (i.e., as indicated by mathematics and mathematics pedagogical content knowledge) is dependent on the following: the selectivity and the type of program preparing future primary teachers (e.g., for mathematics specialists or subject generalists, or for different grade levels)[2]; the background that future teachers bring with them as they enter the program (i.e., individual characteristics such as having an adequate level of mathematics knowledge at program entry); the breadth, depth, and coherence of the OTLs provided by the teacher education program; and the beliefs about mathematics and mathematics teaching held by future teachers (see Fig. 8.1).

[2] One of the unique contributions of TEDS-M was arriving at common definitions of terms to support measurement across the participating countries. One such term is "program type," used to refer to the variety of programs that prepare future primary teachers across the participating countries. Program type refers to clusters of programs that share similar purposes and structural features—such as the credential earned, the range of school grade levels for which teachers are prepared, and the degree of subject-matter specialization for which future in teachers are prepared—that correspond to distinct pathways to becoming qualified to teach (Tatto et al., 2012, pp. 27–28).

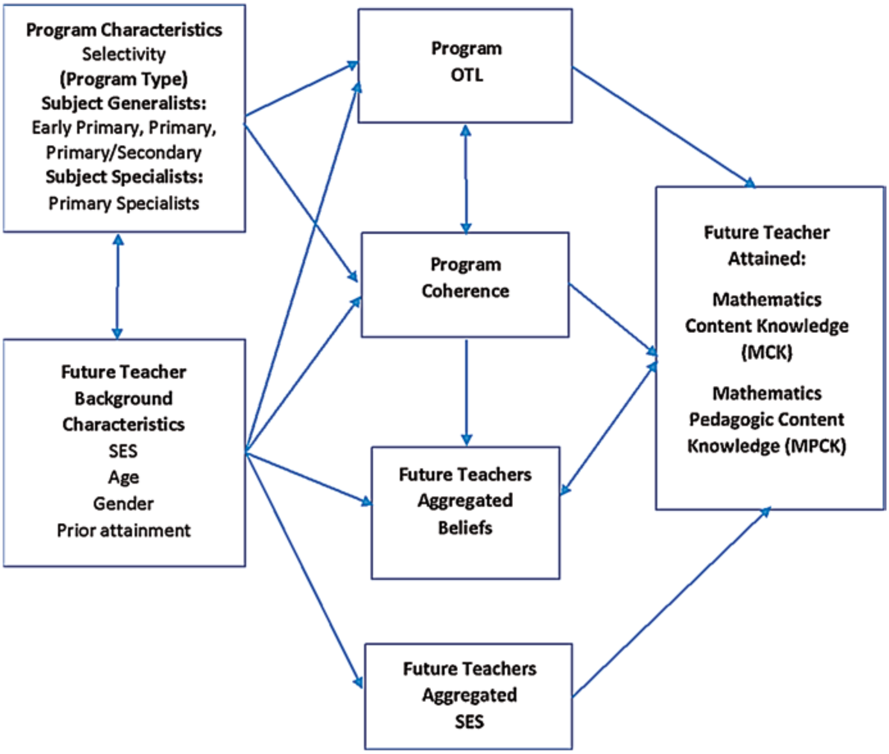


Fig. 8.1 Analytical Framework

Methods

Participants

The data come from the TEDS-M study, which collected data in 2008 and 2009 from nationally representative samples of institutions, and these institutions' future primary teachers in the last year of their teacher preparation. Using a stratified multistage probability sampling design, the study surveyed a sample of 15,163 future primary teachers in programs in 451 institutions across the participating countries. The data analysis takes into account the sampling weights for programs and future teachers, which also provide for a non-response adjustment factor for all the estimates as recommended in the TEDS-M Technical Report (Tatto, 2013). This chapter uses data from 13 countries that meet the representativeness criteria and had acceptable response rates. These include Botswana, Chile, Chinese Taipei, Germany, Malaysia, the Philippines, Poland, Russia, Singapore, Spain, Switzerland (German-speaking cantons only), Thailand, and the United States. Table 8.1 shows the number of future teachers (sample size, valid data, and percent missing) by country and

Table 8.1 Future Primary Teachers Mathematics Content Knowledge (MCK) and Mathematics Pedagogical Content Knowledge (MPCK) by primary program type, country and grade expected to teach

Program Type according to grades future teachers are prepared to teach:	Country	Grade expected to teach	Sample size	Valid data (N)	Percent missing (Weighted)	MCK Scaled Score Mean (SSM) (SE)		MCK-AP1 Percent at or above Anchor Point 1 (SSM = 431) (SE)		MCK-AP2 Percent at or above Anchor Point 2 (SSM = 516) (SE)	
1. Early primary. Prepares generalists to teach the curriculum for the early primary grades up to Grd. 4	Germany	(1–4)	935	907	2.4	501	(3)	86.4	(1.3)	43.9	(2.1)
	Poland[a]	(1–3)	1812	1799	0.9	456	(2)	67.9	(1.3)	16.8	(1.2)
	Russian Fd.[b,g]	(1–4)	2266	2260	0.2	536	(10)	89.7	(2.3)	57.3	(4.6)
	Switzerland[c]	(1–2/3)	121	121	0.0	512	(6)	90.5	(2.7)	44.2	(5.4)
2. Primary. Prepares generalists to teach the curriculum for upper primary grades up to Grd. 6	C. Taipei[g]	(1–6)	923	923	0.0	623	(4)	99.4	(0.3)	93.2	(1.4)
	Philippines	(1–6)	592	592	0.0	440	(8)	60.7	(5.1)	6.3	(0.9)
	Singapore[g]	(1–6)	263	262	0.4	586	(4)	100.0		82.5	(2.3)
	Spain	(1–6)	1093	1093	0.0	481	(3)	83.4	(1.6)	26.2	(1.6)
	Switzerland	(1/3–6)	815	815	0.0	548	(2)	97.2	(0.6)	70.6	(1.4)
	United States[d]	(1–3/4/5)	1310	951	28.6	518	(5)	92.9	(1.2)	50.0	(3.2)
3. Primary/sec. Generalists to teach the curriculum for primary/secondary up to Grd. 10	Botswana[e,g]	(1–7)	86	86	0.0	441	(6)	60.6	(5.3)	7.1	(2.8)
	Chile[f]	(1–8)	657	654	0.4	413	(2)	39.5	(1.8)	4.0	(0.7)
4. Primary spec. Prepares teachers to teach the curriculum for the primary grades with a *specialization* (in mathematics, and in some cases in another related subject+)	Germany	(1–9/10)	97	97	0.0	555	(8)	96.0	(2.1)	71.7	(7.0)
	Malaysia	(1–6)	576	574	0.4	488	(2)	88.7	(1.1)	28.1	(1.3)
	Poland[a,g]	(4–9/12)	300	300	0.0	614	(5)	97.9	(1.0)	91.0	(1.6)
	Singapore[g]	(1–6)	117	117	0.0	600	(8)	98.3	(1.2)	87.3	(2.8)
	Thailand	(1–12)	660	660	0.0	528	(2)	91.7	(0.9)	56.2	(1.4)
	United States[d]	(4–5/8/9)	191	132	33.2	520	(7)	94.9	(1.7)	48.1	(6.5)

Program Type according to grades future teachers are prepared to teach:	Country	Grade expected to teach	Sample size	Valid data (N)	Percent missing (Weighted)	MPCK Scaled Score Mean (SSM) (SE)		MPCK-AP Percent at or above Anchor Point (SSM = 544) (SE)	
1. Early primary. Prepares generalists to teach the curriculum for the early primary grades up to Grd. 4	Germany	(1–4)	935	907	2.4	491	(5)	25.9	(2.0)
	Poland[a]	(1–3)	1812	1799	0.9	452	(2)	11.9	(1.3)
	Russian Fd.[b g]	(1–4)	2266	2260	0.2	512	(8)	31.6	(4.1)
	Switzerland[c]	(1–2/3)	121	121	0.0	519	(6)	31.6	(4.2)
2.Primary prepares generalists to teach the curriculum for upper primary grades up to Grd. 6	C. Taipei[g]	(1–6)	923	923	0.0	592	(2)	77.0	(1.3)
	Philippines	(1–6)	592	592	0.0	457	(10)	5.9	(1.6)
	Singapore[g]	(1–6)	263	262	0.4	588	(4)	74.9	(2.5)
	Spain	(1–6)	1093	1093	0.0	492	(2)	17.5	(1.3)
	Switzerland	(1/3–6)	815	815	0.0	539	(2)	44.0	(1.5)
	United States[d]	(1–3/4/5)	1310	951	28.6	544	(3)	47.6	(1.7)

(continued)

Table 8.1 (continued)

Program Type according to grades future teachers are prepared to teach:	Country	Grade expected to teach	Sample size	Valid data (N)	Percent missing (Weighted)	MPCK Scaled Score Mean (SSM) (SE)		MPCK-AP Percent at or above Anchor Point (SSM = 544) (SE)	
3. Primary/sec. Generalists to teach the curriculum for primary/secondary up to Grd. 10	Botswana[e g]	(1–7)	86	86	0.0	448	(9)	6.2	(2.8)
	Chile[f]	(1–8)	657	654	0.4	425	(4)	4.9	(1.0)
4. Primary spec. Prepares teachers to teach the curriculum for the primary grades with a *specialization* (in mathematics, and in some cases in another related subject+)	Germany	(1–9/10)	97	97	0.0	552	(7)	59.6	(3.4)
	Malaysia	(1–6)	576	574	0.4	503	(3)	23.4	(1.9)
	Poland[a g]	(4–9/12)	300	300	0.0	575	(4)	67.3	(2.3)
	Singapore[g]	(1–6)	117	117	0.0	604	(7)	81.1	(3.9)
	Thailand	(1–12)	660	660	0.0	506	(2)	26.4	(1.5)
	United States[d]	(4–5/8/9)	191	132	33.2	545	(6)	41.4	(6.3)

[a]Poland: Reduced coverage: Institutions with consecutive programs only were not covered. Combined participation rate between 60 and 75%; [b]Russian Federation: Reduced coverage: Secondary pedagogical institutions were excluded; [c]Switzerland: Reduced coverage: The population covered includes only institutions where German is the primary language of use and instruction; [d]United States: Reduced coverage (data were available from less than 85% of respondents): public institutions only. Combined participation rate between 60% and 75%. Although the participation rate for the complete sample meets the required standard, the data contain records that were completed using a telephone interview, when circumstances did not allow administration of the full questionnaire. Of the 1501 recorded as participants, the full questionnaire was administered to 1185. [e]Botswana: The sample size is small (N = 86), but arises from a census of a small population; [f]Chile: combined participation rate between 60% and 75%. [g]Programs that stipulate a specific level of mathematics knowledge as a requirement for entry to the program

Source: Tatto et al., 2012, pp. 29–32, 139 & 143 (with permission) program group (according to grades future teachers are prepared to teach); Tables 8.4, 8.5 and 8.6 in the Appendix show in detail the definition of variables, and the descriptive statistics for the programs, and for the future teachers included in this study.

Data Sources

The data for this study come from the TEDS-M teacher program survey and the primary future teacher survey. The teacher program survey is a questionnaire asking program administrators for information about the organization, the policies, and the curriculum of the programs included in the sample. The primary future teacher survey consisted of a questionnaire and an assessment of mathematics knowledge and mathematics pedagogy content knowledge; future teachers were asked to spend 90 minutes answering the survey with one hour dedicated to the assessment. The TEDS-M instruments were developed in collaboration with the research team in the participating countries using rigorous standards for the development of survey and assessment instruments (see Tatto 2013).

Data from the Program Survey A program was defined as a *specific pathway that exists within an institution requiring students to undertake a set of subjects and experiences and leads to the award of a common credential(s) on completion*. The program data come from questions about program characteristics including program entry requirements and program type (e.g., whether the goal of the program is to prepare future primary teachers as subject specialists or as generalists to teach the earlier or upper grades of primary school). Program type determines in great part future teachers' opportunities to learn. The TEDS-M study used UNESCO's (2007) International Standard Classification of Education (ISCED) to facilitate comparisons of grade levels across countries. The grade levels at which future primary teachers are prepared to teach was information collected via the program survey and is reported in the first column of Table 8.1.

Data from the Future Teachers' Survey Future teachers were defined as *persons enrolled in a teacher preparation program that is explicitly intended to prepare teachers qualified to teach mathematics in any of the grades at the primary school level and who were in the last year of their teacher preparation program*. The data were collected via a questionnaire and knowledge assessments. The questionnaire asked future teachers for their background information, the types of OTL provided by the program, and their beliefs.

Background Future teachers' background was considered a potential moderator of level of attainment at the end of the program. The data on background come from questions about the *background* of respondents, such as their SES (also aggregated in this study at the program level to obtain an indicator of how SES as a program characteristic affects future teachers' attained knowledge), age, gender, and prior attainment. While desirable, no pre-test measure of knowledge was used in TEDS-M because across countries it was not possible to find a clear and comparable entry point to teacher education, and there were no existing valid and reliable indicators of prior attainment that could be used comparatively across countries; instead, respondents were asked to report previous performance to use as a proxy for prior attainment. The strong positive correlations between the previous performance (prior attainment) indicator and the assessment results in this study confirm the reliability of future teachers' self-report.

Opportunities to Learn What respondents report about their OTL is important because, in theory, what future teachers come to know may be in part determined by what they learn in their programs and in part by what they bring with them when they enroll. That is, what future teachers come to know at the end of their programs is moderated by the program's entry requirements (e.g., it is possible that programs' lack of emphasis on academic and school mathematics may be compensated by a mathematics requirement as a condition for admittance into teacher education) and by the actual learning experiences that a program provides. This study, explores the extent to which future teachers have opportunities to learn mathematics (e.g., university-level mathematics and mathematics of the school curriculum) and mathematics content pedagogy including the extent to which they have been given opportunities to read research on teaching and mathematics; and the extent to which these experiences are coherent.

This study uses the OTL *indices* developed by the TEDS-M study team and released in the public database and scaled using the Rasch model (De Ayala, 2009; Wu, Adams, Wilson, & Haldane, 2007) centered at the point on the scale that is associated with the middle of the rating scale (i.e., the neutral position) and given a value of 10.

Opportunities to Learn Mathematics and Mathematics Pedagogy Included in the analysis are geometry topics (GEOM), as important indicators of mathematics topics studied including foundations of geometry or axiomatic geometry, analytic/coordinate geometry, non-Euclidean geometry, and differential geometry; and school-level mathematics topics including topics related to functions (SLMF) such as relations, equations, data representation, probability, statistics, calculus, and validation, structuring, and abstracting. In addition the analysis included scales that serve as indicators of opportunities to learn key pedagogical knowledge such as opportunities to engage in readings on mathematics education and pedagogy (READ), including readings about research on mathematics, on mathematics education, and on teaching and learning mathematics, including analysis of teaching examples.[3]

[3] The final composition of the OTL indicators was done based on the logical organization of courses as judged by experts, after repeated piloting of the questions ultimately used to develop them. The Comparative Fit Index, or CFI, was used to test the degree to which the indicators were internally

Coherence The degree to which programs structured OTL in a coherent manner (COH) was considered an important characteristic with the potential to create conducive learning environments for future teachers (i.e., whether the courses seem to follow a logical sequence of development in terms of content, whether topics are built on what was taught in earlier courses, and whether each of the courses is clearly designed to prepare future teachers to meet a common set of explicit standard expectations), (Tatto, 1996, 1998, 1999a, 1999b).

Beliefs The data on *beliefs* come from questions that ask whether future teachers view mathematics as a formal, structural, procedural (RULE), or an applied subject (ACTV), views that are seen as influential in regulating the learning and teaching of mathematics (Grigutsch et al,. 1998; Op 'T Eynde, De Corte, & Verschaffel, 2002). The belief questions originally used a 6-point scale (i.e., *strongly agree* to *strongly disagree*) and were scaled using the Rasch model (De Ayala, 2009) centered at the point on the scale that is associated with the middle of the rating scale (i.e., the neutral position) and given a value of 10.[4] The beliefs responses from future teachers are aggregated in this study at the program level to obtain an indicator of how beliefs as a program characteristic affect future teachers' attained knowledge.

Knowledge The data on future teachers' knowledge comes from two distinct assessments. The TEDS-M assessment of *Mathematics Content Knowledge* (MCK) consists of four domains: number and operations, algebra and functions, geometry and measurement, and data and chance; and three subdomains: knowing (e.g.,

---

consistent; the CFI depends in large part on the average size of the correlations in the data, where an acceptable model is indicated by a CFI larger than .93, but .85 is acceptable (see Bollen, 1989). Another approximation is the Tucker Lewis index (TLI), which is relatively independent of sample size (Marsh, Balla, McDonald, 1988), where values over .90 or.95 are considered acceptable (e.g., Hu & Bentler, 1999). The Root Mean Square Error of Approximation, or RMSEA, is another test of model fit; good models are considered to have a RMSEA of .05 or less, while models whose RMSEA is .1 or more have a poor fit. Fit Indices provided evidence that the groupings that formed the opportunity to learn indicators make sense (i.e., tertiary level mathematics CFI .911, TLI .954, RMSEA .044, school-level mathematics CFI .97, TLI .973, RMSEA .057). The reliabilities for the OTL scales were unweighted and were estimated using jMetrik 2.1 (Meyer, 2011). The reliability estimates were based on the congeneric measurement model, which allows each item to load on the common factor at different levels, and allows item error variances to vary freely (each item can be measured with a different level of precision). This was considered by the TEDS-M study research team to be the most flexible measurement model and the most appropriate for measures with few items. The reliabilities for the opportunity to do class reading on research on mathematics teaching and learning for the primary sample is .85; for the opportunity to learn in a coherent program, the reliability is .96.

[4] The reliabilities for the beliefs scales were unweighted and were estimated using jMetrik 2.1 (Meyer, 2011). The reliability estimates were based on the congeneric measurement model, which allows each item to load on the common factor at different levels, and allows item error variances to vary freely (each item can be measured with a different level of precision) as described in footnote 3 for the OTL scales. For the international sample, the reliability for the beliefs scale "mathematics as a set of rules and procedures" for future primary teachers is .94 and .92 for "learning mathematics through active involvement."

recall, recognize, compute, retrieve, measure, classify, and order), applying (e.g., select, represent, model, implement, and solve routine problems), and reasoning (e.g., analyze, generalize, synthesize, integrate, justify, and solve non-routine problems). These domains reflect the consensus among the participating countries regarding the knowledge that is essential for teachers to know to effectively teach the primary-level curriculum. These domains align with the TEDS-M analysis of the school curriculum and of the syllabi of the teacher education programs in the participating countries (see Chapter 4 in this book, and Tatto & Hordern, 2017), and with similar analysis of TIMSS data (Mullis, Martin, Foy, & Arora, 2012). The TEDS-M assessment of *Mathematics Pedagogical Content Knowledge* (MPCK) measures pedagogical knowledge in the same four domains of number and operations, algebra and functions, geometry and measurement, and data and chance, and in three subdomains specific to pedagogical concerns: curricular knowledge, knowledge of planning for mathematics teaching and learning (pre-active), and knowledge for enacting mathematics for teaching and learning (interactive). The description of the domains for the two MCK and MPCK assessments can be found in the TEDS-M Conceptual Framework (Tatto et al., 2008, pp. 37–39). Thus, the mathematics and mathematics pedagogy assessments developed by the TEDS-M team required future teachers to demonstrate knowledge beyond the mastering of rules and procedures and toward a more inquiry-based approach. The assessments used a block design and included 70 items distributed across five blocks, with about two-thirds of the items measuring MCK and one-third measuring MPCK; the question types included multiple choice, complex multiple choice, and constructed response. Both assessments benefitted from previous work done by mathematics education scholars (Ball & Bass, 2000; Clements, Bishop, Keitel, Kilpatrick, & Leung, 2013; Hill et al., 2007; Kilpatrick, Swafford, & Findell, 2001; Lappan, 2000). The analysis in this chapter uses the MCK and MPCK scales, which were developed from the assessment results using IRT (De Ayala, 2009) by the TEDS-M study research group. The two scales were calibrated for cross-country comparison to have a mean of 500 and a standard deviation of 100.[5]

Relevance and Validity To achieve relevance and validity the TEDS-M assessments were developed according to established standards for educational and psychological testing (AERA, APA, & NCME, 2014). Extensive consultation with policy makers, mathematicians, mathematics teacher educators, and future teachers occurred prior to, during, and after the study. Each country's national team was asked to describe in their context the mathematical work entailed in teaching to inform instrument development. In addition, as mentioned in the previous paragraph, the

[5] For the international sample, the reliabilities for the mathematics content knowledge and the mathematics pedagogical content knowledge scores were .83 and .76, respectively. Reliabilities tend to be high if there is a great deal of variation in the sample relative to the size of the standard error. The reliability will be low if one of the following occurs: there is a small standard deviation in the sample, or there is a large standard error (e.g., the test was too easy for a particular sample; this was the case for Chinese Taipei and Singapore).

national teams engaged in the analysis of the national school curriculum and standards, as well as in the analysis of the teacher education program curriculum and standards. The curriculum analysis helped verify the content domains of the MCK and MCPK assessments and was useful in collecting content-related validity evidence.

The TEDS-M items were validated in a manner consistent with current notions of validity, which is seen as a process that validates the use of items for a particular purpose (e.g., Reckase, McCrory, Floden, Ferrini-Mundy, & Senk, 2015). Prior to designing the instruments, the TEDS-M International Study Center (ISC) solicited items from all the participating countries.[6] Items that had been developed by other studies and additional items developed by the ISC were mapped onto the domains of knowledge described above—that is, those deemed necessary for future primary teachers to have at the end of their teacher education to teach a demanding curriculum across the participating countries. Throughout the assessment development process, items were translated from the local languages to English and from English to the local languages and back translated to confirm accuracy and consistency. The items were piloted with a small sample from the intended population and were examined for their performance, congruence with participants' knowledge, and familiarity with respect to format; the codebooks were developed at this stage, and extensive scoring training was carried out. Analysis by psychometricians determined which items were included in the assessment and which were discarded at this stage. The selected items were carefully placed in five blocks with link items and piloted with a sample from the intended population. Codebooks were revised, and once again extensive scoring training was carried out. The resulting data were analyzed looking for item fit (see footnote 3), and this analysis determined items that were included in the final assessments and those that were discarded. The results were analyzed according to the internal structure of the test by the ISC and by all the country teams before the instruments were declared final; items that did not measure well in any given country were eliminated. All throughout the instrument development-phase, factor analysis was used to obtain construct-related validity evidence; curriculum analysis and expert review was used to obtain content-related validity evidence. The final assessments were assembled in collaboration with psychometricians from the ISC, the national centers, and the International Association for the Evaluation of Educational Achievement (IEA) Data Processing Center. In sum, the TEDS-M study followed the IEA's standards to ensure validity and reliability in international comparative studies (see Martin & Mullis, 2008).

[6] Several TEDS-M items were provided by other studies, including Study of Instructional Improvement (SII) Learning Mathematics for Teaching/Consortium for Policy Research in Education (CPRE), University of Michigan, School of Education, Ann Arbor, MI, supported by NSF grants REC-9979873, REC- 0207649, EHR-0233456 & EHR 0335411. Developing Subject Matter Knowledge in Math Middle School Teachers (P-TEDS), Michigan State University, supported by NSF Grant REC-0231886. Knowing Mathematics for Teacher Algebra (KAT), Michigan State University, supported by NSF Grant REC-0337595 (TEDS-M received 2006 publication copyright for those items).

Limitations

Because TEDS-M was designed to examine the knowledge outcomes of teacher education close to graduation, the study did not assess whether future primary teachers who are knowledgeable in the MCK and MPCK assessments are also able to teach mathematics effectively to pupils. Since then, however, a number of validation studies have included items similar to those used by the TEDS-M study, and report that measures of Mathematics Knowledge for Teaching (MKT) a composite of MCK and MPCK using these items are strongly related to the mathematical quality of instruction (Delaney, 2012; Delaney et al., 2008; Schilling & Hill, 2007). A new study is using items from TEDS-M to measure novice teachers' knowledge and practices, taking into account previous education and/or preparation (Tatto, Rodriguez, Smith, Reckase, & Pippin, forthcoming). We will have to wait until the study results are available to ascertain whether the most knowledgeable teachers according to TEDS-M assessments are also the most effective.

Analytic Methods

This chapter uses the results of the assessments and anchor points descriptions to analyze the breadth and depth of the mathematical knowledge and mathematical pedagogical content knowledge reached by future primary teachers across programs in the participating countries. Descriptive statistics and HLM and OLS analysis were used to explore the research questions (Appendix Table 8.4 and Appendix Tables 8.5 and 8.6 provide for each analysis a brief description of the variables used in the analysis along with descriptive statistics).

Because this study is based on the hypothesis of a multilevel structure, in which the effects of teacher background may differ according to the programs future teachers are in, a multilevel statistical analysis such as hierarchical linear modeling (HLM[7]) is the optimal approach to investigating the relationship between teacher education program characteristics and mathematics knowledge for teaching (MCK and MPCK). HLM was generally used for countries with large program samples. The number of programs preparing future primary teachers varied widely across countries, with the most extreme cases being Botswana and Singapore, with four

[7] Hierarchical Linear Modeling or HLM (Raudenbush & Bryk, 2002; Raudenbush, Bryk, & Congdon, 2004) is a statistical method that helps compute regressions at multiple levels, estimating a regression within each program and combining them to see if there is a common regression across programs within a given country. If regression slopes vary across programs, it is possible to examine program-level characteristics that may explain such variation, and to explore the program factors that may show a relationship with future teachers' outcomes. The analysis was done using a two-level HLM model in which future teachers were nested within their teacher education programs within countries. The descriptive statistics for the institutions and future teachers are in Appendix Table 8.4.

and one program respectively, and Poland and the United States with 125 and 78 programs respectively, see Appendix Tables 8.5 and 8.6. However, it was necessary to use Ordinary Least Squares (OLS[8]) for some countries because their programs were less likely to differ. In part, this happened for a small sample of programs with a very similar structure—including Singapore, with one institution, Botswana with four, Chinese Taipei with 11, and Germany and Switzerland with 14 each—and, in part, because countries have a strict centralized system (e.g., in Malaysia, although there were 23 programs in the sample, the program data is the same for all programs, as the Ministry of Education requires that all programs operate in the same way effectively reducing the program information to an $n = 1$). Because of the smaller number of institutions and/or because of institutional isomorphism within these countries, variance between institutions is not reliably estimated, so that variance (if it exists systematically in the population) is part of the individual future-teacher variability and not partitioned separately. To proceed with the analysis of the association between teacher background and knowledge, program-level characteristics have been added as additional explanatory variables in OLS models.

The analyses used standardized coefficients to explore individual- and program-level features associated with MCK and MPCK as teacher education outcomes, using the same HLM or OLS model across all countries, but analyzing each country separately.

All participating countries with acceptable response rates on the assessments were included. Each country's data was analyzed across the same variables following the theoretical framework. Variables across countries were examined for missing values across variables of interest, and collinearity; this examination resulted in a number of variables (and programs) being excluded from the analysis. Other variables were excluded because they lacked variability (e.g., all programs reported offering field experiences).

Findings

The Characteristics of Teacher Education Programs

Table 8.1 shows the types of programs that prepare future primary teachers in this study. There are four major categories of programs (see program type column) that prepare future teachers to teach mathematics: (a) early primary, including grades up to Grade 4 (found in Germany, Poland, the Russian Federation, and Switzerland); (b) primary grades up to Grade 6 (found in Chinese Taipei, the Philippines,

[8] Regression analysis (ordinary least squares or OLS) is a method that helps to explore the relationship between a dependent variable, in this case mathematics knowledge for teaching (defined as MCK and MPCK), and one or more explanatory variables, in this case individual and program variables. The descriptive statistics for the institutions and future teachers are in Appendix Tables 8.5 and 8.6.

Singapore, Spain, Switzerland, and the United States); (c) primary and lower secondary combined, with an expectation that teachers will be able to teach up to Grade 10 (found in Botswana and Chile); (d) primary mathematics specialists (who are expected to teach mostly mathematics and who received training as mathematics specialists to teach from Grade 1 to Grade 12 depending on the country expectations).[9] These programs are found in Germany, Malaysia, Poland, Singapore, Thailand, and the United States. Note that a number of countries—notably, Germany, Poland, Singapore, and the United States—have two types of programs that prepare future primary teachers either as generalists or specialists.

Programs vary on the degree to which they require a credential or some mathematics knowledge at the time of entry into teacher education. Countries that stipulate a mathematics requirement as a condition for entry into teacher education programs are indicated in Table 8.1 by the superscript *g* and include in group 1, the Russian Federation; in group 2, Chinese Taipei, and Singapore; in group 3, Botswana; and, in group 4, Poland and Singapore.

Table 8.1 also includes the results of MCK and MPCK assessments used in the study. These results are explained below.

The Level and Depth of the Mathematics and Mathematics Teaching Knowledge Attained by Prospective Primary Teachers

The results of the MCK and MPCK assessments across the different countries and programs preparing primary teachers are shown in scaled score means (SSM) in the column labeled MCK and MPCK in Table 8.1. Note that Germany, Poland, Switzerland, Singapore, and the United States are countries that have two program types with different approaches to preparing primary teachers to teach the early and upper grades. The scores of the future teachers in each of these two different approaches reflect access to different OTL. For instance, Germany has two different program types; future teachers in program type 1 are prepared to teach Early Primary up to Grade 4, while future teachers in program type 4 are prepared to teach primary as mathematics specialists. German future teachers in program type 1 had an MCK score of 501, in contrast with those in program type 4 who had an MCK score of 555; similarly, for MPCK, future teachers in program type 1 had a score of 491, while those in program type 4 had a score of 552 (this means that future primary teachers in program type 4 obtained 54 and 61 more points in the assessments than

[9] The definition of primary grades is "the first stage of basic education which starts normally between the ages of 5–7," according to UNESCO's International Standard Classification of Education (UNESCO-UIS, 2006). Since the age of children enrolled in primary education varies, as does the grade span for which teachers are prepared to teach, countries were asked to define the grade range. For TEDS-M, the grade span for which teachers are prepared to teach in the different countries is included in Table 8.1 and ranges from Grades 1 to 10 for generalists and Grades 1 to 12 for mathematics specialists.

their counterparts in program type 1). A similar pattern is found in Poland (with MCK scores 456 and 614, and MPCK scores 452 and 575), in Switzerland (MCK 512 and 548, and MPCK 519 and 539), and in Singapore (MCK 586 and 600, and MPCK 588 and 604). Future primary teachers in the rest of these countries showed a similar pattern, with the exception of the United States where there was no discernible difference in the scores obtained by future primary teachers in these two program types (the scores for each generalist and specialist program types were respectively in MCK 518 and 520, and in MPCK 544 and 545). In short, future teachers expected to teach the upper grades of primary school (program type 2) generally had significantly higher scores in the assessments than those expected to teach in the earlier primary grades.

Among the countries that rely on one program type to prepare future primary teachers, the higher mean scores in MCK and in MPCK, respectively, were in Chinese Taipei, (623 and 592), the Russian Federation (536 and 512), and Thailand (527 and 507). Lower average scores in MCK and in MPCK were observed in Spain (480 and 492), Malaysia (489 and 503), Botswana (448 and 463), the Philippines (442 and 463), and Chile (414 and 424). Note that the difference in performance between the Chinese Taipei future primary teachers and the Chilean future primary teachers was more than 2 standard deviations (or more than 200 points) in the MCK assessment and more than 150 points in the MPCK assessment. When compared with future primary teachers in Chinese Taipei, the United States' future teachers scored more than 100 points lower in the MCK assessment and close to 50 points lower in the MPCK assessment.

Anchor Points as Indicators of What Teachers Know Table 8.1 columns labeled MCK, MCK-AP1, MCK-AP2, and MPCK, MPCK-AP contain the scale score means for the respective assessments as well as the percentage of future primary teachers who reached the anchor points in each assessment. Anchor points (AP) are reference points on the scales defined by the item response theory analysis of the MCK and MPCK assessments. These APs were selected so that there would be sufficient items measuring skills and abilities for future teachers estimated to be at those points so that good descriptions of the skills and abilities could be developed. The AP descriptions were developed by the TEDS-M study psychometric team and a panel of mathematicians and mathematics educators who analyzed the items that measured well at these APs and who formulated empirically based descriptions of the knowledge that future teachers demonstrated at each AP (see Chapter 19 of this book for a detailed account on how these Anchor Points were developed).

The AP descriptions complement the quantitative information provided by the knowledge assessments. Items used to describe performance at the APs were determined by the probability that a person with a score at that point would get the relevant item correct or incorrect. For MCK, two APs were identified that would have sufficient items to yield good descriptions. For MCK Anchor Point 1 (MCK-AP1) "can do" items had a .70 or greater probability of a correct response among future teachers estimated to be near the AP and "cannot do" items had less than a .50 probability of a correct response at the same point. The contrasting information from the "can do" and "cannot do" items were used to develop the description of skills and knowledge at the AP. The items related to MCK Anchor Point 2 (MCK-AP2) were selected to have the same probability characteristics. MCK-AP1 indicated a lower level of performance (with a scale score mean of 431, or 69 points below the international mean which has a scale score of 500), and MCK-AP2 indicated a higher level of performance (with a scale score mean of 516, or 16 points above the international mean).

For instance, those future primary teachers who reached MCK-AP1 were able to successfully solve basic computations with whole numbers, identify the properties of operations with whole numbers, and demonstrate adequate reasoning about odd or even numbers. Also, they were able to solve straightforward problems using simple fractions, visualize and interpret standard two-dimensional and three-dimensional geometric figures, and solve routine problems about perimeter. In addition, they demonstrated straightforward uses of variables and equivalence of expressions, and solved problems involving simple equations. However, these same teachers had difficulties answering items that included abstract problems and problems requiring multiple steps, proportionality, multiplicative reasoning, and least common multiples. They also had difficulties solving problems that involved coordinates, or problems about relations between geometric figures, finding the area of a triangle drawn on a grid, and identifying an algebraic representation of three consecutive even numbers. In addition, they had trouble demonstrating their reasoning about multiple statements and about relationships among several mathematical concepts (such as understanding that there is an infinite number of rational numbers between two given numbers) (Tatto et al., 2012, p. 136-140).

Those future primary teachers who reached MCK-AP2 were able to demonstrate the knowledge to reach MCK-AP1 and, in addition, were able to use fractions to solve story problems, recognize examples of rational and irrational numbers, find the least common multiple of two numbers in a familiar context, and recognize that some arguments about whole numbers are logically weak. They were also able to determine the areas and perimeters of simple figures, and had a notion of class inclusion among polygons. They demonstrated familiarity with linear expressions and functions and were able to do problems involving proportional reasoning. Even at this level, however, many future primary teachers had trouble reasoning about factors, multiples, and percentages, demonstrating applications of quadratic or exponential functions or applying algebra to geometric situations (such as writing an expression for the reflection image of the point with coordinates (a, b) over the x-axis). They also had difficulty identifying a set of geometric statements that uniquely define a square, and describing properties of a linear function (Tatto et al., 2012, p. 136-140).

For MPCK, one anchor point (MPCK-AP) was identified[10] indicating a proficient level (with a scale score mean of 544, or 44 points above the international mean). Those future primary teachers who reached MPCK-AP were able to recog-

[10]Anchor points are dependent on the number of items included in a particular measure. In the assessment 2/3 of the items measured MCK while 1/3 measured MPCK, therefore for MPCK only one anchor point was defined at the proficient level.

nize whether a teaching strategy was correct for a particular concrete example, evaluate students' work when the content was conventional or typical of the primary grades, and identify the arithmetic elements of single-step story problems that influence the difficulty of these problems. However, they had difficulty interpreting students' work, using concrete representations to support students' learning and recognizing how students' thinking relate to a particular algebraic representation. They had trouble understanding measurement or probability concepts needed to reword or design a task, knowing why a particular teaching strategy made sense and if it would always work, and knowing whether a strategy could be generalized to a larger class of problems. They were unaware of common misconceptions and found it difficult to conceive useful representations of numerical concepts (Tatto et al., 2012, p. 140-142).

Future primary teachers at MCK-AP2 and at MPCK-AP have reached high levels of performance according to the TEDS-M assessments. Examples of items used to describe these anchor points follow.

Example of an MCK Complex Multiple-Choice Item Used to Define MCK-AP1 Figure 8.2 shows an example of a complex multiple-choice item measuring MCK in the algebra domain, and it illustrates MCK-AP1. On average, the proportion of future teachers in the international sample who answered the item correctly was very high for "A," "B," and "C" (ranging from 81 to 92% across countries), but lower for "D" (64%). Thus, future teachers with scores at or above MCK-AP1 were likely to answer parts A, B, and C correctly, while those with scores at MCK-AP1 were not likely to answer part D correctly. Teachers at MCK-AP2, however, are more likely to answer item D successfully).

ID: MFC202ABCD	MS Booklet: PM1, PM2	MS Block: B2PM	Item Format: CMC	Max Points: 4
Outcome: MCK	Domain: Algebra		Sub-domain: Knowing	

Indicate whether each of the following statements is true for the set of all whole numbers a, b and c greater than zero.

Check one box in each row.

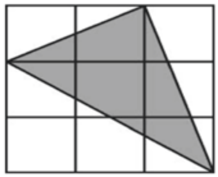
		True	Not True
A.	$a - b = b - a$	o	o
B.	$a \div b = b \div a$	o	o
C.	$(a + b) + c = a + (b + c)$	o	o
D.	$(a - b) - c = a - (b - c)$	o	o

Key (international average): A. 81 % (NT), B. 86% (NT), C. 92% (T), D. 64% (NT)

Fig. 8.2 Example of a primary complex multiple choice MCK algebra item. Source: Tatto et al., 2012, p. 137

ID: MFC408	MS Booklet: PM3, PM4	MS Block: B4PM	Item Format: MC	Max Points: 1
Outcome: MCK	Domain: Geometry		Sub-domain: Applying	

The area of each small square is 1 cm^2.



What is the area of the shaded triangle in cm^2?

Check one box.

A.	3.5 cm^2	☐
B.	4 cm^2	☐
C.	4.5 cm^2	☐
D.	5 cm^2	☐

Key (international average): 60% (A).

Fig. 8.3 Example of a primary multiple choice MCK geometry item. Source: Tatto et al., 2012, p. 138

Example of an MCK Multiple-Choice Item Used to Define MCK-AP2 Figure 8.3 shows an example of a multiple-choice item in the MCK assessment in the geometry domain. On average, across countries, 60% of teachers in the international sample answered this item correctly—accordingly, future teachers with scores at AP1 were not likely to answer this item correctly, but those scoring at or above AP2 were more likely to answer this item correctly.

Example of an MPCK Constructed Response Item Used to Define the MPCK-AP Figure 8.4 shows an example of a constructed response item in the MPCK assessment in the *enacting teaching of number* domain. On average, the proportion of future primary teachers in the international sample who answered this item correctly was quite low. Only 20% received full credit and 12% received partial credit in the part "a" of the item. Concerning part "b" of the item, only 16% received full credit and 16% partial credit.

A summary of what answers count as correct, partially correct, and incorrect taken from the *User Guide for the International Database* (Brese & Tatto, 2012) is included here to provide an illustration. The codebook stipulated that correct responses for "(a) what is [Jeremy's] most likely misconception?" must suggest that

ID: MFC208A	MS Booklet: PM1, PM2	MS Block: B2PM	Item Format: CR	Max Points: 2
Outcome: MPCK	Domain: Number		Sub-domain: Enacting	

[Jeremy] notices that when he enters 0.2 × 6 into a calculator his answer is smaller than 6, and when he enters 6 ÷ 0.2 he gets a number greater than 6. He is puzzled by this, and asks his teacher for a new calculator!

(a) What is [Jeremy's] most likely misconception?

(b) Draw a visual representation that the teacher could use to model 0.2 × 6 to help [Jeremy] understand **WHY** the answer is what it is?

Key (international average): (a) full credit 20%, partial credit 12%; (b) full credit 16%, partial credit 16%

Fig. 8.4 Example of a primary constructed response on enacting teaching of number MPCK Item. Source: Tatto et al., 2012, p. 141

"the misconception is that multiplication always gives a larger answer and that division always gives a smaller answer" (i.e., "he thinks that when you multiply the answer should be larger and when you divide the answer should be smaller"). For a response to be considered partially correct, it must either suggest that "the misconception is that multiplication always gives a larger answer or that division always gives a smaller answer but not both" (i.e., "he thinks that when you multiply, the answer should be larger than either/both numbers") or suggest that "Jeremy considers 0.2 as a whole number" (i.e., "he thinks he is multiplying and dividing by 2 rather than by 0.2"). The codebook stipulated that correct responses for "(b) Draw a visual representation that the teacher could use to model 0.2 × 6 to help [Jeremy] understand WHY the answer is what it is" must include a "suitable visual representation that clearly shows why 0.2 × 6 is 1.2." Area models to represent two-tenths of six are included in the *User Guide for the International Database* (Brese & Tatto, 2012, Supplement 4, pp. 14–16) as well but space limitations preclude inclusion here.

Significant Gaps in Mathematics Knowledge Were Found Across Countries

The comparative importance of the AP definitions can be seen when we turn back to Table 8.1. Table 8.1 column MCK-AP1 shows the proportion of teachers who scored at or above that AP (at or above the scale score mean of 431), and in column MCK-AP2 the proportion of teachers who scored at or above that AP (at or above the scale score mean of 516). Among the programs preparing teachers to teach the upper-primary grades as generalists (program type 2), over 90% of future primary teachers in Chinese Taipei reached Anchor Point 2, but only 50% of those in the United States did. While the scores among the primary mathematics specialists (program type 4) were higher for most countries, in the United States the proportion reaching MCK-AP2 remained the same. Table 8.1 column MPCK-AP shows the proportion of teachers who scored at or above the MPCK anchor point (scale score mean of 544). Overall, these items were more challenging for all future primary teachers, especially for those preparing to teach in the earlier grades. Among those in programs preparing teachers to teach the upper-primary grades as generalists (program type 2), a larger proportion reached the MPCK-AP in Chinese Taipei and Singapore, than in the United States. In program type 4, corresponding to teachers prepared as specialists, about 80% reached the MPCK-AP in Singapore, and close to 70% in Poland compared to about 40% in the United States.

Significant gaps concerning future primary teachers' MCK had to do with their ability to answer correctly abstract, complex, and multiple-steps problems, or to demonstrate relationships among several mathematical concepts. Algebra and geometry problems seemed particularly challenging, while problems that required basic computations, operations with whole numbers, simple figures, and linear expressions were less challenging. Concerning MPCK, future primary teachers had difficulty interpreting students' work, especially in the context of the mathematics they found challenging, such as algebraic representations, and, in general, they had problems supporting students' learning, as they were unable to recognize common misconceptions or conceive of useful representations. The relatively low level of knowledge demonstrated by the more than 50% of U.S. future teachers who were unable to reach the MCK-AP2 level and the MPCK-AP in the assessments suggests the difficulty these teachers will encounter when addressing what U.S. researchers consider essential in mathematics instruction, namely, conceptual understanding, procedural skills and fluency, and application[11] (Hiebert & Grouws, 2007; National Mathematics Advisory Panel, 2008). Furthermore, the results of this study fall far short of the aims stated in documents such as *Principles to Actions* (NCTM, 2014) and raise concerns about future teachers' ability to establish goals to focus learning; implement tasks that promote reasoning and problem solving; use mathematical representations; facilitate meaningful discourse; build conceptual understanding

[11] According to the Common Core State Standards, "To help students meet the standards, educators will need to pursue, with equal intensity, three aspects of rigor in the major work of each grade: conceptual understanding, procedural skills and fluency, and application. Conceptual understanding: The standards call for conceptual understanding of key concepts, such as place value and ratios. Students must be able to access concepts from a number of perspectives in order to see math as more than a set of mnemonics or discrete procedures. Procedural skills and fluency: The standards call for speed and accuracy in calculation. Students must practice core functions, such as single-digit multiplication, in order to have access to more complex concepts and procedures. Fluency must be addressed in the classroom or through supporting materials, as some students might require more practice than others. Application: The standards call for students to use math in situations that require mathematical knowledge. Correctly applying mathematical knowledge depends on students having a solid conceptual understanding and procedural fluency" (CCSS, 2016).

and procedural fluency; and elicit and use evidence of student thinking (NCTM, 2014, p. 3).

An examination of the programs' features that correlate with respondents' levels of knowledge is provided below.

Examining Relationships Among the Variables

The analysis considers MCK and MPCK measured by the TEDS-M assessments as the key outcomes of teacher education programs and explores the association of individual-level (background) variables and program-level variables with these outcomes. Below is the HLM model used for the analysis for countries with large program samples. Table 8.2 shows the results.

HLM Model for Countries with Large Program Samples

STEP 1

Full Unconditional Model (Level 1 & 2)

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}, \text{where } u_{0j} \sim \mathrm{N}\left(0, \tau_{00}\right) \text{and } r_{ij} \sim \mathrm{N}\left(0, \sigma^2\right)$$

$$\mathrm{ICC} = \left(\tau_{00}\right) / \left(\tau_{00} + \sigma^2\right)$$

STEP 2

Level-1 Model

$$Y_{ij} = \beta_{0j} + \beta_{1j}\left(\mathrm{SES}\right)_{ij} + \beta_{2j}\left(\mathrm{MFA001}\right)_{ij} + \beta_{3j}\left(\mathrm{MFA002}\right)_{ij} + \beta_{4j}\left(\mathrm{MFA009}\right)_{ij} + r_{ij}, \text{where } r_{ij} \sim \mathrm{N}\left(0, \sigma^2\right)$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\left(\mathrm{MFB1GEOM}\right)_j + \gamma_{02}\left(\mathrm{MFB2SLMF}\right)_j + \gamma_{03}\left(\mathrm{MFB5READ}\right)_j + \gamma_{04}\left(\mathrm{MFB15COH}\right)_j + \gamma_{05}\left(\mathrm{MeanSES}\right)_j + \gamma_{06}\left(\mathrm{MFD1RULE}\right)_j + \gamma_{07}\left(\mathrm{MFD2ACTV}\right)_j + u_{0j}, \text{where } u_{0j} \sim \mathrm{N}\left(0, \tau_{00}\right)$$

$$\beta_{pj} = \gamma_{p0}$$

for $p = 1$ to 4 for the four level-1 predictors (which are fixed).

Table 8.2 HLM model of future primary teacher mathematics teaching knowledge given background and program characteristics

	Chile		Philippines		Poland		Russian Federation	
Variable	MCK	MPCK	MCK	MPCK	MCK	MPCK	MCK	MPCK
Future teacher background								
SES	0.04	0.01	0.06	−0.04	0.03+	0.03	0.01	−0.02
Age	−0.04	−0.01	−0.11*	−0.03	−0.05+	−0.12***	0.03+	0.04
Gender	−0.11***	−0.08*	−0.09	−0.01	−0.10***	−0.04*	−0.06*	−0.01
Prior attainment	0.17***	0.14**	0.08	0.07	0.14***	0.12***	0.08***	0.09***
Program level characteristics								
OTL: University level mathematics: Geometry	0.06	0.06	0.08	0.10	0.12*	0.11*	0.01	0.03
OTL: School level mathematics: function, probability and calculus	0.07	0.04	0.12+	−0.03	0.33***	0.19**	0.24**	0.03
OTL: Reading research on teaching and mathematics	0.10	−0.01	−0.07	−0.11+	−0.06	−0.03	0.04	0.08
Program coherence	0.00	−0.08	0.12	0.01	0.01	−0.01	0.30**	0.27**
Average SES for each program (aggregated from future teachers SES)	0.06	0.00	0.15+	0.08	0.07+	0.03	0.02	0.02

Future teacher beliefs (aggregated at program level)								
Mathematics is a collection of rules and procedures	−0.14	−0.03	−0.00	−0.14+	−0.12+	−0.10+	−0.28**	−0.24*
Mathematics is better learned through active learning	0.02	0.13**	0.03	−0.10	0.06	0.10*	−0.08	0.04
ICC (from the unconditional model)	21%	5%	16%	11%	59%	37%	46%	38%
% variance explained within programs	4%	1%	3%	0%	5%	3%	0%	0%
% variance explained between programs	7%	44%	61%	37%	85%	80%	36%	18%

	Spain		Thailand		United States	
Variable	MCK	MPCK	MCK	MPCK	MCK	MPCK
Future teacher background						
SES	−0.06	−0.01	0.10**	0.01	0.05+	0.08**
Age	0.07*	−0.04	−0.00	−0.05	−0.06	−0.10**
Gender	−0.21***	−0.07*	−0.18***	−0.10***	−0.16***	−0.04*
Prior attainment	0.29***	0.11**	0.20***	0.10*	0.18***	0.15***
Program level characteristics						
OTL: University level mathematics: Geometry	−0.01	−0.06+	−0.05	−0.07	0.03	−0.03
OTL: School level mathematics: function, probability and calculus	−0.03	0.01	0.12*	0.11+	0.02	0.03

(continued)

Table 8.2 (continued)

	Spain		Thailand		United States	
OTL: Reading research on teaching and mathematics	−0.01	0.04	−0.15**	−0.12*	0.02	0.02
Program coherence	−0.01	−0.05	0.02	0.00	−0.05	−0.05
Average SES for each program (aggregated from future teachers SES)	0.15***	0.07	0.08	0.12	0.21***	0.12***
Future teacher beliefs (aggregated at program level)						
Mathematics is a collection of rules and procedures	−0.06	−0.05	−0.22**	−0.17*	−0.09*	−0.09*
Mathematics is better learned through active learning	0.03	0.02	0.04	0.03	0.10*	0.04
ICC (from the unconditional model)	4%	2%	31%	22%	17%	6%
% variance explained within programs	15%	2%	8%	1%	5%	5%
% variance explained between programs	12%	53%	54%	55%	82%	99%

Note: Coefficients are standardized

Significant at * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; + not significant ($p > .05$ and $< .10$)

The coefficients in the model (which were standardized within each country and expressed in standardized betas, or β) were helpful in explaining variation in the knowledge assessments. For example, for Poland under the MCK column, there was a β equal to 0.33, meaning that a change of one standard deviation in the predictor variable (in this case, the *opportunity to learn functions, probability, and calculus*) was significantly associated ($p < .001$) with a 0.33 standard deviation change (gain) in the outcome variable (mathematics content knowledge or MCK after controlling for background variables). The ICCs (at the bottom of the tables) indicate the variance in MCK and MCPK attributable to differences between programs in each country; for the mathematics content assessment (MCK), the higher ICCs were in Poland, Russia, Thailand, Chile, and the United States (with 59%, 46%, 31%, 21%, and 17%, respectively). For the mathematics pedagogy assessment (MPCK), the higher ICCs were in Russia, Poland, and Thailand (with 38%, 37%, and 22%, respectively). With the exception of Poland for MCK, the proportion of variance explained was higher within programs than between programs.[12]

Below we show the OLS model for future primary teachers and programs in countries with small program samples. The results in Table 8.3 are given in standardized beta coefficients, or β.

OLS Model for Countries with Small Program Samples

The OLS estimation employs the repeated replicate weights to account for the sample complex design. Unlike HLM, OLS models do not include random coefficients between programs.

Partial Model

The model specification is similar from the one presented above (without random terms for the intercept). The model is expressed in the following single equation.

$$Y_i = \beta_0 + \beta_1 (\mathrm{SES})_i + \beta_2 (\mathrm{MFA001})_i + \beta_3 (\mathrm{MFA002})_i + \beta_4 (\mathrm{MFA009})_i + r_i$$

Full Model

The full model is expressed by the following single equation (again, random terms for the intercept are not considered; the beliefs variables are aggregated to the program level).

[12] The ICC reports the percent of variance between programs, where 100-ICC (100 minus the ICC) is the percent of variance within programs. This is always reported simply as the ICC for percent of variance between groups. The % of variance explained in the last two rows then simply states how much variance each model explains. In Table 8.2, in Poland, 59% of the variance in MCK performance is between programs, 41% is within programs. The model including all variables explains 5% of the variance within programs (i.e., student characteristics do not explain much of the variance in their performance within program); that is, the student characteristics explain 5% of the 41% variance within programs. The model also explains 85% of the variance between programs; that is, the program characteristics explain 85% of the 59% variance between programs.

$$Y_i = \beta_0 + \beta_1(\mathrm{SES})_i + \beta_2(\mathrm{MFA001})_i + \beta_3(\mathrm{MFA002})_i + \beta_4(\mathrm{MFA009})_i$$
$$+ \beta_5(\mathrm{MFB1GEOM})_i + \beta_6(\mathrm{MFB2SLMF})_i + \beta_7(\mathrm{MFB5READ})_i$$
$$+ \beta_8(\mathrm{MFB15COH})_i + \beta_9(\mathrm{MeanSES})_i + \beta_{10}(\mathrm{MFD1RULE})_i$$
$$+ \beta_{11}(\mathrm{MFD2ACTV})_i + \beta_{12}(\mathrm{TARGETp\_S})_i + r_i$$

For instance, in Table 8.3 it is possible to explore across countries the level of association between program coherence, MCK, and MPCK. In Malaysia, the coefficients are 0.19 for MCK and 0.17 for MPCK, showing a moderate positive and significant association ($p < .01$) between program coherence and the scores in both assessments after controlling for background characteristics of future teachers.

While findings are only considered significant at the $p \leq .05$, findings with p values >.05 and < .10 are identified by a "+" to make it possible to see whether the trends in the latter group are consistent with those found to be significant.

Future Primary Teachers' Background was associated with MCK and MPCK scores. The results of the analysis for future primary teachers in countries with large and small program samples show that individual characteristics such as SES, age, gender, and previous attainment were associated with higher scores in the MCK and MPCK assessments. Table 8.2 shows a small positive association between higher SES levels and higher levels of performance in the mathematics content knowledge assessment (MCK) in Thailand and in the United States in the mathematics pedagogical content knowledge assessment (MPCK). Regarding age, the analysis indicates that younger future teachers did better in the MCK assessments than their older counterparts in the Philippines, and better in the MCPK assessments in Poland and the United States. Similarly, males did significantly better than females in all countries in both assessments, with the exception of the Philippines and the Russian Federation (on the MPCK assessment). Attainment in school prior to entering the teacher education program had a significant and positive association with the scores obtained in both MCK and MPCK assessments across all countries, with the exception of the Philippines.

Table 8.3 shows the results of the analysis for future primary teachers in countries with small program samples. The association of SES with the level of MCK and MPCK is positive, but in general small and not significant. Across most of the countries, younger future teachers scored higher in the assessments; this association was significant but moderate in Botswana, Singapore, and Switzerland (for MPCK). Male future teachers did significantly better in the mathematics portion of the assessment (MCK) in Botswana, Chinese Taipei, Singapore, and Switzerland; only in Malaysia did females perform better than males. Females in Botswana, Chinese Taipei, and Germany did better on the mathematics pedagogy portion of the assessment (MPCK), but this association is small and not significant; in Malaysia, females did significantly better than their male counterparts. Prior attainment played a positive and significant role on the MCK and MPCK assessments results across all countries, with the exception of Botswana for the MCK.

Table 8.3 OLS model of future primary teacher mathematics teaching knowledge given background and program characteristics

	Botswana		Chinese Taipei		Germany	
Variable	MCK	MPCK	MCK	MPCK	MCK	MPCK
Future teacher background						
SES	0.16+	0.16+	0.01	−0.01	0.00	0.05
Age	0.07	−0.19+	−0.08**	−0.03	−0.03	−0.06+
Gender	−0.25*	0.18+	−0.11**	0.01	−0.05	0.02
Prior attainment	−0.01	0.23*	0.18**	0.16**	0.19**	0.20**
Program level characteristics						
OTL: Tertiary level geometry	−0.31	0.09	0.21	0.04	0.27**	0.18+
OTL: School level mathematics: functions, probability, calculus	0.05	0.04	−0.02	0.02	0.04	0.09
OTL reading research on teaching and mathematics	0.44	−0.10	.00	0.06	−0.02	−0.06
Program coherence	0.22	0.01	0.03	−0.05	−0.04	−0.04
Program SES	−0.49	−0.16	0.05	0.04	0.10*	0.08*
Future teacher beliefs (aggregated at program level)						
Mathematics as rules and procedures	0.06	−0.01	−0.04	−0.01	−0.13*	−0.09+
Mathematics as active learning	0.00	0.00	−0.04	−0.01	−0.02	0.06
R^2	.09	.16	.12	.05	.25	.22
F	0.45	0.88	11.26***	4.34***	26.15***	22.12***

(continued)

Table 8.3 (continued)

	Malaysia		Singapore		Switzerland	
Variable	MCK	MPCK	MCK	MPCK	MCK	MPCK
Future teacher background						
SES	0.02	0.09*	0.05	0.03	0.04	0.02
Age	0.00	−0.03	−0.10*	−0.17**	−0.07**	−0.12**
Gender	0.04	0.08*	−0.20**	−.06	−0.11**	−0.03
Prior attainment	0.10**	0.12**	0.14**	0.13**	0.23**	0.13**
Program level characteristics						
OTL: Tertiary level geometry	0.13**	0.07	0.06	0.04	0.12**	−0.01
OTL: School level mathematics: functions, probability, calculus	−0.11*	−0.03	−0.02	−0.02	−0.15**	−0.12*
OTL reading research on teaching and mathematics	−0.14**	0.03	−0.12	−0.11	−0.09**	−0.12**
Program coherence	0.19**	0.17**	−0.06	0.01	0.16*	0.03
Program SES	0.01	−0.05	0.16	0.26	0.03	0.00
Future teacher beliefs (aggregated at program level)						
Mathematics as rules and procedures	0.02	−0.05	−0.05	−0.09	0.03	−0.11**
Mathematics as active learning	0.12+	0.01	0.14	0.20	−0.10*	−0.11**
R^2	.06	.06	.12	.10	.15	.07
F	3.17***	3.17***	4.51***	3.67***	14.63***	6.24***

Note: Coefficients are standardized betas
Significant at * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; + not significant ($p > .05$ and $< .10$)

Opportunities to learn university and school mathematics in coherent programs were associated with higher scores in the MCK and MPCK assessments. Tables 8.2 and 8.3 indicate that differences across programs in the emphasis placed on OTL have an important and significant association on future primary teachers' performance on the TEDS-M assessments. For instance, in Poland, in type 4 programs where entry into teacher education was conditioned on specific mathematics requirements, and future primary teachers were given additional OTL university-level mathematics (geometry in particular), MCK and MPCK scores were about one-tenth of a standard deviation higher than in those programs that did not have those requirements and opportunities ($\beta = 0.12$, $p < .05$; $\beta = 0.11$, $p < .05$). Future primary teachers in programs that covered topics on school mathematics in the areas of functions, probability, and calculus also scored significantly higher on the MCK assessment in Thailand, Poland, and Russia (with up to one-third of a standard deviation higher in MCK in Poland and one-fourth in Russia). Having the opportunity to read research connected with mathematics teaching and learning, on the other hand, had a non-significant association with future teachers' performance on the assessments, with the exception of Thailand, where the association with both MCK and MCPK was negative and significant (meaning that those with high scores had less opportunity to read mathematics or mathematics education related research). In the Russian Federation, there was a strong and positive association between the level of program coherence and high scores on both assessments, with programs judged coherent by future teachers showing scores of about one-third of a standard deviation higher in MCK and above one-fourth of a standard deviation higher in MPCK ($\beta = 0.30$, $p < .001$; $\beta = 0.27$, $p < .01$).

Among the countries in Table 8.3, the opportunity to learn mathematics, and specifically geometry at the university or tertiary level in Germany was associated with high scores on the MCK and MPCK assessments ($\beta = 0.27$, $p < .01$; $\beta = 0.18$, $p < .10$, n.s.); a more moderate but positive and significant association on the MCK score was seen in Malaysia ($\beta = 0.13$, $p < .01$), and Switzerland ($\beta = 0.12$, $p < .01$). A negative association between the MCK and the MPCK scores and opportunities to learn school-level mathematics existed in Malaysia, Singapore, and Switzerland, presumably because future primary teachers who had a firm knowledge of these domains did not engage in studying these areas during their teacher education program.

The level of program coherence was positively associated with higher scores in both assessments in Malaysia, and with high MCK scores in Switzerland. Future teachers with high MCK scores were also more likely to be in programs that did not provide opportunities to read research on mathematics and mathematics teaching and learning (with βs ranging from −0.14 to −0.09 in Malaysia and Switzerland, respectively).

The aggregated socioeconomic background of future teachers within programs was positively associated with higher scores for individual teachers in the MCK and MPCK assessments. Table 8.2 shows that overall, the aggregated socioeconomic background of future teachers within programs was positively and significantly associated with future teachers' performance. This association was significant in Spain for MCK ($\beta = 0.15$, $p < .001$), and in the United States concerning MCK and MPCK performance in ($\beta = 0.21$, $p < .001$; $\beta = 0.12$, $p < .001$). In Table 8.3, programs' aggregated socioeconomic background was significant in Germany, for both MCK and MPCK ($\beta = 0.10$, $p < .05$; $\beta = 0.08$, $p < .05$).

Beliefs about teaching and learning mathematics as a collection of rules and routines were associated with lower scores in the assessments. Table 8.2 shows a general negative association between performance in both assessments and the view that mathematics can be learned by mastering a collection of rules and procedures; this association was significant for MCK and MPCK in Russia, Thailand, and the United States. There is a positive association between the scores in the assessments and the view that mathematics can be better learned through active inquiry for MPCK in Chile and Poland, and for MCK in the United States. Countries in Table 8.3 show a similar trend, with significant negative associations between the assessment results and the view that mathematics can be learned by mastering a collection of rules and procedures in Germany for MCK and in Switzerland for MPCK. There was a negative and significant association between performance in the MCK and the MPCK assessment and the belief that mathematics is better learned through inquiry and active learning in Switzerland.

Implications of these findings for teacher education policy and practice are discussed below.

Discussion and Conclusions

The research reported in this chapter confirms at a larger scale findings from other studies, highlights emphasis areas for mathematics teacher education, and lends support to a number of policy recommendations that have emerged from the sustained research program maintained by the mathematics education community over the years (Ball, 2003; CBMS, 2012; Kilpatrick & Swafford, 2002; National Commission on Mathematics and Science Teaching for the 21st Century, 2000; NCTM, 2014; NRC, 2010).

The findings from this study highlight the need for teacher education programs to move beyond providing future teachers with the mathematics knowledge that supports their ability to engage with basic computations, operations with whole numbers, simple figures, and linear expressions, toward providing mathematics knowledge that supports teachers' ability to engage with abstract, complex, and multiple-steps problems, and to understand mathematical concepts and how these relate to one another. This, however, may be difficult without raising entry-level standards to demonstrate mathematics competence. University-level knowledge of geometry and school-level knowledge of algebra (functions, probability, and calculus) were particularly strong among those who performed well in the assessments. These findings suggest that these subjects can be used as leverage to support the mathematics knowledge needed to teach the primary school curriculum. This finding is consistent with Baumert et al. (2010), who found that deficiencies in mathematics content knowledge were reflected in weak mathematics pedagogical content knowledge, a situation that impeded future teachers' ability to interpret students' work and support students' learning, because of their failure to recognize common misconceptions, or conceive of useful representations, especially in the context of the mathematics they found challenging, such as algebraic representations and geometry concepts and applications.

The inclusion in TEDS-M of countries that had two types of programs allowed the possibility of exploring different approaches to prepare future primary mathematics teachers. Those future teachers in programs that targeted the upper-primary grades, those in specialist programs, and those that required mathematics knowledge as a condition for entry had significantly better results in the TEDS-M assessments. Thus, even within the same country, different curriculum emphases are associated with significant differences in performance on assessments (e.g., in Poland, 1.5 standard deviation difference in the MCK assessment and close to a 1.25 standard deviation in the MPCK assessment).

Given the findings from this study, teacher education programs that produce highly knowledgeable primary mathematics teachers can be characterized in the following ways:

- They require a specific level of mathematics knowledge as a condition to enter the teacher education program.
- They provide ambitious opportunities to learn over a significant time period of course work, including
  - mathematics topics (such as algebra and geometry) that serve to leverage other mathematics areas to support deep understandings of the mathematics needed to teach the primary school curriculum;
  - school mathematics topics that provide curricular knowledge as well as deeper understanding of the mathematics to enable teachers to address non-routine problems and to access higher levels of complexity and abstraction;
  - and mathematics pedagogy that allows teachers to understand student thinking, plan for instruction, and build fluency and conceptual understanding.
- They challenge future teachers' naïve views about teaching and learning mathematics, moving their understanding of mathematics learning from mastering rules and procedures toward balancing the learning of procedures with ways to engage students with processes of inquiry.

These propositions are addressed briefly below. The conclusion addresses challenges for the future of mathematics teacher education and for the mathematics education research community in an era of increased accountability.

Mathematics Knowledge As a Condition to Enter Teacher Education

A key finding from this study is that future teachers who did well in their previous schooling performed better in the mathematics knowledge assessments than those who did not. Given that previous attainment is an important predictor of future attainment, programs that have less rigorous selection policies are evidently unable to compensate for poor preparation in mathematics; weak previous attainment, as we have seen, also results in inadequate preparation in mathematics pedagogy. Those programs in countries that emphasize a minimum standard of mathematics knowledge as a condition for program entry are better able to prepare their future teachers (as demonstrated by their performance in the assessments) than those programs that have no such requirement. For instance, in Chinese Taipei, candidates need to have one year of tertiary-level studies, plus passing scores in a national examination in order to enter a teacher education program with mathematics as a required subject. Other high performing countries such as Poland (for upper-primary teachers), the Russian Federation, and Singapore require a mathematics credential or equivalent upon graduation from upper-secondary school in order to enter a teacher education program in university. In contrast, graduation from upper-secondary school with no specific mathematics requirement was the only requirement to enter a teacher education program in Chile, Germany, Malaysia, the Philippines, Spain, Switzerland, Thailand, and the United States (Tatto et al., 2012, pp. 41–46). Based on these results, teacher education policy can help improve the quality of primary mathematics teachers by increasing the level of subject matter knowledge required to enter a teacher education program—for instance, by requiring a minor in mathematics, by preparing mathematics specialists for primary schools (e.g., by requiring a mathematics major), or by including more mathematics courses within teacher education programs. Teacher education programs have the choice to increase entry standards or to adjust their programs to provide remedial courses for the weak mathematics preparation future primary teachers may bring with them to their programs; if the latter, knowledge exit requirements should be raised.

Opportunities to Learn Ambitious Mathematics over a Significant Period of Coursework

Opportunity to learn university-level mathematics such as geometry, and school-level mathematics such as algebra (functions, probability and calculus) proved to be important to do well in both assessments. Yet most programs for future primary teachers provide extensive opportunities to learn school mathematics such as number and measurement at the expense of providing opportunities to learn mathematics for instance in the areas of function, probability, and calculus. A recommendation for the teacher education community is that the preparation of future primary teachers include the study of university-level mathematics that supports understanding of mathematics of the school curriculum such as geometry and algebra—that is, understanding that would raise teachers to MCK-AP2 according to the anchor point analysis presented in this chapter. Learning the mathematics of the school curriculum must include particular attention to reaching the cognitive levels that currently elude a large proportion of future primary teachers. Important areas for improving the mathematical and mathematical pedagogical content knowledge of future primary teachers can be identified in the AP descriptions (Tatto et al., 2012, pp. 135–142): for MCK, "reasoning about factors, multiples and percentages to describing properties of a linear function," and for MPCK, "interpretation of students' work," the use of "concrete representations to support students' learning," "knowing why a particular teaching strategy made sense, if it would always work, or whether a strategy could be generalized to a larger class of problems." These are knowledge and skills that can only be acquired through a sustained period of course work. An important observation is needed here. While the *duration* of teacher preparation does not necessarily imply *better* preparation (see Tatto, 2008), becoming a highly knowledgeable mathematics teacher does require extensive study. The countries that did well in the TEDS-M study (i.e., Chinese Taipei, Germany, Poland (for specialists), Russian Federation, Singapore, and the United States) had programs no shorter than 4 years, with most lasting 5 years. In Germany, programs were even longer, from 5.5 to 6.5 years (Tatto et al., 2012, pp. 29–32).

As presented in the results section, with some exceptions (in Chinese Taipei, the Russian Federation, and the United States), those teachers who had high scores in the mathematics content portion of the assessment also reported infrequently engaging in courses that gave them OTL MPCK about research on mathematics and mathematics education, and on teaching and learning. While this finding merits exploration in future work (using the TEDS-M data collected from the teacher education curriculum analysis), it raises the question of time allocation and balance of content in teacher education. This concern finds resonance with Ball, Lubienski, and Mewborn (2001), who argued that while knowledge of mathematics is essential it should not come at the expense of knowledge that may support thoughtful and insightful mathematics practices or opportunities to do systematic investigations on mathematics teaching and learning. This finding also relates to the question of program coherence, defined as the degree to which programs succeed at providing an internally consistent, logical set of courses, and experiences as preparation to meet a common set of standards and expectations when beginning to teach. There is a positive relationship between coherence (especially as reflected in a consistent and sound curriculum and university-school collaboration) and knowledge outcomes—notably, in Malaysia and the Russian Federation as well as Switzerland. This may suggest the need to construct teacher education practices around a core curriculum, as occurs in Finland (see Sahlberg, 2007, 2010), Chinese Taipei, Russia, and other high-performing countries.

Challenging Future Teachers' Beliefs About Teaching and Learning Mathematics

There was a significant association in some countries between future teachers' beliefs and their demonstrated level of knowledge in both assessments. Specifically, lower levels of performance in the assessments were found in those programs in which future teachers, as a group, held the view that learning mathematics consists of memorizing a collection of rules and procedures that prescribe how to solve a problem, and the application of definitions, formulas, mathematical facts, and procedures. The finding that there were programs with very traditional views about the nature of mathematics at the expense of more balanced views (e.g., that learning facts and procedures and learning through active inquiry are equally important) needs more exploration—for instance, to understand whether these are views that the programs are especially promoting in an era of increased testing and accountability, against views that are more consistent with cognitive science research on mathematics learning (Boaler, 2016).

Because the TEDS-M study did not collect entry data, it is not possible to ascertain whether teachers' beliefs were the products of their programs alone. The traditional view that mathematics can be seen as a series of rules and procedures that need mastering, however, is widespread and may reflect how future teachers originally learned to approach mathematics. While it is possible to assume from these results that higher levels of MCK and MPCK contribute to the development of teachers' deeper understandings of mathematics, and how these can be applied to teaching, more research needs to be done in this area to confirm such assumptions, and more importantly to explore the particular OTL more proficient teachers received (see Chapter 6 of this book for an exploration of program influence on future teachers beliefs).

Building the Research Basis for the Self-Study of Teacher Education Programs in an Era of Accountability

The research presented in this chapter contributes to the research evidence of teacher education program outcomes, work that is much needed in the new climate of accountability, and which requires from programs the development of self-study strategies in order to comply with high-stakes accreditation demands. To meet these demands, it is important to understand which of the many possible areas of study should be the focus of program self-study efforts, and what methods to use to collect valid and reliable information. Because the TEDS-M study makes the instruments, methods, and procedures, as well as the data, publicly available, teacher education programs can learn much from it about how to carry out rigorous studies of teacher education programs. Thus, this study contributes to building the research basis for the self-study of teacher education programs by and for teacher educators, and with usable results for the teacher education community.

Appendices

Table 8.4 Variables in the general models: future teacher characteristics and program characteristics

Future teachers characteristics
Socioeconomic status [SES]
A scale was derived from the TEDS-M data using principal components analysis as a composite of a home possessions index (derived using IRT), number of books at home, father's highest level of education and mother's highest level of education. The final SES measure was standardized to have a mean of 0 and a standard deviation of 1.
Age [MFA001] *Mean of age as reported*
Gender [MFA002] *Mean of age as reported (1 = female, 0 = male)*
Prior attainment [MFA009] *Mean for level of grades in secondary school (1 = below average for year level; 5 = Always at top of year level)*
Program characteristics
Opportunities to learn
Tertiary level mathematics – geometry [MFB1GEOM]
A. Foundations of geometry or axiomatic geometry (e.g., Euclidean axioms)
B. Analytic/coordinate geometry (e.g., equations of lines, curves, conic sections, rigid transformations or isometrics)
C. Non-Euclidean geometry (e.g., geometry on a sphere)
D. Differential geometry (e.g., sets that are manifolds, curvature of curves, and surfaces)
This variable represents counts of topics studied. For this study the variable was formed by calculating the mean value for each program aggregated at the program level (range 0–4).
School level mathematics - function probability calculus [MFB2SLMF]
D. Functions, relations, and equations (e.g., algebra, trigonometry, analytic geometry)
E. Data representation, probability, and statistics
F. Calculus (e.g., infinite processes, change, differentiation, integration)
G. Validation, structuring, and abstracting (e.g., Boolean algebra, mathematical induction, logical connectives, sets, groups, fields, linear space, isomorphism, homomorphism)
This variable represents counts of topics studied. For this study the variable was formed by calculating the mean value for each program aggregated at the program level (range 0–4).
Mathematics education pedagogy - class reading [MFB5READ]
H. Read about research on mathematics
I. Read about research on mathematics education
J. Read about research on teaching and learning
K. Analyze examples of teaching (e.g., film, video, transcript of lesson)
This variable was formed by calculating the mean value for each program aggregated at the program level. This variable is a scaled score with mean = 10 and the standard deviation = 1, with 10 representing the neutral point in a Likert type scale (e.g., neither agree nor disagree).
Program coherence [MFB15COH]
A. Each stage of the program seemed to be planned to meet the main needs I had at that stage of my preparation.
B. Latter <courses> in the program built on what was taught in earlier <courses> in the program.
C. The program was organized in a way that covered what I needed to learn to become an effective teacher.

(continued)

Table 8.4 (continued)

D. The <courses> seemed to follow a logical sequence of development in terms of content and topics.
E. Each of my <courses> was clearly designed to prepare me to meet a common set of explicit standard expectations for beginning teachers.
F. There were clear links between most of the <courses> in my teacher education program.
This variable was formed by calculating the mean value for each program aggregated at the program level. This variable is a scaled score with mean = 10 and the standard deviation = 1, with 10 representing the neutral point in a Likert type scale (e.g., neither agree nor disagree).
Program philosophy (views)
Rules and procedures [MFD1RULE]
1. Mathematics is a collection of rules and procedures that prescribe how to solve a problem.
2. Mathematics involves the remembering and application of definitions, formulas, mathematical facts and procedures.
3. When solving mathematical tasks you need to know the correct procedure else you would be lost.
4. Fundamental to mathematics is its logical rigor and preciseness.
5. To do mathematics requires much practice, correct application of routines, and problem solving strategies.
6. Mathematics means learning, remembering and applying.
This variable was formed by calculating the mean value for each program aggregated at the program level. This variable is a scaled score with mean = 10 and the standard deviation = 1, with 10 representing the neutral point in a Likert type scale (e.g., neither agree nor disagree).
Active learning [MFD2ACTV]
1. In addition to getting a right answer in mathematics, it is important to understand why the answer is correct.
2. Teachers should allow pupils to figure out their own ways to solve mathematical problems.
3. Time used to investigate why a solution to a mathematical problem works is time well spent.
4. Pupils can figure out a way to solve mathematical problems without a teacher's help.
5. Teachers should encourage pupils to find their own solutions to mathematical problems even if they are inefficient.
6. It is helpful for pupils to discuss different ways to solve particular problems.
This variable was formed by calculating the mean value for each program aggregated at the program level. This variable is a scaled score with mean = 10 and the standard deviation = 1, with 10 representing the neutral point in a Likert type scale (e.g., neither agree nor disagree).
Program's SES [SES]
This variable was formed by calculating the mean value for each program to get the SES measure aggregated at the program level.

Table 8.5 Descriptive statistics for future teacher characteristics and program characteristics[a] scale mean scores and standard deviations for variables used in the analysis for primary programs (Large Program Samples)

	Chile		Philippines		Poland		Russian Fed.		Spain		Thailand		United States	
	$N_{Level\ 1} = 621$		$N_{Level\ 1} = 550$		$N_{Level\ 1} = 1822$		$N_{Level\ 1} = 2162$		$N_{Level\ 1} = 1021$		$N_{Level\ 1} = 644$		$N_{Level\ 1} = 1073$	
	$N_{Level\ 2} = 31$		$N_{Level\ 2} = 32$		$N_{Level\ 2} = 125$		$N_{Level\ 2} = 49$		$N_{Level\ 2} = 43$		$N_{Level\ 2} = 52$		$N_{Level\ 2} = 78$	
Future teachers primary (level 1)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Mathematics content knowledge score [MCK] (mean = 500, sd = 100)	413.73	64.22	442.42	50.76	476.14	88.91	536.37	94.07	479.76	56.41	527.39	75.39	515.73	69.54
Mathematic pedagogy content knowledge score [MPCK] (mean = 500; sd = 100)	424.20	91.03	462.59	65.14	467.74	95.04	511.87	89.06	491.81	62.45	506.61	69.90	543.43	67.50
Socioeconomic status scale [SES] (mean = 0, sd = 1)	−0.28	0.86	−0.69	0.93	−0.29	0.73	0.62	0.70	−0.37	1.06	−0.90	1.14	0.44	0.84
Age [MFA001] (mean age)	23.59	2.56	20.92	1.79	24.95	5.33	28.88	4.75	23.04	4.14	22.26	0.78	25.46	6.45
Gender proportion [MFA002] 1 = female; 0 = male)	.85	.36	.79	.40	.95	.22	.93	.26	.80	.40	.75	.43	.92	.28
Prior attainment [MFA009]: Average grades in secondary school for year level (1 = below average; 5 = always at top)	3.18	1.15	2.76	0.82	2.80	0.84	3.37	0.92	2.77	1.14	3.34	0.85	3.49	1.00
Teacher education programs primary (level 2)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Number of university level mathematics topics in geometry ever studied [MFB1GEOM] (average number of topics range 0–4)	1.91	0.49	2.82	0.42	2.29	0.75	2.06	0.45	2.20	0.38	3.40	0.41	1.88	0.71
Number of school level mathematics topics in function, probability and calculus studied as part of the TE program [MFB2SLMF] (average number of topics range 0–4)	1.48	0.45	2.36	0.29	1.80	1.40	2.22	0.50	2.00	0.39	3.59	0.34	2.13	0.47

(continued)

Table 8.5 (continued)

	Chile		Philippines		Poland		Russian Fed.		Spain		Thailand		United States	
	$N_{Level\ 1} = 621$		$N_{Level\ 1} = 550$		$N_{Level\ 1} = 1822$		$N_{Level\ 1} = 2162$		$N_{Level\ 1} = 1021$		$N_{Level\ 1} = 644$		$N_{Level\ 1} = 1073$	
	$N_{Level\ 2} = 31$		$N_{Level\ 2} = 32$		$N_{Level\ 2} = 125$		$N_{Level\ 2} = 49$		$N_{Level\ 2} = 43$		$N_{Level\ 2} = 52$		$N_{Level\ 2} = 78$	
Future teachers primary (level 2)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Frequency with which future teachers engaged in reading research on teaching and mathematics [MFB5READ] (average scale mean = 10 or neutral point, sd = 1)	9.38	0.80	10.80	0.49	8.12	0.85	10.70	0.88	8.67	0.94	10.10	0.75	10.40	1.13
Average SES for each program (aggregated from future teachers SES) [SES]	−0.19	0.51	−0.71	0.34	−0.25	0.34	0.61	0.21	−0.37	0.31	−0.88	0.49	0.44	0.33
Degree of program coherence [MFB15COH] (average scale mean = 10 or neutral point, sd = 1)	11.94	1.23	13.86	1.04	11.29	0.97	13.39	0.80	10.19	0.86	13.07	0.98	12.97	1.49
Agreement with the belief that mathematics is a collection of rules and procedures [MFD1RULE] (average scale mean = 10 or neutral point, sd = 1)	10.82	0.46	12.65	0.51	10.86	0.63	10.72	0.35	10.74	0.30	11.93	0.61	11.03	0.47
Agreement with the belief that mathematics is better learned through active learning [MFD2ACTV] (average scale mean = 10 or neutral point, sd = 1)	12.61	0.45	11.80	0.34	12.10	0.60	11.77	0.36	11.75	0.41	11.82	0.36	12.02	0.66

[a]Note: This table shows the descriptive statistics for the HLM analysis. In this study level 1 represents the number of future teachers in a country within programs; and level 2 represents the number of programs. Missing data across measures accounts for the reduced number of cases from those shown in Tables 8.2 and 8.3 representing the sample sizes. See Appendix Table 8.4 for a detailed description of the variables included in the model and how they were constructed

Table 8.6 Descriptive statistics for future teacher characteristics and program characteristics[a] scale mean scores and standard deviations for variables used in the analysis for primary programs (Small Program Samples)

	Botswana		Chinese Taipei		Germany		Malaysia		Singapore		Switzerland	
	$N_{Level 1} = 63$		$N_{Level 1} = 921$		$N_{Level 1} = 875$		$N_{Level 1} = 559$		$N_{Level 1} = 376$		$N_{Level 1} = 924$	
	$N_{Level 2} = 4$		$N_{Level 2} = 11$		$N_{Level 2} = 14$		$N_{Level 2} = 23$		$N_{Level 2} = 1$		$N_{Level 2} = 14$	
Future teachers primary (level 1)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Mathematics content knowledge score [MCK] (mean = 500, sd = 100)	448.94	47.19	623.77	84.10	510.57	86.01	489.31	53.95	590.70	73.97	543.30	65.52
Mathematic pedagogy content knowledge score [MPCK] (mean = 500; sd = 100)	463.69	65.02	592.48	67.91	499.75	89.44	503.33	67.52	593.22	71.20	537.61	63.47
Socioeconomic status scale [SES] (mean = 0, sd = 1)	−0.48	1.33	−0.36	0.93	0.53	0.90	−0.74	0.89	−2.19	2.44	0.23	.90
Age [MFA001] (mean age)	25.25	4.90	23.16	2.11	27.32	4.04	25.88	2.34	26.68	4.73	23.64	3.60
Gender proportion [MFA002] 1 = female; 0 = male)	.62	.49	.72	.45	.92	.26	.64	.48	.74	.44	.85	.35
Prior attainment [MFA009]: Average grades in secondary school for year level (1 = below average; 5 = always at top)	2.92	0.75	3.23	1.11	2.53	0.85	3.64	1.01	3.03	0.90	2.89	0.96
Teacher education programs primary (level 2)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Number of university level mathematics topics in geometry ever studied [MFB1GEOM] (average number of topics range 0–4)	1.58	0.29	2.05	0.15	1.05	0.63	2.35	0.35	1.31	.037	2.18	0.18
Number of school level mathematics topics in function, probability and calculus studied as part of the TE program [MFB2SLMF] (average number of topics range 0–4)	2.37	0.37	1.96	0.20	1.10	0.63	2.36	0.34	1.51	0.33	1.09	0.40
Frequency with which future teachers engaged in reading research on teaching and mathematics [MFB5READ] (average scale mean = 10 or neutral point, sd = 1)	10.10	0.57	8.97	0.64	7.61	1.10	10.61	0.54	9.44	0.44	8.89	1.06
Average SES for each program (aggregated from future teachers SES) [SES]	12.65	0.48	11.46	0.37	9.27	0.49	13.10	0.57	12.68	0.52	10.18	0.42

(continued)

Table 8.6 (continued)

	Botswana		Chinese Taipei		Germany		Malaysia		Singapore		Switzerland	
	$N_{Level\ 1} = 63$		$N_{Level\ 1} = 921$		$N_{Level\ 1} = 875$		$N_{Level\ 1} = 559$		$N_{Level\ 1} = 376$		$N_{Level\ 1} = 924$	
	$N_{Level\ 2} = 4$		$N_{Level\ 2} = 11$		$N_{Level\ 2} = 14$		$N_{Level\ 2} = 23$		$N_{Level\ 2} = 1$		$N_{Level\ 2} = 14$	
Future teachers primary (level 2)	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Degree of program coherence [MFB15COH] (average scale mean = 10 or neutral point, sd = 1)	−0.46	0.05	−0.36	0.10	0.53	0.16	−0.74	0.16	−0.43	0.05	0.23	0.21
Agreement with the belief that mathematics is a collection of rules and procedures [MFD1RULE] (average scale mean = 10 or neutral point, sd = 1)	12.01	0.32	10.73	0.15	9.98	0.25	11.74	0.45	11.04	0.17	9.98	0.20
Agreement with the belief that mathematics is better learned through active learning [MFD2ACTV] (average scale mean = 10 or neutral point, sd = 1)	11.98	0.17	12.12	0.11	12.32	0.42	11.32	0.37	11.75	0.20	12.39	0.34

[a]Note: This table shows the descriptive statistics for the OLS analysis. In this study level 1 represents the number of future teachers in a country within programs; and level 2 represents the number of programs. Missing data across measures accounts for the reduced number of cases from those shown in Tables 8.2 and 8.3 representing the sample sizes. See Appendix Table 8.4 for a detailed description of the variables included in the model and how they were constructed

References

AERA, APA, & NCME. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher, 36*, 369–387.

An, S., Kulm, G., & Wu, Z. (2004). The pedagogical content knowledge of middle school mathematics teachers in China and the U.S. *Journal of Mathematics Teacher Education, 7*, 145–172.

Ball, D. L. (1990a). The mathematical understandings that prospective teachers bring to teacher education. *The Elementary School Journal, 90*, 449–466.

Ball, D. L. (1990b). Prospective elementary and secondary teachers' understanding of division. *Journal for Research in Mathematics Education, 21*, 132–144.

Ball, D. L. (1991). Research on teaching mathematics: Making subject matter part of the equation. In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 2, pp. 1–48). Greenwich, CT: JAI Press.

Ball, D. L. (Chair). (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education, RAND mathematics study panel*. Washington, DC: U.S. Department of Education's Office of Educational Research and Improvement. Retrieved from http://www.rand.org/pubs/research_briefs/RB8023/index1.html

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on the teaching and learning of mathematics* (pp. 83–104). Westport, CT: Ablex.

Ball, D. L., Lubienski, S., & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed.). New York, NY: Macmillan.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., … Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180.

Begle, E. G. (1979). *Critical variables in mathematics education: Findings from a survey of empirical literature*. Washington, DC: Mathematics Association of America and the National Council of Teachers of Mathematics.

Blomeke, S., Suhl, U., Kaiser, G., & Dohrmann, M. (2012). Family background, entry selectivity and opportunities to learn: What matters in primary teacher education? An international comparison of fifteen countries. *Teaching and Teacher Education, 20*(1), 44–55.

Boaler, J. (2016). *Mathematical mindsets*. San Francisco, CA: Jossey-Bass.

Boero, P., Dapueto, C., & Parenti, L. (1996). Didactics of mathematics and the professional knowledge of teachers. In A. J. Bishop, M. A. Clements, C. Keitel, J. Patrick, & F. K. S. Leung (Eds.), *International handbook of mathematics education* (pp. 1097–1122). Dordrecht, The Netherlands: Kluwer.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods Research, 17*, 303–316.

Boyd, D., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., McDonald, M., … Wyckoff, J. (2008). Surveying the landscape of teacher education in New York City: Constrained variation and the challenge of innovation. *Educational Evaluation and Policy Analysis, 30*, 319–343.

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wycoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 416–440.

Brese, F., & Tatto, M. T. (Eds.). (2012). *User guide for the TEDS-M international database.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Cochran-Smith, M., & Zeichner, K. (Eds.). (2015). *Studying teacher education: The report of the AERA panel on research and teacher education*. Mahweh, NJ: Lawrence Erlbaum Publishers.

CCSS (Common Core State Standards). (2016). *Key shifts in mathematics*. Retrieved from http://www.corestandards.org/other-resources/keyshifts-in-mathematics

Clements, M. A., Bishop, A. J., Keitel, C., Kilpatrick, J., & Leung, F. K. S. (2013). *Third international handbook of mathematics education*. New York, NY: Springer.

Coggshall, J. G., Bivona, L., & Reschly, D. J. (2012). *Evaluating the effectiveness of teacher preparation program to support accountability*. National Comprehensive Center for Teacher Quality. Retrieved from http://files.eric.ed.gov/fulltext/ED543773.pdf

Comiti, C., & Ball, D. L. (1996). Preparing teachers to teach mathematics: A comparative perspective. In A. J. Bishop et al. (Eds.), *International handbook of mathematics education* (pp. 1123–1153). Dordrecht, The Netherlands: Kluwer.

Conference Board of the Mathematical Sciences (CBMS). (2012). *The mathematical education of teachers II*. Providence, RI/Washington, DC: American Mathematical Society/Mathematical Association of America. Retrieved from http://cbmsweb.org/MET2/

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification*. Retrieved from http://www.mathematica-mpr.com/publications/pdfs/Education/teacherstrained09.pdf

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives, 8*, 1.

Darling-Hammond, L. (2006). *Powerful teacher education: Lessons from exemplary programs*. San Francisco, CA: Jossey-Bass.

De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: The Guilford Press.

DeCorte, E., Op't Eynde, P., & Verschaffel, L. (2002). Knowing what to believe: The relevance of students' mathematical beliefs for mathematics education. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 297–320). Mahwah, NJ: Erlbaum Associates.

Delaney, S. (2012). A validation study of the use of mathematical knowledge for teaching measures in Ireland. *ZDM, 44*(3), 427–441.

Delaney, S. F., Ball, D. L., Hill, H. C., Schilling, S. G., & Zopf, D. A. (2008). Adapting U.S. measures of "mathematical knowledge for teaching" for use in Ireland. *Journal of Mathematics Teacher Education, 11*, 171–197.

Delors, J. (1996). *Learning, the treasure within. Report of the international commission on education for the twenty-first century*. Paris, France: UNESCO.

Deng, Z. (1995). *Estimating the reliability of the teacher questionnaire used in the Teacher Education and Learning to Teach (TELT)*. (National Center for Research on teacher learning technical series, 95, 1). Retrieved from http://ncrtl.msu.edu/HTTP/TSeries/TS%2095-1.pdf

Develay, M. (1998). Didactique et pédagogie. In J. C. Ruano-Borbalan (Ed.), *Éduquer et former*. Paris, France: Éditions Sciences Humaines.

Even, R., & Ball D. L. (Eds.). (2009). *The professional education and development of teachers of mathematics: The 15th ICMI Study* (New ICMI study series, Vol. 11). New York, NY: Springer.

Even, R., & Tirosh, D. (2002). Teacher knowledge and understanding of students' mathematical learning. In L. English (Ed.), *Handbook of international research in mathematics education* (pp. 219–240). Mahwah, NJ: Laurence Erlbaum.

Fennema, E., & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147–164). New York, NY: Macmillan.

Floden, R. (2002). The measurement of opportunity to learn. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 231–266). Washington, DC: National Academy Press.

Ginsburg, A. S. L., Anstrom, T., & Pollock, E. (2005). *What the United States can learn from Singapore's world-class mathematics system*. Washington, DC: American Institutes for Research.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*, 361–396.

Grigutsch, S., Raatz, U., & Törner, G. (1998). Einstellungen gegenüber Mathematik bei Mathematiklehrern (mathematics teachers' epistemological beliefs about the nature of mathematics). *Journal für Mathematik-Didaktik (Journal of Mathematics Education), 19*, 3–45.

Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal, 45*(1), 184–205.

Grouws, D. A. (Ed.). (1992). *Handbook of research on mathematics teaching and learning: A project of the National Council of teachers of mathematics*. New York, NY: MacMillan.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemologies: The psychology of beliefs about knowledge and knowing* (pp. 169–190). Mahwah, NJ: Erlbaum Associates.

Henry, G. T., Bastian, K. C., & Smith, A. A. (2012). Scholarships to recruit the "best and brightest" into teaching: Who is recruited, where do they teach, how effective are they, and how long do they stay? *Educational Researcher, 41*(3), 83–92.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study* (NCES 2003–013). Washington, DC: U.S. Department of Education.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Charlotte, NC: Information Age Publishing.

Hill, H., & Ball, D. L. (2009). The curious—and crucial—Case of mathematical knowledge for teaching. *Phi Delta Kappan, 91*(2), 68–71.

Hill, H., Sleep, C. L., Lewis, J. M., & Ball, D. L. (2007). Assessing teachers' mathematical knowledge: What matters and what evidence counts? In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 111–155). Charlotte, NC: Information Age.

Hill, H. C. (2007). Mathematical knowledge of middle school teachers: Implications for the no child left behind policy initiative. *Educational Evaluation and Policy Analysis, 29*, 95–114.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Hu, L.-T., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Husen, T. (Ed.). (1967). *International study of achievement in mathematics: A comparison of twelve countries* (Vol. 1). New York, NY: Wiley.

Ingvarson, L., Beavis, A., Danielson, C., Ellis, L., & Elliott, A. (2005). *An evaluation of the Bachelor of Learning Management at Central Queensland University*. Canberra, Australia: Australian Government Department of Education, Science and Technology.

Ingvarson, L., Beavis, A., & Kleinhenz, E. (2007). Factors affecting the impact of teacher education courses on teacher preparedness: Implications for accreditation policy. *European Journal of Teacher Education, 30*(4), 351–381.

Kennedy, M. (2016). Parsing the practice of teaching. *Journal of Teacher Education, 67*(1), 6–17.

Kilpatrick, J., & Swafford, J. (Eds.). (2002). *Helping children learn mathematics*. Washington, DC: National Academy Press. Retrieved from www.nap.edu/books/0309084318/html

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Lappan, G. (2000). A vision of learning to teach for the 21st century. *School Science and Mathematics, 100*(6), 319–326.

Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago, IL: University of Chicago Press.

Luschei, T. F. (2011). In search of good teachers: Patterns of teacher quality in two Mexican states. *Comparative Education Review, 56*, 69–97.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 102*, 391–410.

Martin, M. O., & Mullis, I. V. S. (2008). *Ensuring comparative validity: Quality control in IEA studies*. Retrieved from http://www.iea.nl/fileadmin/user_upload/General_Assembly/49th_GA/GA49_ensuring_comparative_validity.pdf

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis, 17*(3), 305–322.

Mewborn, D., & Tyminski, A. (2006). Lortie's apprenticeship of observation revisited. *For the Learning of Mathematics, 26*(3), 30–32.

Mewborn, D. S., & Stinson, D. W. (2007). Learning to teach as assisted performance. *Teachers College Record, 109*, 1457–1487.

Meyer, J. P. (2011). *jMetrik (version 2.1)* [computer software]. Charlottesville, VA: University of Virginia. Retrieved from http://www.itemanalysis.com

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*(2), 125–145.

Mullens, J. E., Murnane, R. J., & Willett, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review, 40*(2), 139–157.

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Boston College.

National Center for Analysis of Longitudinal Data in Education Research. (2012). *CALDER conversations, topic 1: Evaluating teacher training programs* [online discussion]. Washington, DC: Author. Retrieved from http://www.caldercenter.org/calder-conversations-tpps.cfm

National Commission on Mathematics and Science Teaching for the 21st Century. (2000). *Before it's too late: A report to the nation*. Retrieved from http://www.ed.gov/americacounts/glenn/toc.html

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

National Science Board (2004). *A statement of the national science board: In support of the math and science partnership program at the national science foundation.* Retrieved from http://www.nsf.gov/nsb/documents/2004/nsb_msp_statement2.pdf

NCTM. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: NCTM.

NRC (National Research Council). (2010). *Preparing teachers: Building evidence for sound policy*. Washington, DC: The National Academies Press.

Op 'T Eynde, P., De Corte, E., & Verschaffel, L. (2002). Framing students' mathematics related beliefs: A quest for conceptual clarity and a comprehensive categorization. In G. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A hidden variable in mathematics education* (pp. 13–38). Boston, MA: Kluwer Academic Publishing.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Raudenbush, S.W., Bryk, A. S, & Congdon, R. (2004). *HLM 6 for Windows* [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Reckase, M. D., McCrory, R., Floden, R. E., Ferrini-Mundy, J., & Senk, S. L. (2015). A multidimensional assessment of teachers' knowledge of algebra for teaching: Developing an instrument and supporting valid inferences. *Educational Assessment, 20*(4), 249–267.

Sahlberg, P. (2007). Education policies for raising student learning: The Finnish approach. *Journal of Education Policy, 22*(2), 147–171.

Sahlberg, P. (2010). *The secret to Finland's success: Educating teachers*. Stanford: Stanford Center for Opportunity Policy in Education. Research Brief. Retrieved from http://edpolicy.stanford.edu

Santibañez, L. M. (2002). *Why we should care if teachers get A's: Impact on student achievement in Mexico*. Unpublished doctoral dissertation, Stanford University School of Education, Stanford.

Schilling, S. G., & Hill, H. C. (2007). Assessing measures of mathematical knowledge for teaching: A validity argument approach. *Measurement: Interdisciplinary Research and Perspectives, 5*(2–3), 70–80.

Schmidt, W., Blömeke, S., & Tatto, M. T. (2011). *Teacher education matters. A study of middle school mathematics teacher preparation in six countries*. New York, NY: Teachers College Press.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Simon, M. A., & Blume, G. W. (1994). Building and understanding multiplicative relationships: A study of prospective elementary teachers. *Journal for Research in Mathematics Education, 25*, 472–494.

Staub, F. C., & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology, 94*(2), 344–355.

Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS and TIMSS-R video studies. *Educational Psychologist, 35*, 87–100.

Stigler, J. W., & Hiebert, J. (1997). Understanding and improving classroom mathematics instruction: An overview of the TIMSS video study. *Phi Delta Kappan, 79*(1), 14–21.

Tatto, M. T. (1996). Examining values and beliefs about teaching diverse students: Understanding the challenges for teacher education. *Educational Evaluation and Policy Analysis, 18*, 155–180.

Tatto, M. T. (1998). The influence of teacher education on teachers' beliefs about purposes of education, roles and practice. *Journal of Teacher Education, 49*, 66–77.

Tatto, M. T. (1999a). Improving teacher education in rural México: The challenges and tensions of constructivist reform. *Teaching and Teacher Education, 15*, 15–35.

Tatto, M. T. (1999b). The socializing influence of normative cohesive teacher education on teachers' beliefs about instructional choice. *Teachers and Teaching, 5*, 111–134.

Tatto, M. T. (2008). Teacher policy: A framework for comparative analysis. *Prospects: Quarterly Review of Comparative Education, 38*, 487–508.

Tatto, M. T. (Ed.). (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M). Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries: Technical Report*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Tatto, M. T. (2017). The Role of Comparative and International Research in Developing Capacity to Study and Improve Teacher Education. In M. A. Peters, B. Cowie, & I. Menter (Eds.), *A companion to research in teacher education*. Singapore: Springer.

Tatto, M. T., & Coupland, D. (2003). Teaching and measuring attitudes in teacher education. In J. Raths & A. McAninch (Eds.), *Teacher beliefs and classroom performance: The impact of teacher education (advances in teacher education)* (Vol. 6, pp. 123–181). Greenwich, CT: Information Age Publishing.

Tatto, M. T., & Kularatna, N. G. (1993). The interpersonal dimension of teacher education: Comparing distance education with two other programs in Sri Lanka. *International Journal of Educational Research, 19*, 755–778.

Tatto, M. T., Nielsen, H. D., Cummings, W. C., Kularatna, N. G., & Dharmadasa, D. H. (1993). Comparing the effectiveness and costs of different approaches for educating primary school teachers in Sri Lanka. *Teaching and Teacher Education, 9*, 41–64.

Tatto, M. T., Rodriguez, M., Reckase, M., Smith, W., & Pippin, J. (forthcoming). *The first five years of mathematics teaching (FIRSTMATH): Concepts, methods and strategies for comparative international research*. Dordrecht, Netherlands: Springer.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework*. Teacher Education and Development International Study Center, Michigan State University, East Lansing, MI, and IEA.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA teacher education and development study in mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Tatto, M. T., & Hordern, J. (2017). The configuration of teacher education as a professional field of practice: A comparative study of mathematics education. In J. Furlong & G. Whitty (Eds.), *Knowledge and the Study of Education: an international exploration* (pp. 255–274). Oxford, UK: Oxford Comparative Education Series, Symposium Books.

UNESCO. (1998). *Global monitoring report: Teachers and teaching in a changing world*. Paris, France: Author.

UNESCO. (2014). *Teaching and learning: Achieving quality for all: Global monitoring report*. Paris, France: Author.

UNESCO. (2016). *Education for people and planet: Creating sustainable futures for all*. Global education monitoring report. Paris: UNESCO. Retrieved from http://unesdoc.unesco.org/images/0024/002457/245752e.pdf

UNESCO's International Standard Classification of Education (ISCED). (2007). Paris: UNESCO Institute for Statistics. Retrieved from http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx

Van Dooren, W., Verschaffel, L., & Onghena, P. (2002). The impact of pre-service teachers' content knowledge on their evaluation of students' strategies for solving arithmetic and algebra word problems. *Journal for Research in Mathematics Education, 33*(5), 319–351.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.

Wilkins, J. L. M., & Brand, B. R. (2004). Change in preservice teachers' beliefs: An evaluation of a mathematics methods course. *School Science and Mathematics, 104*(5), 226–232.

Wu, M., Adams, R., Wilson, M., & Haldane, S. (2007). *ACER conquest: Generalised item response modelling software* (Version 2.0). Melbourne, Australia: Australian Council for Educational Research (ACER).

Zeichner, K., & Conklin, H. (2005). Teacher education programs. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher education* (pp. 645–736). New York, NY: Routledge.

Chapter 9
How Primary Future Teachers' Knowledge Is Shaped by Teacher Preparation

Hong Qian and Peter Youngs

Abstract This chapter uses data from the Teacher Education and Development Study in Mathematics (TEDS-M) to examine associations between primary future teachers opportunity to learn (OTL) in mathematics courses and mathematics methods courses in Chinese Taipei, Singapore, and the United States and their mathematics content knowledge and mathematics pedagogical content knowledge. The study found evidence that primary candidates' knowledge seems to be affected by the content of mathematics courses taken and by the number of topics addressed and their OTL in mathematics methods courses.

Introduction

A robust body of evidence indicates that among various school-based factors, teacher quality has the strongest impact on student achievement (Rivkin, Hanushek, & Kain, 2005; Wayne & Youngs, 2003). In addition, several studies have found that teachers' mathematical content knowledge (MCK) and mathematical pedagogical content knowledge (MPCK) are closely associated with students' learning in mathematics (Baumert et al., 2010; Hill, Rowan, & Ball, 2005; Jacob, Kane, Rockoff, & Staiger, 2009). At the same time, prospective primary teachers vary significantly with regard to their MCK and MPCK both within and across countries. Further, there is less evidence in the research literature about how primary candidates' experiences in pre-service teacher education are associated with their knowledge levels.

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

H. Qian (✉)
National Council of State Boards of Nursing, Chicago, IL, USA

P. Youngs
University of Virginia, Charlottesville, VA, USA
e-mail: pay2n@eservices.virginia.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_9

In this chapter, we draw on data from the Teacher Education and Development Study in Mathematics (TEDS-M) to examine how primary candidates' preparation experiences in three countries are related to their levels of MCK and MPCK. TEDS-M is unique in that it was the first cross-national study of mathematics teacher preparation to feature large-scale samples of future teachers (FTs) (Tatto et al., 2012). In this analysis, we focus on primary candidates from Chinese Taipei, Singapore, and the United States because these countries are developed nations with relatively high-performing education systems. We concentrate on primary mathematics teachers because they strongly affect students' academic trajectories and life outcomes (Chetty, Friedman, & Rockoff, 2011).

Our analyses indicated that the *content* of mathematics courses taken by primary FTs has a stronger effect on their MCK than the *number* of math courses that they take. In addition, our results indicate that candidates' opportunity to learn topics related to mathematics instruction was associated with both MCK and MPCK.

In the first section of this chapter, we review the research literature on teacher preparation and teacher knowledge. The second section introduces our theoretical framework, which posits a number of associations between FTs' preparation experiences and their knowledge levels. The third section describes our research methods, including samples, instruments, measures, and analytical strategies. We present our results in the fourth section and the final section discusses our results in relation to other research and identifies some limitations of this analysis.

Literature Review

In a review of research from the United States, Clift and Brady (2005) reported that primary FTs' experiences in mathematics methods courses and student teaching experiences had indeterminate effects on their beliefs and practices. Some studies found that methods courses and field experiences seemed to affect candidates' beliefs about teaching mathematics and the extent to which they incorporated constructivist principles in planning instruction (Langrall, Thornton, Jones, & Malone, 1996; Kim & Sharp, 2000; Mewborn, 1999). At the same time, Vacc and Bright (1999) reported that prospective primary teachers were limited in their ability to employ knowledge of students' mathematical thinking in their planning and instruction even when they believed that it was important to do so. According to Schmidt, Blömeke, & Tatto (2011), one reason for the inconsistent relationship between teacher education and FTs' beliefs and practices was that opportunity to learn (OTL) in teacher preparation had been measured imprecisely in most studies. In other words, Schmidt, Blömeke, and Tatto, argued that the use of better measures of OTL could lead to the consistent identification of relationship between OTL and key outcomes for teacher candidates.

Recently, researchers have employed more precise measures that assess teacher education experiences in a low-inference way (Schmidt, Blömeke, & Tatto, 2011) to investigate how specific components of teacher preparation affect student achievement. For example, Boyd, Grossman, Lankford, Loeb, and Wyckoff (2009) reported that several aspects of student teaching were associated with effectiveness (as measured by student achievement gains) among 1st-year primary teachers in New York City (NYC): experience with student teaching, having a supervisor who provided oversight of their student teaching, and close alignment between their current school context and the school context in which they completed student teaching. In a different analysis using the same dataset, Ronfeldt (2012) found that teachers who completed student teaching in easier-to-staff schools were more likely to continue teaching in NYC schools and to have increased student scores on state standardized tests during their first five years of teaching. Both of these studies used student achievement as a dependent variable; neither included measures of teacher knowledge.

Schmidt, Blömeke, and Tatto (2011) examined precursors to mathematics content knowledge and mathematics pedagogical content knowledge for lower secondary[1] mathematics teachers in six countries. They reported that lower secondary candidates' opportunities to study calculus and advanced mathematics were related to their knowledge of number, geometry, algebra, function, and data. They also found that MPCK was associated with three measures of practical experience: opportunity to engage in instructional interactions in mathematics, the number of different types of practical experiences, and the number of weeks during which they had primary responsibility for mathematics instruction during student teaching. In an analysis using TEDS-M data, Schmidt, Cogan, and Houang (2011) examined country-level correlations between primary and lower secondary mathematics candidates' MCK and the average number of courses taken. They reported that primary candidates took a significantly higher number of mathematics courses in the countries with the highest scaled scores of MCK than in lower-performing countries. They also found that among lower secondary candidates, those in countries with the highest MCK scores took almost twice as many mathematics courses and significantly more mathematics methods courses than their counterparts in lower-performing countries.

The findings from these two studies provide a foundation for the present study. For lower-secondary mathematics candidates, they indicate that the number and content of mathematics content courses taken are related to MCK, while MPCK is related to more practical experiences during teacher preparation. However Schmidt, Cogan, and Houang's analysis considered primary candidates, focusing only on their MCK, while neither study investigated associations between primary candidates' preparation experiences and their MPCK. The analysis presented in this chapter advances the work of Schmidt, Blömeke, Tatto (2011), and colleagues in three ways. First, our study explores how primary candidates' MCK and MPCK are shaped by their teacher preparation experiences. Second, we focus opportunities to learn in mathematics education courses, general pedagogy courses. and student teaching; these measures were not included in the earlier studies. Third, we employ more precise variables to control for differences in candidates' mathematics knowledge prior to entering teacher preparation.

[1]Lower Secondary level refers to middle school, usually serving students whose age are around 11–14.

Theoretical Framework

It is important for FTs to possess and be able to use content knowledge and pedagogical content knowledge in their classroom teaching (Shulman, 1987). Mathematical content knowledge (MCK) is the knowledge of mathematics; MPCK is knowledge specific to teaching mathematics. Previous studies have found that MCK and MPCK are essential for future mathematics teachers to be effective (Hill, Ball, Goffney, & Rowan, 2008; Hill et al., 2005; Jacob et al., 2009). In the present study, we used measures of both MCK and MPCK as defined by the TEDS-M study (see Tatto, 2013, p. 32–36).

TEDS-M defined opportunity to learn (OTL) as an experience with an anticipated or intended learning outcome, "OTL for teachers can occur at any point in the continuum of teacher learning, from the opportunities associated with schooling before entry into a formal teacher preparation program to the opportunities given to experienced teachers throughout their careers… TEDS-M concentrated on the opportunities that future teachers have to learn mathematics, mathematics pedagogy, and the general pedagogy provided by their pre-service preparation programs" (Tatto et al., 2008, p. 23). In the present study, our framework considers the association of future teachers' MCK and MPCK levels with their opportunity to learn in (a) mathematics courses, and (b) mathematics methods courses.

Mathematics Courses

At the primary level, Hill (2010) reported that for a nationally representative sample of teachers in the United States taking additional coursework was associated with higher levels of mathematical knowledge for teaching (MKT). MKT includes *both* MCK and MPCK; it features the common mathematical knowledge that is held by a well-educated adult and the specialized mathematical knowledge that is only held by teachers (such as the ability to represent content in ways that are accessible to young learners and to anticipate student errors). Among Chinese primary candidates, Youngs and Qian (2013) found that completion of courses in number theory and mathematical reasoning was associated with significantly higher levels of MKT in number and operations. As noted above, Schmidt, Blömeke, and Tatto (2011) reported that lower secondary mathematics candidates' opportunity to study calculus and advanced mathematics was significantly related to their MCK in number, algebra, geometry, function, and data. At the same time, these learning opportunities did not influence their MPCK. The findings from these studies indicate a need to examine how the mathematics content courses taken by primary candidates in multiple countries are associated with their knowledge levels. Thus, in our study, we used TEDS-M data to test the following hypotheses:

Hypothesis 1A: *The number of the university-level mathematics content courses taken by primary future teachers is positively related to their level of knowledge.*

Hypothesis 1B: *The content of the university-level mathematics content courses taken by primary future teachers is positively related to their level of knowledge.*

Mathematics Methods Courses

In their 2013 study, Youngs and Qian found that Chinese primary candidates' exposure to certain topics and learning experiences in mathematics methods courses and general pedagogy courses was related to their MKT levels. As noted above, Schmidt, Blömeke, and Tatto (2011) reported lower secondary candidates' MPCK was associated with engaging in math instructional interactions, the number of different types of practical experiences they had, and the number of weeks during which they had primary responsibility for mathematics instruction during student teaching. Building on these findings, we used TEDS-M data to test the following hypotheses:

Hypothesis 2A: *The number of topics addressed in mathematics methods courses taken by primary future teachers is positively related to their level of knowledge.*

Hypothesis 2B: *The opportunities to learn (OTL) in mathematics methods courses for primary future teachers are positively related to their level of knowledge.*

In sum, our theoretical framework contends that primary FTs' knowledge (including MCK and MPCK) will be associated with their opportunities to learn in mathematics courses and mathematics methods courses.

Method

Samples

TEDS-M study focused on primary and lower secondary mathematics teacher education in 17 countries. The FTs who participated in the study were all in their final year of teacher preparation. The study presented in this chapter focuses on programs that prepared individuals to work as primary generalist teachers up through grade 6. Among the six countries in TEDS-M that prepared primary generalist teachers, we concentrate in this chapter on three: Chinese Taipei, Singapore, and the United States. The sample size for programs and future teachers for this group are in Table 9.1. The sampling plan in the TEDS-M study was designed to produce nationally representative data for each of the participating countries. See Tatto et al. (2012) for more details on sampling.

Table 9.1 Sample sizes for preparation programs and future teachers in primary generalist program group

Country	Number of preparation programs	Number of future teachers
Chinese Taipei	11	923
Singapore	6	263
United states	71	1,310
Total	88	2,496

Context

In this section, we provide some contextual information about the status of the teaching profession in each of the three countries. The status of teaching in a given country can affect FTs' MCK and MPCK and impact their experiences in teacher preparation. The contextual information included below is based on Schwille, Ingvarson, and Holdgreve-Resendez (2013).

Chinese Taipei Teaching has traditionally been an attractive profession in China, characterized by high levels of status and prestige. Teacher salaries in Chinese Taipei are funded by the government and are stable and generous. During the work day, teachers are allocated time to plan instruction and collaborate with colleagues in addition to fulfilling their instructional duties. Since teaching is an attractive occupation, it is also quite competitive. During formal preparation, candidates must pass a series of rigorous examinations to obtain entry to a program, earn their teaching licenses, and obtain employment. Guidelines are set by the government regarding recruitment, pre-service preparation, licensure, and in-service professional development (Hsieh, Lin, Chao, & Wang, 2012).

Singapore As in China, teaching in Singapore has traditionally been an attractive profession. The nation recruits FTs from among the top one-third of each cohort by academic ability. The government funds teacher salaries, which are stable and competitive with other professions. In Singapore, the National Institute of Education (NIE) maintains a high degree of control over teacher training and certification. The Ministry of Education recruits FTs and they are trained by the NIE. When FTs graduate from the NIE, they are automatically qualified to teach in Singapore schools (Wong et al., 2012).

United States The United States has shifted toward centralization of teacher licensure policy at the state and the national level. Yet there is still notable variation within and across states with regard to teacher education program-types and licensure requirements for primary and lower secondary mathematics. Along with more traditional university-based preparation program types, alternative routes to licensure have grown significantly in the United States. Compared to other professionals, teachers in the United States usually have lower status and are paid less and they

Table 9.2 Composition of future teacher questionnaire

Section	Focus	Time (min)
A	General background	5
B	Opportunity to learn in teacher preparation program	15
C	MCK and MPCK	60
D	Beliefs about mathematics and teaching	10

generally express less satisfaction than other groups with respect to pay, benefits, and promotion opportunities. Primary FTs tend to have lower college aptitude test scores than the average college graduate (Youngs & Grogan, 2013).

Instruments

Three main surveys were used in the TEDS-M study: questionnaires of future teachers, educators, and institutional program representatives. The future teacher questionnaire had four parts: (a) general background, (b) opportunity to learn in teacher education, (c) MCK and MPCK, and (d) beliefs about teaching and mathematics. The time allotted for completing each is shown in Table 9.2. Parts A, B, and D featured rating-scale items while part C consisted of items that assessed participants' knowledge. For part C, to measure breadth and depth of knowledge, the TEDS-M study used a rotated block design; each participant responded to questions in two blocks out of a total of five blocks.

Teacher educators who taught courses in (a) mathematics and mathematics pedagogy and (b) general pedagogy were asked to complete the educator questionnaire. This survey included questions about their academic and teaching backgrounds, their professional and research experiences, their participation in field-based instruction, and their own beliefs about mathematics. In addition, it included questions about the learning opportunities they provided in their courses for primary candidates. The institutional program questionnaire included questions about candidates' backgrounds, the admission process, course content, field placements, program standards and accountability, and program resources.

Measures

In this section, we describe the main variables used in our analyses. The independent variables included (a) number and content of university-level mathematics content courses and (b) topics addressed and OTL in mathematics pedagogy courses; the dependent variables were MCK and MPCK scores; and the control variables included previous mathematics achievement and parents' education).

Number of Mathematics Content Topics We calculated the number of mathematics content courses offered and taken at the institutional and individual levels. At the institutional level, this variable measures the number of university-level mathematics content topics at the program level while at the individual level it reports the number of such topics taken. In particular, FTS were asked whether they had studied each of 19 distinct mathematics topics. Including this variable at both levels allowed us to examine the effects of institutional policies (i.e., some institutions require all students to take certain types of courses while others do not) as well as candidates' choices (i.e., in some programs, some individual FTs elected to take certain types of topics (Tatto et al., 2012).

Content of Mathematics Content Courses The 19 distinct mathematics topics included the following:

- foundations of geometry or axiomatic geometry
- analytic/coordinate geometry
- non-Euclidean geometry
- differential geometry
- topology
- linear algebra
- set theory
- abstract algebra
- number theory
- beginning calculus
- calculus
- multivariate calculus
- advanced calculus or real analysis or measure theory
- differential equations
- theory of real functions and theory of complex functions or functional analysis
- discrete mathematics
- graph theory, game theory, combinatorics, or Boolean algebra
- probability, theoretical or applied statistics
- mathematical logic

These 19 topics can be conceptually grouped into four broader categories representing university-level mathematics: geometry, discrete structure and logic, continuity and functions, and probability and statistics.

Topics Addressed in Mathematics Methods Courses This variable is also defined at the institutional and individual levels. The institutional program questionnaire asked about the number of mathematics pedagogy courses required for the program, while the future teacher questionnaire asked FTs if they had studied one or more of the following topics during teacher preparation:

- foundations of mathematics
- context of mathematics education

- development of mathematics ability and thinking
- mathematics instruction
- developing lesson plans
- observation, analysis and reflection
- mathematics standards and curriculum
- affective issues in mathematics

For each candidate, we counted the number of different topics that had been addressed in their mathematics methods courses.

OTL in Mathematics Methods Courses Since measures of the number of courses taken cannot fully capture primary candidates' experiences within courses, we use this group of variables to closely examine their experiences in mathematics methods courses. The future teacher questionnaire Items 5 and 6 asked candidates to indicate how frequently they engaged in a list of activities in mathematics pedagogy courses, such as using students' misconceptions to plan instruction or analyzing examples of teaching (see Table 9.3 for the full list of activities). The response categories were coded as 1 (*never*), 2 (*rarely*), 3 (*occasionally*), and 4 (*often*). The TEDS-M study reported seven scaled scores for OTL in mathematics methods courses based on Items 5 and 6: class participation, class reading, solving problems, instructional practice, instructional planning, assessment uses and assessment practice. The activities used to construct each scaled score are listed in Table 9.3. The reliabilities for these seven scales are .85, .83, .78, .89, .90, .91, and .87, respectively.

Previous Mathematics Achievement This individual-level control was constructed based on responses to two questions on the future teacher questionnaire: (a) What was the highest grade level at which you studied mathematics in secondary school? And (b) In secondary school, what was the usual level of grades that you received?

Parents' Education This control variable was also at the individual level; it was constructed based on responses to two questions on the future teacher questionnaire: (a) What is the highest level of education completed by your mother, and (b) What is the highest level of education completed by your father?

MCK Score Future teachers' responses to the mathematics content knowledge assessment were used to construct one of the dependent variables, *MCK* score. This assessment was designed to measure advanced mathematics knowledge and mathematics curricular knowledge related to primary-level school mathematics (Tatto et al., 2012). The MCK assessment consisted of 74 items across four content subdomains: number and operations, algebra and functions, geometry and measurement, and data and chance. Individual-scaled scores were created for each FTs using Rasch scaling. The international mean for the MCK scale was 500, and the standard deviation was 100. The reliability for the MCK measure was .8.

Table 9.3 Lists of items used to construct each OTL variable in mathematics education courses

OTL	In the mathematics education<pedagogy/teaching methods> courses that you have taken or are currently taking in your teacher preparation program, how frequently did you do any of the following? (Never, Rarely, Occasionally, Often)
	Ask questions during class time
	Participate in a whole class discussion
	Make presentations to the rest of the class
	Teach a class session using methods of my own choice
	Teach a class session using methods demonstrated by the instructor
Class reading	Read about research on mathematics
	Read about research on mathematics education
	Read about research on teaching and learning
	Analyze examples of teaching (e.g., film, video, transcript of lesson)
Solving problems	Write mathematical proofs
	Solve problems in applied mathematics
	Solve a given mathematics problem using multiple strategies
	Use computers or calculators to solve mathematics problems
Instructional practice	Explore how to apply mathematics to real-world problems
	Explore mathematics as the source for real-world problems
	Learn how to explore multiple solution strategies with pupils
	Learn how to show why a mathematics procedure works
	Make distinctions between procedural and conceptual knowledge when teaching mathematics concepts and operations to pupils
	Integrate mathematical ideas from across areas of mathematics
Instructional planning	Accommodate a wide range of abilities in each lesson
	Create learning experiences that make the central concepts of subject matter meaningful to pupils
	Create projects that motivate all pupils to participate
	Deal with learning difficulties so that specific pupil outcomes are accomplished
	Develop games or puzzles that provide instructional activities at a high interest level
	Develop instructional materials that build on pupils' experiences, interests and abilities
	Use pupils' misconceptions to plan instruction
Assessment uses	Give useful and timely feedback to pupils about their learning
	Help pupils learn how to assess their own learning
	Use assessment to give effective feedback to parents or guardians
	Use assessment to give feedback to pupils about their learning
	Use classroom assessments to guide your decisions about what and how to teach
Assessment practice	Analyze and use national or state standards or frameworks for school mathematics
	Analyze pupil assessment data to learn how to assess more effectively
	Assess higher-level goals (e.g., problem-solving, critical thinking)
	Assess low-level objectives (factual knowledge, routine procedures and so forth)
	Build on pupils' existing mathematics knowledge and thinking skills

MPCK Score Primary candidates' responses to the mathematics pedagogical content knowledge assessment were used to construct the other dependent variable, *MPCK* score. This assessment included 32 MPCK items across three subdomains: curricular knowledge, planning for teaching and learning, and enacting teaching and learning. The assessment items included three formats: multiple-choice (MC), complex multiple-choice (CMC), and constructed response (CR). Future teachers' MPCK scores were reported in scaled scores generated through the use of item response theory, with a mean of 500 and standard deviation of 100. The reliability for the MPCK measure was .7 (Tatto et al., 2012).

Analytical Strategies

Because FTs were nested within programs, we employed multilevel linear models to examine the effects of preparation programs on primary FTs' knowledge. The independent variables were defined at both the individual and institutional levels while the dependent variables, MCK and MPCK, were defined at the individual level. Variables defined at the individual level of the analysis highlight the effects of individual differences or choices while variables defined at the institutional level reflect differences between institutions with regard to policies and practices (Tatto et al., 2012). A population model for MCK at the future teacher level is as follows:

$$\text{MCK}_{ij} = \beta_{0j} + \beta_{1j}\left(\text{Parent education}_{ij}\right) + \beta_{2j}\left(\text{Previous mathematics achievement}_{ij}\right) + \beta_{3j}\left(\text{independent variable}_{ij}\right) + u_{ij}$$

A population model for MCK at the program level is as follows:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\left(\text{independent variable}_{j}\right) + v_{0j}$$

in which MCK_{ij} represents the MCK score for FT *i* in program *j*; independent variable$_{ij}$ represents an independent variable at the individual level, such as content of mathematics courses taken by FT *i* in program *j*; independent variable *j* represents an independent variable at the institutional level, such as the number of mathematics courses required by program *j* which may have an effect on the intercept for the first-level model.

Although these independent variables represent our central theoretical concerns, other factors could affect FTs' MCK scores. Therefore, we include control variables for candidates' previous mathematics achievement, parents' levels of education, and the number of required mathematics courses. Controlling for these measures helps us account for differences among FT's that were evident prior to their beginning teacher education programs. A similar model for MPCK at the FT level is as follows:

$$\mathrm{MCK}_{ij} = \beta_{0j} + \beta_{1j}\left(\text{Parent education}_{ij}\right) + \beta_{2j}\left(\text{Previous mathematics achievement}_{ij}\right) + \beta_{3j}\left(\text{independent variable}_{ij}\right) + u_{ij}$$

A population model for MCK at the program level is as follows:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\left(\text{independent variable}_{j}\right) + v_{0j}$$

Results

Effect of Mathematics Content Topics

Table 9.4 reports the standardized coefficients from a multilevel linear regression model for Chinese Taipei, Singapore, and the United States. The independent variables are the number of mathematics content topics at the individual and institutional levels. The number of mathematics content topics at the individual level refers to the number of mathematics content topics the FT has taken, we are defining these as 'courses'; while the number of mathematics content topics at the institutional level refers to the number of aggregated mathematics content topics at the program level, we are defining this as the 'courses required' by the institution. The dependent variable is FTs' MCK scores. The model also included two control variables: FTs' previous mathematics achievement and parents' education.

The results in column 4 of Table 9.4 show that for Chinese Taipei, the coefficient for the number of mathematics content topics taken by FTs is positive and statistically significant, which indicates an association between the number of mathematics content topics taken by candidates and their MCK score even when controlling for the number of mathematics content courses required by the institution, FTs'

Table 9.4 Estimated effects of number of mathematics content courses on MCK score

	Previous achievement	Parent education	Number of math topics (FT)	Number of topics aggregated at the program level/required (institution)
Chinese Taipei	0.30***	0.01	0.16***	−0.02
Singapore	0.22***	−0.01	0.10	0.15*
United States	0.22***	0.21***	−0.04	−0.12*

*$p < .05$; **$p < .01$; ***$p < .001$

previous mathematics achievement, and parents' education. In other words, the more mathematics content courses a candidate in Chinese Taipei has taken, the higher their MCK score, assuming that the other three variables remain constant. But this is not the case in Singapore or the United States. In these two countries, the coefficient for the number of mathematics content courses taken by FTs is not statistically significant.

The average number of topics and standard deviation of mathematics content required by the institution in each country are as follows: Chinese Taipei: 3.24 (standard deviation is 4.428); Singapore: 3.04 (5.228); and the United States: 2.61 (2.985). The effect of the number of mathematics content courses required by the institution is mixed, as shown in Table 9.4 column 5. In Singapore, the coefficient for the number of mathematics content courses required by the institution is positive, which indicates that the greater the number of math content courses required by a given institution, the higher the MCK scores for candidates at that institution. However, in the United States, the coefficient is negative, which means that the greater the number of mathematics content courses an institution required, the lower the MCK scores for FTs, assuming the other three variables remain constant. One explanation is that when previous mathematics achievement for all candidates in an institution is generally low, the institution requires that they take a greater number of mathematics content courses (the negative correlation between the two from data confirmed this hypothesis). However, the additional required courses do not put these candidates on par with their peers in programs requiring fewer courses; even when controlling for their lower previous mathematics achievement, these students' MCK at the completion of the program is still lower than the MCK of students in other institutions requiring a smaller number of required mathematics content courses.

Table 9.5 focuses on MPCK scores. As shown in column 4 of Table 9.5 for Chinese Taipei and Singapore, there is a positive association between the number of mathematics content courses taken by candidates and their MPCK scores, while in the United States, there is a negative association between these two variables. For the institution-level variable, in Singapore, the number of mathematics content courses required by the institution is positively associated with candidates' MPCK scores.

The results for the control variables from Tables 9.4 and 9.5 are worth noting. FTs' previous mathematics achievement (in column 2) has a significant association

Table 9.5 Estimated effects of number of mathematics content courses on MPCK score

	Previous achievement	Parent education	Number of math topics (candidate)	Number of required courses (institution)
Chinese Taipei	0.25***	−0.03	0.11***	−0.05
Singapore	0.14**	−0.01	0.12*	0.05*
United States	0.22***	0.13***	−0.07*	−0.05

*$p < .05$; **$p < .01$; ***$p < .001$

with their MCK and MPCK across all countries. Therefore, it is important to include this variable to control for differences in candidates' mathematics ability before they begin teacher preparation. When it comes to parents' education (in column 3), there is no effect on candidates' MCK or MPCK except in the United States, where there is a strong relationship between candidates' parents' education levels and both MCK and MPCK. For MCK, the effect size of parents' education is as large as candidates' previous mathematics achievement. This indicates that for U.S. candidates, MCK and MPCK are significantly associated with their parents' education levels, which means that preparation programs are unlikely to eliminate the inequity produced by family background. This is not the case in the other two countries.

While the overall number of mathematics content courses did not have a consistent impact on MCK score, we also examined whether taking particular mathematics content courses was consequential. As noted, the 19 distinct math content courses or topics in the TEDS-M study can be conceptually grouped into four broader categories geometry, discrete structure and logic, continuity and functions, and probability and statistics. Using the number of courses FTs took in each category as an independent variable, we estimated the effects of each category on candidates' MCK scores in a model that included each of the independent variables and control variables (previous mathematics achievement, parents' education, number of mathematics content courses required by the institution). Table 9.6 reports the unstandardized coefficients for each of the four independent variables in three countries (the coefficients for the control variables are not reported).

The results in Table 9.6 column 2 show that, for FTs in Chinese Taipei, taking one more geometry course (range from 0 to 4), would increase their MCK score by 5.85 points compared to candidates with the same previous mathematics achievement, the same level of parents' education, and the same number of total math content courses required by their institution. Note that the international mean for MCK score is 500 and, thus, the effect is significant. However, for U.S. candidates, taking one more course in geometry would lower their MCK score – by 6.04 points compared to other candidates, when holding the other three variables remain constant. This does not necessarily mean that the geometry course is harming them. It is possible that, when controlling for the number of mathematics content courses the institution requires, the candidates who chose to take more courses in geometry earned lower MCK scores for some reason related to their propensity to take geometry. The effect of the number of mathematics content courses in discrete structure and logic, continuity and functions, and probability and statistics on MCK is more

Table 9.6 Estimated effects of content of mathematics content courses on MCK score

	Geometry	Discrete structures and logic	Continuity and functions	Probability and statistics
Chinese Taipei	5.85*	5.78**	14.22***	12.55**
Singapore	2.36	3.70	5.31	7.83
United States	−6.04***	−1.78	6.12***	−1.77

*$p < .05$; **$p < .01$; ***$p < .001$

consistent (i.e., it is positively significant in at least one country). When the total number of mathematics content courses required by the institution is the same, it seems that taking a greater number of mathematics courses in continuity and functions helps candidates acquire more knowledge. For Chinese Taipei candidates, taking one more course in continuity and functions (range from 0 to 6) would lead to an increase of 14.22 in their MCK score, while for FTs in the United States, the increase would be 6.12.

Effect of Mathematics Education Courses

The results in column 4 of Table 9.7 show that for Chinese Taipei, the coefficient for the number of topics addressed in mathematics education courses (at the individual student level) is positive and statistically significant at the .05 level. This indicates an association between the number of topics addressed in these courses and FTs' MCK score. Candidates in Chinese Taipei who addressed more of the eight topics in these courses were more likely to have higher levels of MCK. However, there is no positive effect for the number of mathematics education courses required by the institution. For MPCK, there are no significant associations between the dependent variable and two independent variables.

Paralleling our analysis of the content of mathematics *content* courses, we examined which topics addressed in mathematics *education* courses were most important. In the TEDS-M study, the eight mathematics education topics can be conceptually grouped into two categories: instruction and foundations. Instruction includes five topics: mathematics instruction; developing lesson plans; observation, analysis and reflection; mathematics standards and curriculum; and affective issues in mathematics. Foundations include three topics: foundations of mathematics, context of mathematics education, and development of mathematics ability and thinking. We counted the number of topics that candidates addressed in each category and used this as an independent variable for MCK and MPCK. All models included three control variables: previous mathematics achievement, parents' education, and the number of mathematics education courses the institution required. Table 9.8 reports the coefficients in four models for the three countries (the coefficients for the control variables are not reported).

Table 9.7 Estimated effects of number of topics addressed in mathematics education courses on MCK score

	Previous achievement	Parent education	Number of math topics (candidate)	Number of required courses (institution)
Chinese Taipei	0.32***	0.01	0.06*	−0.01
Singapore	0.27***	0.01	−0.01	−0.04
United States	0.22***	0.20***	−0.01	0.02

*$p < .05$; **$p < .01$; ***$p < .001$

Table 9.8 Estimated effects of number of topics addressed in mathematics education courses on MPCK score

	Previous achievement	Parent education	Number of math topics (candidate)	Number of required courses (institution)
Chinese Taipei	0.27***	−0.03	0.03	−0.05
Singapore	0.19***	0.01	−0.01	−0.06
United States	0.21***	0.13***	0.03	−0.01

*$p < .05$; **$p < .01$; ***$p < .001$

Table 9.9 Estimated effects of content of mathematics education topics on MCK and MPCK scores

	Foundation		Instruction	
	MCK	MPCK	MCK	MPCK
Chinese Taipei	2.25	−0.13	4.09*	2.29
Singapore	−10.12*	−9.79**	6.62*	6.30*
United States	−2.20	−0.03	1.30	2.60

*$p < .05$; **$p < .01$; ***$p < .001$

The results in Table 9.9 columns 2 and 3 show that the association of the number of topics addressed in *foundations of mathematics* with both MCK and MPCK is negative and statistically significant in Singapore. In other words, when the number of mathematics education courses required by the institution is the same, the candidate who chose to take more courses in foundations of mathematics had lower MCK and MPCK scores. On the other hand, the effect of the number of topics addressed in *mathematics instruction* is more consistent. There was a positive and statistically significant effect of the number of math instruction topics addressed in at least one country for both MCK and MPCK.

Since measures of the number of topics addressed cannot fully capture experiences within mathematics education courses, we used OTL in mathematics education courses to closely examine candidates' experiences in these courses. The TEDS-M study reported seven scaled scores for OTL in mathematics methods:

- class participation
- class reading
- solving problems
- instructional practice
- instructional planning
- assessment uses
- assessment practice

We used the OTL variables as independent variables and put them into the model one at a time. The control variables for each model are: previous mathematics, parents' education, and the number of topics learned by the candidate. The dependent variable is MPCK. Table 9.10 reported the coefficient for each OTL variable in each country (the coefficients for the control variables are not reported).

Table 9.10 Estimated effects of OTL in mathematics education courses on MPCK score

	Class participation	Class reading	Solving problems	Instructional practice	Instructional planning	Assessment uses	Assessment practice
Chinese Taipei	3.02*	4.23***	3.72**	3.11*	2.92**	1.23	3.82**
Singapore	4.85	0.88	4.09	1.69	−1.95	0.72	0.65
United States	−2.96	−2.24	−6.74**	−2.85	−0.95	−1.81	0.37

*$p < .05$; **$p < .01$; ***$p < .001$

The results in Table 9.10 show that the coefficients for class participation, class reading, instructional practice, instructional planning and assessment practice are positive and statistically significant in Chinese Taipei, which indicates that candidates who have more opportunities to engage in these activities in math methods courses are more likely to have higher MPCK scores even if they have studied the same number of topics in mathematics education courses. The coefficient for solving problems is also positive and statistically significant in Chinese Taipei, but in the United States, the coefficient for this variable is negative and statistically significant. Chinese Taipei is the only country where candidates' experiences in mathematics methods classes (OTL) were positively associated with MPCK; we did not find the same effects in other countries.

Summary of Main Findings

This study used data from TEDS-M to explore associations between primary FTs' knowledge and their OTL in teacher preparation in Chinese Taipei, Singapore, and the United States. Previous research has found that primary teachers' knowledge is associated with student learning in mathematics (Hill et al., 2005; Jacob et al., 2009). But there is little understanding in the research literature of how primary candidates acquire MCK and MPCK in different countries. In this study, we hypothesized that a number of aspects of teacher education might be relevant to the acquisition of MCK and MPCK, such as the number and content of university-level mathematics content courses, exposure to topics in mathematics methods courses and opportunities to learn in coursework. In Table 9.11, we present a summary of the statistically significantly positive effects of these independent variables on MCK and/or MPCK. Hypotheses 1A, 1B, 2A, 2B are supported in at least one country in Table 9.11.

Table 9.11 Summary of the statistically significant positive relationships of teacher preparation components to teacher knowledge

		Chinese Taipei		Singapore		United States	
Hypothesis	Independent variable	MCK	MPCK	MCK	MPCK	MCK	MPCK
1A	number of mathematics content courses	F*	F*	I*	F*/I*	---	---
1B	Geometry	*		---		---	
	discrete structure & logic	*		---		---	
	continuity & functions	*		---		*	
	probability & statistics	*		---		---	
2A	number of topics addressed in mathematics methods courses	F*	---	---	---	---	---
	Foundation	---	---	---	---	---	---
	Instruction	*	---	*	*	---	---
2B	class participation		*		---		---
	class reading		*		---		---
	solving problems		*		---		---
	instructional practice		*		---		---
	instructional planning		*		---		---
	assessment uses				---		---
	assessment practice		*		---		---

Table 9.12 Standard deviations of MCK and MPCK and the number of significantly positive relationships between teacher preparation components and knowledge for five countries

Country	MCK standard deviation	MPCK standard deviation	The number of relationships
Chinese Taipei	84.23	68.39	16
Singapore	72.39	73.32	14
United States	70.01	67.60	1

Table 9.11 indicates that there are more significantly positive relationships between preparation components and candidates' knowledge in some countries than in other countries.[2] For example, in Chinese Taipei, the relationships between preparation components and candidates' knowledge are very persistent, while in the United States, there is only one significantly positive relationship between preparation components and candidates' knowledge.

One possible explanation for these differences between countries is related to selection effects associated with entering teacher education programs. In some countries, such as Chinese Taipei, preparation programs are very selective, which means that candidates' MCK and MPCK levels are uniformly high and less variable before they begin teacher education. In countries like the United States, though, programs are less selective, so candidates' MCK and MPCK levels can vary to a large degree before they begin teacher education. While we included control variables in our models that could account for some of these differences, there may be other factors connecting program selectivity with candidates' incoming capacity for learning that are not accounted for, which may make various teacher preparation components seem to be ineffective in some countries, where effects are confounded with program selectivity. Another possibility is that the reliability of the MCK and MPCK measures is higher in high-performance countries than in other countries, which could make it easier to observe a significant effect (Table 9.12).

Discussion

Based on prior research (Goldhaber & Brewer, 1997; Monk & King, 1994), we expected that the number and content of the university-level mathematics content courses taken by primary FTs would influence their level of knowledge, especially their MCK. We found that the number of mathematics content courses taken does have an effect on candidates' levels of MCK and MPCK in Chinese Taipei and Singapore. With regard to the content of mathematics courses, taking more courses

[2] In order to exclude the possibility that this may be related to the variance of MCK/MPCK in these countries, we have checked the relationships between the variance (standard deviation) of MCK/MPCK and the number of significant associations between teacher preparation components and MCK/MPCK. We found that there is no relationship between them. Table 9.12 lists the result.

in discrete structure and logic had an effect on MCK in Chinese Taipei, and taking more courses in continuity and functions had an effect on MCK in Chinese Taipei and the United States (See Table 9.11 for a summary of the statistically significantly positive effects of the independent variables on MCK and/or MPCK in the three countries.)

Our study found that in mathematics methods courses, primary candidates' exposure to topics related to instruction (such as developing lesson plans) was associated with MCK in Chinese Taipei and MPCK in Singapore. At the same time, the number of topics encountered in mathematics methods courses was only associated with MCK in Chinese Taipei and was not associated with MPCK in any country. Further, the number of opportunities to learn about several topics in mathematics methods courses were related to MPCK in Chinese Taipei.

For U.S. candidates, only the number of mathematics content courses in continuity and function was associated with their MCK. Other proposed relationships – hypotheses that were supported for Chinese Taipei and Singapore – were not evident in the U.S. sample. One possible cause is the effects of the control variables. In particular, with regard to the effect of previous mathematics achievement, there is a statistically significant association between prior mathematics achievement and FTs' MCK and MPCK across the three countries. It is not surprising that candidates' mathematics achievement before they begin teacher education influences their development of MCK and MPCK while they are in teacher preparation. This study provides strong evidence for this association.

The research presented in this chapter research can inform theory regarding opportunity to learn in teacher preparation and suggests that mathematics content and methods courses can influence primary FTs' professional knowledge. Our findings suggest that the *content* of mathematics courses taken has a stronger effect on candidates' MCK than the number of math courses taken. In addition, our results indicate that candidates' opportunity to learn topics related to mathematics instruction was associated with both MCK and MPCK. By helping to discern how preparation experiences can shape teachers' knowledge, the findings presented in this chapter can inform theory as well as future research regarding ways to structure and improve primary mathematics teacher preparation.

References

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 416–440.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood* (working paper no. 17699). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w17699

Clift, R. T., & Brady, P. (2005). Research on methods courses and field experiences. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 309–424). Washington, DC: American Educational Research Association.

Goldhaber, D. D., & Brewer, D. J. (1997). Evaluating the effect of teacher degree level on educational performance. In W. J. Fowler (Ed.), *Developments in school finance, 1996* (pp. 197–210). Washington, DC: National Center for Education Statistics/U.S. Department of Education.

Hill, H. C. (2010). The nature and predictors of elementary teachers' mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 41*(5), 513–545.

Hill, H. C., Ball, D. L., Goffney, I. M., & Rowan, B. (2008). Validating the ecological assumption: The relationship of measure scores to classroom teaching and student learning. *Measurement: Interdisciplinary Research and Perspectives, 5*(2–3), 107–117.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Hsieh, F. J., Lin, P. J., Chao, G., & Wang, T. Y. (2012). Preparing teachers of mathematics in Chinese Taipei. In J. Schwille, L. Ingvarson, & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in seventeen countries* (pp. 69–85). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Jacob, B. A., Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2009). *Can you recognize an effective teacher when you recruit one?* (*CLOSUP* Working Paper Series No. 11). Ann Arbor, MI: University of Michigan, Center for Local, State, and Urban Policy.

Kim, M. K., & Sharp, J. (2000). Investigating and measuring preservice elementary mathematics teachers' decision about lesson planning after experiencing technology-enhanced methods instruction. *Journal of Computers in Mathematics and Science Teaching, 19*(4), 317–338.

Langrall, C. W., Thornton, C. A., Jones, G. A., & Malone, J. A. (1996). Enhanced pedagogical knowledge and reflective analysis in elementary mathematics teacher education. *Journal of Teacher Education, 47*(4), 271–282.

Mewborn, D. S. (1999). Reflective thinking among preservice elementary mathematics teachers. *Journal for Research in Mathematics Education, 30*(3), 316–341.

Monk, D. H., & King, J. A. (1994). Multilevel teacher resource effects on pupil performance in secondary mathematics and science: The case of teacher subject-matter preparation. In R. Ehrenberg (Ed.), *Choices and consequences: Contemporary policy issues in education* (pp. 29–58). Ithaca, NY: ILR Press.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Ronfeldt, M. (2012). Where should student teachers learn to teach? Effects of field placement school characteristics on teacher retention and effectiveness. *Educational Evaluation and Policy Analysis, 34*(1), 3–26.

Schmidt, W. H., Blömeke, S., & Tatto, M. T. (2011). *Teacher education matters: A study of middle school mathematics teacher preparation in six countries*. New York, NY: Teachers College Press.

Schmidt, W. H., Cogan, L., & Houang, R. (2011). The role of opportunity to learn in teacher preparation: An international context. *Journal of Teacher Education, 62*(2), 138–153.

Schwille, J., Ingvarson, L., & Holdgreve-Resendez, R. (Eds.). (2013). *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in seventeen TEDS-M countries*. East Lansing, MI: Michigan State University.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Tatto, M. T. (Ed.). (2013). *The teacher education and development study in mathematics (TEDS-M). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report*. Amsterdam: International Association for the Evaluation of Student Achievement.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDSM): Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, Michigan State University, IEA.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries*. East Lansing, MI: Teacher Education International Study Center, College of Education, Michigan State University.

Vacc, N. N., & Bright, G. W. (1999). Elementary preservice teachers' changing beliefs and instructional use of children's mathematical thinking. *Journal for Research in Mathematics Education, 30*(1), 89–110.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.

Wong, K. Y., Lim-Teo, S. K., Lee, N. H., Boey, K. L., Koh, C., Dindyal, J., … Cheng, L. P. (2012). Preparing teachers of mathematics in Singapore. In J. Schwille, L. Ingvarson, & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in seventeen countries* (pp. 193–207). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Youngs, P., & Grogan, E. (2013). Preparing teachers of mathematics in the USA. In J. Schwille, L. Ingvarson, & R. Holdgreve-Resendez (Eds.), *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in seventeen countries* (pp. 255–272). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Youngs, P., & Qian, H. (2013). The influence of university courses and field experiences on Chinese elementary candidates' mathematical knowledge for teaching. *Journal of Teacher Education, 64*(3), 244–261.

Chapter 10
Opportunities to Learn Mathematics Pedagogy and Connect Classroom Learning to Practice: A Study of Future Teachers in the United States and Singapore

Traci Shizu Kutaka, Wendy M. Smith, and Lorraine M. Males

Abstract In this study, we conducted secondary analyses using the TEDS-M database to explore future mathematics specialists teachers' opportunities to learn (OTL) how to teach mathematics. We applied latent class analysis techniques to differentiate among groups of prospective mathematics specialists with potentially different OTL mathematics pedagogy within the United States and Singapore. Within the United States, three subgroups were identified: (a) *Comprehensive OTL*, (b) *Limited OTL*, and (c) *OTL Mathematics Pedagogy*. Within Singapore, four subgroups were identified: (a) *Comprehensive OTL*, (b) *Limited Opportunities to Connect Classroom Learning with Practice*, (c) *OTL Mathematics Pedagogy*, and (d) *Basic OTL*. Understanding the opportunities different prospective teachers had to learn from and their experiences with different components of instructional practice in university and practicum settings has implications for teacher preparation programs.

Introduction

Around the world, well-intentioned people disagree about how primary teachers should be prepared to teach mathematics effectively. Whereas the United Kingdom seems to be moving from university-based to school-based teacher preparation,

---

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

T. S. Kutaka (✉)
University of Denver, Denver, CO, USA
e-mail: traci.kutaka@du.edu

W. M. Smith · L. M. Males
University of Nebraska, Lincoln, NE, USA
e-mail: wsmith5@unl.edu; lmales2@unl.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_10

other countries, like the Philippines, have recently increased university-based requirements for teacher preparation. In the United States, some alternative teacher-preparation programs minimize preparation and believe teachers can learn what they need to know by teaching (e.g., Teach for America). Research suggests teacher preparation matters in two ways. First, preparation can enhance the initial effectiveness of novice teachers who graduate from university-based undergraduate programs, particularly in comparison to teachers who come from alternative certification programs (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; 2007; 2009; Darling-Hammond, Chung, & Frelow, 2002; Darling-Hammond, Holtzman, Gatlin, & Heilig, 2005). Second, preparation reduces the well-documented attrition that occurs within the first five years of teaching (Henke, Chen, & Geis, 2000; National Commission on Teaching and America's Future, 1996), increasing the likelihood of remaining in the profession long enough to become a more skilled professional—particularly after the third year (Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2007; Clotfelter, Ladd, & Vigdor, 2007).

Documenting the types and quality of opportunities prospective teachers have to learn on the path to certification gives researchers the chance to study the extent to which programmatic visions of the knowledge and skills prospective teachers need to master classroom tasks are realized. Additionally, if the goal is to develop teachers who are prepared to address the complexities inherent within the tasks of teaching mathematics as well as increase the likelihood of retaining them, we need to determine which coursework and field experiences are central to cultivating prospective teachers' professional knowledge and skills for teaching mathematics. Some countries prepare mathematics teachers at all levels as mathematics specialists; others prepare mostly primary generalists and secondary specialists. In the United States, mathematics specialists have become more in demand in the past decade as states have created primary mathematics specialist licensure (Association of Mathematics Teacher Educators [AMTE], 2013). Although what is essential for teachers to learn and the optimal timing of these learning experiences is debatable, there is consensus regarding the importance of opportunities to learn the foundations of mathematics pedagogy and instructional practice as well as to connect classroom learning to instructional practice. Indeed, prospective teachers with differential learning opportunities exit preparation programs with disparate levels of knowledge and skills, which has enormous implications for student learning and achievement. Thus, in this study, we identify subgroups of future primary mathematics specialists teachers characterized by specific patterns of opportunities to learn mathematics pedagogy.

The Teacher Education and Development Study in Mathematics (TEDS-M)

The data for this study come from the Teacher Education and Development Study in Mathematics (TEDS-M), an international comparative study of the preparation of primary and lower-secondary mathematics teachers. Data were collected from institutions, teacher educators, and future teachers from 17 developed and developing countries. The conceptual framework, design, and methodology of this study are thoroughly documented in various other reports and can be found online: https://www.ilsa-gateway.org/studies/factsheets/64.

Theoretical Framework

We frame this study with both theories of cultural contexts and theories connecting child development to the psychology of caregivers. Super and Harkness' (1986) developmental niche theory describes how cultural contexts shape child learning and development. The niche is composed of three subsystems: (a) the physical and social settings in which the child lives, (b) culturally regulated customs of child care and child rearing, and (c) the psychology of the caretakers and educators. For the purposes of this study, the latter subsystem, the psychology of the caregivers and educators, may prove to be instructive. Super and Harkness theorize the psychology of the caregiver organizes child care strategies (pp. 556–557), while recognizing the influence of constraints within the physical environment, customs of child care, and the demands of caregiver activities. We extend this logic to teacher preparation: We believe the psychology of future teachers—composed of beliefs about mathematics teaching and learning as well as professional bodies of knowledge germane to the tasks of teaching—serve as organizational influences that are related to future classroom practices.

Goodnow (2010) proposes four ways of specifying cultural contexts for empirical study: (a) multiplicity and context, (b) ideologies, values, and norms, (c) practices, activities, and routines, and (d) paths, routes, and opportunities. These approaches are not mutually exclusive of each other, but "paths, routes, and opportunities" (p. 10) are the lenses through which we study the intended and achieved outcomes of teacher preparation programs. "Paths," in Goodnow's view, refer to the stages or steps individuals are expected to follow as they move through social institutions. The concept of "paths/pathways" gives rise to questions regarding expected timetables (Neugarten, 1979), including the way one step is related to another, the skills needed for each step, and the flexibility afforded to those in need of alternative routes. Certainly, variability in path "access" and "availability," or opportunities to learn, may in part account for heterogeneity in outcomes (Goodnow, 2005) within teacher preparation programs and is the focus of the current study.

Thus, taken together, we consider multiple influences on outcomes, including academic achievement. Teachers' knowledge and beliefs about mathematics teaching and learning frame their future classroom practices. Understanding teachers' paths (opportunities to learn) in turn frame the development of their knowledge and beliefs, within the cultural contexts of their teacher preparation programs.

Review of Relevant Literature

Professional Knowledge for Teaching

Understanding the knowledge used in teaching can help stakeholders in mathematics education to develop a sense of what it means to teach mathematics well and how to prepare prospective teachers. Teachers need to cultivate knowledge, competencies, and skills that will help them analyze and understand student thinking to provide the appropriate support and strategies for learning mathematics (Ball, Thames, & Phelps, 2008; Dalgarno & Colgan, 2007; Hill & Lubienski, 2007; Kelly, Luke, & Green, 2008). In fact, mathematics content knowledge is necessary but not sufficient – teachers need subject-matter expertise (Schwab, 1978; Warfield, 2001), as well as mathematics pedagogical content knowledge *for* teaching (Ball, 1993; Ball et al., 2008; Lampert, 1990, 2001). Mathematical pedagogical content knowledge is a body of knowledge composed of what Ma (1999) refers to as "profound" mathematical knowledge that teachers draw upon as they calibrate what are appropriate learning goals, anticipate and analyze student misconceptions and errors, select and present representations of central mathematical concepts, and respond to student thinking and reasoning (Thames & Ball, 2010). Future teachers with a strong background in mathematics have a solid foundation to develop mathematics pedagogical content knowledge for teaching – if they are provided an appropriate set of preparation experiences.

Mathematics Specialists

Primary mathematics specialists are "teachers, teacher leaders, or coaches who are responsible for supporting effective mathematics instruction and student learning at the classroom, school, district, or state levels" (AMTE, 2013, p. 1). Within the TEDS-M database, primary mathematics specialists are prepared to teach one or two subjects (including mathematics), whereas their primary generalist peers are prepared to teach three or more subjects (Tatto et al., 2012). In general, mathematics specialists are expected to take more mathematics content courses on the path to certification. In seeking to study the influence of teachers' opportunities to learn on their mathematical pedagogical content knowledge, this study focuses on a group of teachers who, by virtue of their pathway to certification, had sufficient opportunities to learn mathematics. Thus, this paper focuses on primary mathematics specialists.

It is the norm in many East Asian countries that all students learn mathematics from mathematics specialists starting in first grade (e.g., China and Japan). Internationally, countries such as Singapore have a history of producing effective teachers and specialists, as evidenced by student performance on the Trends in International Mathematics and Science Study (TIMSS) of the International Association for the Evaluation of Educational Achievement (IEA) (Mullis, Martin, Foy, & Arora, 2012). Primary mathematics specialists are able to focus their energies on developing and teaching mathematics lessons, whereas primary generalists must also prepare many other lessons, including language arts, science, and social studies.

Within the United States, multiple stakeholders in mathematics education have released federal reports making the case that in-service primary teachers are not adequately prepared to meet the demands for increasing student achievement in mathematics (National Council of Teachers of Mathematics, 2000; National Mathematics Advisory Panel, 2008), given the poor mathematical preparation endemic to early childhood and primary educators (Graven, 2004; Grootenboer & Zevenbergen, 2008; Ginsburg, Lee, & Boyd, 2008; Hodgen & Askew, 2007; Lerman, 2012). Primary mathematics specialists have been identified as a promising strategy for improving early childhood mathematics teaching and learning (Reys & Fennell, 2003). Indeed, the AMTE (2013) and the Conference Board of Mathematical Sciences (CBMS, 2012) have each published position statements advocating for the establishment of a primary specialist license in the United States. There is growing evidence of the effectiveness of primary mathematics specialists for increasing student mathematics achievement from the Vermont Mathematics Initiative (Meyers & Harris, 2008) as well as the states of Ohio and Virginia (Brosnan & Erchick, 2010; Campbell & Malkus, 2011; Campbell, Ellington, Haver, & Inge, 2013).

Theory, empirical studies, and wisdom of practice suggest mathematics content knowledge is necessary but not sufficient for high-quality mathematics teaching. Thus, mathematics specialists need more than just knowledge of mathematics content. Recently, Campbell et al. (2013) released a handbook focusing on primary mathematics specialists, outlining requisite knowledge-based skills and abilities that included mathematical content knowledge and mathematical pedagogical content knowledge described earlier in this section. Campbell et al. (2013) also suggested specialists need: coaching strategies and skills, knowledge of mathematics curricula, knowledge of special populations of students, knowledge of assessment, and knowledge of research and resources. The foundation for the development of the aforementioned skills can be laid down in preparation programs, but must be animated through field experiences. It may be the case that prospective mathematics specialists benefit from field experiences in school/classroom settings where they are given opportunities to observe and participate in the daily work of teaching, as well as encounter and attempt to make sense of student thinking and reasoning.

Opportunities to Learn

The concept of opportunity to learn (OTL) was introduced by the IEA (e.g., the First and Second International Mathematics Studies) in the 1960s and was considered to be a technical concept conceived as a means to ensure the validity of cross-national comparisons in mathematics achievement. OTL captured curricular differences as "…a measure of whether or not students have had an opportunity to study a particular topic or learn to solve a particular type of problem presented by the test" (Husen as cited in Burnstein, 1993, p. xxxiii).

McDonnell (1995) outlines the evolution of the use of the OTL as a technical concept for research and its utility in policy debates in the 1990s. OTL entered policy debates under the premise that schools needed to provide students with “adequate” opportunities to learn before schools could be held accountable for meeting achievement standards. As a research tool, OTL was envisioned as an indicator that could help unpack the proverbial “black box” connecting school inputs and student outcomes.

Ingvarson, Beavis, & Kleinhenz (2007) approached the question of OTL in the context of teacher education in their attempt to identify the characteristics of effective teacher-preparation programs, as reported by novice teachers who had just completed their first year of teaching. The purpose of this study was to provide guidance for policymakers regarding the standards that might be appropriate for assessing and accrediting teacher education programs to ensure graduates were well prepared to meet the demands of classroom teaching. Ingvarson and colleagues postulated there were three main factors associated with novice teachers’ preparedness to teach: personal background characteristics, pre-service courses and coursework (OTL), and the characteristics of the school where graduates had their first teaching position.

To assess the extent to which novice teachers felt prepared to teach, Ingvarson et al. (2007) administered the Teacher Preparedness Survey to teachers beginning in their second year of teaching. In this study, OTL refers to both the *form* and *substance* of learning experiences in teacher preparation programs in four domains (pp. 357–359): (a) “opportunity to learn content knowledge and how it is taught,” (b) “opportunity to learn the practice of teaching,” (c) “opportunity to learn via feedback from university staff,” and (d) “opportunity to learn assessment and planning.” The OTL variables were regressed onto the Australian Council for Educational Research Teacher Preparedness Inventory (TPI). The TPI is composed of three factors (and their subscales): professional knowledge (*professional knowledge and how to teach it* and *professional knowledge about students and how they learn*); professional practice (*professional practice to do with curriculum, professional practice to do with classroom management*, and *professional practice to do with assessment*); and professional engagement (*reflection on teaching* and *work with parents and others*).

Significant relationships were found between professional knowledge and the OTL domains of *content knowledge and how it is taught* and *assessment and planning*. The OTL via *feedback from university staff* was also significant, but these coefficients were smaller. When the outcome was defined as perceptions of preparedness to teach, the OTL domain *the practice of teaching* had a strong effect, whereas the OTL domains *content knowledge and how it is taught* and *assessment and planning* had moderate effects. OTL variables, as defined in this study, had the strongest and most consistent effects on TPI scores and teacher perceptions of their preparedness to teach in their first year. The effects of this group of OTL variables were independent of the background characteristics of the teacher, the teacher’s in-school experiences during pre-service courses, and the school in which the teacher worked during his or her first year as a teacher. All of this suggests that better understanding of OTL can allow us to make practical policy recommendations for improving teacher education practices. Additionally, OTL in areas that can be considered *connecting theory to practice* seem to be particularly important for predicting teacher professional knowledge.

Opportunity to Connect Classroom Learning to Practice

The definition and conceptual argument regarding how theory relates to and can be used in practice have been topics of debate with respect to teacher preparation—notably, in the United States (e.g., Shulman, 1998), the United Kingdom (e.g., Carr, 1992, 1995, 2003), the Netherlands (e.g., Korthagen & Kessels, 1999), and Asian countries (e.g., Deng, 2004). Resolving this debate is outside the scope of this study. We thus subscribe to its most basic definition as described by the TEDS-M framework (Tatto et al., 2008): theory is a body of empirical findings that can be used to anchor prospective teachers' interpretation of classroom events as they arise, make instructional decisions specific to the context of their classrooms, and assess and evaluate the outcomes of those decisions.

The importance of the connection between pedagogical theory and practice can be understood through the lens of situated cognition theory, which suggests professional knowledge, competencies, and skills are situated in and inseparable from the activities, context, and culture in which they are constructed (Brown, Collins, & Duguid, 1989). Situated cognition is connected to Goodnow's (2010) paths, as teachers reflect upon their opportunities to learn within their cultural contexts. Learning to teach, therefore, is a process of enculturation: prospective teachers are apprenticed into particular practices and modes of thinking (Lortie, 1975) aligned with local cultural contexts (Goodnow). Field experiences are a context where future teachers have opportunities to cultivate sound professional judgment stemming from "…a coherent, enlightened, integrated body of knowledge that will inform, and in turn be informed by, classroom practice" (Calderhead & Robson, 1991, p. 1). Indeed, field experiences are a context in which prospective teachers' mathematical pedagogical content knowledge for teaching can develop.

Pedagogical content knowledge is composed of two key components (Shulman, 1986): (a) knowledge of student thinking, understanding, and difficulties with particular topic strands and concepts and (b) knowledge of strengths and weaknesses of particular strategies and representations for teaching these topics. Crespo (2000) focuses on the first strand of pedagogical content knowledge by examining how prospective Canadian teachers in the middle of their two-year preparation programs interpreted fourth-grade students' mathematical thinking and reasoning through a mathematics letter exchange program. Early analysis of the letters and interviews suggested prospective teachers were fixated on whether students generated the correct answers and were quick to make inferences about students' mathematics abilities and dispositions toward learning. However, after four or five rounds of correspondence, prospective teachers began to focus less on answers and more on students' mathematical thinking. Moreover, prospective teachers began to question and revise claims about students' mathematics abilities and attitudes, more skillfully distinguishing between describing and making inferences about student thinking. Crespo suggests the latter finding emerged in light of prospective teachers being faced with contradictory data gathered from letter correspondence coupled with meeting their letter partners and spending time in their classrooms. This study highlights how acquiring access to students' mathematical thinking and reasoning, in coursework and in the field, can alter how prospective teachers see, talk, listen, and act toward their students.

Although the theory underlying pedagogical content knowledge seems intricately connected with practice, in actuality, teachers do not always have the necessary OTL or time to connect theory to practice. Allen and Wright (2014) followed and interviewed one group of prospective teachers regarding the factors that enabled or hindered their abilities to integrate classroom theory and practice during a three-week field experience in the first year of their teaching programs in Australia. The authors report three central themes from semi-structured follow-up interviews with 11 teachers. First, prospective teachers valued both theoretical and practical components of their graduate-level preparation programs—not privileging one at the expense of the other (contrary to other empirical studies that find practice being privileged over theory—e.g., Allen, 2009; Hartocollis, 2005). Second, teachers' opportunities to connect classroom learning to practice varied as a function of the clarity of stakeholders' roles and responsibilities. Third, prospective teachers supported the notion of linking university coursework assessment to field experience as a means of bridging the gap between theory (the university classroom) and practice (field experience). Together, these themes reflect prospective teachers' recognition that their competence as educators is in part reliant upon the development of what Cochran-Smith and Lytle (1999) refer to as the knowledge-for-practice (i.e., formal knowledge generated by university-based scholars for teachers to use in order to improve practice) and knowledge-in-practice (i.e., knowledge that is embedded within classroom practice and teacher reflection on practice).

Imre and Akkoç (2012) examine the link between professional knowledge and field experiences more directly. Their case study closely examines the development of pedagogical content knowledge for number patterns in three prospective teachers (in the last year of their four-year programs) through a school field experience course in Turkey. The authors used prospective teachers' lesson plans, videos of micro-teaching lessons, and follow-up interviews to examine the extent to which prospective teachers took student understanding and difficulties during micro-teaching. Analysis suggested observations in real classroom settings and discussions of those observations with university faculty and peers were responsible for improvement of prospective teachers' pedagogical content knowledge. The authors further postulate that observing students in classrooms helped prospective teachers identify students' understanding of patterns, the difficulties students encounter, and specific strategies mentors use in real time. Thus, field placements are where prospective teachers have the opportunity to encounter, attend, and respond to student thinking, fertilizing the ground in which pedagogical content knowledge grows.

Latent Class Analysis

The extent to which individually varying patterns of university- and field-based OTL exist and contribute to differential levels of professional knowledge associated with high-quality teaching is unclear. However, within the framework for linear models, we are not able to observe whether some groups of prospective teachers have different patterns of OTL. Indeed, it may be the case that knowledge does not vary as a function of greater or fewer opportunities to learn—it may be that some patterns of OTL are more consequential to the development of knowledge than others. If this is the case, latent class analysis may leverage our ability to investigate this hypothesis.

Latent class analysis (LCA)[1] is a type of latent variable mixture modeling—a flexible, person-centered analytic tool focused on similarities and differences among individuals—standing in contrast to statistical modeling that focuses on relations among variables (Berlin, Williams, & Parra, 2013; Muthén & Muthén, 1998). The goal of LCA is to identify homogeneous subgroups of individuals who possess a unique set of characteristics that differentiates them from other subgroups. Thus, within the LCA framework, subgroup membership is inferred from, not observed in, the data. This method empirically subdivides individuals and places them in groups that are characterized by sharing similar "domains" of OTL. Here we use *domains* to refer to related sets of opportunities to learn (cf. Ingersoll, Merrill, & May, 2014). Thus, the latent class analysis looked for distinct patterns of OTL shared by subgroups of prospective teachers within each country.

Research Questions

Is there a latent subgroup structure that adequately represents the heterogeneity of opportunities to learn among mathematics specialists across the United States and Singapore? If so, what are the types and their corresponding prevalence?

Hypotheses: We expect to find more latent subgroups within the United States, where there are multiple pathways to certification that have extremely different OTL about connecting theory and classroom practice, than in Singapore, which has only one centralized institution that prepares teachers.

Method

The data used for this study were part of the larger TEDS-M study, in which 22,078 future teachers from 17 countries are represented. However, for the purpose of the current study, we focus on the sub-sample of future primary mathematics specialists from two countries with complete data: the United States and Singapore. We restricted our sample to future primary mathematics specialists because in studying the associations among OTL mathematics pedagogy domains, we know mathematics pedagogy is in some sense dependent on mathematical content

[1] Readers interested in more information about Latent Class Analysis may explore the extensive materials available from The Methodology Center at Pennsylvania State University's College of Health and Human Development: https://methodology.psu.edu/ra/lca

Table 10.1 Mean professional knowledge scores by primary mathematics specialists by country

Professional knowledge	United States (*n* = 191)	Singapore (*n* = 117)
Mathematical content knowledge	555 (7)	600 (8)
Pedagogical content knowledge for teaching mathematics	534 (7)	604 (7)

knowledge: teachers do not typically have strong pedagogy related to mathematics content they do not understand deeply. By focusing on mathematics specialists, we hoped the sample would contain teachers with adequate mathematical content knowledge, enabling us to focus on the OTL associations. We chose to include Singapore in the present analysis for two reasons. First, we wanted to choose a country with high mathematics content knowledge scores for primary math specialists, in order to clarify the how OTL mathematics pedagogy relate to each other. As can be observed in Table 10.1, both countries have mathematical content knowledge and pedagogical knowledge scores that are above the international mean of 500; Table 10.1 illustrates country means and standard deviations (in parentheses). Second, Singapore has different qualifications for entry into the teaching profession and routes to certification than the United States. Although the models we specify to answer our research question are not intended to be used for direct comparison across countries, interpreting findings descriptively can fortify our discussion with respect to how different "paths" and "routes" made accessible through OTL are associated with different preparation program outcomes.

Measures

Opportunity to Learn Latent Class Analysis Variables

We selected three types of OTL factors to test the existence of latent subgroups. Two of these factors, *opportunity to learn instructional practice* and *opportunity to connect classroom learning to practice,* had categorical response formats, whereas *opportunity to learn mathematics instruction* had a binary response format. The item responses for the variables that composed the *opportunity to learn instructional practice* and *opportunity to connect classroom learning to practice* factors were recoded to binary responses, consistent with Blömeke (2012). We acknowledge that this recoding results in the loss of variability. Yet, this recoding makes it possible to distinguish more clearly between OTL profiles.[2] In the TEDS-M survey, "opportunity to learn mathematics pedagogy" is a categorical variable where the response options were coded as 1 (*never*), 2 (*rarely*), 3 (*occasionally*), and 4 (*often*). Such response options focus on *frequency* of OTL; but by capturing mainly frequency, it is

[2] Although a Latent Profile approach would allow for a greater number of responses, results of such analyses are not easily interpretable.

assumed all opportunities are of equivalent quality. We focus our attention on whether prospective teachers report having had any one particular learning opportunity.

Opportunity to Learn Mathematics Instruction

This factor is composed of five binary response items with answers 1 (*did not study*) or 2 (*did study*), which were included in the LCA. Future mathematics specialists were asked to indicate whether they studied a particular topic as part of their teacher preparation program, such as:

- Mathematics instruction (e.g., representation of a mathematical concept);
- Developing teaching plans (e.g., selection and sequencing of mathematics content);
- Observation, analysis, and reflection;
- Mathematics standards and curriculum; or
- Affective issues in mathematics (e.g., anxiety).

Opportunity to Learn Instructional Practice

This factor was composed of six items that used a 4-point ordinal response format, coded as 1 (*never*), 2 (*rarely*), 3 (*occasionally*), and 4 (*often*). Since LCA is based on categorical data, the ratings were transformed into binary codes with answers 1 (*never/rarely*) or 2 (*occasionally/often*). Future mathematics specialists were asked to indicate how frequently they engaged in activities such as:

- Explore how to apply mathematics to real-world problems;
- Explore mathematics as the source for real-world problems;
- Learn how to explore multiple solution strategies with pupils;
- Learn how to show why a mathematics procedure works;
- Make distinctions between procedural and conceptual knowledge when teaching mathematics concepts and operations to pupils; or
- Integrate mathematical ideas from across areas of mathematics.

Opportunity to Connect Classroom Learning to Practice

This factor is composed of eight items that used a 4-point ordinal response format coded as 1 (*never*), 2 (*rarely*), 3 (*occasionally*), and 4 (*often*). Again, the ratings were transformed into binary codes with answers 1 (*never/rarely*) or 2 (*occasionally/often*). Future mathematics specialists were asked to indicate how frequently they engaged in activities such as:

- Observe models of teaching strategies you were learning in your courses;
- Practice theories for teaching mathematics that you were learning in your courses;

- Receive feedback about how well you had implemented teaching strategies you were learning about in your courses;
- Collect and analyze evidence about pupil learning as a result of your teaching methods;
- Develop strategies to reflect upon your professional knowledge;
- Demonstrate that you would apply the teaching methods you were learning in your courses;
- Complete assessment tasks that asked you to show how you were applying ideas you were learning in your course; or
- Test out findings from educational research about difficulties pupils have in learning.

For more information, refer to the technical report by Tatto (2013), which is also available on the TEDS-M website.

Covariates

Based on the TEDS-M results more generally (Tatto, Rodriguez, Reckase, Rowley, & Lu, 2013), we included the following variables as covariates: gender, the number of books in home (as a proxy for socioeconomic status), and grades in high school (as a proxy for prior achievement). Given the TEDS-M results for countries, it is reasonable to expect all of these variables to interact significantly with OTL, and thus we controlled for these in our analyses. By restricting our sample to prospective mathematics specialists, we thus did not control for mathematical content knowledge, since as a group, specialists have higher content knowledge.

Analytical Method

We used latent class analysis (Hagenaars & McCutcheon, 2002; Lanza, Dziak, Huang, Wagner, & Collins, 2015; McCutcheon, 1987) in Mplus (Version 6.11, Muthèn and Muthèn 1998–2012) to identify subgroups of future teachers with specific patterns of opportunities to learn mathematics education pedagogy. This is a person-centered analytic approach focused on similarities and differences among individuals instead of relations among variables (Muthén & Muthén, 1998–2012). This particular person-centered approach has been used before on the TEDS-M database in Blömeke (2012; also Blömeke, Hsieh, Kaiser, & Schmidt, 2014). Not all items were used, as some items did not demonstrate any variability of OTL within subgroups. This is an acceptable practice within the LCA framework (e.g., Kim, Wang, Orozco-Lapray, Shen, & Murtuza, 2013; Weaver & Kim, 2008).

To determine the optimal number of latent subgroups, one would ideally apply a bootstrap as a dimension of fit criteria to consider. However, this option was not available to us if we wanted to include the TEDS-M sampling weights. We determined that it was important to include the sampling weights because they enable us to make observations about the latent subgroup composition that are generalizable to prospective mathematics specialists who are prepared within the same country.

For each country, we specified alternative models ranging from two to five subgroups. Model assessment and selection were also based on a variety of other fit criteria, including the log likelihood, Akaike's Information Criterion (AIC; Akaike, 1974), Bayesian Information Criterion (BIC; Schwarz, 1978), sample-size adjusted BIC (SSBIC; Sclove, 1987), and entropy. Smaller AIC, BIC, and SSBIC values indicate better fit; BIC in particular is an optimal indicator for LCA classes.[3] The entropy statistic ranges from 0 to 1 and is a standardized summary measure of the classification accuracy of placing respondents into subgroups based on their model-based posterior probabilities. Thus, entropy values closer to 1 reflect better classification of individuals (Ramaswamy, DeSarbo, Reibstein, & Robinson, 1993). Using a combination of model fit indices strengthens the reliability of latent subgroup enumeration (Muthén, 2003). Lanza, Collins, Lemmon, and Schafer (2007) also suggest model interpretability should be considered: each latent subgroup should be distinguishable from others based on item-response probabilities; latent subgroups should not be trivial in size (i.e., with a near-zero probability of membership); and it should be possible to assign a meaningful label to each subgroup.

Results

Latent Class Analysis

Tables 10.4 and 10.5 in the Appendix show the distribution of all variables used to select the base model for each country.

Baseline Model Selection

For all selected optimal solutions derived from latent class analyses, the AIC and BIC were the lowest, or the decline between two sequential models leveled off. The optimal solutions for each country are presented in Tables 10.2 and 10.3. In the discussion that follows, the number of subgroup profiles are described and labeled.

Within each country, a latent subgroup profile was labeled according to how it compares with other subgroup profiles on the three dimensions of OTL (mathematical instruction, instructional practice, and connecting classroom learning to practice). Figures 10.1, 10.2, 10.3, 10.4, 10.5 and 10.6 depict future mathematics specialists'

[3] The Latent variable mixture modeling discussion group on the Mplus webpage devotes considerable discussion to this topic, and responses favoring BIC include some by Muthén, author of Mplus. For more information, see http://www.statmodel.com/discussion/messages/13/13.html?1462022592

Table 10.2 Goodness of fit criteria for various latent class models for United States (n = 191)

Number of classes	# of parameters	Log likelihood	AIC	BIC	SSBIC	Entropy
1	25	−1834	3715	3799	3720	–
2	42	−1136	2356	2481	2348	.878
3	**65**	**−1076**	**2282**	**2475**	**2270**	**.901**
4	82	−1045	2254	2498	2239	.916
5	111	−1021	2265	2596	2245	.893

Note: Dashes indicate criterion was not calculated for the model. Bold indicates the selected model.

Table 10.3 Goodness of fit criteria for various latent class models for Singapore (n = 117)

Number of classes	# of parameters	Log likelihood	AIC	BIC	SSBIC	Entropy
1	25	−1583	3215	3284	3205	–
2	42	−1074	2233	2349	2216	.852
3	62	−1034	2192	2363	2167	.904
4	**88**	**−1004**	**2184**	**2427**	**2149**	**.930**
5	102	−981	2167	2449	2126	.907

Note: Dashes indicate criterion was not calculated for the model. Bold indicates the selected model

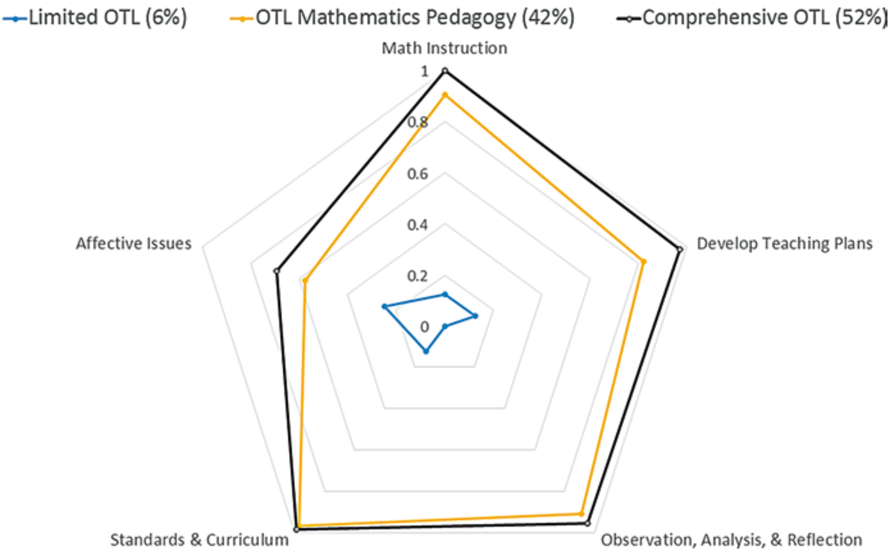


Fig. 10.1 Opportunity to learn mathematics pedagogy in the United States

opportunities to learn conditional on latent subgroup membership. Please note the items are discrete; the lines connecting one OTL variable to another are present to more easily see the differences between subgroups. We applied a probability of .75 to determine whether subgroups had OTL each item; groups reporting an average OTL

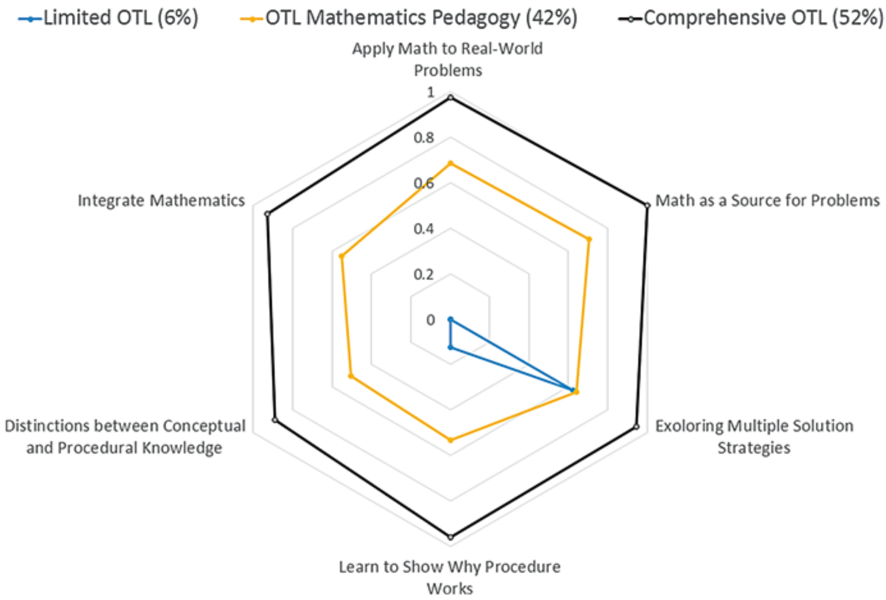


Fig. 10.2 Opportunity to learn instructional practice in the United States

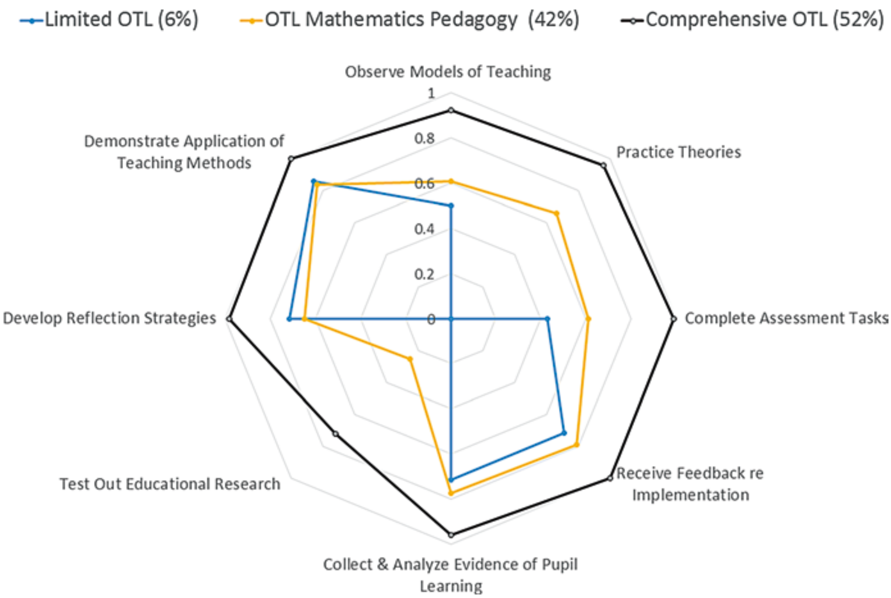


Fig. 10.3 Opportunity to connect classroom learning to instructional practice in the United States

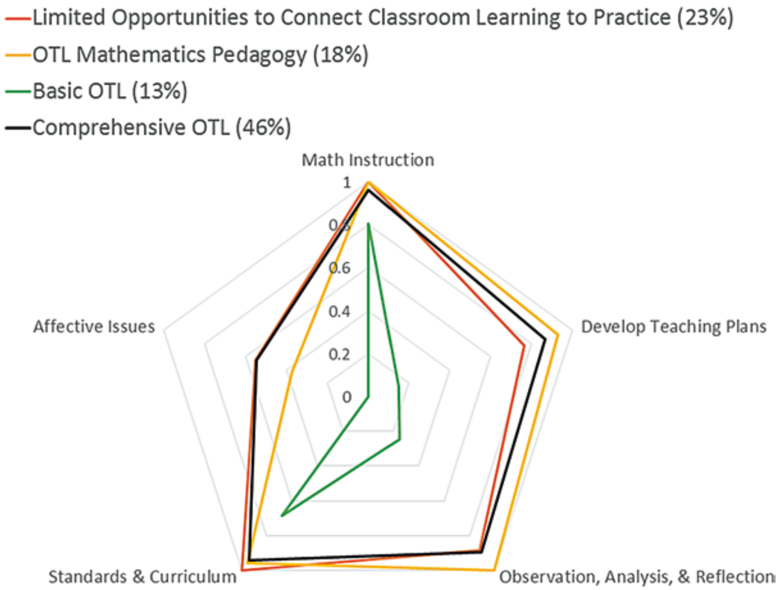


Fig. 10.4 Opportunity to learn mathematics pedagogy in Singapore

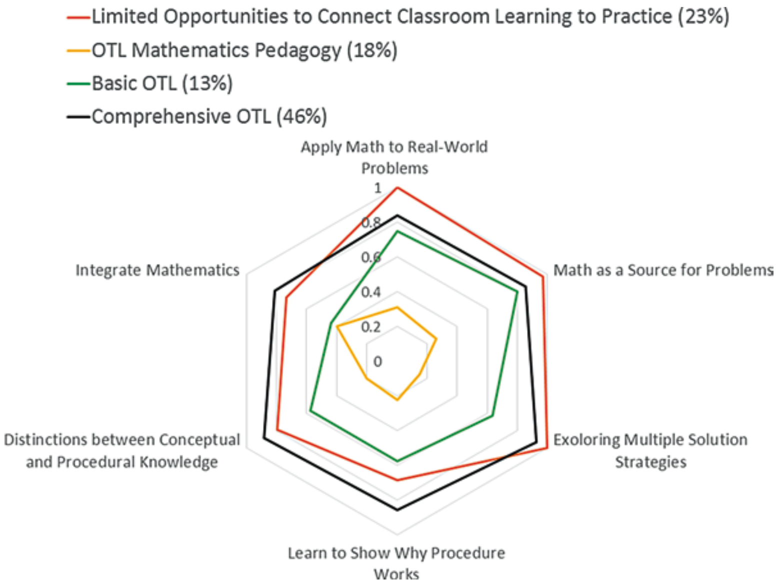


Fig. 10.5 Opportunity to learn instructional practice in Singapore

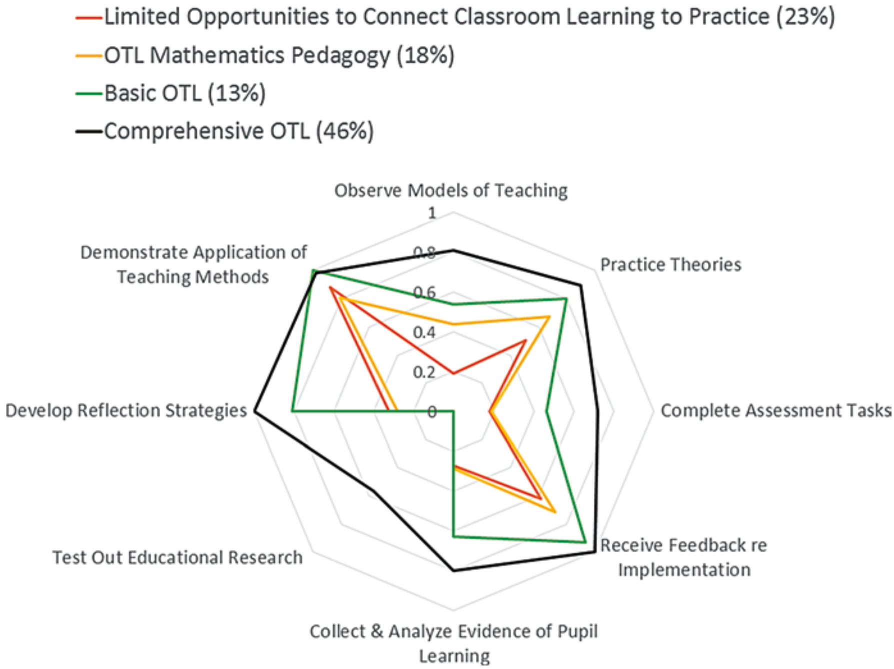


Fig. 10.6 Opportunity to connect classroom learning to instructional practice in Singapore

of more than .75 were considered to have had sufficient opportunities to learn that particular domain. Tables 10.6 and 10.7 (in the Appendix) depict the parameter estimates for the optimal latent subgroup solution for each country. We interpret the model parameters as the probability of any subgroup of prospective mathematics specialists reporting having had the OTL. For example, in the first row for Table 10.6, there is a 91% and 100% probability that prospective mathematics specialists in the *Mathematics Pedagogy* and *Comprehensive OTL* subgroups, respectively, report having had OTL mathematics instruction. However, there is only 13% probability that prospective specialists in the *Limited OTL* would report having had the same OTL.

In the United States, three latent OTL subgroup profiles emerged. The first latent subgroup comprises 6% of prospective U.S. teachers and is depicted by blue lines in Figs. 10.1, 10.2 and 10.3. Members of this group, which we refer to as *Limited OTL*, report few opportunities to learn any of the mathematical pedagogical skills of interest. Whereas 6% may seem small, it represents a non-trivial proportion of a representative sample of pre-service teachers. Indeed, approximately one out of 20 teachers report limited OTL across all three domains. The second subgroup comprises 42% of prospective teachers and is depicted by yellow lines. We characterize this group as having *OTL mathematics pedagogy*. This group had lower probabilities of reporting OTL instructional practice. This group also had lower probabilities of reporting having opportunities to connect classroom learning to instructional practice, with the exception of collecting and analyzing evidence of pupil learning as a result of their teaching methods; to demonstrate that they could apply the teaching methods they were learning about in coursework; and to receive feedback about how well they had implemented teaching strategies they were learning about in coursework. The third subgroup comprises 52% of prospective teachers. Depicted by black lines, this subgroup is characterized as having *comprehensive OTL*, although members report lower probabilities of both covering affective issues in mathematics and testing out findings from educational research about difficulties pupils have in learning in their coursework.

Figures 10.1, 10.2 and 10.3 are radar graphs that depict the profiles of OTL among the three subgroups. The vertices of each figure represent items within one of the OTL domains. The lines within the shape depict probability levels for each item. For example, at the top of Fig. 10.1, the probability of subgroups reporting the OTL mathematics pedagogy is nearly 100% for Subgroups 2 (*OTL Mathematics Pedagogy*) and 3 (*Comprehensive OTL*), but 12% for Subgroup 1 (*Limited OTL*).

In Singapore, four latent OTL subgroup profiles emerged. The first latent subgroup comprises 23% of prospective mathematics specialists and is depicted by orange lines in Figs. 10.4, 10.5 and 10.6. This subgroup can be characterized as having *limited opportunities to connect classroom learning to instructional practice,* although they do report being expected to demonstrate their ability to apply teaching methods they were learning about in coursework. Additionally, these prospective specialists had relatively low probabilities of reporting opportunities to study affective issues in mathematics and opportunities to learn how to show why a procedure works. The second subgroup comprises 18% of prospective mathematics specialists and is depicted by yellow lines. This subgroup was characterized as having *OTL mathematics pedagogy* but limited OTL instructional practice and OTL connecting classroom learning to instructional practice. The third subgroup comprises 13% of prospective specialists and is depicted by green lines. This subgroup was characterized as having *basic OTL*. Prospective teachers in this group reported experiencing what could be considered a fundamental set of opportunities to learn to teach mathematics from each of the three OTL domains, which included OTL math instruction exploring how to apply mathematics to real-world problems and seeing math as a source for real-world problems. They also reported some opportunities to connect theory to instructional practice, including the opportunity to practice theories for teaching mathematics they were learning about in coursework, demonstrate that they could apply the teaching methods they were learning about in coursework, receive feedback about how well they implemented teaching strategies they were learning about in coursework, and develop reflection strategies. The fourth subgroup comprises 46% of prospective specialists and is depicted by black lines. This subgroup is characterized as having *comprehensive OTL*, although they did not report covering affective issues in mathematics, completing assessments tasks that required them to apply ideas they were learning about through coursework, or testing out findings from educational research about difficulties pupils have in learning. Figures 10.4, 10.5 and 10.6 are radar graphs that depict the profiles of OTL between the four subgroups.

Discussion

The goal of this study was to identify distinct profiles of OTL within the United States and Singapore. Since the TEDS-M data encompasses weighted samples, the prospective math specialists included in this analysis can be considered to be representative of mathematics specialists in their countries. Multiple profiles of OTL were found in each country, even after controlling for the effect of gender and proxies for socioeconomic status and prior achievement. These subgroups can be labeled with respect to OTL mathematics instruction, instructional practice, and opportunities to connect classroom learning to practice. In the United States, three subgroups existed: *Comprehensive OTL* (52%), *OTL Mathematics Pedagogy* (42%), and *Limited OTL* (6%). These groups did not overlap much in their relative OTL the different domains of mathematics pedagogy. Relative to the *Comprehensive OTL* subgroup, the *OTL Mathematics Pedagogy* subgroup has slightly fewer OTL mathematics pedagogy (specifically, affective issues and developing teaching plans), but distinctly fewer OTL connect classroom learning to practice and OTL instructional practice.

In Singapore, on the other hand, four subgroups existed. Unlike those in the United States, these subgroups varied in which one reported the fewest opportunities to learn the different mathematical pedagogical domains. The Singapore subgroups are *Comprehensive OTL* (46%), *Limited Opportunities to Connect Classroom Learning to Instructional Practice* (23%), *Basic OTL* (13%), and *Limited OTL* (18%). The *Basic OTL* group presents an interesting pattern, with respondents reporting adequate opportunities to learn foundational pedagogy and develop the skills to participate in the most "basic" parts of the teaching cycle, such as opportunities to demonstrate their ability to enact teaching practices that are grounded in classroom theory, receive feedback on the quality of their implementation of teaching methods, and develop the capacity to reflect upon how these experiences have shifted their professional knowledge and understanding of teaching and learning.

Our hypothesis that the United States, with more pathways (Goodnow, 2010) to certification, would have more subgroups, was not confirmed by the data. Prospective teachers in the United States have numerous options for becoming teachers and specialists, including public and private institutions, consecutive and concurrent routes, and widely varying course and field requirements. Teachers in the United States are prepared at more than 1300 institutions in all 50 states, and although the United States has moved toward more centralized certification policies at the state level (Ingvarson et al., 2013), there is still great variation. We had thought that, given the singular teacher preparation institution in Singapore, prospective teachers there would be more uniform in their reported OTL. However, within the National Institute of Education in Singapore, there are 11 different teacher preparation programs. Primary math specialists can be trained via either a concurrent or consecutive program. The *TEDS-M Encyclopedia* (Schwille, Ingvarson, & Holdgreve-Resendez, 2013) reports great variation in the qualifications of supervisors in Singapore. There is also extensive variation in the required courses and durations of the different types of programs. Future research could look more closely at the Singapore teacher variation in OTL and explore connections to specific preparation programs with the National Institute of Education. Although the purpose of this study is not to statistically compare differential OTL between future mathematics specialists in the United States and Singapore, our findings may be instructive for program and thought leaders concerned with the extent to which programmatic visions are being achieved.

Differential OTL naturally raises issues related to teaching quality and equity. Certainly, differential preparation of teachers has significant implications for student access to highly qualified teachers. Within the United States, disadvantaged children living in urban or poor rural areas are disproportionally taught by teachers with lower qualifications: they have less teaching experience, fewer certifications and advanced degrees, and come from preparation institutions with lower levels of selectivity (e.g., Darling-Hammond, 2000; Jerald, 2002). International comparisons of programs (including descriptive, exploratory studies such as this one) enable reflection on other possibilities for a given country. What does Singapore—whose specialist programs contain greater variability than that of the United States and whose students have historically and presently done well in assessments such as the PISA and TIMSS—do to ensure equitable allocation of highly qualified teachers?

Opportunities to Connect Classroom Learning to Practice

In both the United States and Singapore, approximately half of future mathematics specialists report comprehensive OTL (52% and 46%, respectively). However, the other half of future specialists in both countries report limited opportunities to connect classroom learning to instructional practice. We wonder about what happens in the classrooms of novice teachers who have strong mathematical content knowledge, but report limited opportunities to observe other teachers in action, to experiment with and explore teaching methods in ways that serve to organize their professional bodies of knowledge and skills, or to encounter student thinking and reasoning from one moment to the next. This is particularly consequential for the United States, which is shifting toward developing mathematics specialists: Are future mathematics specialists really given the best possible professional start toward developing the skills to enact the tasks of teaching (Thames & Ball, 2010), including those outlined by Campbell et al. (2013), if nearly half of them report not having opportunities to translate classroom learning to instructional practice?

If we subscribe to situated learning theory (Brown et al., 1989) and recognize the power of learning *in* and *from* practice (Cochran-Smith & Lytle, 1999; Darling-Hammond, 1998; 2009), then, in order to address limited opportunity to translate theory to practice, preparatory institutions may need to re-examine specific intended and achieved programmatic inputs as they relate to bridging this gap. Alternatively, it may be the case that some prospective specialists have found it difficult to connect field experiences with course content, for a variety of possible reasons. For example, there may have been a mismatch between course content and the field experiences being offered, or it may be that the connection between theory and practice was not facilitated by the course instructor. It may simply be the case that some students did not self-advocate and request particular learning opportunities or simply overlooked them. Primary math specialists may enter preparation programs already trained as primary generalists, in which case, they may not have the same OTL in some areas, such as math pedagogy, insofar as programs would assume prospective specialists had already acquired some basic knowledge. Particularly for consecutive routes to specialist certification, programs may require a bachelor's degree focused on primary mathematics, and thus would only include OTL in more specialized aspects of teaching mathematics. Nevertheless, field experiences are a place where the tension between classroom theory and practice can be made productive, particularly when questions about teaching and learning arise in the context of interacting with real students and work in progress. Indeed, well-designed clinical experiences are a setting that can "…empower [future] teachers with greater understanding of complex situations rather than seek to control them with simplistic formulas or cookie cutter routines" (Darling-Hammond, 1998, p. 170).

Limitations

The findings of this study need to be considered in light of the following limitations. First and foremost, selecting the optimal number of subgroups is not straightforward, as it requires the triangulation of fit statistics along with consideration of model interpretability. Further, whereas the fit indices for weighted and un-weighted samples both indicated the same number of latent classes, we could not perform LCA bootstrap on the weighted sample, because of limitations in statistical software packages. Consequently, the optimal number of latent subgroups present within the analyzed sample of each country is open to interpretation. Although our decisions align with our research question and related literature, others could make different decisions and also provide support for those decisions (e.g., to allow subgroups that capture smaller proportions of the sample, select a different subgroup solution). Additionally, model fit indices do not perform optimally with fewer than 100 observations, and a minimum of 200 observations is preferred (Nylund, Asparouhov, & Muthén, 2007). The standard, but not the preference, was met for both countries.

The latent subgroups are specific to those about to be certified as math specialists at the primary level. These participants are potentially different from those being certified as primary generalists. A future study should determine whether these same latent subgroups are present in other populations, including those from other countries and earning different types of certification. For our purposes, we were looking for associations among those with potentially high mathematical knowledge, so the restriction to math specialists was reasonable.

Further, the data are self-reported. Participants were asked to complete a survey and report whether they had opportunities to learn each of 19 topics. Self-reports of opportunities to learn how to connect classroom learning and practice are not the same as direct observation of teachers connecting classroom learning to their practices, through classroom observations and interviews. Furthermore, knowing whether participants had the opportunities to learn particular topics does not give us insight into the quality of these learning experiences. However, the novice teacher questionnaire utilized by TEDS-M does have good psychometric properties (Tatto et al., 2013), and research shows students' perceptions of learning are related to their overall evaluation of courses and to "actual" learning (Centra & Gaubatz, 2005).

Because of the differences in the items on the survey instrument, all participant responses were coded using a forced binary response. Whereas the LCA models binary responses, forcing 4-point scales into binary responses reduces the variability of the data. Although Latent Profile Analysis can handle responses with more than two categories, results of such analyses are not easily interpretable. Thus, LCA with constrained binary responses was considered preferable, in order to interpret the results.

Despite these potential limitations, this study provides us with a way to describe potential differences in OTL. More research is needed to investigate OTL, particularly examining the quantity and quality associated with different OTL. Coupling self-report data with additional measures such as document and observational data from programs would aid in producing a more robust description of OTL and its potential influences.

Conclusions

This study utilized a person-centered approach to identify different subgroups of prospective teachers who share OTL. The findings highlight significant differences in patterns of OTL that would not have been identified using variable-centered methods. This approach allows for meaningful distinctions to be made among opportunities to learn common across teacher preparation programs.

The results of this study inform institutional policies by providing a more complete and complex understanding of the reported OTL of prospective mathematics specialists. In both the United States and Singapore, distinct groups emerge with markedly different reported OTL mathematics pedagogy. Future studies can more closely examine the alignment between the OTL that pre-service teachers perceive and the OTL institutions see their preparation programs as encompassing. Further research can also examine the associations among OTL, mathematical content knowledge, and mathematical pedagogical content knowledge. Teacher preparation institutions can examine their curricula to determine whether the OTL they are providing for pre-service teachers are lacking in some of the key areas of mathematics pedagogy.

Appendix

Table 10.4 Frequency distributions for seventeen observed variables from the TEDS-M future teacher survey: Percent of future teachers who report opportunities to learn in the United States

	% Studied
Opportunity to learn mathematics instruction	
Mathematics instruction	90.9%
Develop teaching plans	85.6%
Observation, analysis, and reflection	89.4%
Mathematics standards and curriculum	93.2%
Affective issues in mathematics	62.1%
Opportunity to learn instructional practice	
Explore how to apply mathematics to real-world problems	79.4%
Explore mathematics as the source of real-world problems	80.2%
Learn how to explore multiple solution strategies with pupils	78.6%
Learn how to show why a mathematics procedure works	72.5%
Make distinctions between procedural and conceptual knowledge when teaching mathematics concepts and operations to pupils	65.6%
Integrate mathematics ideas from across areas of mathematics	71.0%
Opportunity to connect classroom learning to practice	
Observe models of teaching strategies you were learning in coursework	76.2%
Practice theories for teaching mathematics you were learning in coursework	77.0%
Receive feedback about how well you had implemented teaching strategies you were learning in coursework	91.2%
Collect and analyze evidence about pupil learning as a result of your teaching methods	86.4%
Develop strategies to reflect upon your professional knowledge	83.9%
Demonstrate that you could apply the teaching methods you were learning in coursework	93.5%
Complete assessment tasks that asked you to show how you were applying ideas you were learning in your courses	80.0%
Test out findings from educational research about difficulties pupils have in learning	48.0%

Note: All indicators were coded as 1 (Studied) = Occasionally/Often, 2 (Not Studied) = Never/Rarely for *OTL Instructional Practice* and *OTL Connect Classroom Learning to Practice*
Note: Percentage Studied indicates the percentage of people who responded to an item who selected "studied."
Note: Data missing for 44 future teachers for *OTL Mathematics Instruction*, 45 for *OTL Instructional Practice*, and ≥ 50 for *OTL Connect Classroom Learning to Practice*

Table 10.5 Frequency distributions for seventeen observed variables from the TEDS-M future teacher survey: Percent of future teachers who report opportunities to learn in Singapore

	% Studied
Opportunity to learn mathematics instruction	
Mathematics instruction	95.7%
Develop teaching plans	76.1%
Observation, analysis, and reflection	82.8%
Mathematics standards and curriculum	92.3%
Affective issues in mathematics	42.2%
Opportunity to learn instructional practice	
Explore how to apply mathematics to real-world problems	76.1%
Explore mathematics as the source of real-world problems	76.1%
Learn how to explore multiple solution strategies with pupils	76.1%
Learn how to show why a mathematics procedure works	66.7%
Make distinctions between procedural and conceptual knowledge when teaching mathematics concepts and operations to pupils	69.2%
Integrate mathematics ideas from across areas of mathematics	66.7%
Opportunity to connect classroom learning to practice	
Observe models of teaching strategies you were learning in coursework	56.9%
Practice theories for teaching mathematics you were learning in coursework	75.9%
Receive feedback about how well you had implemented teaching strategies you were learning in coursework	85.3%
Collect and analyze evidence about pupil learning as a result of your teaching methods	56.0%
Develop strategies to reflect upon your professional knowledge	69.0%
Demonstrate that you could apply the teaching methods you were learning in coursework	93.1%
Complete assessment tasks that asked you to show how you were applying ideas you were learning in your courses	46.6%
Test out findings from educational research about difficulties pupils have in learning	25.9%

Note: All indicators were coded as 1 (Studied) = Occasionally/Often, 2 (Not Studied) = Never/Rarely for *Opportunities to Learn Instructional Practice* and *Opportunities to Connect Classroom Learning*

Note: Percentage Studied indicates the percentage of people who responded to an item who selected “studied.”

Table 10.6 Parameter estimates for model of three latent opportunities to learn and effect of latent subgroup membership on MPCK scores for mathematics specialists in the United States

	Limited OTL (6%)	OTL mathematics pedagogy (42%)	Comprehensive OTL (52%)
OTL mathematics education - instruction			
Math instruction	.125	.906	1.000
Develop teaching plans	.125	.820	.970
Observation, analysis, and reflection	.000	.911	.956
Standards and curriculum	.124	.969	.986
Affective issues	.249	.575	.693
Opportunity to connect classroom learning to practice			
Observe models of teaching strategies you learned in coursework	.500	.606	.924
Practice theories for teaching mathematics that you learned in coursework	.000	.662	.958
Complete assessment tasks that asked you to show how you were applying ideas you learned in coursework	.429	.610	.991
Receive feedback about how well you implemented teaching strategies you learned in coursework	.714	.792	1.000
Collect and analyze evidence of pupil learning as a result of your teaching methods	.714	.773	.960
Test out findings from educational research about difficulties pupils have in learning	.000	.256	.723
Develop strategies to reflect upon your professional knowledge	.714	.646	.983
Demonstrate that you could apply the teaching methods you were learning about in your coursework	.857	.837	1.000
OTL instructional practice			
Explore how to apply mathematics to real-world problems	.000	.687	.974
Explore mathematics as the source for real-world problems	.000	.704	1.000
Learn how to explore multiple solution strategies with pupils	.625	.641	.946
Learn how to show why a mathematics procedure works	.124	.534	.958
Make distinctions between procedural and conceptual knowledge when teaching mathematics concepts and operations to pupils	.000	.504	.888
Integrate mathematical ideas from across areas of mathematics	.000	.552	.928

Table 10.7 Parameter estimates for model of three latent opportunities to learn and effect of latent subgroup membership on MPCK scores for mathematics specialists in Singapore

	Limited opportunities to connect classroom learning to practice (23.03%)	OTL mathematics pedagogy (17.62%)	Limited OTL (13.35%)	Comprehensive OTL (46%)
OTL mathematics education - instruction				
Math instruction	1.000	1.000	.806	.962
Develop teaching plans	.763	.930	.150	.866
Observation, analysis, and reflection	.886	1.000	.246	.898
Standards and curriculum	1.000	.956	.684	.943
Affective issues	.549	.371	.000	.546
Opportunity to connect classroom learning to practice				
Observe models of teaching strategies you learned in coursework	.192	.438	.539	.809
Practice theories for teaching mathematics that you learned in coursework	.508	.675	.800	.898
Complete assessment tasks that asked you to show how you were applying ideas you learned in coursework	.180	.189	.462	.721
Receive feedback about how well you implemented teaching strategies you learned in coursework	.619	.714	.932	1.000
Collect and analyze evidence of pupil learning as a result of your teaching methods	.271	.286	.627	.799
Test out findings from educational research about difficulties pupils have in learning	.000	.000	.000	.566
Develop strategies to reflect upon your professional knowledge	.325	.281	.810	1.000
Demonstrate that you could apply the teaching methods you were learning about in your coursework	.882	.809	1.000	.980

(continued)

Table 10.7 (continued)

	Limited opportunities to connect classroom learning to practice (23.03%)	OTL mathematics pedagogy (17.62%)	Limited OTL (13.35%)	Comprehensive OTL (46%)
OTL instructional practice				
Explore how to apply mathematics to real-world problems	1.000	.309	.748	.837
Explore mathematics as the source for real-world problems	.968	.261	.799	.857
Learn how to explore multiple solution strategies with pupils	1.000	.150	.633	.929
Learn how to show why a mathematics procedure works	.687	.223	.576	.859
Make distinctions between procedural and conceptual knowledge when teaching mathematics concepts and operations to pupils	.794	.199	.572	.884
Integrate mathematical ideas from across areas of mathematics	.735	.401	.437	.811

References

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on Automatic Control, 19*(6), 716–723.

Allen, J. M. (2009). Valuing practice over theory: How beginning teachers re-orient their practice in the transition from the university to the workplace. *Teaching and Teacher Education, 25*, 647–654.

Allen, J. M., & Wright, S. E. (2014). Integrating theory and practice in the preservice teacher education field. *Teachers and Teaching: Theory and Practice, 20*(2), 136–151.

Association of Mathematics Teacher Educators. (2013). *Standards for elementary mathematics specialists: A reference for teacher credentialing and degree programs*. San Diego, CA: Author.

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *Elementary School Journal, 93*(4), 373–397.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Berlin, K. S., Williams, N. A., & Parra, G. R. (2013). An introduction to latent variable mixture modeling (part I): Overview and cross-sectional latent class and latent profile analyses. *Journal of Pediatric Psychology, 39*(2), 174–187.

Blömeke, S. (2012). Content, professional preparation, and teaching methods: How diverse is teacher education across countries? *Comparative Education Review, 56*(4), 684–714.

Blömeke, S., Hsieh, F. J., Kaiser, G., & Schmidt, W. (Eds.). (2014). *International perspectives on teacher knowledge, beliefs and opportunities to learn*. Dordrecht, England: Springer.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in the entry requirements alter the teacher workforce and affect student achievement. *Education and Finance Policy, 1*(2), 176–216.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2007). Who leaves? Teacher attrition and student achievement. *Teacher Policy Research.* Retrieved from https://cepa.stanford.edu/sites/default/files/Do%20Effective%20Teachers%20Leave.pdf

Boyd, D., Lankford, S., Loeb, S., Rockoff, J., & Wyckoff, J. (2007). *The narrowing gap in New York City qualifications and its implications for student achievement in high poverty schools*. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Education Evaluation and Policy Analysis, 31*(4), 416–440.

Brosnan, P., & Erchick, D. B. (2010). Mathematics coaching and its impact on student achievement. Brief research report. In P. Brosnan, D. B. Erchick, & L. Flevares (Eds.), *Proceedings of the 32nd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 6, pp. 1362–1370). Columbus, OH: The Ohio State University.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32–42.

Burnstein, L. (1993). Studying learning, growth, and instruction cross-nationally: Lessons learned about why and why not to engage in cross-national studies [prologue]. In L. Burnstein (Ed.), *The IEA study of mathematics III: Student growth and classroom processes* (pp. xxii–xlii). London, UK: Permagon.

Calderhead, J., & Robson, M. (1991). Images of teaching: Student teachers' early conceptions of classroom practice. *Teaching and Teacher Education, 7*(1), 1–8.

Campbell, P. F., Ellington, A. J., Haver, W. E., & Inge, V. L. (Eds.). (2013). *The elementary mathematics specialist's handbook*. Reston, VA: National Council of Teachers of Mathematics.

Campbell, P. F., & Malkus, N. N. (2011). The impact of elementary mathematics coaches on student achievement. *Elementary School Journal, 111*, 430–454. https://doi.org/10.1086/657654

Carr, D. (1992). Practical enquiry, values and the problem of educational theory. *Oxford Review of Education, 18*(3), 241–251.

Carr, D. (1995). Is understanding the professional knowledge of teachers a theory-practice problem? *Journal of Philosophy of Education, 29*(3), 313–331.

Carr, D. (2003). *Making sense of education: An introduction to the philosophy and theory of education and teaching*. London, England: Routledge Falmer.

Centra, J. A., & Gaubatz, N. B. (2005). *Student perceptions of learning and instructional effectiveness in college courses* (SIR II Research). Retrieved from https://www.ets.org/Media/Products/perceptions.pdf

Clotfelter, C. T., Ladd, H. F. & Vigdor, J. L. (2007). How and why do teacher credentials matter for student achievement? *National Bureau of Economic Research Working Paper No. 12828*. Retrieved from http://www.nber.org/papers/w12828

Cochran-Smith, M., & Lytle, S. L. (1999). Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education, 24*, 249–305.

Conference Board of the Mathematical Sciences. (2012). *The mathematical education of teachers II*. Providence, RI/Washington, DC: American Mathematical Society and Mathematical Association of America.

Crespo, S. (2000). Seeing more than right and wrong answers: Prospective teachers' interpretations of students' mathematical work. *Journal of Mathematics Teacher Education, 3*(2), 155–181.

Dalgarno, N., & Colgan, L. (2007). Supporting novice elementary mathematics teachers' induction in professional communities and providing innovative forms of pedagogical content knowledge through information and communication technology. *Teaching and Teacher Education, 23*(7), 1051–1065.

Darling-Hammond, L. (1998). Teachers and teaching: Testing policy hypotheses from a national commission report. *Educational Researcher, 27*(1), 5–15.

Darling-Hammond, L. (2000). Reforming teacher preparation and licensing: Debating the evidence. *Teachers College Record, 102*(1), 28–56.

Darling-Hammond, L. (2009). Teacher quality definition debates: What is an effective teacher? In T. L. Good (Ed.), *21st century education: A reference handbook* (Vol. 2, pp. 12–22). Thousand Oaks, CA: Sage Publications.

Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? *Journal of Teacher Education, 53*(4), 286–302.

Darling-Hammond, L., Holtzman, D., Gatlin, S. J., & Heilig, J. V. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 13*, 42. Retrieved from http://epaa.asu.edu/epaa/v13n42/

Deng, Z. (2004). The role of theory in teacher preparation: An analysis of the concept of theory application. *Asia-Pacific Journal of Teacher Education, 32*(2), 143–157.

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. Social policy report. *Society for Research in Child Development, 22*(1), 1–24.

Goodnow, J. J. (2005). Contexts, diversity, pathways: Advances and next steps. In C. R. Cooper, C. Garcìa Coll, W. T. Bartko, H. M. Davis, & C. Chatman (Eds.), *Developmental pathways through middle childhood* (pp. 295–312). Mahwah, NJ: Erlbaum.

Goodnow, J. J. (2010). Culture. In M. H. Bornstein (Ed.), *Handbook of cultural developmental science* (pp. 3–19). New York, NY: Taylor and Francis Group.

Graven, M. (2004). Investigating mathematics teacher learning within an in-service community of practice: The centrality of confidence. *Educational Studies in Mathematics, 57*(2), 177–211.

Grootenboer, P., & Zevenbergen, R. (2008). Identity as a lens to understand learning mathematics: Developing a model. In M. Goos, R. Brown, & K. Makar (Eds.), *Proceedings of the 31st annual conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 243–249). Brisbane, Australia: MERGA.

Hagenaars, J. A., & McCutcheon, A. L. (Eds.). (2002). *Applied latent class analysis*. Cambridge, England: Cambridge University Press.

Hartocollis, A. (2005, July 31). Who needs education schools? What colleges teach. What teachers need to know. And why they're not the same. *New York Times Education Supplement* (pp. 1–6).

Henke, R. R., Chen, X., & Geis, S. (2000). *Progress through the teacher pipeline: 1992–93 college graduates and elementary/secondary school teaching as of 1997*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Hill, H. C., & Lubienski, S. T. (2007). Teachers' mathematical knowledge for teaching: A study of California teachers. *Educational Policy, 21*, 747–768. https://doi.org/10.1177/0895904807307061

Hodgen, J., & Askew, M. (2007). Emotion, identity and teacher learning: Becoming a primary mathematics teacher. *Oxford Review of Education, 33*(4), 469–487. Retrieved from http://www.uvm.edu/~vmi/index_files/2008%20VMI%20Evaluation.pdf

Imre, S. Y., & Akkoç, H. (2012). Investigating the development of prospective mathematics teachers' pedagogical content knowledge of generalizing number patterns through school practicum. *Journal of Mathematics Teacher Education, 15*(3), 207–226.

Ingersoll, R., Merrill, L., & May, H. (2014). *What are the effects of teacher education and preparation on beginning teacher attrition?* Research report (#RR-82). Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania.

Ingvarson, L., Beavis, A., & Kleinhenz, E. (2007). Factors affecting the impact of teacher education programmes on teacher preparedness: Implications for accreditation policy 1. *European Journal of Teacher Education, 30*(4), 351–381.

Ingvarson, L., Tatto, M. T., Peck, R., Schwille, J., Rowley, G., & Senk, S. L. (2013). *An analysis of teacher education context, structure, and quality-assurance arrangements in TEDS-M* *countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, the Netherlands: IEA.

Jerald, C. D. (2002). *All talk, no action: Putting an end to out-of-field teaching*. Retrieved from http://www2.edtrust.org/NR/rdonlyres/8DE64524-592E-4C83-A13A-6B1DF1CF8D3E/0/AllTalk.pdf

Kelly, G. J., Luke, A., & Green, J. (2008). Introduction: What counts as knowledge in educational settings: Disciplinary knowledge, assessment, and curriculum. *Review of Educational Research, 32*, vii–vix. https://doi.org/10.3102/0092732X07311063

Kim, S. Y., Wang, Y., Orozco-Lapray, D., Shen, Y., & Murtuza, M. (2013). Does "Tiger Parenting" exist? Parenting profiles of Chinese Americans and adolescent developmental outcomes. *Asian American Journal of Psychology, 4*(1), 7–18.

Korthagen, F., & Kessels, J. P. A. (1999). Linking theory and practice: Changing the pedagogy of teacher education. *Educational Researcher, 28*(4), 4–17.

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*(1), 29–63.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.

Lanza, S. T., Collins, L. M., Lemmon, D. R., & Schafer, J. L. (2007). PROC LCA: A SAS procedure for latent class analysis. *Structural Equation Modeling, 14*(4), 671–694.

Lanza, S. T., Dziak, J. J., Huang, L., Wagner, A., & Collins, L. M. (2015). *PROC LCA & PROC LTA users' guide* (Version 1.3.2). University Park: The Methodology Center, Penn State. Retrieved from http://methodology.psu.edu

Lerman, S. (2012). Mapping the effects of policy on mathematics teacher education. *Educational Studies in Mathematics*. Springer On-Line. https://doi.org/10.1007/s10649-012-9423-9.

Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago, IL: University of Chicago Press.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Earlbaum Associates.

McCutcheon, A. L. (1987). *Latent class analysis (No. 64)*. London, United Kingdom: Sage.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and policy instrument. *Educational Evaluation and Policy Analysis, 17*(3), 305–322.

Meyers, H., & Harris, D. (2008, April). *Evaluation of the VMI through 2008*. Retrieved from http://www.wou.edu/~burtonl/courses/math612W16/AMTE_EMSStandards.pdf

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling: Comment on Bauer and Curren (2003). *Psychological Methods, 8*, 369–377.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus: The comprehensive modeling program for applied researchers; User's guide; [Version 1.0]*. Los Angeles, CA: Authors.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Authors.

National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future*. New York, NY: Teachers College Press.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Mathematics Advisory Panel (NMAP). (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

Neugarten, L. (1979). Time, age, and the life cycle. *American Journal of Psychiatry, 136*(7), 887–894.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535–569.

Ramaswamy, V., DeSarbo, W. S., Reibstein, D. J., & Robinson, W. T. (1993). An empirical pooling approach for estimating marketing mix elasticities using PIMS data. *Marketing Science, 12*, 103–124.

Reys, B., & Fennell, F. (2003). Who should lead mathematics instruction at the elementary level? A case for mathematics specialists. *Teaching Children Mathematics, 9*, 277–282.

Schwab, J. J. (1978). *Science, curriculum, and liberal education*. Chicago, IL: University of Chicago Press.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.

Schwille, J., Ingvarson, L., & Holdgreve-Resendez, R. (Eds.). (2013). *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika, 52*, 333–343.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14.

Shulman, L. S. (1998). Theory, practice, and the education of professionals. *The Elementary School Journal, 98*(5), 511–526.

Super, C. M., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioral Development, 9*(4), 545–569.

Tatto, M. T. (Ed.). (2013). *The teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Tatto, M. T., Rodriguez, M. C., Reckase, M., Rowley, G., & Lu, Y. (2013). Scale development and reporting: Opportunities to learn, beliefs and mathematics knowledge for teaching. In M. T. Tatto et al. (Eds.), *The teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report* (pp. 161–174). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Thames, M. H., & Ball, D. L. (2010). What mathematical knowledge does teaching require? Knowing mathematics in and for teaching. *Teaching Children Mathematics, 17*(4), 220–225.

Warfield, J. (2001). Where mathematics content knowledge matters: Learning about and building on children's mathematical thinking. In T. Wood, B. S. Nelson, & J. Warfield (Eds.), *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 135–155). Mahwah, NJ: Lawrence Erlbaum Associates.

Weaver, S. R., & Kim, S. Y. (2008). A person-centered approach to studying the linkages among parent– child differences in cultural orientation, supportive parenting, and adolescent depressive symptoms in Chinese American families. *Journal of Youth and Adolescence, 37*, 36–49.

Chapter 11
Preparing Primary Mathematics Teachers to Learn to Work with Students from Diverse Backgrounds

Elizabeth B. Dyer

Abstract This study investigates how teacher education prepares primary teachers to teach mathematics to students from diverse backgrounds. Hierarchical linear modeling is used to investigate the relationship between opportunities to learn to teach students from diverse backgrounds during teacher preparation and teachers' mathematical knowledge for teaching using the TEDS-M international dataset. In general, primary mathematics specialist teachers with more opportunities to learn to teach students from diverse backgrounds had lower levels of mathematical knowledge for teaching. Primary generalist teachers do not consistently show the same results across all countries, with some showing higher and other showing lower levels of mathematical knowledge for teaching. These results suggest that teachers who are better prepared for the mathematical aspects of teaching tend to be less prepared for addressing the needs of diverse learners.

Introduction

Inequality in student educational outcomes by socioeconomic status and race, particularly in mathematics, is a great challenge to promoting equal opportunity worldwide (Levin, 2007; Organisation for Economic Co-operation and Development, 2014). Efforts to reduce inequities in educational systems typically focus on ways to improve and/or equalize students' opportunities to learn in classrooms (Simon, Malgorzata, & Beatriz, 2007). As research begins to highlight the crucial role teachers play in students' opportunities to learn and resultant learning (Rivkin, Hanushek, & Kain, 2005), more international efforts have worked to provide access to and

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

E. B. Dyer (✉)
STEM Program, WestEd, Redwood City, CA, USA
e-mail: edyer@wested.org

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_11

support for teachers. In keeping with this movement, the Organisation for Economic Co-operation and Development recommends that access to high-quality teachers be provided equally to students in all locations (Simon et al., 2007). However, research has documented similar differences in access to high-quality teachers and student achievement by socioeconomic status (Akiba, LeTendre, & Scribner, 2007; Kang & Hong, 2008; Levin, 2007; Little & Bartlett, 2010; Luschei et al., 2013). Therefore, equity in access to high-quality teachers may help reduce inequalities in student learning by student background. This change would be particularly important for children in primary grades, as the strength of their educational preparation can have large consequences for future educational opportunities.

With this unequal distribution of teachers seen in many countries, it is unclear what role teacher preparation might play in this unequal distribution. In many ways, teacher preparation programs are an essential step in the process of creating greater equity in access to high-quality teachers. Many programs place teachers in particular types of schools or in particular locations, meaning an improvement in certain preparation programs could result in improving the quality of instruction for populations of interest. Additionally, teacher preparation programs can prepare teachers for the unique demands of teaching students from diverse backgrounds. At the same time, teacher preparation programs could instead exacerbate or have a neutral effect on existing inequality of teachers entering their programs, essentially contributing to inequity in access to high-quality teachers.

This study explores the role of teacher preparation in developing primary teachers who are well prepared to engage in mathematics instruction for students from diverse backgrounds. In particular, the study examines the relationship between teachers' opportunities to learn to teach students from diverse backgrounds and their mathematical knowledge for teaching. By examining this relationship, this study takes a first look at whether primary teachers are equally prepared for the mathematical aspects of teaching and for teaching students from diverse backgrounds. Given the differences in student mathematics achievement based on racial, socioeconomic, and linguistic background, similar differences may exist in future teachers according to their background. If future teachers from diverse backgrounds, with lower mathematics achievement, are more likely to seek preparation to teach students from similar backgrounds, these differences could appear as a negative relationship between mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds. That is, the teachers most likely to be prepared to teach students from diverse backgrounds may be those with the weakest mathematical knowledge for teaching. However, teacher preparation programs may tend to be high quality (or low quality) in multiple aspects of preparation, leading teachers with more opportunities to learn to teach students from diverse backgrounds also having higher levels of mathematical knowledge for teaching. Additionally, opportunities to learn to teach students from diverse backgrounds could strengthen teachers' mathematical knowledge for teaching, leading to a positive relationship between these two aspects of preparation. In short, there are arguments to be made for expecting both positive and negative relationships, and not enough is known to favor one hypothesis over the other. Therefore, this study is primarily exploratory in nature, examining whether any relationship exists between mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds, and what type of relationship it is.

This study makes use of the TEDS-M international dataset of mathematics teacher preparation to explore these associations. The data on future teachers of mathematics is nested within teacher preparation programs, within countries. The analysis uses hierarchical linear modeling to account for this nested nature of the data. Additionally, the within-group (i.e., within teacher preparation program) and between-group relationships are separated to reduce bias in the estimates. The results of this study can help determine whether a positive relationship, a negative relationship, or no relationship exists between mathematical knowledge for teaching for teachers at the end of their teacher preparation programs. These findings could help inform teacher preparation programs in understanding how well they currently prepare teachers for teaching mathematics to students from diverse backgrounds. Additionally, the findings could inform how programs determine emphasis on particular aspects of preparation, or suggest particular groups of teachers who may need additional preparation in certain areas.

Theoretical Background

Mathematical Knowledge for Teaching and Effective Mathematics Instruction

To teach mathematics effectively, teachers need to understand mathematics. However, exactly what mathematics teachers need to understand in order to teach mathematics effectively is still an open question. At the primary level, this issue is particularly important as many people who choose to become primary teachers do not have strong mathematical backgrounds, and may not be most interested in teaching the subject of math as part of their jobs (Conference Board of the Mathematical Sciences [CBMS], 2001, 2012; Tatto et al., 2008).

One theory suggests that teachers have a body of knowledge they draw on when teaching mathematics, which is referred to as *mathematical knowledge for teaching* (Ball, Thames, & Phelps, 2008; Hill, Sleep, Lewis, & Ball, 2007). This knowledge includes both knowledge that is mathematical in nature and knowledge that is pedagogical in nature. Mathematical knowledge includes mathematics commonly taught in schools, including math taught by teachers (common content knowledge), and math in levels beyond the grade taught by teachers (horizon content knowledge). In addition, there is specialized content knowledge, which is defined as the knowledge specific to the profession of teaching, which is mathematical in nature. Pedagogical content knowledge encompasses knowledge of how students think about mathematics (knowledge of content and students), how to support students in learning mathematics (knowledge of content and teaching), and what curriculum materials support student learning (knowledge of content and curriculum).

Studies have shown that high levels of mathematical knowledge for teaching are related to more effective teaching and higher levels of students learning (Hill et al., 2008; Hill, Rowan, & Ball, 2005; Hill, Umland, Litke, & Kapitula, 2012). While this aspect of teaching expertise explains a small part of the variation in the quality of mathematics teaching practice and student learning, it is one of the few teacher characteristics that have shown any relationship with student outcomes. Mathematical knowledge for teaching is a particularly important metric for the quality of teachers at the end of teacher preparation programs, as more direct measures of teacher quality such as observational measures of instructional practice and student learning gains are often impossible to measure before teachers start teaching. Additionally, more direct measures of teacher quality are highly variable across lessons and raters (Hill, Charalambous, et al., 2012).

Teaching Students from Diverse Backgrounds

Across much of the world, there is strong evidence that students' background is often predictive of their level of achievement in school. For example, students' race and ethnicity, linguistic backgrounds, and socioeconomic status are associated with differences in achievement (Levin, 2007; Organisation for Economic Co-operation and Development, 2014). Typically, students from non-dominant backgrounds tend to have lower levels of achievement, which is typically interpreted as inequity in educational experiences. These findings have raised questions about whether all students have similar access to high-quality teachers, as differences in teacher quality may drive differences in student learning. However, this concern is often raised without considering that the quality of teachers may be relative to the types of students they are teaching. In other words, if students from different populations benefit from some types of teachers or some types of teaching more than others, access to high-quality teachers becomes a question of whether students and teachers are well matched, not just having access to the same types of teachers.

Research is beginning to suggest teachers from racial backgrounds similar to their students are more effective than teachers with different racial backgrounds (Dee, 2005; Villegas & Irvine, 2010). This result could be interpreted as suggesting that it is important for teachers to be able to understand the communities from which students come and how knowledge of students' backgrounds could be useful in making instructional decisions. This interpretation aligns well with research on effective teaching for students from diverse backgrounds, which suggests teachers need to leverage the resources (i.e., knowledge, skills, and experiences) students bring with them to the classroom to support student learning (Aguirre et al., 2012; Delpit, 1995; Hand, 2012; Irvine, 2003; Turner & Celedón-Pattichis, 2011). In this conception of teaching that promotes equity, teachers do not simply connect with students in a general way or act as a role model, but, instead, incorporate their understanding of students into the ways they teach specific subjects (González, Andrade, Civil, & Moll, 2001; González, Moll, & Amanti, 2005; Moll, Amanti, Neff, & González, 1992). For example, a teacher could use young children's experiences making small purchases with coins at corner stores, a common practice among students with low socioeconomic status, to help students understand place value ideas (Taylor, 2009).

This view of effective teaching is not specific to teaching students from diverse backgrounds; it can be generalized, suggesting that effective teaching for any group of students needs to leverage student resources to support student learning (Robertson, Scherr, & Hammer, 2015). Therefore, if teachers are more effective in noticing and leveraging the resources that students from dominant or majority backgrounds bring into the classroom (perhaps because this background matches their own), there will be fewer high-quality teachers available for students from non-dominant backgrounds. However, this conception of effective teaching suggests this skill can be learned by teachers, such that teacher preparation can help teachers learn how to teach students from backgrounds that are different from their own.

Primary Mathematics Teacher Preparation

Programs that train and certify teachers are found around the world. While the length, structure, and focus of these programs differ by country, there is some overlap in their goals (Schwille, Ingvarson, & Holdgreve-Resendez, 2013). Often the goals of teacher preparation are to help teachers develop the pedagogical knowledge and skills to be effective teachers in the classroom.

Teacher preparation for the primary grades is often done in a way that prepares teachers to be generalists, or to teach multiple subjects. In addition to preparing generalist teachers, some countries provide specialist teacher preparation in mathematics at the primary grades, which prepares teachers only to teach mathematics. The push to train more primary mathematics specialists has been supported by the idea that teaching mathematics at the primary level requires a great deal of mathematical knowledge that many primary teachers do currently not have (Gerretson, Bosnick, & Schofield, 2008; Reys & Fennell, 2003). The subject matter knowledge primary generalists need spans numerous subjects, including mathematics, which some scholars believe is unrealistic to expect of teachers (Reys & Fennell, 2003).

Because of the diversity of knowledge and skills teacher preparation must develop in teachers and the limited time available, programs vary considerably in their structure and the opportunities provided to teachers. Differences in the eventual teaching positions taken by participants in teacher preparation programs, including generalist and specialist positions, may dictate different programmatic structures and goals. This variation in programs leads to differences between programs in teachers' opportunities to learn. Additionally, pre-service teachers may have flexibility in the courses they take or in the aspects of teaching they try to focus on, which leads to differences in teachers' opportunities to learn within the same program. Therefore, different programmatic structure and teacher choice lead to different opportunities to learn, which result in different outcomes for teachers (i.e., levels of knowledge, skills, and beliefs).

In order to prepare mathematics teachers at the primary level, many teacher preparation programs focus on strengthening the mathematical knowledge of teachers, as teachers in the primary grades often have weak mathematical knowledge (CBMS, 2001, 2012; Swars, Hart, Smith, Smith, & Tolar, 2007; Tatto et al., 2008). Efforts also have been made to help teachers develop pedagogical knowledge and skills specific to teaching mathematics, not just general pedagogical knowledge (Ball, Lubienski, & Mewborn, 2001; Mewborn, 2001). Therefore, future teachers' mathematical knowledge for teaching is a key outcome targeted by teacher preparation programs.

Teacher preparation programs often provide opportunities for teachers to learn other aspects of pedagogy, including pedagogy for teaching students from diverse backgrounds. Teachers' opportunities to learn to teach students from diverse backgrounds could depend on numerous factors. Teacher preparation programs may place little emphasis on how to teach students from diverse backgrounds if most of their teachers plan to work with students from dominant or privileged backgrounds. Alternatively, teacher preparation programs could specialize in preparing teachers in ways consistent with social justice methods or to teach students from non-dominant or less privileged backgrounds. Teachers also may have flexibility in the courses they take, which could allow some teachers to seek out opportunities to learn to teach students from diverse backgrounds beyond basic requirements.

Opportunities to learn to teach students from diverse backgrounds have shown a number of positive effects on teacher beliefs and quality of instruction, particularly at the primary level. Teacher preparation programs can shift teachers' beliefs about students from diverse backgrounds away from deficit-oriented view, supporting more equitable teaching (Castro, 2010; Foote et al., 2013). Efforts to this end may include not only courses focused on teaching students from diverse backgrounds (Freedman & Appleman, 2009), but also cultural and community experiences with students from diverse backgrounds (Adams, Bondy, & Kuhel, 2005; Bartell et al., 2010; Garmon, 2004, 2005; Whipp, 2013). Recently, new equity-oriented courses have been implemented to help teachers develop pedagogical skills and knowledge for teaching mathematics to students from diverse backgrounds. For example, opportunities to analyze mathematics classroom videos from an equity-oriented perspective (Aguirre et al., 2012; McDuffie, Foote, Bolson, et al., 2014; McDuffie, Foote, Drake, et al., 2014) and developing mathematics lesson plans accounting for students' backgrounds (Aguirre et al., 2013) have been shown to lead teachers to consider the diversity of students in ways that inform mathematics teaching specifically. In all, this body of research shows opportunities to learn to teach students from diverse backgrounds during teacher preparation can lead teachers to be better prepared to teach students from diverse backgrounds in the classroom, including during mathematics lessons.

Relationship Between Preparation in Mathematics and for Student Diversity

This study focuses on how teachers' opportunities to learn to teach students from diverse backgrounds relate to their mathematical knowledge for teaching. Examining this relationship helps to investigate more broadly whether teachers who are better prepared to teach students from diverse backgrounds show differences in their mathematical preparation. Conversely, this line of inquiry examines whether teachers who are better prepared in the mathematical aspects of teaching show differences in their preparedness to teach students from diverse backgrounds. Therefore, this study sheds light on whether teachers are well prepared to teach mathematics in equitable ways based on their preparation to handle the mathematical demands and the demands of teaching students from diverse backgrounds.

There are several reasons why a negative relationship might be expected—that is, why teachers with more opportunities to learn to teach students from diverse backgrounds might have lower levels of mathematical knowledge for teaching. First, these teachers may differ in some way from teachers who have fewer opportunities to learn to teach students from diverse backgrounds. Even if all teachers learned the same amount of mathematical knowledge for teaching during teacher preparation, differences at the end of programs likely would be found that are due to differences existing at the beginning. Teachers with more opportunities to learn to teach students from diverse backgrounds may differ from other teachers in a number of ways, including lower levels of mathematical knowledge when starting teacher preparation, lower interest in mathematics, or less confidence in their mathematical abilities. Teachers who seek more opportunities to learn to teach students from diverse backgrounds could come from diverse backgrounds themselves, as teachers often choose to work in communities near where they grew up (Boyd, Lankford, Loeb, & Wyckoff, 2005; Reininger, 2012). Based on well-documented differences in achievement in K-12 mathematics, these teachers will also have lower levels of mathematical knowledge at the beginning of teacher preparation on average. Thus, opportunities to learn to teach students from diverse backgrounds and mathematical knowledge for teaching may be associated with one another because of their connection to coming from a diverse background, that is associated with both. Background characteristics that could indicate future teachers come from diverse or non-dominant backgrounds include socioeconomic status, racial or ethnic background, and being a language learner.

Alternatively, certain types of people may be more drawn to teaching students from diverse backgrounds—for instance, people who want to make a difference, in contrast to people who are drawn to teaching because of their passion for mathematics. Finally, teachers with more opportunities to learn to teach students from diverse backgrounds could have lower levels of mathematical knowledge for teaching because of differences in what occurs within programs. For example, those teachers could have fewer opportunities to learn mathematics because they spend more time learning how to teach students from diverse backgrounds during preparation, which could lead them to learn less mathematics.

A positive relationship could also be expected, where teachers with more opportunities to learn to teach students from diverse backgrounds have higher levels of mathematical knowledge for teaching. First, if program quality is typically high in all aspects of teacher preparation (i.e., some schools are generally more high quality than others), high levels in both aspects of teacher preparation would likely be found. Finally, if teacher preparation developed teachers' mathematical knowledge for teaching during opportunities to learn to teach students from diverse backgrounds (Turner & Drake, 2016; Turner et al., 2012), increased opportunities also would lead to increased mathematical knowledge for teaching.

This study is not attempting to determine which explanation for any relationship found is most appropriate or supported by the data, particularly not in a causal way. Instead, this study is purely descriptive and attempts to uncover whether any relationship exists. Based on the arguments made above, it also explores whether there is evidence that the relationship could be driven by differences in teacher background characteristics or opportunities to learn mathematics. Therefore, this study may help future researchers identify factors that would be fruitful for investigations using research designs appropriate for causal inference or mediation models.

This study aims to answer the following main research question, along with three sub-questions that explore potential factors that could drive any relationship seen:

How are opportunities for primary pre-service mathematics teachers to learn to teach students from diverse backgrounds associated with mathematical knowledge for teaching?

(a) Is this association found between teachers in the same program and between different teacher preparation programs?
(b) Do teacher background characteristics or opportunities to learn mathematics partly account for any association found?
(c) Does this association vary by country and between primary generalists and specialists?

Methods

Data

The data used in this study come from the TEDS-M international research study that collected data on mathematics teacher preparation (Tatto, 2013; Tatto et al., 2008). Fourteen countries preparing primary mathematics teachers are included in the study. Data were collected from pre-service primary teachers at the end of their teacher preparation programs, so these data are cross-sectional in nature.

Participants The TEDS-M project gathered data from nationally representative probability samples of future teachers of mathematics in the final year of their training in participating countries. The sampling was done in accordance with the International Education Association's quality standards (see Dumais & Meinck, 2013; Dumais et al., 2013; IEA, 2007, for sampling details). The project used stratified, multistage probability sampling. First, institutions containing teacher preparation programs were sampled from complete lists of institutions provided by each country. Each different program that prepared teachers in the institutions was included in the sample. Some institutions had multiple programs, including programs at the primary and secondary levels, while others only had one program. Additionally, programs varied in the number of teachers within them. The sampling of future teachers

Table 11.1 Pathways for generalist and specialist primary mathematics teachers

	Lower-primary generalist	Primary generalists	Primary/ lower-secondary generalists	Primary specialists
Botswana			X	
Chile			X	
Chinese Taipei		X		
Georgia	X			
Germany	X			X
Malaysia				X
Philippines		X		
Poland	X			X
Russia	X			
Singapore		X		X
Spain		X		
Switzerland	X	X		
Thailand				X
United States		X		X

aimed to include at least 30 teachers from each program. In small programs with fewer than 30 teachers all teachers were sampled. This variability in the number of teachers sampled from each program creates an unbalanced panel. Dumais and Meinck (2013) used balanced repeated replication based on the sampling design to create estimation weights. These sampling weights are used in the analyses in this study, which allows for making inferences about population estimates.

The data for primary future teachers include both primary generalists and specialists who are trained to teach mathematics. The 14 countries included in this study all had a response rate of at least 76% from future teachers. Four groups of teachers were identified: (a) generalist teachers being certified to teach Grade 4 or below (lower-primary generalists), (b) generalist teachers being certified to teach Grade 6 or below (primary generalists), (c) generalist teachers being certified to teach Grade 10 or below (primary/lower-secondary generalists), and (d) primary specialists. The three groups of primary generalist teachers and the group of primary specialist teachers make up the four distinct groups of pre-service teachers investigated in this study. Table 11.1 shows the different primary certification pathways for countries in the TEDS-M data.

As differences between types of primary teachers due to programmatic differences are likely, all analyses are completed separately for each of the four types. However, not all countries have programs in each of the groups. In fact, most countries have programs in only one or two of the groups. Therefore, generalizations about comparisons between these four groups should be interpreted with caution, as any differences could be due to differences among particular countries, rather than differences in the target certification. Numbers of participants and teacher preparation programs included in the analyses are shown in Table 11.2.

Table 11.2 Number of pre-service primary teachers and teacher preparation programs by type

	Number of programs	Number of teachers
Lower-primary generalist	171	5,640
Primary generalists	170	4,996
Primary/lower-secondary generalists	61	1,294
Primary mathematics specialists	143	1,941

Table 11.3 Reliabilities for scaled scores using congeneric measurement model

Scale	Reliability
Mathematics Content Knowledge (MCK)[a]	.83
Mathematics Pedagogical Content Knowledge (MPCK)[a]	.66
Opportunity to learn to teach students from diverse backgrounds[a]	.89
Teaching for impact and change	.80

[a]Tatto et al. (2013)

Variables of Interest The TEDS-M study collected a variety of information about primary-pre-service mathematics teachers, including information on their background, opportunities to learn, beliefs about teaching and learning, and mathematical knowledge for teaching. Two main components of these data are used in the following analysis: assessments of mathematical knowledge for teaching and teachers' reported opportunities to learn to teach students from diverse backgrounds. Reliabilities for the scaled scores used in this study can be found in Table 11.3. Details about item development, assessment frameworks, and scaling can be found in Tatto et al. (2008, 2013).

Mathematical Knowledge for Teaching The TEDS-M study developed two separate assessments one for primary teachers and one for secondary teachers. Two different domains of mathematical knowledge for teaching were assessed: mathematical content knowledge (MCK) and mathematical pedagogical content knowledge (MPCK). For the primary assessment, the first domain, mathematical content knowledge, includes mathematical knowledge that would be taught in schools, including mathematics taught in primary grades, as well as university-level mathematics. Mathematics pedagogical content knowledge includes both knowledge of content and students (considered pedagogical knowledge) and specialized content knowledge, or knowledge that is mathematical in nature that is specific to the profession of teaching. Both domains were assessed with separate assessments, and, as such, composite scores were developed for each. Tatto et al. (2013) created composite scores for each assessment using the standard Rasch model for dichotomous items and the partial credit model for polytomous items. The scores for each assessment were set at a mean of 500 and standard deviation of 100. Using a congeneric measurement model, Tatto et al. (2013) found a reliability of .83 for MCK and .66 for MPCK.

Table 11.4 Items in the opportunity to learn to teach students from diverse backgrounds scaled score

In your current teacher preparation program, how frequently did you engage in activities that gave you the opportunity to learn how to do the following?
Develop specific strategies for teaching students with behavioral and emotional problems
Develop specific strategies and curriculum for teaching pupils with learning disabilities
Develop specific strategies and curriculum for teaching gifted pupils
Develop specific strategies and curriculum for teaching pupils from diverse cultural backgrounds
Accommodate the needs of pupils with physical disabilities in your classroom
Work with children from poor or disadvantaged backgrounds

Opportunities to Learn A main component of the TEDS-M data collection was about teachers' reported opportunities to learn (OTL) different topics in order to better understand teachers' experiences during teacher preparation. This analysis focuses on the opportunities for teachers to learn to teach students from diverse backgrounds (*OTL DIVERSITY*). This opportunity was measured with six items focused on particular types of students (listed in Table 11.4). The response scale for these items had four options: *never*, *rarely*, *occasionally*, and *often*. Tatto et al. (2013) used Rasch modeling to develop the composite score from these items, which are used in this study. This composite is centered at 10, which corresponds to the middle of the rating scale (i.e., between *rarely* and *occasionally*). Using a congeneric measurement model, Tatto et al. (2013) calculated a reliability estimate of .90 for this composite.

In addition to opportunities to learn to teach students from diverse backgrounds, variables for opportunities to learn mathematics are included in the analysis. These opportunities were measured by sets of items asking whether teachers had opportunities to learn a wide variety of mathematics, including mathematics from the grades teachers will be teaching as well as more advanced mathematics. These items had two response choices: *studied* or *not studied*. In this analysis, the two domains of school-level mathematics are used: functions, probability and calculus; and numbers, measurement, and geometry. A list of items for each domain is given in Table 11.5. These OTL domains were chosen because they are most aligned with the mathematics found on the assessments of mathematical knowledge for teaching. Composites for the two domains used in this study were created by Tatto et al. (2013) by summing the number of topics marked as studied in each domain.

Teacher Background The TEDS-M data collected a wide variety of information on teachers' backgrounds. This analysis uses the aspects of teachers' backgrounds that may either help explain variation in teachers' mathematical knowledge of teaching or help account for part of the relationship of interest. These variables include age, gender, self-reported typical level of grades obtained in secondary school, whether the language of the test was typically spoken in the respondent's home, and mother's education. Although race and socioeconomic status are two teacher background characteristics related to math achievement (Barton & Coley, 2009, 2010;

Table 11.5 Items in the opportunity to learn school-level mathematics composites

Consider the following list of mathematics topics that are often taught at the secondary school level. Please indicate whether you have studied each topic as part of your current teacher preparation program.
(options: *studied, not studied*)
Numbers, measurement, and geometry composite
Numbers (e.g., whole numbers, fractions, decimals, integer, rational, and real numbers; number concepts; number theory; estimation; ratio and proportionality)
Measurement (e.g., measurement units; computations and properties of length, perimeter, area, and volume; estimation and error)
Geometry (e.g., 1-D and 2-D coordinate geometry, Euclidean geometry, transformational geometry, congruence and similarity, constructions with straightedge and compass, 3-D geometry, vector geometry)
Functions, probability and calculus composite
Functions, relations, and equations (e.g., algebra, trigonometry, analytic geometry)
Data representation, probability, and statistics
Calculus (e.g., infinite processes, change, differentiation, integration)
Validation, structuring, and abstracting (e.g., Boolean algebra, mathematical induction, logical connectives, sets, groups, fields, linear space, isomorphism, homomorphism)

Reardon, 2011; Reardon & Galindo, 2009; Reardon, Robinson-Cimpian, & Weathers, 2015) and could account for the relationship of interest, this information was not included in the TEDS-M dataset.

A variable for gender is included, as previous research has shown there is often an association between math achievement and gender (Robinson & Lubienski, 2011). Teachers were asked to report being male or female, and an indicator variable was created for selecting female.

Teachers older in age may have started teacher preparation later or taken longer than other teachers, which is likely to lead to reduced retention of knowledge. Therefore, teacher's age is included in the analysis. Teachers wrote in their age on the survey, and this age is used directly in the variable *AGE*.

Being a non-native language speaker is likely to indicate teachers are from a non-dominant background, and is associated with differences in math achievement (Reardon & Galindo, 2009). As such, an indicator for non-native language status was created. Teachers were asked about the frequency with which they spoke the language of the test at home. Teachers who reported they spoke the language of the test *sometimes* or *never* were coded as 1, while responses of *always* or *almost always* were coded as 0.

Teachers were asked about their mother's education level, which can be interpreted as a proxy for socioeconomic status. Two indicator variables were created for the highest level of education completed by teachers' mothers: completing secondary school and completing a post-secondary degree.

Teachers' previous achievement in mathematics is likely to impact their scores in mathematical knowledge for teaching at the end of their teacher preparation

Table 11.6 Items in the teaching for impact scaled score

To what extent does each of the following identify your reasons for becoming a teacher?
(options: not a reason, a minor reason, a significant reason, a major reason)
I believe that I have a talent for teaching.
I like working with young people.
I want to have an influence on the next generation.

programs. While there is no direct measure of previous math achievement in the TEDS-M data, teachers were asked about their previous grades in secondary school. An indicator variable was created based on responses to this item. Teachers who reported that their grades were always or usually near the top of their year level were coded as 1, while those reporting their grades were generally above, about, or below average for their year level were coded as 0.

Finally, a new scaled score was created for the reasons the participants chose to become a teacher. Teachers' reasons for becoming a teacher are likely to correlate with seeking opportunities to learn to teach students from diverse backgrounds. For example, teachers who become teachers to make an impact or bring about change could be more likely to want opportunities to learn to work with students from diverse backgrounds. Teachers were asked to rate the importance of several reasons for becoming a teacher on a scale with the options of *not a reason*, *a minor reason*, *a significant reason*, and *a major reason*. A composite variable for entering the teaching force in order to bring about impact and change was created for this study using Rasch modeling with a partial credit structure. This composite variable is not included in the released TEDS-M data. Because of limited items on teachers' reasons for choosing teaching, only three items make up this composite, and are given in Table 11.6. The scaled score has a reliability of .80 using a congeneric measurement model with the jMetrik software.

Analysis

To examine the relationship between mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds, two-level hierarchical linear modeling was used. Each country was run as a separate model because of the small sample of countries included in the dataset. Additionally, running separate models avoids creating one estimate for the average value of the relationships across all countries. All analysis was completed using the HLM7 software with the full maximum likelihood estimation method. All relationships were modeled linearly because adding non-linear terms did not increase model fit and would make the interpretation of the results less straightforward. In the models, teachers were nested within teacher preparation programs, by country. Each type of program (i.e., lower-primary generalist, primary generalist, primary/lower secondary generalist, and specialist) was analyzed with a separate model. Additionally, the two assessment scores of teacher knowledge, mathematical content knowledge and mathematical pedagogical content knowledge, were analyzed separately for each program type. Mathematical knowledge for teaching was modeled as the outcome, so results from the two different outcomes are presented separately (referred to as *MKT* when referring to both outcomes). Sampling weights from the TEDS-M data were used in all analyses.

The between-program and within-program relationships were disaggregated using the centered within context (i.e., a CWC(M)) approach, which was accomplished by including individual teacher variables at level 1 and introducing the mean of each of those variables for each program in the equation for the intercept at level 2. At each of these levels, the variables were group-mean centered. This approach reduces bias in the estimation of the relationship (Enders & Tofighi, 2007; Preacher, Zyphur, & Zhang, 2010; Zhang, Zyphur, & Preacher, 2009). This separation is especially important because program admission policies and teacher preferences likely contribute a considerable amount of selection bias into different teacher preparation programs. This non-random selection can bias the estimates at the teacher-level. For this reason, the estimates of the relationship at the program level should be treated with caution.

Four different specifications were tested. The first specification was an unconditional model, which provides a baseline for understanding where variance in the outcome is concentrated. Next, the basic specification included the variable for opportunity to learn to teach students from diverse backgrounds (*OTL DIVERSITY*) as a covariate. The coefficients at level 1 are modeled as random effects. In the third and fourth specifications, additional covariates were added to the model at both the teacher level and the program level to control for teacher characteristics and opportunities to learn that might influence the main relationship. In the third specification, teacher background characteristics are added. In the fourth specification, variables for opportunities to learn mathematics are added. All coefficients for additional covariates were modeled as fixed effects at level 2.

These additional covariates were included in the model as control variables rather than as mediators for several reasons. First, this method simplifies the interpretation of the coefficients in the model, which is particularly important for this type of exploratory analysis. Additionally, it would be misleading to interpret these coefficients as explaining how or why the main relationship was found. These covariates are not effects or events so much as they are variables that likely influence the main variables of interest or their relationship. The addition of the covariates allows for investigating whether the simpler models suffer from omitted variables bias, as indicated if the coefficient for the main relationship changes. For these reasons, the covariates are included as control variables rather than using a mediation model.

The fourth and most comprehensive model is shown below. The subscripts ij denote the value for the ith teacher in the jth program in the country. T_{ij} is used as a vector of variables of teacher characteristics and M_{ij} is used as a vector of variables of opportunity to learn school-level mathematics at the teacher level, and $\overline{T_j}$ and $\overline{M_j}$ are used at the program level.

$$\text{Level 1}: MKT_{ij} = \beta_{0j} + \beta_{1j} OTL\ \ DIVERSITY_{ij} + \beta_{2j} T_{ij} + \beta_{3j} M_{ij} + r_{ij}$$

$$\text{Level 2}: \beta_{0j} = \gamma_{00} + \gamma_{01}\overline{OTL\ \ DIVERSITY_j} + \gamma_{02}\overline{T_j} + \gamma_{03}\overline{M_j} + u_{0j}$$

The single equation mixed model representation is given below:

$$\begin{aligned} MKT_{ij} = {} & \gamma_{00} + \gamma_{01}\overline{OTL\ DIVERSITY_j} + \gamma_{02}\overline{T_j} + \gamma_{03}\overline{M_j} + \gamma_{10} OTL\ DIVERSITY_{ij} + \gamma_{20} T_{ij} \\ & + \gamma_{30} M_{ij} + u_{0j} + u_{1j} OTL\ DIVERSITY_{ij} + r_{ij} \end{aligned}$$

The model fit for each specification was investigated by calculating the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) with similar three-level HLM. Combining the countries into one model allows for looking at model fit for all countries together to determine general trends in model fit. Results in Tables 11.7 and 11.8 show that the full model, or fourth specification, generally has the best fit. The purpose of this analysis is to examine the relationship between mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds. Because there were small differences in model fit between the third and fourth specification, only the results from the second (basic) and fourth (full) specifications are reported.

Table 11.7 Model fit statistics for 3-level models predicting mathematics content knowledge

Model specification	Lower-primary generalists	Primary generalists	Primary/lower-secondary generalists	Primary specialists
AIC Unconditional	63,905.31	52,056.61	14,502.73	21,354.43
Basic	60,659.27	50,987.61	13,646.40	20,877.42
Teacher Background	52,155.41	47,215.99	11,854.28	18,224.79
Math Opportunity	51,870.82	47,184.64	11,857.90	18,130.38
BIC Unconditional	63,931.83	52,082.37	14,523.39	21,376.59
Basic	60,744.80	51,071.08	13,712.73	20,949.14
Teacher Background	52,328.97	47,387.42	11,988.42	18,370.24
Math Opportunity	52,070.05	47,381.46	12,011.90	18,297.28

Note: Decreases in value between models indicate better fit.

Table 11.8 Model fit statistics for 3-level models predicting mathematics pedagogical content knowledge

Model specification	Lower-primary generalists	Primary generalists	Primary/lower-secondary generalists	Primary specialists
Akaike information criterion (AIC)				
Unconditional	65,149.22	52,071.85	14,930.09	21,345.88
Basic	61,841.73	51,007.29	14,062.27	20,848.61
Teacher Background	53,073.17	47,327.69	12,261.59	18,271.79
Math Opportunity	52,865.87	47,302.02	12,266.50	18,209.55
Bayesian information criterion (BIC)				
Unconditional	65,175.74	52,097.62	14,950.74	21,368.04
Basic	61,927.26	51,090.76	14,128.60	20,920.33
Teacher Background	53,246.73	47,499.12	12,395.72	18,417.25
Math Opportunity	53,065.10	47,498.84	12,420.50	18,376.45

Note: Decreases in value between models indicate better fit.

Because of small population sizes of teacher preparation programs in some countries, some modifications were made to specifications including teacher background characteristics. In countries where the number of programs was smaller than the number of variables included at level 2, the intercept equation at level 2 does not include any of the additional variables at the program level. This intercept equation only includes the variable for the mean opportunity to learn to teach students from diverse backgrounds. Therefore, the within and between relationships are still separated. Although the between relationship is estimated in these cases, it is not reported because it is not comparable to the estimates for the other countries including additional covariates at level 2.

Results from Germany should be treated with caution as well, because the within and between relationships cannot be determined. Although several different teacher preparation programs were sampled as part of the data collection, data about the program each teacher was in was not included in the dataset. Therefore, results for Germany are only presented for the within-program columns in tables. Program assignment was used to develop sampling weights in Germany, so population estimates can still be determined.

Results

Across the different specifications, there is evidence of both positive and negative relationships between opportunity to learn to teach students from diverse backgrounds and mathematical knowledge for teaching. Negative relationships indicate that teachers with more opportunities to learn to teach students from diverse backgrounds tend to have *lower* levels of mathematical knowledge for teaching, while positive relationships indicate these teachers tend to have *higher* levels of mathematical knowledge for teaching. There was not a consistent relationship found across all countries, although more negative relationships were found than positive relationships. The results for the mathematical content knowledge outcome are discussed first, followed by mathematical pedagogical content knowledge. Finally, results regarding the variation in observed relationships are presented.

Mathematics Content Knowledge (MCK)

Evidence of a relationship between mathematical content knowledge and opportunities to learn to teach students from diverse backgrounds was found in several countries. First, looking at the basic model (Table 11.9), there is evidence of a positive relationship for lower-primary generalist teachers in Georgia within programs (3.94) and for primary generalists in Chinese Taipei between programs (77.47). Similarly, there is a positive relationship within programs for primary specialists in the United States (6.56), but there is also a negative relationship found between programs (−16.49). In the latter case, programs that on average provide more opportunities to learn to teach students from diverse backgrounds tend to have teachers with lower levels of mathematical content knowledge. A negative relationship was found in Poland for lower-primary generalists within programs (−3.05) and primary specialists between programs (−17.65), and a negative relationship as was found both within and between programs in Thailand for primary specialists (−3.57 and −12.99, respectively).

When looking at the full model (see Table 11.10), many of the results change, suggesting that the results are generally not robust for mathematics content. Teacher background and opportunity to learn mathematics may account for some of the relationships previously seen. The results from the basic model found in Germany, Chinese Taipei, the United States and the between relationship in Thailand are no longer significant. In these cases, teacher background and opportunity to learn mathematics seem to account for the relationship seen between opportunities to learn to teach students from diverse backgrounds and mathematical content knowledge. New statistically significant relationships also are found in the full model, including a negative relationship within programs for primary generalists in Malaysia (−3.83). Positive relationships between programs were found for primary generalists for both the Philippines and Spain (10.08 and 6.49).

Mathematics Pedagogical Content Knowledge (MPCK)

In the basic specification (Table 11.11), the results show several negative relationships between programs and a mix of negative and positive relationships within programs between mathematical pedagogical content knowledge and opportunities

Table 11.9 Two-level model results by country and preparation type for the estimates of the association of mathematics content knowledge with opportunity to learn to teach students from diverse backgrounds in the basic specification (β_{1j} and γ_{01})

	Lower-primary generalist		Primary generalists		Primary/lower-secondary generalists		Primary mathematics specialists	
	Within program	Between programs	Within program	Between programs	Within program	Between programs	Within program	Between programs
Botswana					0.59 (3.65)	20.96 (13.51)		
Chile					−1.34 (1.63)	−14.62⁺ (8.54)		
Chinese Taipei			−3.22 (2.18)	77.47* (27.93)				
Georgia	3.94* (1.45)	9.78 (6.07)						
Germany	2.56 (2.08)						−3.78 (9.24)	
Malaysia							−2.28⁺ (1.14)	−4.18 (7.50)
Philippines			−1.99⁺ (1.11)	3.70 (8.57)				
Poland	−3.05* (1.32)	−9.51⁺ (5.29)					1.39 (3.70)	−17.65* (8.39)
Russia	−1.19 (1.26)	17.42⁺ (9.06)						
Singapore			2.85 (2.38)	−30.62 (12.80)			−3.66 (5.09)	561.47 (2961.01)

Spain			0.18 (0.96)	4.99 (3.92)				
Switzerland	−7.20⁺ (3.50)	−10.71 (10.00)	−1.29 (1.70)	5.54 (9.97)				
Thailand							−3.57* (1.75)	−12.99* (5.63)
United States			0.33 (1.08)	2.03 (7.12)			6.56* (2.35)	−16.49* (7.21)

Note: Robust standard errors are reported in parentheses and clustered at the program level. Within and between relationships are separated in all countries except Germany where program assignment data was not available.

$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table 11.10 Two-level model results by country and preparation type for the estimates of the association of mathematics content knowledge with opportunity to learn to teach students from diverse backgrounds in the full specification (β_{1j} and γ_{01})

	Lower-primary generalist		Primary generalists		Primary/lower-secondary generalists		Primary mathematics specialists	
	Within program	Between programs	Within program	Between programs	Within program	Between programs	Within program	Between programs
Botswana					−4.56	4.32		
					(24.71)	(3.75)		
Chile					−1.83	−10.17		
					(1.87)	(6.21)		
Chinese Taipei			−2.08	−110.47				
			(1.48)	(410.69)				
Georgia	4.05⁺	40.54						
	(2.14)	(19.69)						
Germany	−0.07						−4.13	
	(2.34)						(9.65)	
Malaysia							−3.83**	−4.60
							(1.28)	(12.24)
Philippines			−0.72	10.08*				
			(−0.72)	(4.51)				
Poland	−3.30*	−8.65⁺					−0.71	−17.64*
	(1.59)	(4.86)					(4.04)	(6.65)
Russia	−1.54	−14.73⁺						
	(1.26)	(8.24)						
Singapore			3.33				−0.83	
			(2.87)				(5.13)	

Spain			1.11	6.49**				
			(1.20)	(2.35)				
Switzerland	−6.98		−1.78	−9.14				
	(5.61)		(2.12)	(10.77)				
Thailand							−3.34*	−11.22⁺
							(1.61)	(6.16)
United States			1.12	−2.92			2.47	−8.52
			(1.24)	(4.26)			(4.24)	(9.10)

Note: Robust standard errors are reported in parentheses and clustered at the program level. All estimates are from regressions that control for variables at the teacher level and mean values of teacher variables at the program level. Countries without reported between relationship estimates did not include control variables at the program level due to small program sample size, but still separated the within and between relationships in the model. Controls include: age, gender, self-reported typical level of grades obtained in secondary school, whether the language of the test was typically spoken in their home, mother's education, opportunity to learn functions, probability and calculus, and opportunity to learn numbers, measurement, and geometry. Within and between relationships are separated in all countries except Germany where program assignment data was not available.

$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table 11.11 Two-level model results by country and preparation type for the estimates of the association of mathematics pedagogical content knowledge with opportunity to learn to teach students from diverse backgrounds in the basic specification (β_{1j} and γ_{01})

	Lower-primary generalist		Primary generalists		Primary/lower-secondary generalists		Primary mathematics specialists	
	Within program	Between programs	Within program	Between programs	Within program	Between programs	Within program	Between programs
Botswana					8.83*	−18.02		
					(2.27)	(13.64)		
Chile					2.62	−8.62		
					(1.77)	(7.44)		
Chinese Taipei			−1.83	19.28				
			(1.03)	(13.34)				
Georgia	7.83*	2.55						
	(2.58)	(8.58)						
Germany	4.82*						−3.77	
	(2.25)						(10.28)	
Malaysia							−1.22	9.51
							(1.37)	(7.25)
Philippines			−3.64	−7.00				
			(2.24)	(15.06)				
Poland	−3.82*	−16.74*					0.84	−14.03⁺
	(1.57)	(7.40)					(2.33)	(7.49)
Russia	−2.06	6.79						
	(1.82)	(6.82)						
Singapore			4.39⁺	−35.20			−1.00	−1084.21
			(1.44)	(13.56)			(4.40)	(2557.91)

Spain			1.42	−3.85				
			(1.22)	(5.32)				
Switzerland	−2.96	−31.10	1.02	1.82				
	(5.72)	(17.93)	(2.44)	(8.21)				
Thailand							−3.33*	−13.90***
							(1.57)	(3.56)
United States			−1.34	4.97			12.33	−21.24*
			(1.70)	(5.60)			(9.60)	(8.36)

Note: Robust standard errors are reported in parentheses and clustered at the program level. Within and between relationships are separated in all countries except Germany where program assignment data was not available.

$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

to learn to teach students from diverse backgrounds. Lower-primary generalists in Poland, as well as primary specialists in Thailand and the United States show a negative relationship between programs (−16.74, −13.90 and−21.24, respectively). In fact, for these groups of teachers in Poland and Thailand there is also evidence of a negative relationship within teachers from the same program (−3.82 and−3.33, respectively). In addition to these negative relationships within programs, positive relationships within programs were found for lower-primary generalists in Georgia and Malaysia, along with primary/lower-secondary generalists in Botswana (7.83, 4.82, and 8.8, respectively).

In the full specification (Table 11.12), many of the relationships remain statistically significant, and no new relationships are found. In particular, the positive relationship within programs for lower primary generalists in Poland, the negative relationship within programs for primary specialists in Thailand, and the negative relationship between programs in the United States for primary specialists are no longer statistically significant. In these cases, teacher background and opportunity to learn mathematics may account for the relationships found in the basic specification. The remaining significant relationships found in the basic specification remain statistically significant and of the same sign.

Variation in the Within-Program Relationship for Different Programs

There is evidence of variation in the within-program relationship between opportunities to learn to teach students from diverse backgrounds and mathematical knowledge for teaching for different programs in some countries. Looking at the Level 2 OTL Diversity columns in Tables 11.13 and 11.14, there are several statistically significant variances found. For the relationship with mathematical content knowledge in Table 11.13, the relationship was found to vary for different programs for primary/lower-secondary generalists in Chile, lower-primary generalists in Russia, and both lower-primary generalists and primary specialists in Poland. For the relationship with mathematical pedagogical content knowledge in Table 11.14, there is less evidence of variation of the relationship among programs. Statistically significant variation was found for lower-primary generalists in Russia and primary specialists in the United States. These variances indicate that while the models estimated the average within-program relationship for these types of teachers in these countries, the within-program relationship may vary widely by program. In fact, some programs in these countries may see a positive relationship while other programs may see a negative relationship within teachers from the same program.

Table 11.12 Two-level model results by country and preparation type for the estimates of the association of mathematics pedagogical content knowledge with opportunity to learn to teach students from diverse backgrounds in the full specification (β_{1j} and γ_{01})

	Lower-primary generalist		Primary generalists		Primary/lower-secondary generalists		Primary mathematics specialists	
	Within program	Between programs	Within program	Between programs	Within program	Between programs	Within program	Between programs
Botswana					18.82*	−35.69		
					(4.94)	(32.86)		
Chile					3.32	1.91		
					(2.21)	(7.12)		
Chinese Taipei			−2.76+	−54.55				
			(1.25)	(344.65)				
Georgia	9.43**	−6.15						
	(2.85)	(26.21)						
Germany	3.17						−6.62	
	(2.56)						(6.73)	
Malaysia							−2.58	10.85
							(1.67)	(10.51)
Philippines			−1.74	4.11				
			(3.69)	(7.94)				
Poland	−4.47**	−14.77*					2.69	−22.05*
	(1.63)	(6.82)					(2.83)	(9.32)
Russia	−2.34	−14.99+						
	(1.56)	(8.47)						
Singapore			4.09				0.20	
			(2.69)				(4.32)	

(continued)

Table 11.12 (continued)

	Lower-primary generalist		Primary generalists		Primary/lower-secondary generalists		Primary mathematics specialists	
	Within program	Between programs	Within program	Between programs	Within program	Between programs	Within program	Between programs
Spain			2.29	−3.33				
			(1.35)	(2.94)				
Switzerland	−4.24		1.28	−12.19				
	(6.90)		(2.51)	(9.38)				
Thailand							-2.97^{+}	−9.46*
							(1.54)	(4.31)
United States			−0.75	0.38			10.51	−13.96
			(1.48)	(3.62)			(9.43)	(10.97)

Note: Robust standard errors are reported in parentheses and clustered at the program level. All estimates are from regressions that control for variables at the teacher level and mean values of teacher variables at the program level. Countries without reported between relationship estimates did not include control variables at the program level due to small program sample size, but still separated the within and between relationships in the model. Controls include: age, gender, self-reported typical level of grades obtained in secondary school, whether the language of the test was typically spoken in their home, mother's education, opportunity to learn functions, probability and calculus, and opportunity to learn numbers, measurement, and geometry. Within and between relationships are separated in all countries except Germany where program assignment data was not available.

$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table 11.13 Two-level model results by country and preparation type for the variance components and standard deviation of random effects in the full specification predicting mathematics content knowledge

	Lower-primary generalist			Primary generalists			Primary/lower-secondary generalists			Primary mathematics specialists		
	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2*OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*
Botswana							1352.44***	17.05				
							(36.78)	(4.13)				
Chile							3091.56***	184.07***	34.52*			
							(55.60)	(13.57)	(5.88)			
Chinese Taipei				6138.76***	0.15	0.09						
				(78.35)	(0.39)	(0.31)						
Georgia	4348.76***	0.24	0.08									
	(65.95)	(0.49)	(0.29)									
Germany												
Malaysia										2602.90***	161.24***	5.26
										(51.02)	(12.70)	(2.29)
Philippines				2267.31***	0.72	0.14						
				(47.62)	(0.85)	(0.37)						
Poland	3601.00***	260.02***	19.08**							5103.77***	1279.31***	221.25**
	(60.01)	(16.13)	(4.37)							(71.44)	(35.77)	(14.87)
Russia	4149.20***	1573.32***	12.35**									
	(64.41)	(39.67)	(3.51)									
Singapore				4548.23***	101.83*					5228.08***	0.17	
				(67.44)	(10.09)					(72.31)	(0.42)	

(continued)

Table 11.13 (continued)

	Lower-primary generalist			Primary generalists			Primary/lower-secondary generalists			Primary mathematics specialists		
	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2*OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*
Spain				2604.91***	3.30	3.51						
				(51.04)	(1.82)	(1.87)						
Switzerland	2876.65***	567.94***		3424.40***	91.22***	11.74						
	(53.63)	(23.83)		(58.52)	(9.55)	(3.43)						
Thailand										3721.56***	788.58***	3.60
										(61.00)	(28.08)	(1.90)
United States				3692.04***	260.69***	8.49				2601.31***	0.29	0.31
				(60.76)	(16.15)	(2.91)				(51.00)	(0.54)	(0.55)

Note: Standard deviations are reported in parentheses. All estimates are from regressions that control for variables at the teacher level and mean values of teacher variables at the program level. Countries with eight or less programs sampled did not include program level control variables. Controls include: age, gender, self-reported typical level of grades obtained in secondary school, whether the language of the test was typically spoken in their home, mother's education, opportunity to learn functions, probability and calculus, and opportunity to learn numbers, measurement, and geometry. Within and between relationships are separated in all countries except Germany where program assignment data was not available.

$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table 11.14 Two-level model results by country and preparation type for the variance components and standard deviation of random effects in the full specification predicting mathematics pedagogical content knowledge

	Lower-primary generalist			Primary generalists			Primary/lower-secondary generalists			Primary mathematics specialists		
	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*
Botswana							2617.51***	1.75				
							(51.16)	(1.32)				
Chile							7287.11***	1.11	0.51			
							(85.36)	(1.06)	(0.71)			
Chinese Taipei				4322.90***	0.11	0.12						
				(65.75)	(0.33)	(0.35)						
Georgia	7705.13***	0.44	0.40									
	(87.78)	(0.67)	(0.63)									
Germany												
Malaysia										4345.53***	85.54**	4.91
										(65.92)	(9.25)	(2.22)
Philippines				4068.47***	144.30***	12.84						
				(63.78)	(12.01)	(3.58)						
Poland	5999.56***	649.72***	23.19							4386.53***	1195.54***	9.84
	(77.46)	(25.49)	(4.82)							(66.23)	(34.58)	(3.14)
Russia	4135.12***	1262.79***	9.55**									
	(64.30)	(35.54)	(3.09)									
Singapore				4658.95***	93.81*					3693.92***	0.12	
				(68.26)	(9.69)					(60.78)	(0.35)	
Spain				3311.99***	26.06*	10.30						
				(57.55)	(5.11)	(3.21)						

(continued)

Table 11.14 (continued)

	Lower-primary generalist			Primary generalists			Primary/lower-secondary generalists			Primary mathematics specialists		
	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*
Switzerland	3854.37***	745.22***		3526.50***	42.78***	30.74^{+}						
	(62.08)	(27.30)		(59.38)	(6.54)	(5.54)						
Thailand										3696.25***	308.78***	15.12
										(60.80)	(17.57)	(3.89)
United States				3959.48***	44.48^{+}	26.26				3724.73***	21.25	390.47***
				(62.92)	(6.67)	(5.12)				(61.03)	(4.61)	(19.76)

Note: Standard deviations are reported in parentheses. All estimates are from regressions that control for variables at the teacher level and mean values of teacher variables at the program level. Countries with eight or less programs sampled did not include program level control variables. Controls include: age, gender, self-reported typical level of grades obtained in secondary school, whether the language of the test was typically spoken in their home, mother's education, opportunity to learn functions, probability and calculus, and opportunity to learn numbers, measurement, and geometry. Within and between relationships are separated in all countries except Germany where program assignment data was not available.

$^{+}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

Discussion

The results presented show that primary teacher preparation in some countries does not develop teachers who are as well prepared for the mathematical aspects of teaching as they are to work with students from diverse backgrounds, while in other countries they are more equally prepared for both aspects of teaching. While it is difficult to generalize across the different program types, the highest proportion of negative relationships found was for primary mathematics specialists, while the other program types tended to show more mixed results. Also, in many cases, different relationships were found by type of teacher preparation, such as generalists versus specialists, suggesting that there may be important differences between program types in this relationship.

The concentration of negative relationships in programs for primary mathematics specialists suggests that programs that are more mathematically demanding may be less so in preparing future teachers for student diversity and for the mathematical aspects of teaching. Additionally, teachers may need to choose between programs that focus on teaching students from diverse backgrounds or concentrate on mathematics, which could create large differences between program type. Further research to better understand these types of programs and explore the reasons for seeing these differences in program type is warranted.

However, these tentative conclusions about program type should be taken with caution as there was wide variation among countries within each program type. For example, it may be wiser not to focus on the program type and instead focus on the individual countries making up the different types. The results showed wide differences between countries, even in groups that did not show a consistent trend in the relationship. This variation suggests any relationship seen may be as much an artifact of the cultural context of particular countries, such as their conceptions of teaching, how teacher preparation is implemented, or even the role diversity in student backgrounds plays in the educational system. With these varying results between countries, further work including case-study comparisons between countries is warranted.

There are several reasons why this analysis would find negative relationships, including differences in the types of teachers who attend programs and programs' emphases on topics. Having different types of teachers who are more likely to attend programs with different opportunities to learn to teach students from diverse background could bring about the relationships seen. Although the analysis controlled for teacher background characteristics, there may be other characteristics that are associated with mathematical knowledge for teaching and programs with different opportunities to learn to teach students from diverse backgrounds. Future work would benefit from including additional teacher background characteristics, especially teachers' racial or ethnic background. Additionally, pre-test measures of teachers' mathematical knowledge for teaching could help determine whether the differences exist in the teachers who chose to attend particular programs and/or the differences increased or decreased by the end of program. While this type of study design is difficult to implement in pre-service programs because of the widely differing lengths of programs, and the different routes into teaching, even pre-post measures of outcomes during the final year of teacher preparation would be a substantial contribution to the field.

Beyond the context of teacher education, these results also speak to the equity in access to mathematics that future teachers at the primary level encounter. These results indicate that inequalities are apparent at the end of teacher preparation programs. Additionally, it is reasonable to assume that teachers who choose programs with greater opportunities to learn to teach students from diverse backgrounds would be more likely to actually teach these populations of students. It is puzzling however that these teachers seem to be less qualified in the mathematical aspects of teaching, a situation that may result in less effective teachers. However, making this claim would be much better supported by data looking at these teachers' eventual effectiveness and instruction in the classroom. Additionally, it is not clear whether mathematical preparation or preparation for student diversity is most important for teacher effectiveness. As teacher preparation programs often have limited time and many goals to meet, particularly in programs that focus on more advanced mathematics, these programs may be focusing on the aspects of teaching that will have the greatest impact on student learning when working with students from diverse backgrounds. Future research in this area would be helpful for informing teacher preparation about where efforts are best spent.

Conclusion

This study examined whether primary teachers are equally prepared for student diversity and the mathematical aspects of teaching from teacher preparation. The results found evidence that primary mathematics specialist programs tend not to prepare teachers equally well in these two areas. Therefore, the teachers with better mathematical knowledge for teaching tend to have less preparation in learning to teach students from diverse backgrounds and vice versa. As these teachers enter the teaching force, it is likely that these differences will lead to inequity in access to high-quality mathematics teachers. These results suggest that mathematically strong programs may wish to place more emphasis on preparing for student diversity, while programs with more emphasis on student diversity may benefit from additional mathematical preparation.

Although the cross-sectional nature of the TEDS-M data limits our ability to distinguish among the reasons behind the differences seen, there are still clear implications for teacher preparation programs. First, programs would benefit from examining whether their teachers have sufficient opportunities to learn both mathematical knowledge for teaching and how to teach students from diverse backgrounds. Additionally, programs may wish to consider how well prepared their teachers are mathematically when entering the program and ways of increasing the mathematical background of teachers as necessary. Programs with strong emphasis on teaching students from diverse backgrounds could consider integrating mathematics into these courses. In other words, programs could provide opportunities for teachers to learn simultaneously to teach mathematics to students from diverse backgrounds, making some of these course elements specific to the subject of mathematics. Research on these integrated types of courses at the primary level show promise for developing mathematics teachers who are prepared to teach in ways that bring about more equitable student outcomes (Aguirre et al., 2013; Foote et al., 2013; McDuffie, Foote, Bolson, et al., 2014).

Acknowledgments This material is based upon work supported in part, by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-0824162, and the NAEd/Spencer Dissertation Fellowship Program. Its contents are solely the responsibility of the author and do not necessarily represent the official views of the previous organizations.

References

Adams, A., Bondy, E., & Kuhel, K. (2005). Preservice teacher learning in an unfamiliar setting. *Teacher Education Quarterly, 32*(2), 41–62.

Aguirre, J. M., Turner, E. E., Bartell, T. G., Drake, C., Foote, M. Q., & McDuffie, A. R. (2012). Analyzing effective mathematics lessons for English learners: A multiple mathematical lens approach. In N. Ramirez & S. Celedón-Pattichis (Eds.), *Beyond good teaching: Advancing mathematics education for ELLs* (pp. 207–222). Reston, VA: National Council of Teachers of Mathematics.

Aguirre, J. M., Turner, E. E., Bartell, T. G., Kalinec-Craig, C., Foote, M. Q., McDuffie, A. R., & Drake, C. (2013). Making connections in practice how prospective elementary teachers connect to children's mathematical thinking and community funds of knowledge in mathematics instruction. *Journal of Teacher Education, 64*(2), 178–192. https://doi.org/10.1177/0022487112466900

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher, 36*(7), 369–387. https://doi.org/10.3102/0013189X07308739

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. *Handbook of Research on Teaching, 4*, 433–456.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Bartell, T. G., Foote, M. Q., Aguirre, J. M., Roth McDuffie, A., Drake, C., & Turner, E. E. (2010). Preparing preK-8 teachers to connect children's mathematical thinking and community based funds of knowledge. In P. Brosnan, D. B. Erchick, & L. Flevares (Eds.), *Proceedings of the 32nd annual meeting of the North American chapter of the international group for the psychology of mathematics education* (pp. 1183–1191).

Barton, P. E., & Coley, R. J. (2009). *Parsing the achievement gap II. Policy information report*. Princeton, NJ: Educational Testing Service Retrieved from http://eric.ed.gov/?id=ED505163

Barton, P. E., & Coley, R. J. (2010). *The black-white achievement gap: When progress stopped*. Policy information report. Educational Testing Service. Retrieved from http://eric.ed.gov/?id=ED511548.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management, 24*(1), 113–132. https://doi.org/10.1002/pam.20072

Castro, A. J. (2010). Themes in the research on preservice teachers' views of cultural diversity Implications for researching millennial preservice teachers. *Educational Researcher, 39*(3), 198–210. https://doi.org/10.3102/0013189X10363819

Conference Board of the Mathematical Sciences. (2001). *The mathematical education of teachers* (Vol. 11). Providence, RI/Washington, DC: American Mathematical Society/Mathematical Association of America Retrieved from http://www.cbmsweb.org/MET_Document/

Conference Board of the Mathematical Sciences (Ed.). (2012). *The mathematical education of teachers II* (Vol. 17). Providence, RI/Washington, DC: American Mathematical Society/ Mathematical Association of America Retrieved from http://cbmsweb.org/MET2/

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *The American Economic Review, 95*(2), 158–165.

Delpit, L. D. (1995). *Other people's children: cultural conflict in the classroom*. New York, NY: The New Press.

Dumais, J., & Meinck, S. (2013). *Chapter 10: Estimation of weights, participation rates, and sampling error* (The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries. Technical Report). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Dumais, J., Meinck, S., Tatto, M. T., Schwille, J., & Ingvarson, L. (2013). *Chapter 6: Sampling design* (The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries. Technical Report). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Foote, M. Q., Roth McDuffie, A., Turner, E. E., Aguirre, J. M., Bartell, T. G., & Drake, C. (2013). Orientations of prospective teachers toward students' family and community. *Teaching and Teacher Education, 35*, 126–136. https://doi.org/10.1016/j.tate.2013.06.003

Freedman, S. W., & Appleman, D. (2009). "In It for the long haul"—How teacher education can contribute to teacher retention in high-poverty, urban schools. *Journal of Teacher Education, 60*(3), 323–337. https://doi.org/10.1177/0022487109336181

Garmon, M. A. (2004). Changing preservice teachers' attitudes/beliefs about diversity: What are the critical factors. *Journal of Teacher Education, 55*(3), 201–213. https://doi.org/10.1177/0022487104263080

Garmon, M. A. (2005). Six key factors for changing preservice teachers' attitudes/beliefs about diversity. *Educational Studies, 38*(3), 275–286. https://doi.org/10.1207/s15326993es3803_7

Gerretson, H., Bosnick, J., & Schofield, K. (2008). A case for content specialists as the elementary classroom teacher. *The Teacher Educator, 43*(4), 302–314. https://doi.org/10.1080/08878730802249866

González, N., Andrade, R., Civil, M., & Moll, L. (2001). Bridging funds of distributed knowledge: Creating zones of practices in mathematics. *Journal of Education for Students Placed at Risk, 6*(1–2), 115–132. https://doi.org/10.1207/S15327671ESPR0601-2_7

González, N., Moll, L. C., & Amanti, C. (Eds.). (2005). *Funds of knowledge: Theorizing practices in households and classrooms*. New York, NY: Routledge.

Hand, V. (2012). Seeing culture and power in mathematical learning: Toward a model of equitable instruction. *Educational Studies in Mathematics, 80*(1–2), 233–247. https://doi.org/10.1007/s10649-012-9387-9

Hill, H. C., Blunk, M., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430–511.

Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M. A., Beisiegel, M., … Lynch, K. (2012). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment, 17*(2–3), 88–106. https://doi.org/10.1080/10627197.2012.715019

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371.

Hill, H. C., Sleep, L., Lewis, J. M., & Ball, D. L. (2007). Assessing teachers' mathematical knowledge: What knowledge matters and what evidence counts. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, 2nd ed., pp. 111–155). Charlotte, NC: Information Age Publishing Inc.

Hill, H. C., Umland, K., Litke, E., & Kapitula, L. R. (2012). Teacher quality and quality teaching: Examining the relationship of a teacher assessment to practice. *American Journal of Education, 118*(4), 489–519. https://doi.org/10.1086/666380

International Association for the Evaluation of Educational Achievement. (2007). *TEDS-M sample preparation manual*. Hamburg, Germany: Author.

Irvine, J. J. (2003). *Educating teachers for diversity: Seeing with a cultural eye*. New York, NY: Teachers College Press.

Kang, N.-H., & Hong, M. (2008). Achieving excellence in teacher workforce and equity in learning opportunities in South Korea. *Educational Researcher, 37*(4), 200–207. https://doi.org/10.3102/0013189X08319571

Levin, B. (2007). Schools, poverty, and the achievement gap. *Phi Delta Kappan, 89*(1), 75–76. https://doi.org/10.1177/003172170708900115

Little, J. W., & Bartlett, L. (2010). The teacher workforce and problems of educational equity. *Review of Research in Education, 34*(1), 285–328. https://doi.org/10.3102/0091732X09356099

Luschei, T. F., Chudgar, A., Rew, W., Klugman, J., Francois, C., & Mehta, J. (2013). Exploring differences in the distribution of teacher qualifications across Mexico and South Korea: Evidence from the teaching and learning international survey. *Teachers College Record, 115*(5).

McDuffie, A. R., Foote, M. Q., Bolson, C., Turner, E. E., Aguirre, J. M., Bartell, T. G., … Land, T. (2014). Using video analysis to support prospective K-8 teachers' noticing of students' multiple mathematical knowledge bases. *Journal of Mathematics Teacher Education, 17*(3), 245–270. https://doi.org/10.1007/s10857-013-9257-0

McDuffie, A. R., Foote, M. Q., Drake, C., Turner, E., Aguirre, J., Bartell, T. G., & Bolson, C. (2014). Use of video analysis to support prospective K-8 teachers' noticing of equitable practices. *Mathematics Teacher Educator, 2*(2), 108–140.

Mewborn, D. (2001). Teachers content knowledge, teacher education, and their effects on the preparation of elementary teachers in the United States. *Mathematics Teacher Education and Development, 3*, 28–36.

Moll, L. C., Amanti, C., Neff, D., & González, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice, 31*(2), 132–141.

Organisation for Economic Co-operation and Development. (2014). *Education at a glance 2014*. Paris, France: OECD Publishing Retrieved from http://www.oecd-ilibrary.org/education/education-at-a-glance-2014_eag-2014-en

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233. https://doi.org/10.1037/a0020141

Reardon, S. F. (2011). The widening academic–achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity: Rising inequality, schools, and children's life chances* (pp. 91–115). New York, NY: Russell Sage Foundation.

Reardon, S. F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal, 46*(3), 853–891. https://doi.org/10.3102/0002831209333184

Reardon, S. F., Robinson-Cimpian, J. P., & Weathers, E. S. (2015). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In H. F. Ladd & M. G. Goertz (Eds.), *Handbook of research in education finance and policy* (2nd ed., pp. 491–509). New York, NY: Routledge.

Reininger, M. (2012). Hometown disadvantage? It depends on where you're from: Teachers' location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis, 34*(2), 127–145. https://doi.org/10.3102/0162373711420864

Reys, B. J., & Fennell, F. (2003). Who should lead mathematics instruction at the elementary school level? A case for mathematics specialists. *Teaching Children Mathematics, 9*(5), 277–282.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Robertson, A. D., Scherr, R., & Hammer, D. (Eds.). (2015). *Responsive teaching in science and mathematics*. New York, NY: Routledge.

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in Mathematics and reading during elementary and middle school examining direct cognitive assessments and reacher ratings. *American Educational Research Journal, 48*(2), 268–302.

Schwille, J., Ingvarson, L., & Holdgreve-Resendez, R. (Eds.). (2013). *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Simon, F., Malgorzata, K., & Beatriz, P. (2007). *Education and training policy no more failures ten steps to equity in education: Ten steps to equity in education*. Paris, France: OECD Publishing.

Swars, S., Hart, L. C., Smith, S. Z., Smith, M. E., & Tolar, T. (2007). A longitudinal study of elementary pre-service teachers' mathematics beliefs and content knowledge. *School Science and Mathematics, 107*(8), 325–335. https://doi.org/10.1111/j.1949-8594.2007.tb17797.x

Tatto, M. T. (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Tatto, M. T., Rodríguez, M. C., Reckase, M., Rowley, G., & Lu, Y. (2013). *Chapter 11: Scale development and reporting: Opportunities to learn, beliefs, and mathematics knowledge for teaching* (The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries. Technical Report). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics: Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.

Taylor, E. (2009). The purchasing practice of low-income students: The relationship to mathematical development. *Journal of the Learning Sciences, 18*(3), 370–415. https://doi.org/10.1080/10508400903013462

Turner, E. E., & Celedón-Pattichis, S. (2011). Mathematical problem solving among Latina/o kindergartners: An analysis of opportunities to learn. *Journal of Latinos and Education, 10*(2), 146–169. https://doi.org/10.1080/15348431.2011.556524

Turner, E. E., & Drake, C. (2016). A review of research on prospective teachers' learning about children's mathematical thinking and cultural funds of knowledge. *Journal of Teacher Education, 67*(1), 32–46. https://doi.org/10.1177/0022487115597476

Turner, E. E., Drake, C., McDuffie, A. R., Aguirre, J., Bartell, T. G., & Foote, M. Q. (2012). Promoting equity in mathematics teacher preparation: a framework for advancing teacher learning of children's multiple mathematics knowledge bases. *Journal of Mathematics Teacher Education, 15*(1), 67–82. https://doi.org/10.1007/s10857-011-9196-6

Villegas, A. M., & Irvine, J. J. (2010). Diversifying the teaching force: An examination of major arguments. *The Urban Review, 42*(3), 175–192. https://doi.org/10.1007/s11256-010-0150-1

Whipp, J. L. (2013). Developing socially just teachers: The interaction of experiences before, during, and after teacher preparation in beginning urban teachers. *Journal of Teacher Education, 64*(5), 454–467. https://doi.org/10.1177/0022487113494845

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models problems and solutions. *Organizational Research Methods, 12*(4), 695–719. https://doi.org/10.1177/1094428108327450

Chapter 12
Differences in Beliefs and Knowledge for Teaching Mathematics: An International Study of Future Teachers

Traci Shizu Kutaka, Wendy M. Smith, and Anthony D. Albano

Abstract Mathematical content knowledge and beliefs about teaching and learning interact in complex ways that, in turn, influence the quality of mathematics instruction and, therefore, are important teacher-preparation program outcomes. We used the Teacher Education Development Study in Mathematics (TEDS-M) database to study the relationship between future primary teachers' beliefs about the nature of mathematics, beliefs about learning mathematics, and mathematical content knowledge (MCK) within and between teacher preparation programs across 15 countries. A series of multilevel models were fit to four program groups (lower primary, primary, primary/secondary, and primary mathematics specialists) with future teachers nested within institutions. Confirming our hypotheses, procedural beliefs were associated with lower MCK scores, and inquiry beliefs were associated with higher MCK scores. We hypothesized that endorsing fixed-ability beliefs would be associated with lower MCK scores, but this was only confirmed in some countries and program-types. The chapter concludes with possible explanations for programmatic differences between and within countries, grounded in a discussion of program features such as entry requirements, program-types, and credentials, as well as curriculum organization and content.

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

T. S. Kutaka (✉)
University of Denver, Denver, CO, USA
e-mail: traci.kutaka@du.edu

W. M. Smith · A. D. Albano
University of Nebraska, Lincoln, NE, USA
e-mail: wsmith5@unl.edu; albano@unl.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_12

Introduction

Shulman (1986) identified subject-matter knowledge as one of three bodies of sophisticated, professional knowledge necessary to carry out the tasks of teaching. Mathematical content knowledge (MCK) is part of an instructional triangle, where the relationships between the teacher, the content, and the student interact in complex ways (Hawkins, 1974; Lampert, 2001). Beliefs about the nature of mathematics, teaching, and learning are "predispositions to action" (Rokeach, 1968, p. 113), channeling how mathematical content knowledge is used in practice (Ambrose, 2004; Weinstein, 1989). Indeed, deep mathematical content knowledge combined with productive beliefs about teaching and learning are important outcomes for teacher preparation programs. This study uses the IEA Teacher Education and Development Study in Mathematics (TEDS-M) database to explore the extent to which the belief patterns about teaching and learning endorsed by future primary teachers are related to different levels of mathematical content knowledge within and between preparation program-types across 15 developing and developed countries.

The first section of this chapter describes the purpose of the TEDS-M study and why studying cross-national teacher-preparation programs may be instructive. Next is a summary of the literature on the role of mathematical content knowledge in mathematics teaching, what is known about productive beliefs in mathematics teaching and learning, and, as theoretical context for this study, the significance of the relationship between knowledge and beliefs. The next section outlines the details of the design of the TEDS-M study, including our analytic strategy. The results are presented followed by a discussion of the findings. The chapter concludes with recommendations for teacher preparation programs and policymakers.

The Teacher Education and Development Study in Mathematics (TEDS-M)

The purpose of the TEDS-M study was to collect cross-national data on the knowledge that future primary and secondary teachers have acquired upon completing their mathematics teacher education programs. The TEDS-M database presents an opportunity to study the variation in the nature and impact of primary teacher education programs within and between 15 countries/municipalities: Botswana, Chile, Chinese Taipei, Georgia, Germany, Malaysia, Norway, Philippines, Poland, Russia, Spain, Singapore, Switzerland, Thailand, and the United States. Data were collected in 2008 from representative samples of pre-service teacher education programs and from their teacher educators with the goal of informing policy and practice regarding the professional preparation of teachers.

TEDS-M identified two types of teacher preparation routes across the participating countries: consecutive routes and concurrent routes. The consecutive route includes two phases: academic study, with specialization in the subject matter to be taught, followed by pedagogy and practical/field experiences. The concurrent route is defined by the first and second phases representing a single program of study resulting in a single teaching credential. Occasionally, program-types would be categorized as mixed, with elements of both concurrent and consecutive routes. In the whole of the TEDS-M study, there are 22,078 future teachers from 504 institutions and 775 programs in 17 countries across six types of teacher education programs (four types of primary programs and two secondary). The program types and the grades at which future primary-level teachers are qualified to teach are in Table 12.1; the number of candidates in each program group by country can be seen in Table 12.2 in the Methods section.

Table 12.1 Program-types represented in the TEDS-M database

Level	Program-type	Grades eligible
Primary	Lower-Primary	Generalists up to Grade 4
	Primary	Generalists up to Grade 6
	Primary/secondary	Generalists up to Grade 10
	Mathematics specialist	Specialists
Lower secondary	Lower-Secondary	Up to Grade 10
	Lower & upper secondary	To Grade 11 and above

Table 12.2 Primary program groups by country, represented in the TEDS-M database

	Lower primary	Primary	Primary/ secondary	Math specialist	Total
Botswana			86 (4)		86 (4)
Chile			657 (31)		657 (31)
Chinese Taipei		923 (11)			923 (11)
Georgia	506 (10)				506 (10)
Germany	935 (19)			97 (8)	1,032 (27)
Malaysia				576 (12)	576 (12)
Norway			551 (32)		551 (32)
Philippines		592 (33)			592 (33)
Poland	1,812 (86)			300 (39)	2,112 (125)
Russian Federation	2,266 (45)				2,266 (45)
Singapore		263 (4)		117 (2)	380 (6)
Spain		1,093 (48)			1,093 (48)
Switzerland	121 (7)	815 (14)			936 (21)
Thailand				660 (51)	660 (51)
United States		1,310 (56)		191 (15)	1,501 (71)
Total	5,640 (167)	4,996 (166)	1,294 (67)	1,941 (127)	13,871 (527)

Note: Counts are the number of future teachers in a given program group, with the number of programs in parentheses

The overarching TEDS-M research questions are in Tatto et al. (2012), along with an in-depth description of the two-stage sampling design and procedures that resulted in a representative sample of future teachers and teacher educators. The conceptual framework, design, and methodology for the study have also been rigorously documented in various reports (see Tatto et al., 2008; Tatto et al., 2012). Data were collected from four sources: (a) teacher institutions and programs, (b) teacher educators, (c) future primary teachers prepared to teach mathematics, and (d) future secondary teachers also prepared to teach mathematics.

Why Use Internationally Comparable Data?

Data from IEA's *Trends in International Mathematics and Science Study* (TIMSS, Gonzales et al., 2008) show considerable variation in fourth- and eighth-grade students' performance across 37 countries. In an effort to understand and give meaning to these findings, one possible path is to study factors that influence student performance. Certainly, teachers and teaching influence student performance, but teachers differ in their professional knowledge and beliefs, competencies, and capacities. For this reason, one avenue of investigation is teacher education programs, where the recruitment and preparation of future teachers have been shown to vary in important ways (Organisation for Economic Co-operation and Development [OECD], 2005). The TEDS-M project database contains data concerning how 17 countries prepare future primary and secondary mathematics educators. This database also includes information on the types of programs, which consist of multiple pathways to certification.

Teacher certification has not been consistently linked to higher-quality instruction nor to improved student outcomes in the Western world, with a particular lack of connection in the United States (Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2008; Darling-Hammond, Berry, & Thoreson, 2001; Guarino, Santibañez, & Daley, 2006; Phillips, 2010). In fact, given the variation in program quality, certification alone may not be adequate as a single predictor. Programs that lead to different kinds of certification may attract different types of future teachers as well as have different compulsory coursework that reflects different ideas about what kinds and amount of knowledge are sufficient and/or necessary to engage in the work of teaching. Tatto's (1998) research using the Teacher Education and Learning to Teach Study from the National Center for Research on Teacher Education (1985–1990) showed teacher education programs do not have coherent or unified goals for teacher education. When programs send mixed messages to participants, it is no surprise that programs do not substantially change participants' beliefs.

Therefore, examining who is qualified to teach and how they are prepared may facilitate an "understanding of how different ways of doing things may be used to accomplish similar goals, or how similar ways of doing things may serve different goals" (Rogoff, 2003, p. 34). Thus, the purpose of this study is to understand regularities and patterns of variation across primary-teacher preparation programs in multiple countries in (a) teaching-related beliefs, (b) mathematical content knowledge, and (c) the relationship between teaching-related beliefs and mathematical content knowledge.

Review of Relevant Literature

Mathematical Content Knowledge

How best to identify and measure the characteristics of highly qualified teachers remain open questions. In the 1980s, Shulman provoked a national conversation in the United States around the kinds of professional knowledge necessary for the work of teaching. Importantly, Shulman (1986) also set the stage for discipline-based educational research by posing the missing paradigm question (p. 6): "How does the subject-matter knowledge of the teacher transform into the content of instruction?" This call was answered by theorists, most notable of which includes the work of Deborah Ball, Heather Hill, and colleagues in mathematics education research. Ball proposed and validated multidimensional domains of teacher knowledge (Ball, Thames, & Phelps, 2008; Hill, Schilling, & Ball, 2004), arguing that there exists a specialized body of knowledge that is unique to teaching mathematics: pedagogical content knowledge for teaching. This knowledge domain has been linked to higher levels of instructional quality (Hill, Ball, Blunk, Goffney, & Rowan, 2007; Learning Mathematics for Teaching Project, 2011) as well as greater gains in student achievement (Baumert et al. 2010; Hill, Rowan, & Ball, 2005; Hill et al., 2007).

Mathematical content knowledge—or subject-matter knowledge—is the cornerstone of pedagogical content knowledge for teaching. Subject-matter knowledge, or disciplinary content knowledge, is composed of substantive and syntactic knowledge structures (Schwab, 1978). Substantive knowledge structures are defined by how basic concepts and principles are organized, including what ideas are central or peripheral. Syntactic knowledge structures are the set of ways that truths, falsehoods, or invalidity are established. Before teachers are able to respond to errors in student thinking, anticipate misconceptions, and respond to student propositions and questions in ways that are generative of future learning, teachers themselves "need not only understand that something is so; the teacher must further understand why something is so; on the grounds its warrants can be asserted and under what circumstances our belief in its justification can be weakened and even denied" (Shulman, 1986, p. 9).

Ball (1990) used a longitudinal questionnaire and interview data to explore both primary and secondary mathematics pre-service teachers' mathematical content knowledge and their perceptions of the role of content knowledge in the work of teaching. Careful analysis of future teacher knowledge of division with fractions indicated both groups of candidates had difficulty unpacking (Ma, 1999) the meaning of dividing with fractions, as evidenced by the proportion of candidates who did not have the ability to select or generate an appropriate representation for 1 ¾ ÷ ½. Moreover, three common assumptions about teaching primary and secondary mathematics emerged. First, candidates operated under the assumption that teaching school-level mathematics is not difficult: if one can *do* first-grade math topics, one can teach first-grade math topics. Second (and related to the first assumption), secondary-level mathematical education provides future teachers with sufficient levels of mathematical knowledge necessary for teaching. Third, majoring in mathematics ensures sufficient subject-matter knowledge for teaching. Contrary to these assumptions, Ball suggests the mathematical knowledge future teachers bring is inadequate for teaching mathematics for understanding.

This returns us to Shulman's question concerning the transformation of subject matter knowledge into instructional content, which is elaborated upon by Barnes (1989, p. 20) in relation to the professional preparation of teachers: "How do novices come to understand teaching (both constraints and possibilities) and how do they acquire the dispositions and capacities needed for principled thought and action?" The principled thought and action to which Barnes refers is the employment of sound judgment of the teacher that is grounded in cycles of observation and reflection coupled with what Ma (1999) refers to as a profound mathematical knowledge that enables teachers to engage in the tasks of teaching mathematics outlined by Ball et al. (2008). How novice teachers achieve this—how they organize, justify, and use their mathematical knowledge to "create and accomplish their intentions for learners" (Barnes, 1989, p. 14)—is an important teacher-preparation program outcome.

TEDS-M Findings

The TEDS-M framework for mathematical content knowledge was derived from IEA's TIMSS. Four domains were used to assess the mathematics content knowledge of future teachers: numbers and operations, geometry and measurement, algebra and functions, and data and chance. Two anchor points that provide descriptions of the performance of future teachers at particular score values were generated from the data and can be found in *Policy, Practice, and Readiness to Teach Primary and Secondary Mathematics in 17 Countries* (Tatto et al., 2012). Two findings are relevant for this study, which justify our interest in patterns of variability between program-types.

First, there is a wide range of achievement across program-types within each country. Even the highest achieving countries had some future teachers with relatively low scores (below the first anchor point); conversely, even the lowest achieving countries had some future teachers with scores above the first anchor point (internationally scaled mean score = 431). Second, for programs within each program-type, the difference between the highest and lowest scaled mean score is 100 points, which is more than one standard deviation. This suggests that some teachers at the primary grade within the same programs graduate with considerably more mathematical content knowledge than others.

Teacher Beliefs

Researchers have reached a general consensus that beliefs are "psychologically held understandings, premises, or propositions about the world that are felt to be true" (Richardson, 1996, p. 103) and are distinct from attitudes (McLeod, 1992; Tourangeau, Rips, & Rasinski, 2000), values (Dewey, 1933; Pajares, 1992, 1996), and knowledge (Connelly & Clandinin, 1986; Nespor, 1987; Nisbett & Ross, 1980; Thompson, 1992). Mathematics beliefs are "personal judgments about mathematics formulated from experiences in mathematics, including beliefs about the nature of mathematics, learning mathematics, and teaching mathematics" (Raymond, 1997, p. 552). Consistent with Rokeach (1968), Raymond's definition describes beliefs as having an organizational structure (similar to atoms), implying a systematic order of relationships between types of beliefs.

The beliefs of future teachers are influenced by their own experiences as learners. Following from Lortie's (1975) theory of the apprenticeship of observation, candidates enter preparation programs with powerful images and ideas (usually implicit, unexamined, and unarticulated) about teaching and student learning. Similarly, Pajares (1992) states beliefs are "formed early and tend to self-perpetuate, persevering even in the face of contradictions caused by reason, time, schooling, and experience" (p. 324). Pajares goes further, describing future teachers as insiders in a strange land: school and classrooms are familiar environments and contain familiar actors. This familiarity, alongside deeply rooted memories as students, act as the filters through which new information and experiences offered in training programs are processed. However, these earlier beliefs and memories may be irrelevant or counterproductive to the development of professional judgment.

Research has documented consistencies (e.g., Peterson, Fennema, Carpenter, & Loef, 1989; Polly et al., 2013) and inconsistencies (e.g., Raymond, 1997; Thompson, 1984) between beliefs and enacted instructional practices. Nevertheless, teacher beliefs that are productive (e.g., enable positive student outcomes; see Ambrose, 2004 or National Council of Teachers of Mathematics, 2014) and resist simplistic representations of the work of teaching can help novice teachers navigate the realities of classroom teaching. This conclusion is consistent with the observations of one of Raymond's (1997) case subjects, Joanna, who stated it would be helpful if preparation programs facilitated the examination and development of a personal belief system, or philosophy of teaching before the trials and errors encountered during her first years of teaching (p. 572). Raymond notes that the conflict Joanna describes is typical for novice teachers.

Our working definition of mathematics beliefs is consistent with the TEDS-M and TIMSS studies and is composed of lower-order beliefs about the nature of mathematics, beliefs about teaching mathematics, and beliefs about learning mathematics. These lower-order beliefs are related but have distinct features. We readily acknowledge there are many different ways to conceptualize and categorize beliefs about mathematics, teaching, learning and students. Since this study draws upon TEDS-M data, we take the belief scales from the TEDS-M questionnaires as our major belief categories.

Beliefs About the Nature of Mathematical Knowledge

Teacher beliefs about the nature of mathematics, or conceptions of the discipline of mathematics, are perhaps more clearly summarized by the question "What is mathematics?" Educated individuals hold different and varied answers to this question. Ball (1999) provides evidence that novice teachers hold unexamined, diminished views of mathematics. These views, in turn, are the basis for teachers' views of the content of instruction, including how material is presented and the means necessary to learn and engage in mathematical tasks.

Thompson (1992) cites four definitive theories of the nature of mathematics, which are dichotomized into two types within the TEDS-M framework. The first type generally characterizes mathematical knowledge as a unified and immutable body of knowledge, the sum of which is contained in facts, rules, and procedures. In contrast, the second type generally characterizes mathematical knowledge as a human creation that is the product of inquiry and therefore open to revision. Although Thompson may have envisioned the belief categories as distinct, these belief categorizations are not opposite ends of a spectrum, but instead can be viewed as orthogonal. Teachers may well have strong beliefs about knowledge construction in mathematics, but also see mathematics as a collection of rules and procedures.

Teachers' beliefs are a crucial filter through which teachers make instructional decisions, thus impacting the content and quality of instruction (Polly et al., 2013) and their use of curricula (Remillard, 2005; Remillard & Bryans, 2004). As teachers are learning to teach, "they apprehend and enact new instructional policies in light of inherited knowledge, belief, and practice" (Cohen & Ball, 1999, p. 335).

Beliefs About Mathematics Achievement

Teachers' beliefs about what it takes to learn mathematics also impact student learning. There is growing evidence that beliefs that foster student-centered pedagogies are associated with improved student learning outcomes (Dweck, 2000; Polly et al., 2013). Teachers with student-centered beliefs take the stance that students can make sense of problems and participate actively in their own learning (Capraro, 2001; Şahin & Yılmaz, 2011). This contrasts with more teacher-centered beliefs in which teachers think they must tell students information in order for students to learn that information. However, as with beliefs about the nature of mathematics, student-centered and teacher-centered beliefs should also be viewed as orthogonal rather than opposite ends of a linear continuum (e.g., Blömeke, Hsieh, Kaiser, & Schmidt, 2014). Teachers may respond positively to survey items about teacher-centered beliefs (such as the important of good explanations and telling students content) and about student-centered beliefs (such as the importance of students actively constructing knowledge and communicating their reasoning). This orthogonality is a reason both of these categories are scales on the TEDS-M questionnaire; were they opposites on a single continuum, one scale (with reverse coded items) would suffice.

TEDS-M Findings

Tatto et al. (2012, Chapter 6) performed descriptive analyses of the variation across and within countries of percentages of candidates who endorsed the procedural, inquiry, and fixed-ability beliefs. On a descriptive level, there are substantial systematic differences across countries, but generally much smaller differences among program groups within countries (e.g., lower primary versus primary within a country). However, we do not know how these beliefs are linked to other program outcomes.

Respondents in all countries generally endorsed the view that mathematics is a process of inquiry as well as a set of rules and procedures. The latter image of mathematics was most strongly endorsed in Botswana, Georgia, Malaysia, Oman, the Philippines, and Thailand, whereas the strongest rejections of this image came from Germany, Switzerland, and Norway. Teachers in Chilé, Chinese Taipei, Poland, the Russian Federation, Singapore, and Spain generally endorsed inquiry beliefs, but also strongly endorsed procedural beliefs. The latter finding is consistent with Blömeke et al. (2014), who asserted procedural and inquiry beliefs were not inversely related, but orthogonal. This finding suggests endorsing both inquiry and procedural beliefs is not indicative of contradictory or internally inconsistent belief systems. Instead, teachers are inherently sensible beings (Leatham, 2006), who have internally consistent beliefs systems, although they may not appear so to outsiders (Philipp, Clement, Thanheiser, Schappelle, & Sowder, 2003). Only a small number of countries saw the majority of respondents endorse the view that mathematics was a fixed ability. This view of learning was most strongly endorsed in Botswana, Georgia, Malaysia, the Philippines, and Thailand, and most strongly rejected in Germany, Norway, Switzerland, and the United States.

Why Productive Beliefs and Mathematical Content Knowledge Are Important Program Outcomes

Future teachers tend to underestimate the complexity of the tasks of teaching and the subject-matter knowledge they will need in order to be successful (Ball, 1990; Grossman, Wilson, & Shulman, 1989). Ambrose (2004) describes the findings from *Children's Mathematical Thinking Experience*, where 15 future teachers were interviewed after participating in an experimental pedagogy course paired with a mathematics course, and after four one-on-one sessions with 10-year-old children working with fractions. During these interviews, participants noted that *doing* math was distinct from *teaching* math, and conceptual understanding (operationalized in this study as multiple representations and solutions to problems) would be essential to their success as teachers. Ambrose suggests that one-on-one interactions with students helped this group of future teachers to realize teaching mathematics is more complex than they expected, highlighting the value of the material in the mathematics content courses. This study also highlights a prominent theme in the work of previous scholars—specifically, Weinstein's (1989) *optimism bias*, a disposition characterized by the novice teacher's belief that teaching entails straightforward work, consisting primarily of offering clear explanations to children.

Knowledge and beliefs are distinct but interdependent psychological resources (Ernest, 2006; McLeod, 1992; Philipp, 2007) teachers draw upon when making pedagogical decisions from one moment to the next during instruction (Campbell et al., 2014; Holm & Kajander, 2012; Ma, 1999; Wilkins, 2008). There is evidence that teachers' mathematical content knowledge has positive influences on the quality of mathematical instruction (Fennema & Franke, 1992; Lloyd & Wilson, 1998), but this does not guarantee that teachers with high levels of knowledge will teach in ways that facilitate understanding for all children (e.g., Ball, Lubienski, & Mewborn, 2001). Such findings substantiate Ball's (1990) claim that teachers with similar kinds and levels of mathematical knowledge may teach very differently from one another. One potential source of this variation is teacher beliefs about teaching and how learning is promoted. For example, different epistemic beliefs about the nature of mathematical knowledge and the depth of mathematical content knowledge lead to different instructional practices (e.g., Ma, 1999; McLeod & McLeod, 2002; Thompson, 1984, 1992).

Indeed, what novice teachers believe about what teaching entails and what they believe about the knowledge essential to engage in the tasks of teaching—such as those proposed by Raymond (1997) and Ball et al. (2008)—potentially influence the knowledge and skills teachers have upon completion of teacher preparation coursework. Hence, an understanding of what teachers believe about the nature of mathematics and learning in connection with levels of mathematical content knowledge may be instructive for short- and long-term programmatic evaluation and ensuing policy recommendations.

Research Question

The research question motivating this investigation centers on the relationship between beliefs about the nature of mathematics, beliefs about mathematics ability, and mathematical content knowledge. To unpack this broad question, we delineate three sub-questions: Within and between teacher preparation programs in the TEDS-M study,

1. To what extent do MCK scores vary as a function of pre-service teachers' *procedural beliefs about mathematics* in different teacher preparation programs in TEDS-M countries?
2. To what extent do MCK scores vary as a function of pre-service teachers' *inquiry beliefs about mathematics* in different teacher preparation programs in TEDS-M countries?

3. To what extent do MCK scores vary as a function of pre-service teachers' *fixed-ability beliefs about mathematics achievement* in different teacher preparation programs in TEDS-M countries?

We offer three associated hypotheses. First, future teaches who endorse the belief that learning mathematics is a procedural process will have lower MCK scores relative to their peers who do not endorse procedural beliefs as strongly. Second, future teachers who endorse the belief that learning mathematics is a process of inquiry will have higher MCK scores relative to their peers who do not endorse inquiry beliefs as strongly. Finally, future teachers who endorse the belief that mathematics is a fixed ability will have lower MCK scores relative to their peers who do not endorse fixed-ability beliefs as strongly.

Methods

TEDS-M Data

In the TEDS-M study, teacher preparation programs were categorized into different groups based on similarities among the programs. The TEDS-M documentation defines a *program* as "a prescribed course of study leading to a teaching credential," and a *program-type* as a group of programs sharing similar purposes and organizational features, such as the range of grade levels for which teachers are prepared, the duration and degree of specialization, and whether the program is concurrent or consecutive (see Schwille, Ingvarson, & Holdgreve-Resendez, 2013; Tatto et al., 2012, p. 20).

Programs were organized into six groups. At the primary level, these included: lower primary, teaching up to Grade 4; primary, teaching up to Grade 6; primary-lower secondary, teaching up to Grade 10; and primary mathematics specialist. At the secondary level, groups included: lower secondary, teaching up to Grade 10; and upper secondary, up to Grade 11 and above. This study focused only on primary programs. Table 12.2 contains the number of students in each program group and the number of programs (in parentheses) for each of the TEDS-M countries. A total of 13,907 future primary teachers participated across 15 countries and 525 programs. Complex sampling methods were used in data collection, and sampling weights are included in the TEDS-M datasets to adjust statistical estimates of population effects.

At the end of their programs, future teachers completed a survey that assessed a variety of constructs, including mathematical content knowledge (MCK), mathematical pedagogical content knowledge, opportunity to learn, and beliefs about mathematics (Tatto, 2013). This study utilized the MCK and beliefs measures. MCK was measured across the four content subdomains of number and operations, geometry and measurement, algebra and functions, and data and chance. These subdomains are labeled here as *number*, *geometry*, *algebra*, and *data*. Three item formats were used: complex multiple-choice, traditional multiple-choice, and constructed-response. Rasch item response theory models were used to scale and score the MCK measure across countries. The final MCK score distribution was rescaled to have a mean across participating countries of $M = 500$ with a standard deviation $SD = 100$. For additional details about study design, sampling, and modeling procedures, see Tatto, Rodriguez, Reckase, Rowley, and Lu (2013).

Future teacher beliefs about mathematics were measured across multiple domains, including beliefs about the nature of mathematics, beliefs about learning mathematics, beliefs about mathematics achievement, and beliefs about a teacher's own mathematics program. These domains were further divided into scales; for details on the validation of these scales, see the TEDS-M technical report (Tatto et al., 2013). Based on the research literature and this study's working hypotheses, this study utilized three of these TEDS-M scales: the nature of mathematics, with scales on rules and procedures (labeled as *procedures*); process of inquiry (labeled *inquiry*); and mathematics achievement as fixed ability (labeled *fixed ability*). Item response theory was again used for scoring, and measurement invariance was established (Tatto et al., 2013). For each scale, the score of 10 was set to be neutral in terms of balancing the dichotomy being measured (Tatto et al., 2013). For example, scores above 10 on the procedures scale indicate a belief that mathematics is fundamentally procedural (e.g., "Mathematics is a collection of rules and procedures that prescribe how to solve a problem"); scores below 10 indicate teachers did not endorse this belief, and scores near 10 indicate teachers were neutral toward this belief. Scores above 10 on the inquiry scale represented a belief that inquiry is integral to mathematics (e.g., "Mathematics involves creativity and new ideas"). Finally, on the fixed-ability scale, scores above 10 indicated a belief that mathematics ability is fixed (e.g., "Some people are good at mathematics and some aren't").

Analysis

Prior to modeling the relationships between MCK and the beliefs measures, descriptive statistics and correlations among these variables were examined for the full primary-teacher data set and by program. Internal consistency reliability (coefficient alpha) for the three beliefs scales was also examined at the program level.

A series of multilevel models were fit to each of the four program group data sets, with future teachers nested within institutions. Some countries only had a single teacher-preparation institution, so the model could not extend to nest institutions within countries. Thus, country was modeled as a fixed effect. Variables were entered sequentially in increasingly complex models, beginning with random effects for institutions in an unconditional base model, and then including main effects for country, and the beliefs scales *inquiry*, *procedures*, and *fixed ability*. Random effects for institutions provided an estimate of the variability in mean MCK by institution. Main effects then provided estimates of the mean MCK by country, and the predicted changes in MCK for one-point increases in beliefs. Two-way interactions between country and the beliefs scales were included next, first for inquiry, then procedures, and then for fixed ability. These interactions estimated differences in the MCK-beliefs relationships *by country*. Finally, random effects for the beliefs scales were included at the institution level, which allowed for variability in the MCK-beliefs relationships *by institution*.

The multilevel model below demonstrates how MCK was modeled for the primary/secondary program group, with teachers (*i*) nested within institutions (*j*). This program group contained three countries: Botswana, Chile, and Norway. Botswana served as the reference group when estimating country effects; thus, indicator variables are only included for Chile and Norway.

$$\begin{aligned} MCK_{ij} = \beta_{0j} &+ \beta_{1j}Chile_{ij} + \beta_{2j}Norway_{ij} + \beta_{3j}Procedures_{ij} \\ &+ \beta_{4j}Rules_{ij} + \beta_{5j}Fixed_{ij} + \beta_{6j}Chile_{ij}{}^{*}Procedures_{ij} \\ &+ \beta_{7j}Chile_{ij}{}^{*}Rules_{ij} + \beta_{8j}Chile_{ij}{}^{*}Fixed_{ij} + \beta_{9j}Norway_{ij}{}^{*}Procedures_{ij} \\ &+ \beta_{10j}Norway_{ij}{}^{*}Rules_{ij} + \beta_{11j}Norway_{ij}{}^{*}Fixed_{ij} + r_{ij} \end{aligned}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\vdots$$

$$\beta_{11j} = \gamma_{110}$$

In this model, all effects besides the intercept β_{0j} are fixed at level 2. The residual term at level 1 is r_{ij} (distributed normally with mean 0 and estimated variance), and the random intercept by institution at level 2 is u_{0j} (also distributed normally with mean 0 and estimated variance). *Chile* and *Norway* are indicator variables for each country. As a result, the fixed main effects γ_{10} and γ_{20} estimate the mean *MCK* for Chile and Norway as a difference from the mean for Botswana, controlling for the remaining covariates. The fixed main effects for the three future teacher variables, γ_{30}, γ_{40}, and γ_{50}, estimate the change in *MCK* for one-point changes in *Procedures*, *Rules*, and *Fixed*, respectively, for Botswana. The remaining interaction effects then estimate these same changes but for *Chile* and *Norway*, relative to the corresponding interaction for Botswana. The other multilevel models for the other program-types followed this same structure, with the countries in each case included when they had data about a program-type.

Statistical significance for a given model term or set of terms was determined by comparing model fit based on AIC for models with and without that term or set of terms. Belief variables were included as individual terms, whereas country indicators, and interactions between them and a given belief variable, were always included as sets of terms. For example, when testing for whether the relationship between *procedures* and *MCK* differed by country, fit for a model containing all *procedures* by country interactions was compared to fit for a model without those interactions; this would be considered a set of terms, as there were always multiple countries. When the addition of a term or set of terms resulted in a significant improvement in model fit, the term was retained in subsequent models; otherwise, it was removed. Models were fit and compared using Mplus version 7.11 (Muthén & Muthén, 2013). TEDS-M sampling weights, referred to as final sampling weights in the TEDS-M documentation, were utilized at both the individual and institution levels. Model results can thus be interpreted as estimating population parameters by country.

Results

Table 12.3 contains descriptive statistics for all groups combined and separate for each program group. The mean MCK for all groups combined was 505.8, just above the mean (500) across all teachers in the TEDS-M study. Mean MCK for lower-primary and primary-to-secondary groups were slightly lower, and means for primary and primary math specialists were slightly higher than the overall mean. Means on the belief scales were similar across program groups. For each belief scale, means were highest for primary math specialists. Note that none of the mean differences for program groups were tested for statistical significance. Instead, they were merely used descriptively to explore potential differences in performance by program group.

Reliabilities for the belief domains by program group tended to be at or just above .80 for all groups combined, lower primary groups, and primary groups. They were at or slightly below .80 for primary/secondary and slightly higher than .80 for primary math specialists.

Correlations among the MCK and belief measures, also contained in Table 12.3, differed somewhat across program groupings. MCK and inquiry showed a positive correlation for all groups and lower primary, but zero or nearly zero correlations for the other groups. MCK and procedures showed a negative correlation for all groups, a slightly smaller negative correlation for lower primary, and a stronger negative correlation for the remaining groups. MCK and fixed ability showed a similar negative correlation across all groups, except for primary math specialists, which showed a slightly more negative correlation than the other groups. Correlations among the belief scales were either near zero, for inquiry and fixed ability, or positive and moderately strong, for inquiry and procedures and procedures and fixed ability. None of the correlations were tested for statistical significance; rather, these were used to explore the potential relationships among variables by program group. The multilevel model results presented below were used to test for statistically significant main effects and relationships within program groups. However, as recommended in the TEDS-M documentation (Tatto et al., 2012), no statistical tests were performed to compare results across program groups.

Table 12.4 contains the fixed effect estimates from the final models selected for each program group. Main effects for country and each belief scale were all found to be statistically significant. Thus, within each program group, countries were found to differ from one another in their mean MCK performance, and changes in the beliefs scores were found to predict significant changes in MCK. In some cases,

Table 12.3 Descriptive statistics by program group

					Correlation		
		M	*SD*	Reliability	Inquiry	Proc	Fixed
All Groups	MCK	505.8	97.0		.12	−.18	−.14
	Inquiry	11.7	1.6	.83		.32	−.06
	Proc	10.9	1.3	.80			.38
	Fixed	9.8	1.0	.82			
Lower Primary	MCK	487.4	101.1		.18	−.13	−.15
	Inquiry	11.0	1.4	.80		.30	.04
	Proc	10.7	1.3	.80			.44
	Fixed	10.0	0.9	.80			
Primary	MCK	528.3	90.2		.00	−.29	−.12
	Inquiry	12.0	1.5	.80		.27	−.01
	Proc	10.9	1.3	.80			.39
	Fixed	9.5	1.0	.80			
Primary/ Secondary	MCK	461.7	85.6		.00	−.33	−.13
	Inquiry	12.2	1.6	.80		.14	−.15
	Proc	10.6	1.2	.75			.28
	Fixed	9.3	0.9	.72			
Primary Math Specialists	MCK	535.0	84.5		.04	−.26	−.24
	Inquiry	12.4	1.6	.85		.46	.05
	Proc	11.4	1.5	.83			.37
	Fixed	10.1	1.0	.81			

Notes: Proc stands for the procedures scale, *M* is the mean, *SD* is the standard deviation, Reliability is internal consistency estimated with coefficient alpha, and Correlation is the bivariate correlation

interaction effects were also statistically significant, that is, for some program groups, the estimated change in MCK for an increase in beliefs scores was found to differ by country.

The Country column of Table 12.4 contains the mean MCK estimate for each country, by program group, at the mean score for each belief scale. Thus, in the lower-primary group, Switzerland was estimated to have the highest mean MCK at 514.77. The remaining columns in the table contain a combination of main effects and interactions (when statistically significant) for each beliefs scale. A lack of significant interaction effect for a given belief scale is indicated by a constant value across countries in the column for the given belief scale. For example, the change in MCK for an increase in the inquiry scale was not found to differ by country for any program group. Instead, MCK was estimated to increase by 6.13, 8.13, 5.15, and 5.72 points in the four program groups, respectively, for a one-point increase on the inquiry scale (i.e., when there is an increase in the strength of the belief that mathematics are fundamentally procedural), and this relationship did not differ significantly by country. The relationship between MCK and procedures was found to differ by country in all program groups, and all of these effects were negative, indicating that MCK scores tended to decrease as belief about the procedural nature of math increased. Finally, the interaction for fixed ability and country was found to be

Table 12.4 Fixed effect estimates by program group

		Country	Inquiry	Procedures	Fixed
Lower Primary	Georgia	375.46	6.13	−1.01	−4.28
	Germany	498.93	6.13	−5.32	−3.04
	Poland	459.51	6.13	−9.19	−8.34
	Russian Federation	522.59	6.13	−2.87	−11.77
	Switzerland	514.77	6.13	−18.89	9.11
Primary	Chinese Taipei	598.22	8.31	−8.35	−1.07
	Philippines	431.05	8.31	−3.45	−1.07
	Singapore	590.12	8.31	−10.04	−1.07
	Spain	478.67	8.31	−11.02	−1.07
	Switzerland	539.59	8.31	−13.47	−1.07
	United States	517.48	8.31	−12.95	−1.07
Primary/ Secondary	Botswana	449.23	5.15	−5.98	0.02
	Chile	420.27	5.15	−5.76	0.02
	Norway	513.72	5.15	−21.89	0.02
Math Specialist	Germany	498.61	5.72	−40.07	10.05
	Malaysia	485.86	5.72	−7.35	−1.06
	Poland	593.87	5.72	−24.24	10.63
	Singapore	601.92	5.72	−10.21	7.17
	Thailand	531.15	5.72	−2.73	−17.44
	United States	525.73	5.72	−4.73	−4.47

Note: The first country listed in each program grouping served as the reference group. Estimates shown are all unstandardized

statistically significant in the lower-primary and primary math specialist groups; some of these were negative (e.g., all lower-primary countries but Switzerland) and a few were positive. The relationship between MCK and fixed ability was not found to differ by country in the primary and primary/secondary program groups, with fixed slopes of −1.07 and 0.02 respectively.

The majority of the belief scale slopes represent small to medium changes in MCK for one-point increases in the corresponding belief measure. For example, the largest effect for the inquiry scale was 8.31 for future teachers in the primary group; this indicates that for a one-point increase in beliefs about the importance of inquiry in mathematics, MCK was estimated to increase by 8.31 points, or just under one tenth of a standard deviation (the standard deviation for MCK in the primary group was 90.2). The remaining inquiry slopes were all positive and above 5 points.

Whereas effects for the procedures scale were all negative, some of these negative effects were noticeably larger than the rest. For Norway, in the primary/secondary group, MCK was estimated to decrease by about 22 points for a one-point increase on the procedures scale, roughly one fourth of a standard deviation (the MCK standard deviation for primary/secondary was 85.6). For Germany, in the primary math specialist group, MCK was estimated to decrease by about 40 points, nearly one half a standard deviation (84.5 for primary math). Lower-primary teachers in Switzerland and math specialist teachers in Poland also had large negative procedures effects.

Beliefs about the fixed nature of mathematics were estimated to have both positive and negative relationships with MCK in different countries, though the majority were negative. In the primary groups, a slight decrease of −1.07 was estimated across countries. In the primary/secondary group, the fixed slope across countries was close to zero at 0.02. In the lower-primary group, all countries but Switzerland were estimated to have negative fixed ability by MCK relationships. The negative effect of −11.77 for the Russian Federation was just over one tenth of a standard deviation. And in the primary math specialist group, the relationship was negative for three countries (Malaysia, Thailand, and the United States) and positive for the others (Germany, Poland, and Singapore), with the largest effect being −17.44, or .21 standard deviations, for Thailand.

Models including random slopes by institution did not fit significantly better than models without these terms. Thus, the beliefs scale slopes were not estimated to vary by institution. The models all included random intercepts by institution; however, all other effects were fixed.

Discussion

We had the opportunity to use the TEDS-M database to examine the statistical relationships between candidate teachers' mathematical content knowledge and beliefs, which varied between program preparation-types (lower primary, primary, lower primary/secondary, and math specialists). In the output from the four statistical models, we noticed five of the 17 countries within the TEDS-M database had multiple programs (Germany, Poland, Switzerland, Singapore, and the United States; e.g., Germany had a lower-primary and mathematics specialist program), presenting us with the opportunity to make descriptive observations within countries between program-types.

Overall, our three hypotheses were mostly confirmed. The following discussion highlights program features such as entry requirements, program-types, and credentials, as well as curriculum organization, content, and assessment, within and between countries. This information serves to contextualize our findings and fortify our discussion regarding why patterns of beliefs are linked to higher and lower mathematical content knowledge scores.

The aforementioned program features that we will use to contextualize and interpret the pattern of research findings are from the *TEDS-M Encyclopedia: A Guide to Teacher Education, Context, Structure, and Quality Assurance in 17 Countries* (Schwille et al., 2013). This document contains a discussion of the history, nature, philosophical underpinnings, and program characteristics of mathematics education of each country at the time of data collection. Importantly, these descriptions were authored by teams of researchers who are native members of each participating country.

Recall that programs within countries were the unit of analysis and that our analysis does not permit causal explanation for findings. The following discussion is organized by program-type. Within each program-type, we briefly describe the relationship between MCK scores and belief-type. We report beliefs individually, working under the theoretical assumption that procedural and inquiry beliefs are roughly orthogonal (as opposed to inversely related; Blömeke et al., 2014) beliefs about the nature of mathematics, while mathematics as fixed ability is a type of belief about student learning.

It was not possible to find comparable "pre" measurement occasions across different teacher preparation program-types; data were collected at one point in time, in the last year of teacher candidates' teacher preparation programs. Thus, we do not know about the extent or the direction of change in MCK scores and other related variables. A future study would be necessary to collect longitudinal data about the iterative changes of beliefs and MCK over time, to better examine the relationships among these variables.

Between Program Differences

Table 12.5 shows overall summaries of the findings by program-type that we discuss in more detail in this section. Generally, our hypotheses about the relationships between belief-types and MCK were confirmed.

Lower-Primary programs

Consistent with our first hypothesis, teacher candidates who endorsed procedural beliefs tended to have lower MCK scores. This effect seemed magnified for Switzerland. This effect was less pronounced for Germany, Poland, Georgia, and the Russian Federation.

Poland and the Russian Federation share historical contexts and teaching conditions, including five-year programs where the graduation requirements for candidates are similar to the standards held for majors in mathematics. This emphasis on mathematics during a two-cycle course of study may be associated with the rela-

Table 12.5 Summary of findings by program-type

Program-type	Inquiry	Procedures	Fixed
Lower primary	Positive for all program-types	Negative (some large) for all four program-types	Negative (mostly)
Primary			Negative
Primary/secondary			Zero
Math specialist			Mixed

tionship between endorsing procedural beliefs and MCK scores seen in Table 12.4. Yet future Russian teachers outperform their Polish counterparts on the measure of MCK. Additionally, endorsing procedural beliefs seems to be associated with a slightly more negative influence on MCK scores for future Polish teachers.

In contrast, German and Swiss teaching programs presently share similar features, but the relationship between procedural beliefs and MCK scores in Table 12.4 are more dissimilar in magnitude (compared to Poland and the Russian Federation). German and Swiss teacher candidates come from programs that have no federal regulation; the 16 German federal states and many cantons in Switzerland fall under more localized ministry guidance. The most significant similarity between Swiss and German candidates are the relatively high entry requirements: although both countries have an open-entry policy to university (i.e., every student who has successfully passed the high school final examination has the right to enroll at a university), academic requirements within programs are relatively high. In fact, Blömeke, Suhl, Kaiser, and Döhrmann (2012) and Blömeke, Suhl, and Kaiser (2011) report candidate performance on the German secondary school exit examination is one of the most important predictors of success in the teacher education program.

Within the lower-primary programs, Georgia seems to be an exceptional case, as the relationship between procedural beliefs and MCK is near zero. In contrast to the features of other countries with lower-primary programs, the most significant differences are, first, the absence of math-specific content knowledge for matriculation into preparation programs, and second, the absence of field coursework and experiences. Additionally, during the TEDS-M data collection period, Georgia was experiencing active conflict with the Russian Federation, including both military and political battles. Given the very low scores, it is not completely clear the MCK measure is adequately discriminating the teachers' mathematical content knowledge. If this floor effect is present, one would expect anomalous correlations among MCK and other teacher variables. Taken together, these interrelated factors may account for Georgia's pattern of results relative to lower-primary programs in other countries.

Consistent with our second hypothesis, teacher candidates who endorsed inquiry beliefs tended to have higher MCK scores. This finding aligns with Ambrose (2004), suggesting that we should expect teachers who endorse inquiry beliefs to have higher MCK scores. A deep knowledge of mathematics is necessary in order for teachers to successfully support student inquiry (e.g., Hill et al., 2005; Ma, 1999).

Also consistent with our hypotheses, teacher candidates who endorsed the belief that mathematics is a fixed ability tended to have lower MCK scores, with the exception of Switzerland. However, only 4.8% of the candidates endorsed fixed-ability beliefs, suggesting this small group of candidates had extreme scores and held views that did not mirror those of their peers. This effect was strongest for Poland and the Russian Federation and less pronounced for Georgia and Germany.

Overall, although there is no significant between-country variability for the relationship of inquiry beliefs with mathematical content knowledge, there is variability in the relationships among mathematical content knowledge and procedural and fixed beliefs for the lower-primary teacher candidates. Also worth noting is that variability in the relationship between procedural beliefs and knowledge emerge even in the presence of: shared historical contexts and teaching conditions (Poland and the Russian Federation), similar program features (Germany and Switzerland), and countries with MCK scores above the international average (the Russian Federation and Switzerland).

Primary Programs

Consistent with our first hypothesis, teacher candidates (up to Grade 6) who endorsed procedural beliefs about the nature of mathematics had lower MCK scores. This effect was strongest for the Switzerland, the United States, Spain, and Singapore and the weakest for the Philippines. Interestingly, secondary school graduates in the Philippines can be immediately certified to teach primary levels. This lack of formal teacher preparation may be related to this attenuated relationship. Without formal teacher preparation, future teachers' beliefs are shaped by their own primary and secondary experiences, which may not include explicit opportunities to reflect on theories of teaching and learning. Additionally, since nearly all Filipino primary teachers endorsed procedural beliefs, although exhibiting variance in MCK, we would not expect to find a strong correlation between procedural beliefs and MCK.

Although the relationship between procedural beliefs and MCK scores varied by country, this was not the case for inquiry and fixed-ability beliefs. Consistent with our second and third hypotheses (respectively), endorsing inquiry beliefs was associated with higher MCK scores, whereas endorsing fixed-ability beliefs was associated with lower MCK scores. Differences between countries were expected, but similarities between countries are especially interesting given the variability in routes to becoming a teacher in the United States and Switzerland relative to the more tightly controlled and more selective pathways of Chinese Taipei, Singapore, and Spain.

Both the United States and Switzerland have programs where multiple states and German-speaking cantons (respectively) have regulation and control over teacher licensure requirements (although there are still national guidelines in place). Chinese Taipei has a strong, centralized system, whereas Singapore has only one teacher education institution, resulting in rigorous and competitive programs. Spain also has a tightly controlled and selective pathway to the classroom. Although Spanish candidates come from public and private institutions that are fairly autonomous in their required programs of study, a teacher is considered to be a civil servant. Thus, candidates must complete a competitive civil servant exam after completing coursework. Selectivity comes into play because there are a fixed number of vacancies for which candidates may be hired. Given these fairly different country contexts and MCK scores, the consistency of connections between fixed-ability and inquiry beliefs and MCK is surprising.

Primary/Secondary Programs

The pattern of relationships found for primary teacher candidates were also found for primary/lower secondary generalists (up to Grade 10), but with some differences in the directions of the effects. Consistent with our first hypotheses, endorsing procedural beliefs was associated with lower MCK scores. This effect was particularly strong for Norway, but less pronounced for Botswana and Chile. The relationship between inquiry and fixed-ability beliefs did not vary by country; however, endorsing fixed-ability beliefs did not seem to be strongly associated with MCK scores.

With the exception of procedural beliefs for Norwegian candidates, the relative homogeneity in the pattern of relationships between all belief-types and MCK scores is striking. Although these countries are located on different continents and are arguably at different stages of development according to the OECD, they share similar program features that may be associated with the pattern of findings. For example, programs are concurrent with internship and field experiences interspersed throughout the program, with the greatest time commitment and most teaching responsibilities occurring during the final semester.

These primary/lower-secondary programs produce generalists who are trained to teach a non-trivial range of students and mathematical content (up to Grades 7, 8, and 10 for Botswana, Chile, and Norway, respectively). It is counterintuitive to think preparing candidates to teach a wider range of grade levels results in similar ideas and images related to the kinds of professional knowledge and dispositions required to be effective. Then again, it may because programs recognize the complexity inherent in teaching students that differ widely in age and experience that there is an explicit effort to articulate and share an image of the student as a learner in a way the transcends age. In other words, a shared image of the learner is what allows programs to prepare candidates to teach such a wide range of grade levels. Perhaps this same logic can be extended down to the primary level, where we observed the same pattern of findings.

Despite similarities in patterns of beliefs, one of the most striking differences among country profiles is who is attracted to the teaching profession. In Botswana, the procedure for program entry is competitive, attracting above-average students who, among other typical requirements, such as secondary school marks and coursework completed, are assessed in a half-day interview. In contrast, as of 2006, more than half of Chile's candidates graduate from private colleges of education and have been documented to score much lower on university entry exams relative to students who attend public universities. We note that, in part due to TEDS-M findings, in more recent years, Chile has initiated major changes to teacher preparation and certification (e.g., Tatto, 2015). Similar to Chile, Norwegian candidates are typically average or below average in secondary school records. Since 1982, the Norwegian Council of Mathematics has published reports on the average mathematics achievement scores of all secondary school graduates who will need to study mathematics during their time at university. In the 2009 report, the average mathematics achievement scores for teacher candidates relative to the sample was over 10% lower, suggesting few prospective generalists had a particular interest in mathematics. Future research on this topic might consider who is attracted to the teaching profession and their academic profiles in juxtaposition with the role of concurrent program timing of field experiences and the professional knowledge and dispositions thought necessary to be successful at teaching a wide range of students and mathematics.

Mathematics Specialists Programs

Consistent with our first hypothesis, mathematics specialist candidates who endorsed procedural beliefs tended to have lower MCK scores. This effect was particularly strong for German and Polish candidates. Consistent with our second hypothesis, endorsing inquiry beliefs was associated with higher MCK scores among mathematics specialists. This pattern of relationships was also found in lower primary, primary and primary/lower-secondary candidates.

Our third hypothesis was partially supported: endorsing the belief that math was a fixed ability was associated with lower MCK scores for Thailand, the United States, and Malaysia. Endorsing fixed-ability belief was associated with higher MCK scores in Germany, Poland, and Singapore. These three countries are particularly interesting in that endorsing procedural beliefs was associated with strong negative impacts on MCK scores. It seems that future specialists in these countries view mathematics as a process of inquiry, but the extent to which any individual can effectively participate is based on their innate abilities. It may be the case that since mathematics specialist programs tend to have higher entry requirements for mathematical content knowledge than do other primary programs, candidates in these programs may have fewer opportunities to learn MCK from their specialist programs. These opportunities (or lack thereof) may in turn influence the degree to which MCK and beliefs (which may be nurtured by these programs) are correlated. Another possible explanation may be that future specialists were already more competent and motivated to do well in math and they generally perceive these characteristics to be stable aspects of individual personality and character, having had no reason to need or think about a growth mindset (Dweck, 2006a, 2008) in the context of their own learning experiences.

Thailand seems to represent an exceptional case with respect to the strength of the negative relationship between MCK and fixed-ability beliefs. At the time of the TEDS-M study, the organization of the Thai system was undergoing change. As of 2004, candidates who wished to teach mathematics in Grades 1 through 12 needed to complete a five-year program that included coursework for a Bachelor of Science and one year of graduate study in education. The first cohort of students, all of whom were considered to be math specialists in the TEDS-M study, were in the fourth year of their programs and had yet to take their education coursework. Thus, the Thai pattern of findings may not represent final program outcomes.

Summary for Between Program Differences

We highlight three broad observations. First, the greatest range in the pattern of relationships was for lower-primary generalists (up to Grade 4) and for mathematics specialists. Second, there was significant variability in the effect endorsing procedural beliefs had on MCK scores, but this effect was consistently negative. Third, endorsing inquiry beliefs was associated with higher MCK scores, regardless of program-type.

Descriptive Differences in Multiple Programs Within Country

The relationships between MCK, beliefs, and their interaction were not statistically modeled. The observations in this section are descriptive and highlight interesting relationships.

Lower-Primary and Math Specialist Programs: Poland and Germany

Mathematics specialists in Poland had higher average MCK scores relative to their peers preparing to teach at the lower-primary level. German lower-primary and math specialist candidates had similar MCK scores. Endorsing procedural beliefs was associated with lower MCK scores in Polish and German candidates (with the effect more pronounced for specialists), whereas endorsing inquiry beliefs was associated with higher MCK scores. Interestingly, in both countries, endorsing beliefs about mathematics as a fixed, or innate ability was associated with higher MCK scores in the specialists, but lower MCK scores for their lower-primary peers. The latter finding might be better understood in light of the content-knowledge preparation and field experiences of Polish candidates.

Polish teacher education is not an independent field of study; it is a field of study within *other* fields of study. In fact, teachers who teach at the lower-primary level (Grades 1 through 3) are trained for "integrated teaching," and tend not to have content-specific training, as most of their coursework is centered on pedagogy. In contrast, candidates who teach Grades 4 and above are trained as math specialists and must meet mathematics requirements to complete the major while taking pedagogy courses that require a minimum amount of hours dedicated to teaching and field experiences that varies by institution. It may be the case that having less contact with children accounts for fixed-ability beliefs, or that pre-existing beliefs are based on their own learning experiences. In line with this interpretation, the TEDS-M headquarters in Poland has stated Poland suffers from academic drift, suggesting an imbalanced emphasis on content knowledge relative to pedagogical knowledge and practice (Tatto et al., 2012).

What accounts for these findings in German candidates is less clear. Candidates go through two phases of preparation. The first phase is academic, which typically takes 42 months to complete; the second phase is the field experience, which takes 18–24 months to complete. During the second phase, candidates teach part-time while also continuing to attend courses that cover general and subject-specific pedagogy. Candidates who intend to teach at the upper-primary grades are considered math specialists. What accounts for the pattern of findings in Poland does not translate over into Germany, which has the balanced program of study we infer to be partly responsible for candidates developing beliefs about the malleability of mathematical ability.

Lower-Primary and Primary Programs: Switzerland

Swiss primary level teaching candidates had slightly higher MCK scores relative to their peers preparing to teach in the lower-primary levels. Endorsing procedural beliefs was strongly associated with lower MCK scores for both groups, whereas endorsing inquiry beliefs was associated with higher MCK scores. However, there were also some noteworthy differences between these groups of teaching candidates. Endorsing fixed-ability beliefs was associated with higher MCK scores for lower-primary candidates, with the effect in the opposite direction for their primary peers.

Thinking about differences between these two Swiss programs is further complicated by differences between cantons. Although only German-speaking cantons were part of the TEDS-M study, responsibility for governance of teacher education lies with individual cantonal parliaments, although teacher education institutions have gained greater autonomy and integration into higher education institutions. Future research could examine Swiss results by canton to better understand the relationships among beliefs and MCK, relative to canton educational policies and contexts.

Primary and Math Specialist Programs: Singapore and the United States

In Singapore and the United States, math specialists had higher average MCK scores relative to their peers preparing to teach in primary schools. Endorsing inquiry beliefs was associated with higher MCK scores, whereas endorsing procedural beliefs was associated with lower MCK scores for primary and math specialists in both Singapore and the United States. The only difference in the pattern of effects lies in fixed-ability beliefs: Endorsing fixed beliefs was associated with lower MCK scores for primary level candidates in Singapore and the United States, as well as specialists in the United States. However, the opposite is true for math specialists in Singapore. Wilkins (2008), who studied the mathematical content knowledge, attitudes, beliefs, and practices of nearly 500 primary teachers found "not only were beliefs found to have the strongest direct effect on instructional practice but they also played a role in mediating the effects of teachers' knowledge and attitudes" (p. 156). It is possible that people who choose to become math specialists in Singapore are those who believe themselves to be good at math (i.e., holding fixed beliefs about their own abilities) and thus see the same in children. Such a fixed mindset is quite prevalent among mathematics teachers in general (Boaler, 2013).

Summary for Between-Program Descriptive Differences

Mathematics specialists generally had higher MCK scores relative to their peers preparing to teach at the lower-primary and primary level, with the exception of Germany. Additionally, candidates who were part of programs that would qualify them to teach at higher grade levels demonstrated increasingly higher average MCK scores. Mindset literature (Boaler, 2013; Dweck, 2006a, 2006b, 2008) suggests such fixed-ability beliefs are detrimental to achievement. Wilkins (2008) found beliefs are the strongest mediator for instructional practices, so it is worth understanding teachers' beliefs and how congruent those are to cultural norms for teaching. Mathematics specialists who endorsed fixed-ability beliefs had higher mathematics content knowledge compared to their counterparts who were preparing to teach at the lower-primary level in Singapore and Poland. This was not the case for the candidates who were preparing to teach at the primary level in the United States and Singapore.

Conclusion

Teacher knowledge and beliefs about mathematics teaching and learning are important teacher-preparation program outcomes. The TEDS-M database presented us with an opportunity to observe patterns among beliefs about the nature of mathematics and learning and mathematical content knowledge for teachers preparing to teach at the lower-primary, primary, or primary/lower-secondary grade levels as generalists, or mathematics specialists. Beliefs are multi-faceted, and although categories like student-centered and teacher-centered seem to be opposites, they in fact act in a more orthogonal manner, with teachers potentially holding strong beliefs along multiple dimensions.

Our analysis paints a picture both of program outcome variability and homogeneity among 15 countries. It appears candidates preparing to specialize in teaching mathematics or teach at the lower-primary levels exit their programs with highly variable beliefs about the nature of mathematics and mathematics learning, which are associated with differential effects on mathematical content knowledge. This finding stands in contrast to the slightly more homogeneous relationships among beliefs and mathematical content knowledge found for their peers preparing to teach at upper-primary and primary/secondary grade levels. Additionally, across program-types, endorsing the belief that mathematics is a process of inquiry from which learners can construct meaning was associated with higher mathematical content knowledge. However, the endorsement of the belief that mathematics is a fixed body of knowledge composed of sets of rules and procedures was associated with lower mathematical content knowledge. Although programs have varying traditions and cultural ideas about what teachers need to know and be able to do, the relationship between inquiry beliefs and program-types appear to be similar across programs-types and countries.

Since the TEDS-M data allow for correlational analyses but not causational, one policy recommendation we make is for programs to focus both on teaching MCK and on providing opportunities (both in coursework and field experiences) for participants to develop beliefs that mathematics is a process of inquiry from which learners can construct meaning. Further research is needed to determine potential causality. Similar research in the United States (e.g., Fennema & Nelson, 1997) suggests such causality may be different for different groups of teachers—for some, beliefs influence mathematical learning, whereas for others, their knowledge base in mathematics then influences their beliefs.

This chapter has described patterns of teacher beliefs across countries. When we made observations among different programs within a single country, we noticed candidates who were preparing to teach at higher grade levels or who were trained as specialists tended to have stronger mathematical content knowledge. The correlations of endorsing procedural and inquiry beliefs on mathematical content knowledge was consistent between programs within countries. However, endorsing mathematics as a fixed body of knowledge was a point of divergence.

These findings may provoke conversation about the extent to which preparation programs are achieving their intended goals. Future research is needed to connect preparation program outcomes to intended goals, with a focus on specific program inputs. Are programs surprised by the types of beliefs candidates endorsed? To what extent do these beliefs align with the intended images of teaching and learning structured by program experiences? Are these dispositions desirable and productive once candidates enter the profession?

Programs may also consider what other ideologies, practices, and strategies might spark innovation or revision to their national or local programs. For example, what kinds of benefits—material or in the form of social capital—need to be in place to attract candidates with strong backgrounds in mathematics or strong enough potential for learning enough mathematics to enter the classroom prepared to teach? Should there be multiple routes to certification? If so, how many is too many? Who are other stakeholders in teacher preparation programs who should be brought into the decision-making process on local and national levels?

This is intended to be neither a comprehensive nor exhaustive list of questions sparked by our analysis. It is not the intent of this paper to impose a single vision of a model preparation program. Ideas about the professional knowledge and dispositions essential to prepare candidates for classrooms of children are deeply rooted in national frames of minds, or beliefs systems, and therefore not something to gloss over. Instead, we offer this cross-national comparison as an invitation to widen the lens of programs and their stakeholders. Teacher preparation programs benefit from understanding their candidates' beliefs, and the connections among those beliefs and mathematical content knowledge, as well as to program goals.

References

Ambrose, R. (2004). Initiating change in prospective elementary school teachers' orientations to mathematics teaching by building on beliefs. *Journal of Mathematics Teacher Education, 7*, 91–119.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The Elementary School Journal, 90*, 449–466.

Ball, D. L. (1999). Crossing boundaries to examine the mathematics entailed in elementary teaching. In T. Lam (Ed.), *Contemporary Mathematics* (pp. 15–36). Providence, RI: American Mathematical Society.

Ball, D. L., Lubienski, S. T., & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed.). New York: Macmillan.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Barnes, H. (1989). Structuring knowledge for beginning teachers. In M. C. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 13–22). Oxford, UK: American Association of Colleges for Teacher Education, Pergamon Press.

Blömeke, S., Hsieh, F. J., Kaiser, G., & Schmidt, W. (Eds.). (2014). *International perspectives on teacher knowledge, beliefs and opportunities to learn*. Dordrecht, The Netherlands: Springer.

Blömeke, S., Suhl, U., & Kaiser, G. (2011). Teacher education effectiveness: Quality and equity of future primary teachers' mathematics and mathematics pedagogical content knowledge. *Journal of Teacher Education, 62*(2), 154–171.

Blömeke, S., Suhl, U., Kaiser, G., & Döhrmann, M. (2012). Family background, entry selectivity and opportunities to learn: What matters in primary teacher education? An international comparison of fifteen countries. *Teaching and Teacher Education, 28*, 44–55.

Boaler, J. (2013). Ability and mathematics: The mindset revolution that is reshaping education. *Forum, 55*(1), 143–152.

Boyd, D., Lankford, H., Loeb, S., Rockoff, J., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools. *Journal of Policy Analysis & Management, 27*(4), 793–818.

Campbell, P. F., Nishio, M., Smith, T. M., Clark, L. M., Conant, D. L., Rust, A. H., … Choi, Y. (2014). The relationship between teachers' mathematical content and pedagogical knowledge, teachers' perceptions, and student achievement. *Journal for Research in Mathematics Education, 45*(4), 419–459.

Capraro, M. M. (2001, November). *Construct validation and a more parsimonious mathematics beliefs scales*. Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR.

Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity, and improvement* (CPRE research report no. RR-043). Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.

Connelly, F. M., & Clandinin, J. (1986). On narrative methods, personal philosophy, and narrative unities in the study of teaching. *Journal of Research in Science Teaching, 23*, 293–310.

Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis, 23*(1), 57–77.

Dewey, J. (1933). *How we think*. Boston, MA: Heath & Co.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press, a Taylor & Francis Group.

Dweck, C. S. (2006a). *Mindset: The new psychology of success*. New York, NY: Ballantine Books.

Dweck, C. S. (2006b). Is math a gift? Beliefs that put females at risk. In S. J. Ceci & W. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence*. Washington, DC: American Psychological Association.

Dweck, C. S. (2008). *Mindsets and math/science achievement*. New York, NY: Carnegie Corporation of New-York-Institute for Advanced Study.

Ernest, P. (2006). The knowledge, beliefs and attitudes of the mathematics teacher: A model. *Journal of Education for Teaching, 15*(1), 13–33.

Fennema, E., & Franke, M. L. (1992). Teachers' knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147–164). New York, NY: Macmillan.

Fennema, E., & Nelson, B. S. (Eds.). (1997). *Mathematics teachers in transition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context* (NCES 2009–001 Revised). Washington, DC: IES.

Grossman, P. L., Wilson, S. M., & Shulman, L. S. (1989). Teachers of substance: Subject matter knowledge for teaching. In M. C. Reynolds (Ed.), *Knowledge base for the beginning teacher* (pp. 23–36). Oxford, UK: American Association of Colleges for Teacher Education, Pergamon Press.

Guarino, C. M., Santibañez, L., & Daley, G. A. (2006). Teacher recruitment and retention: A review of the recent empirical literature. *Review of Educational Research, 76*(2), 173–208.

Hawkins, D. (1967/1974). *I, thou, and it. The informed vision: Essays on learning and human nature* (pp. 48–62). New York, NY: Agathon.

Hill, H. C., Ball, D. L., Blunk, M., Goffney, I. M., & Rowan, B. (2007). Validating the ecological assumption: The relationship of measure scores to classroom teaching and student learning. *Measurement, 5*(2–3), 107–118.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal, 105*(1), 11–30.

Holm, J., & Kajander, A. (2012). Interconnections of knowledge and beliefs in teaching mathematics. *Canadian Journal of Science, Mathematics and Technology Education, 12*(1), 7–21.

Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.

Learning Mathematics for Teaching Project. (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education, 14*(1), 25–47.

Leatham, K. (2006). Viewing mathematics teachers as sensible systems. *Journal of Mathematics Teacher Education, 9*, 91–102.

Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago, IL: University of Chicago Press.

Lloyd, G. M., & Wilson, M. (1998). Supporting innovation: The impact of a teacher's conceptions of functions on his implementations of a reform curriculum. *Journal for Research in Mathematics Education, 29*(3), 248–274.

Ma, L. (1999). *Knowing and teaching elementary mathematics*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575–598). New York, NY: Macmillan.

McLeod, D. B., & McLeod, S. H. (2002). Synthesis—beliefs and mathematics education: Implications for learning, teaching, and research. In G. C. Leder, E. Pehkonen, & G. Törner

(Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 115–123). Dordrecht, The Netherlands: Kluwer Academic Publishers.
Muthén, L. K., & Muthén, B. O. (1998–2013). *Mplus user's guide* (7th ed.). Los Angeles, CA: Authors.
National Council of Teachers of Mathematics (NCTM). (2014). *Principles to Actions*. Reston, VA: Author.
Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies, 19*, 317–328.
Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
Organisation for Economic Co-operation and Development (OECD). (2005). *Teachers matter: Attracting, developing and retaining effective teachers*. Paris, France: OECD Publishing.
Pajares, F. M. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307–332.
Pajares, F. M. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578.
Peterson, P. L., Fennema, E., Carpenter, T. P., & Loef, M. (1989). Teachers' pedagogical content beliefs in mathematics. *Cognition and Instruction, 6*, 1–40.
Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257–315). Reston, VA: National Council of Teachers of Mathematics.
Philipp, R. A., Clement, L. L., Thanheiser, E., Schappelle, B., & Sowder, J. T. (2003). *Integrating mathematics and pedagogy: An investigation of the effects on elementary preservice teachers' beliefs and learning of mathematics*. Paper presented at the research pre-session of the annual meeting of the National Council of Teachers of Mathematics, San Antonio, TX.
Phillips, K. J. R. (2010). What does "highly qualified" mean for student achievement? Evaluating the relationships between teacher quality indicators and at-risk students' mathematics and reading achievement gains in first grade. *The Elementary School Journal, 110*(4), 464–493.
Polly, D., McGee, J. R., Wang, C., Lambert, R. G., Pugalee, D. K., & Johnson, S. (2013). The association between teachers' beliefs, enacted practices, and student learning in mathematics. *The Mathematics Educator, 22*(2), 11–30.
Raymond, A. M. (1997). Inconsistency between a beginning elementary school teacher's mathematics beliefs and teaching practice. *Journal for Research in Mathematics Education, 28*, 550–576.
Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211–246.
Remillard, J. T., & Bryans, M. B. (2004). Teachers' orientations toward mathematics curriculum materials: Implications for teacher learning. *Journal of Research in Mathematics Education, 35*, 352–388.
Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. P. Sikula, T. J. Buttery, & E. Guyton (Eds.), *Handbook of research on teacher education: A project of the Association of Teacher Educators* (pp. 102–119). New York, NY: Macmillan.
Rogoff, B. (2003). *The cultural nature of cognitive development*. New York, NY: Oxford University Press.
Rokeach, M. (1968). *Beliefs, attitudes, and values*. San Francisco, CA: Jossey-Bass.
Şahin, S., & Yılmaz, H. (2011). A confirmatory factor analysis of the Teaching and Learning Conceptions Questionnaire (TLCQ). *Journal of Instructional Psychology, 38*(3–4), 194–200.
Schwab, J. J. (1978). *Science, curriculum, and liberal education*. Chicago, IL: University of Chicago Press.
Schwille, J., Ingvarson, L., & Holdgreve-Resendez, R. (Eds.). (2013). *TEDS-M encyclopedia: A guide to teacher education context, structure, and quality assurance in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).*

Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14.

Tatto, M. T. (1998). The influence of teacher education on teachers' beliefs about the purposes of education, roles, and practice. *Journal of Teacher Education, 49*(1), 66–77.

Tatto, M. T. (Ed.). (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Tatto, M. T. (2015). The role of research in the policy and practice of quality teacher education: An international review. *Oxford Review of Education, 41*(2), 171–201. https://doi.org/10.1080/03054985.2015.1017405

Tatto, M. T., Rodriguez, M. C., Reckase, M., Rowley, G., & Lu, Y. (2013). Scale development and reporting: Opportunities to learn, beliefs and mathematics knowledge for teaching. In M. T. Tatto (Ed.), *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report* (pp. 161–174). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Peck, R., & Rowley, G. (2008). *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M., & Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Thompson, A. (1984). The relationship of teachers' conceptions of mathematics and mathematics teaching to instructional practice. *Educational Studies in Mathematics, 15*, 105–127.

Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of research. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 209–239). New York, NY: Macmillan.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, UK: Cambridge University Press.

Weinstein, C. S. (1989). Teacher education students' preconceptions of teaching. *Journal of Teacher Education, 4*, 31–40.

Wilkins, J. L. M. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education, 11*, 139–164. https://doi.org/10.1007/s10857-007-9068-2

Chapter 13
Future Teachers' and Teacher Educators' Perceptions of Learning Mathematics Instruction and Relationships to Knowledge

Rachel A. Ayieko

Abstract This study used the Teacher Education and Development Study in Mathematics (TEDS-M) to examine the relationships between opportunities to learn (OTL) mathematics instruction for conceptual understanding and primary future teachers' (PSTs) knowledge for teaching mathematics in three countries: Poland, Russia, and the United States. The frequencies of opportunities to learn (OTL) mathematics instruction for conceptual understanding varied between PSTs and teacher educators. A comparison of the teacher educators' and PSTs' responses suggests that the PSTs had fewer opportunities to learn mathematics instruction for conceptual understanding than were intended by the teacher educators at the program level in the three countries. The patterns of relationships from a multilevel regression analysis in each of the selected countries show variations across contexts and categories of knowledge. In particular, the OTL how to (a) make distinctions between procedural and conceptual knowledge and (b) show why a procedure works, were significantly related to PSTs' knowledge for teaching mathematics between programs in the United States and Russia, respectively. The OTL how to show why procedures work was significantly related to PSTs' knowledge for teaching mathematics within the programs in the three countries. Policy implications for mathematics teacher education are discussed.

This research was supported by a grant from the American Educational Research Association, which receives funds for its "AERA Grants Program" from the National Science Foundation under Grant #DRL-0941014.

This report is based on my doctoral dissertation study done while at Michigan State University.

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

R. A. Ayieko (✉)
Department of Instruction and Leadership in Education, Duquesne University, Pittsburgh, PA, USA
e-mail: ayiekor@duq.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_13

Introduction

Reforms in mathematics education require teachers to use instructional approaches that many of them did not experience when learning mathematics in school or learning how to teach in their teacher preparation programs. The reform methods for teaching emphasize teaching mathematics for understanding rather than having students memorize the rules or use prescribed procedures to solve given tasks. Additionally, these instructional methods stress the use of open-ended tasks and real-life situations, problem-solving, and collaboration (Boaler, 2002, 2008; Common Core State Standards [CCSS], 2010; National Council of Teachers of Mathematics [NCTM], 2000, 2014). The reform approaches for teaching mathematics have been shown to improve student learning (e.g., Boaler, 2008). However, to be successful in implementing these instructional methods, teachers need certain knowledge for teaching mathematics.

Teacher educators have introduced strategies to expand on future teachers' (PSTs) knowledge for teaching mathematics, with notable success. Such opportunities include developing skills for providing explanations (e.g., Charalambous, Hill, & Ball, 2011), differentiating procedural and conceptual knowledge (e.g., Bartell, Webel, Bowen, & Dyson, 2012), and using multiple strategies to solve given tasks (e.g., Crespo, 2000; Ryken, 2009). Mathematics educators have found these opportunities to be successful in expanding PSTs' knowledge for teaching. Most of the research, however, has consisted of small-scale studies of promising interventions in teacher preparation using varied instruments, analyses, and data sources. As a result, the insights gathered from these studies are limited in scope and are not generalizable to a wider population. A large-scale study can provide empirical evidence of generalizable relationships between these opportunities to learn mathematics for conceptual understanding and PSTs' knowledge for teaching mathematics insofar as it (a) uses variables representing similar interventions across different contexts and (b) uses the same analytical approach.

In this study, I use the Teacher Education and Development Study in Mathematics (TEDS-M), which is a cross-national study of teacher preparation programs in 17 countries. The purpose of TEDS-M was to investigate how primary and lower secondary teachers are prepared to teach mathematics and the influence of these opportunities on their knowledge development. This study presents a comparison of the intended and experienced opportunities to learn (OTL) mathematics instruction for conceptual understanding using teacher educators' and future teachers' perspectives, respectively. Additionally, it examines the relationships between the experienced OTL mathematics instruction for conceptual understanding and PSTs' mathematics knowledge for teaching in three countries: Poland, Russia, and the United States.

The three countries were selected based on the countries meeting at least two of the following criteria:

1. PSTs' high content knowledge and pedagogical content knowledge scores. Of the selected countries, Poland and Russia were identified as being among the top 10% of teacher preparation programs based on PSTs' mathematics content

knowledge scores, while the United States is among the top 10% in mathematics pedagogical content knowledge scores (Schmidt, Burroughs, & Cogan, 2013).

2. Availability of more than one type of teacher preparation program within a country, to enable comparing the OTL between the generalist and specialist programs.
3. Primary students' performance in international assessments. Poland is also one of the countries identified as having shown a significant improvement in students' mathematics test scores on international assessments, such as the Program for International Student Assessment (Organization for Economic Co-operation and Development [OECD], 2011).
4. Opportunity for comparison of different primary grade specializations offered in teacher preparation programs. The selection of Russia and the United States can provide information on the differences or similarities between lower primary (Grade 4 maximum and Grade 6 maximum) specializations.

Theoretical Perspective

The question used to frame this study is "What knowledge do teachers need to teach mathematics for conceptual understanding, and how do teacher preparation programs support them to develop this specialized knowledge?" This study draws from the teacher knowledge and opportunities to learn frameworks. In the first theoretical section, I discuss the literature on the knowledge domains for teaching. In the second section, I provide different conceptualizations of opportunities to learn. I then discuss studies that have examined relationships of opportunities to learn and PSTs' knowledge. A final section of the theoretical discussion examines earlier studies done using the TEDS-M data. Based on the literature presented, I discuss what is missing and how this study provides another lens for examining OTL, PSTs' knowledge for teaching mathematics, and the relationships that exist between them using a large-scale data analysis.

Knowledge for Teaching Mathematics

Teacher knowledge has been identified as one of the key factors for the improvement of students' learning of mathematics. The interest in teacher knowledge has led scholars to propose categories of teacher knowledge and subsequent redefinitions of teacher knowledge. Shulman (1986, 1987) proposed a *knowledge base for teaching* composed of three domains: content knowledge, pedagogical content knowledge, and curriculum knowledge. Additionally, Shulman (1987) included knowledge of the learners and their characteristics, knowledge of the educational context, and knowledge of educational foundations as part of the knowledge base. Other scholars have reexamined and re-defined the knowledge for teaching mathematics (e.g., Ball, 1993; Ball, Thames, & Phelps, 2008; Fan & Cheong, 2002; Grossman, 1990; Hill, Rowan, & Ball, 2005; Ma, 1999). The suggested categories of knowledge for teaching mathematics are briefly described in the following sections.

Content Knowledge Shulman (1986, 1987) defined content knowledge, using Schwab's (1978) definition, as knowledge of the substantive and syntactic structures of subject matter. Substantive structures, on the one hand, are the "basic concepts and principles organized to incorporate facts" (Shulman, 1987, p. 9). Syntactic structures, on the other hand, include the rules of the subject comprising the language, symbols, and the axioms that ascertain the truisms and validity in the discipline (Shulman, 1986, 1987). The *procedural knowledge* domain of knowledge, which Grossman (1990) and Carpenter (1986) defined as knowing how to perform procedures or follow predetermined steps to solve a problem, would also fall within the content knowledge domain.

Pedagogical Content Knowledge Pedagogical content knowledge comprises knowing the subject matter and how to teach it. Shulman described pedagogical content knowledge as including the representation of ideas and the use of metaphors, analogies, and strategies that teachers draw on to make learning accessible to students (1986, 1987). Additionally, pedagogical content knowledge includes the ability to anticipate students' errors, conceptions, and misconceptions (Shulman, 1986). Dewey (1902) used the term "psychologizing" the curriculum or transforming it to suit the child's level. In particular, it is changing the subject matter in a way that is familiar and appealing to the students (Dewey, 1902).

Curricular Knowledge Curricular knowledge is knowledge of the topics of a subject as well as the different curricula that are available (Shulman, 1986, 1987). Further, knowledge of the *lateral* as well as the *vertical* curriculum is included in this knowledge domain. Lateral curriculum knowledge is essential for knowing the connections of the topics in one lesson to those "lessons or topics or issues being discussed simultaneously in other classes" (Shulman, 1986, p. 10). Vertical curriculum knowledge is necessary for teachers to know the relationships of future topics with those already taught (Shulman, 1986), as well as knowledge of the ever-changing curriculum materials (Shulman, 1987).

Other Conceptualizations of Knowledge for Teaching Mathematics Shulman's knowledge domains have been critiqued, built upon, and redefined by other mathematics education scholars. Ma (1999) associated knowledge for teaching mathematics with a profound understanding of fundamental mathematics. According to Ma, *fundamental* refers to foundations, and *profound* is the understanding of mathematics that is "deep, broad, and thorough" (p. 120). Ma's definition of content knowledge included the knowledge of multiple perspectives for solving mathematical tasks and the longitudinal coherence of the curriculum. Thus, Ma's definition combines two domains of knowledge that Shulman proposed: content knowledge and curricular knowledge.

Other scholars (e.g., Ball et al., 2008) have argued that Shulman's knowledge domains are not well understood and need more development. They have suggested *common content knowledge* and *specialized content knowledge* as subcategories of content knowledge. Common content knowledge is the "mathematical knowledge and skills used in settings other than teaching" (Ball et al., 2008, p. 399). Specialized content knowledge, on the other hand, is the unique knowledge and skill needed to teach mathematics effectively (Ball et al., 2008). Additionally, Ball et al. (2008) divided pedagogical content knowledge into two domains: (a) knowledge of content and students and (b) knowledge of content and teaching. The first domain—"knowing about students and knowing about mathematics" (p. 401)—incorporates knowing students' thought processes when learning mathematics and anticipating students' errors, misconceptions, and difficulties with particular mathematical concepts (Ball et al., 2008). The second domain, knowledge of content and teaching, combines knowing (a) how to select examples, (b) which representations to use, and (c) how to select students' ideas that can be expanded (Ball, et al., 2008). These additional domains of pedagogical content knowledge give explicit attention to the distinct knowledge categories that underlie this knowledge dimension.

Fan and Cheong (2002) suggested other subcategories of pedagogical knowledge, including pedagogical curricular knowledge, knowledge of methods of instruction, and pedagogical content knowledge (as cited in Tatto et al., 2008). The pedagogical curriculum knowledge and pedagogical content knowledge domains are similar to those proposed by Shulman. However, knowledge of methods of instruction is added to the pedagogical knowledge categories and is similar to the specialized content knowledge domain proposed by Ball et al. (2008). The knowledge domains proposed by Ball et al. (2008), Shulman (1986), Ma (1999), and Fan and Cheong (2002) framed the knowledge domains in the TEDS-M knowledge questionnaire.

Opportunity to Learn

OTL is one of the five variables that Carroll (1963) proposed to explain the differences in students' learning. Carroll classified OTL as an external condition that can be manipulated to increase students' success. OTL can refer to the "school schedule… the time allowed for learning" (p. 26). Scholars have reconceptualized the initial definition of OTL given by Carroll (1963). Schmidt et al. (2001), for example, conceptualized OTL frameworks as the instructional time assigned for a topic, the time intended in national curriculum guides for teaching a subject, or the time teachers report they spend teaching a topic. Similarly, Floden (2002) stated that OTL is a measure of the relative emphasis of the topic in relation to other topics and can be inferred from reports about whether the topic has been taught or will be taught. In addition, Husén (1967) proposed understanding OTL as the relationship between what has been taught and what is assessed. Consequently, OTL can be assessed with respect to curriculum, teachers' instructional time on content or subjects (Floden, 2002; Schmidt et al., 2001; Törnroos, 2005), or coverage of topics assessed by various tests Husén, 1967). In all these definitions, the OTL focuses on what the student experiences and is a basis for creating equity in students' learning.

Opportunities to Learn Mathematics Instruction for Conceptual Understanding Teachers must negotiate, transform, and adapt their knowledge for teaching mathematics to align with reforms in methods of instruction. Specifically, teachers are expected to involve their students in problem-solving, encourage multiple strategies to solve tasks, support students' reasoning and proof, and use appropriate representations to communicate their ideas (Conference Board of the Mathematical Sciences [CBMS], 2000, 2012; Common Core State Standards [CCSS], 2010; NCTM, 2000, 2014). To prepare PSTs to be able to teach effectively using these reform methods, OTL classroom instruction specific to learning school mathematics for conceptual understanding must be emphasized during their teacher preparation.

Some approaches used by teacher educators emphasize the ways PSTs learn to identify and teach for conceptual understanding, along with reasoning mathematically. The instructional approaches include providing opportunities to develop skills for providing explanations (Charalambous et al., 2011; Chick, 2003), ways of learning to differentiate procedural and conceptual knowledge (Bartell et al., 2012; Chinnappan & Forrester, 2014), and approaches that emphasize the affordances of multiple strategies to solve problems (e.g., Crespo, 2000; Grant & Lo, 2009; Ryken, 2009). Research on the use of these approaches indicated that PSTs were able to recognize conceptual understanding from students' responses (Crespo, 2000) and to perceive teaching and learning mathematics as focused on reasoning about mathematics (Bartell et al., 2012; Chinnappan & Forrester, 2014). Furthermore, PSTs appreciated the influence of the problem context on students' reasoning (Grant & Lo, 2009), were able to provide mathematical explanations (Charalambous et al., 2011; Chick, 2003), could focus on student thinking, and could critically analyze different representations (Ryken, 2009). In some cases, at the beginning of the mathematics-related courses implementing these approaches, PSTs focused more on right and wrong answers and had a rule-bound perception of mathematics, but by the end of the courses, the PSTs were more critical in analyzing their students' responses. PSTs' focus also shifted toward some of the recommendations made by the CBMS (2000, 2012) and the National Standards for School Mathematics (CCSS, 2010; NCTM, 2000, 2014). Together, these studies point out the positive influence of these interventions on different dimensions of PSTs' knowledge for teaching mathematics.

These reports, however, were mostly small-scale interpretive case studies in which different analytical methods and data analyses were used. The current study's purpose is to investigate these promising relationships on a large scale and to provide empirical evidence of relationships between the interventions used in teacher preparation and the development of a broader knowledge base for teaching mathematics for understanding. Note that there are earlier studies on OTL and PSTs' knowledge using the TEDS-M data. A consideration of these studies highlights how OTL has been conceptualized and how the conceptualizations relate to PSTs' knowledge for teaching.

Initial Studies on OTL and Knowledge for Teaching Mathematics Using TEDS-M Data

Knowledge for Teaching Mathematics In the TEDS-M studies, knowledge for teaching is categorized into content knowledge (CK) and pedagogical content knowledge (PCK). PCK in the TEDS-M study is a composite of "curricular knowledge, knowledge about the planning for teaching, and knowledge related to enacting teaching" (Tatto et al., 2012, p. 131). The TEDS-M categorizations of teacher knowledge borrow and modify Shulman's (1986, 1987) knowledge domains and Fan and Cheong's (2002) proposed knowledge for teaching dimensions. Initial studies using the TEDS-M data suggest there are significant differences in the knowledge for teaching mathematics both within and among teacher education programs across the selected countries (e.g., Blömeke & Kaiser, 2014; Senk et al., 2012).

OTL to Teach Mathematics The earlier studies that used TEDS-M data mainly focused on time PSTs spent in teacher preparation, topics covered, practical experiences, and instructional time. The researchers found differing emphases within teacher preparation programs across contexts, program types, and content areas. Notably, specialist programs covered more tertiary mathematics topics than generalist programs (Blömeke & Kaiser, 2014; Tatto et al., 2012). The topical area most PSTs focused on was *numbers*, while *calculus* was the least studied (Blömeke & Kaiser, 2014). Additionally, the study of school mathematics and tertiary mathematics varied across the countries, but learning of teaching methods was found to be common (Blömeke & Kaiser, 2014). Further, the PSTs in the higher-achieving countries, such as Taiwan, took more formal mathematics courses. In sum, the focus of the research on OTL is on the patterns of course-taking at various levels of mathematics and using various mathematics methods. Missing from these studies is the extent of OTL mathematics teaching for conceptual understanding; an important component of teaching and learning mathematics in the twenty-first century.

Relationships Between OTL and Knowledge for Teaching Mathematics From the earlier studies on teacher preparation that used TEDS-M data, some significant findings on the relationships between OTL and teachers' knowledge for teaching mathematics have been reported. The number of content courses taken and the level of focus on mathematics instruction were significantly related to PSTs' CK and PCK (e.g., Blömeke & Kaiser, 2014; Wong, Boey, Lim-Teo, & Dindyal, 2012). Missing from these studies were discussions of specific OTL reform-oriented mathematics instruction or mathematics instruction for conceptual understanding. This study builds on these initial predictive findings from the TEDS-M data by focusing on particular OTL mathematics instruction that relates to learning to teach mathematics for conceptual understanding at the primary level.

In sum, the studies about OTL that have discussed how to explore multiple solution strategies (Crespo, 2000; Grant & Lo, 2009; Ryken, 2009), how to distinguish between

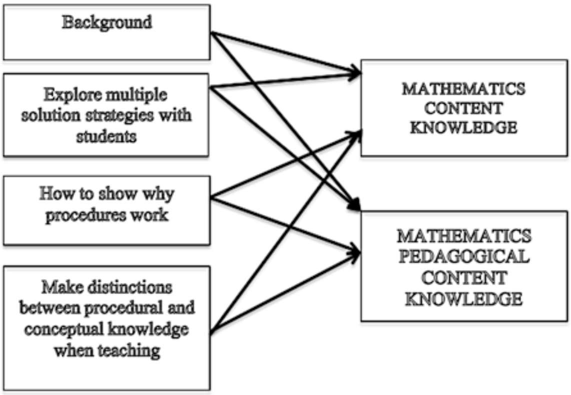


Fig. 13.1 Conceptual framework of OTL mathematics instruction and knowledge for teaching mathematics

procedural and conceptual knowledge (Bartell et al., 2012; Chinnappan & Forrester, 2014), and how to show why a procedure works (Charalambous et al., 2011; Chick, 2003) align with particular items in the TEDS-M data. These specific OTL, that have been shown to expand PSTs knowledge for teaching mathematics, form the basis of this study. Figure 13.1 is a diagrammatic representation of the hypothesized relationships between (a) OTL and PSTs' background and (b) PSTs' knowledge for teaching mathematics as measured by variables examined in this study.

The questions guiding the study are as follows: (a) What are the differences in the OTL mathematics instruction for conceptual understanding in the intended and experienced curriculum in the teacher preparation programs in Poland, Russia, and the United States? (b) Are there significant differences in the OTL mathematics instruction and knowledge for teaching mathematics across the teacher programs in the three selected countries? (c) What are the relationships between the OTL mathematics instruction for conceptual understanding and PSTs' knowledge for teaching mathematics within and among the teacher preparation programs in the three countries?

The following hypotheses were used to guide the study: (a) There are no significant differences in the OTL and knowledge for teaching mathematics across the three selected countries. (b) There is a positive relationship between the OTL mathematics instruction for conceptual understanding and PSTs' knowledge for teaching mathematics within and among the teacher preparation programs in the three selected countries.

Methods

This study used the TEDS-M country-level data. TEDS-M was supported by the International Association for the Evaluation of Educational Achievement (IEA) and the national centers of the participating countries. Russia, Poland, and the United States were the three countries selected for this investigation, which allowed for

Table 13.1 Description of PSTs and programs in the three countries

Country	Number of PSTs	Number of institutions	Number of teacher educators	Program types	% Female	Grade-level
Poland	2,112	78	734	Specialist, generalist	94.8	1–4
Russia	2,266	49	1,212	Generalist	92.2	1–4
United States	1,501	51	665[a]	Specialist, generalist	88.6	1–6

[a]This is the number of teacher educators at the program level who participated in the study. The participation rate for teacher educators was below 60%

within-country comparisons between specialist and generalist programs (in the United States and Poland), and also provided the possibility for comparison between different grade specializations in the programs within the three countries—Poland and Russia compared to the United States. All the countries have concurrent programs, or programs that offer content-specific courses, pedagogy courses, and education courses as a single requirement for the teaching credential, all taken during one course of study. Consecutive programs, in contrast, require candidates to have already received a specialization in the content they intend to teach before taking a second degree that is a teacher preparation course (Tatto et al., 2012). A summary describing the number of institutions and individual PSTs from the three countries is presented in Table 13.1.

The PSTs in the United States had a wider grade span for their grade specialization than the Polish and Russian teacher preparation programs. Additionally, two types of teacher preparation programs are available for full-time study in the United States and Poland, whereas all Russian primary PSTs are from generalist programs. Finally, there are few male PSTs in the teacher preparation programs in all three countries.

Units of Analysis

The relationships between OTL to teach mathematics and PSTs' knowledge for teaching mathematics within and among the teacher preparation programs in the three countries were examined. The units of analysis were the PSTs and teacher educators within and among the institutions.

Variables

Independent Variables for PSTs' OTL mathematics instruction for conceptual understanding are based on responses from teacher educators and PSTs in three OTL categories: (a) how to explore multiple solution strategies (*Multiple strategies*),

(b) how to show why a mathematical procedure works (*Why procedures work*), and (c) how to make distinctions between procedural and conceptual knowledge when teaching mathematics concepts and operations to pupils (*Make distinctions*) (Brese & Tatto, 2012, p. 7). The responses for these items in the future teacher and teacher educator questionnaires were made on the following ordinal scale: (1) *never,* (2) *rarely,* (3) *occasionally,* or (4) *often* (Brese & Tatto, 2012).

Covariates The TEDS-M study used gender and socioeconomic status (SES) as the control variables that describe the characteristics of the PSTs, but these variables are not of interest in this study. The parental level of education and the number of books in the home have been identified in recent studies as providing an adequate measure of SES. Carnoy (2015) referred to these variables as family academic resources.

Highest Level of Education in the Household The highest level of parental education in the household was based on a categorical scale using the International Standard Classification of Education (ISCED). Using the UNESCO Institute of Statistics mapping, the parents' educational level was recoded to give the cumulative years of schooling of the parent with the highest level of education. This mapping was based on the 1997 ISCED because at the time the data was collected, the new 2011 ISCED had not been released.

Resources in the Home The number of books was selected to represent this measure. Across the countries, this variable showed variation across the PSTs. This variable was initially ordinal. However, in this study the variable was dichotomized such that 1 represented more than 100 books in the home and 0 represented 100 books or fewer in the home. As stated by Cowan et al. (2012), family possessions may not be an accurate measure of SES. Indeed, a descriptive analysis of the resources in the home, such as possession of a calculator, desk, DVD player, or computer, did not show a significant variation across all countries that participated in the study.

Dependent Variables The knowledge for teaching mathematics composed of mathematical content knowledge (MCK) and mathematical pedagogical content knowledge (MPCK) were the outcome variables. The mathematical content knowledge domain comprised the following subdomains: numbers and operations, geometry and measurement, algebra and functions, and data and chance. The cognitive domains of mathematical content knowledge items were knowing, applying, and reasoning (see Tatto, 2013). The mathematical pedagogical content knowledge included knowledge of planning for teaching and learning mathematics, knowledge of the curriculum, and knowledge of mathematics instruction. These variables were from questionnaires that included multiple-choice and constructed-response questions the PSTs responded to (Tatto et al., 2008).

Weights The final future teacher weight variable was used in the descriptive and predictive analysis. This weight is a product of the "institutional weight, session group weight, PSTs' base weight, non-response adjustment factor, and the PSTs' level weight" (Tatto, 2013, p. 136). TEDS-M used a weighted sampling technique in order to better generalize from each country sample to the country's population (Tatto, 2013).

Analysis

Descriptive analysis was used to analyze the extent of the OTL mathematics instruction for conceptual understanding. The analysis provides information about the OTL *intended* by the teacher educators and the OTL *experienced* by the future teachers. The teacher educators and the PSTs were sampled separately; the PSTs are linked to the teacher educators through their common institutional affiliations.

Next, the models of the relationships between the extent of OTL mathematics instruction for conceptual understanding and the knowledge for teaching mathematics are reported at the country level. The models represent the variation of the relationships among the variables by institution within the three countries. A multilevel regression was used to examine the relationships between OTL to teach mathematics for conceptual understanding and PSTs' knowledge for teaching mathematics. This model is suitable because it takes the cluster sampling of the PSTs within institutions into consideration and computes the correct standard errors, and because the software (HLM) allows sampling weights in the analysis.

The model equation is shown below (adapted from Lee & Bryk, 1989). Below is the representation of the future teacher level (within the institution).

$$\begin{aligned}\text{MCK}_{ij} = {} & \beta_{0j} + \beta_{1j}\left(\text{Gender}\right)_{ij} + \beta_{2j}\left(\text{Years of schooling}\right)_{ij} + \beta_{3j}\left(\text{Books}\right)_{ij} \\ & + \beta_{4j}\left(\text{OTL} - \text{Why procedures work}\right)_{ij} + \beta_{5j}\left(\text{OTL} - \text{Make distinctions}\right)_{ij} \\ & + \beta_{6j}\left(\text{OTL} - \text{Multiple strategies}\right)_{ij} + r_{ij}\end{aligned}$$

The equation is for the future teacher i in institution j. MCK and MPCK are analyzed separately. The average MCK or MPCK in institution j is represented by β_{0j}, and the beta coefficients are the slopes of the equation that represent the change in MCK or MPCK for every unit change in the predictors. The model includes covariates, including gender, years of schooling of the parent with the highest level of education, and whether there are over 100 books in the home. The OTL variables include how to explore multiple solution strategies, make distinctions between procedural and conceptual knowledge, and show why procedures work. The OTL variables are group-mean centered. The error of prediction in the equation is represented by r_{ij}.

At the institution level:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\left(\text{OTL} - \text{Why procedures work}\right)_j + \gamma_{02}\left(\text{OTL} - \text{Make distinctions}\right)_j + \gamma_{03}\left(\text{OTL} - \text{Multiple strategies}\right)_j + u_{0j}$$

The coefficients γ_{01} to γ_{03} provide information on the relationships between the given OTL mathematics instruction (as experienced by the PSTs at the institution level) and the institution-level means for the outcome variables MCK and MPCK. The level 2 variables are grand mean centered. In addition, γ_{00} is the value of the grand mean of MCK and MPCK considering all institutions in the country, and u_{0j} is the random error or deviation of the group intercept from the grand mean. These relationships are compared for the three countries below.

Results

The results for the first research question provide information about the role of OTL mathematics instruction for conceptual understanding in the intended and experienced curriculum in the three countries. The differences between the *intended* and *experienced* curriculum can be seen in bar graphs representing PSTs' and teacher educators' responses about the extent of OTL mathematics instruction for conceptual understanding within programs. Figures 13.2 and 13.3 provide summaries of responses for each type of OTL in each of the three countries. The findings indicate the PSTs' perceptions of the curriculum as they experience it are different from the teacher educators' perceptions of the intended curriculum.

Russian and Polish Programs Across the two countries, the patterns of responses indicate that PSTs' perceptions of their experiences were lower for OTL mathematics instruction for conceptual understanding than the teacher educators' perceptions of the activities they provided or intended (see Fig. 10.2). Specifically, about 90% of the teacher educators at the program level in Russia reported having activities for OTL *multiple strategies* often, while only a quarter of the PSTs at the classroom level reported experiencing these opportunities often. About 64% of the teacher educators in the Polish specialist group reported having activities for OTL *multiple strategies* often at the program level, but about 29% of the PSTs at the classroom level reported experiencing these opportunities often. Similarly, about 55% of the teacher educators at the program level in the Polish generalist program reported participating in these activities often, but only about 20% of the PSTs at the classroom level reported experiencing these OTL. These patterns of contrast between teacher educators' perceptions at the program level PSTs' perceptions at the classroom level are also seen for the OTL *how procedures work*, as well as for OTL *make distinctions*. However, the percentage of teacher educators at the program level who reported having these two OTLs often was lower in both countries than the percentage that reported having OTL *multiple strategies*—that is, teacher educators in both

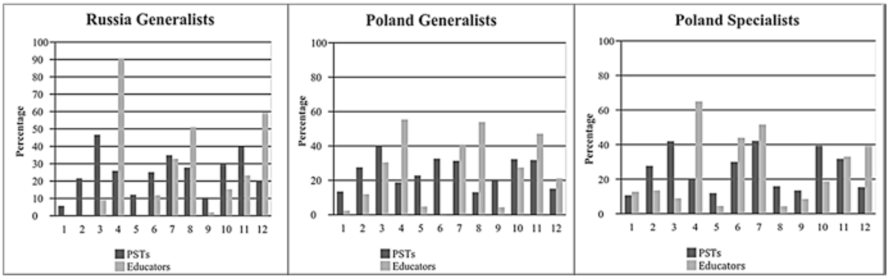


Fig. 13.2 A comparison of PSTs' (N = 1,812; N = 191; N = 2,266 for Polish generalist, Polish specialist, and Russia, respectively) and mathematics teacher educators' perspectives (N = 87; N = 39; N = 49 for Polish generalist, Polish specialist, and Russia, respectively) on the extent of OTL mathematics instruction for conceptual understanding in the Polish and Russian programs

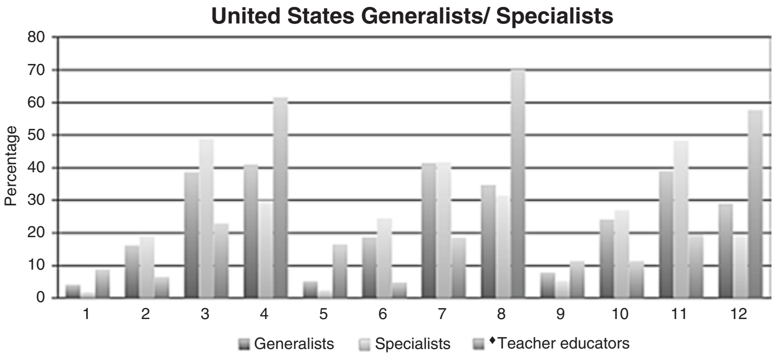


Fig. 13.3 A comparison of the perspectives of generalist (N = 1,310) and specialist (N = 191) PSTs, and mathematics teacher educators (N = 414) in the United States on the extent of OTL mathematics instruction for conceptual understanding

countries reported OTL *multiple strategies* as the most frequent OTL. Furthermore, when compared to the Polish programs, a considerably higher percentage of teacher educators at the program level in Russia reported having OTL *make distinctions*.

Notably, in comparison with Poland, PSTs at the classroom level in the Russian programs had higher percentages of reports that they experienced OTL mathematics instruction for conceptual understanding (across all three OTL types) frequently. Within Poland, PSTs at the classroom level in the Polish generalist programs had lower percentages of OTL mathematics instruction for conceptual understanding. Among the three types of OTL mathematics instruction for conceptual understanding, the PSTs' reports indicate that the OTL *multiple strategies* was experienced often.

The United States The PSTs' reports in the United States programs indicate the OTL mathematics instruction for conceptual understanding was frequently experienced at the classroom level. In particular, a higher percentage of PSTs reported they experienced OTL *multiple strategies* and OTL *making distinctions* often when compared to experiencing OTL *why procedures work* in the specialist programs. In the generalist programs in the United States, the OTL to distinguish between procedural and conceptual understanding was experienced less frequently when compared to other OTL mathematics instruction for conceptual understanding. Taken together, about 60–70% (often and occasionally) of the PSTs in the teacher education programs in the United States had frequent OTL mathematics instruction for conceptual understanding.

Although the mathematics educators in the United States are not a representative sample, the reports of those who participated in the study indicate that they provided these OTLs more frequently than what the PSTs reported having experienced. This pattern of substantial differences between PSTs' reports at the classroom level and teacher educators' reports at the program level differing is consistent across the programs and countries selected in this study.

Relationships Between OTL Mathematics Instruction and Knowledge for Teaching Mathematics

The second analysis examined the relationships between the OTL mathematics instruction and PSTs' content knowledge and pedagogical content knowledge. A multilevel regression analysis was used to answer the second research question: What are the relationships between the OTL mathematics instruction for conceptual understanding and PSTs' knowledge for teaching mathematics in the three countries? This section includes the results for the two-level hierarchical linear model of the relationships and a brief discussion. The results include the unconditional model, models of the relationships of OTL and MCK, and models of the relationships between OTL mathematics instruction for conceptual understanding and MPCK for the three countries.

The differences in PSTs' knowledge for teaching at the country level and the grade specialization level were examined to justify the need for comparing the relationships. The results show that there is a significant difference in PSTs' MCK across the countries, $F(2, 33{,}295) = 579.89$, $p < .001$, and a significant difference across the PSTs' grade specializations, $F(2, 33{,}295) = 486.51$, $p < .001$. Similarly, there is a significant difference in PSTs' MPCK across the countries, $F(2, 33{,}295) = 1{,}740.27$, $p < .001$ and a significant difference across the PSTs' grade specializations, $F(2, 33{,}295) = 2{,}046.45$, $p < .001$. Additionally, post hoc tests confirm the differences are significant between the three countries and across the grade level specializations. The unconditional models, presented in Table 13.2, show the variation that exists within and between the teacher preparation programs in the three countries.

MCK The Polish specialist programs had the highest average MCK score (612.23), with 18.7% of the MCK variation occurring among institutions, and 81.3% of the variation occurring within institutions. In contrast, the average MCK score of the

Table 13.2 Unconditional models for the PSTs' MCK and MPCK in the three countries

	U.S. (Gen)		U.S. (Spec)		Russia		Poland (Gen)		Poland (Spec)	
Variables	MCK	MPCK	MCK	MPCK	MCK	MPCK	MCK	MPCK	MCK	MPCK
Intercept	513.78***	541.07***	524.77***	544.74***	532.82***	510.00***	453.95***	450.40***	612.23***	573.53***
	(4.70)	(3.24)	(7.78)	(7.08)	(9.25)	(7.62)	(3.14)	(4.60)	(9.40)	(7.25)
Variance components										
Intercept u_0	778.84	254.28	291.77	15.24	4,102.63	3,802.17	383.27	938.29	1,601.36	912.93
Level 1 r	4,130.05	4,284.98	3,688.13	5,259.79	4,857.10	4,645.22	4,047.16	7,120.72	6,977.78	4,794.11
ICC	.16	.06	.07	.02	.46	.37	.09	.11	.19	.15

This value in the multi-level output is the variation of the intercept. It is the variance component of the intercept. In this study it is the variation of the average MCK and MPCK score.

The intra class correlation (ICC) is the value that gives the percentage of total variation that is between groups. In this case it gives the percentage of variation between the teacher education institutions (Raudenbush & Bryk, 2002). If the value is high that indicates the variation is mostly composed of the differences that exist between the groups, while a small ICC indicates most of the variation is due to individual differences within groups. MCK ICC at 95%-CI is (.081, .236), while at 99%-CI is (.057, .261) for the United States. MPCK-ICC at 95% -CI is (.018, .094), while at 99%-CI is (.007, .105) for the United States; Russia at 95% CI MCK is (.368, .548) and MPCK is (.273, .627); Poland generalist MCK CI is (.044, .129) and MPCK is (.075, .158) at 95%; Polish specialist MCK CI is (.067, .306) and MPCK is (.029, .291) at 95%

***$p < .001$

PSTs in the Polish *generalist* program was the lowest (453.95), with 8.7% of the MCK variation occurring among institutions and 91.3% of the variation occurring within institutions. These results further confirm there is a difference between the Polish specialist and generalist programs PSTs' MCK. Taken together, the PSTs in Russia had the largest variation in MCK among programs, while the Polish generalist programs had the lowest variation in MCK scores among programs. Considering the confidence intervals (CI) at 95% and 99% across the selected countries, the ICC values suggest conducting multilevel analyses is permissible with the MCK as the outcome variable based on the *duality principle* (Raykov & Marcoulides, 2012) of CI and hypothesis testing. However, ICC values at 95% and 99% CI for the United States specialist program MCK score indicate the multi-level analysis is not permissible. The study, therefore, includes models of the OTL at level 1, but not at level 2 for the United States specialist programs.

MPCK The PSTs in the Polish specialist programs had the highest average MPCK score (573.53), with 15.4% of the proportion of variance among the institutions and 84.6% within the institutions. However, the PSTs in the Polish generalist programs had the lowest average MPCK score (453.95), with 11% of the MPCK variation among the institutions and 86.4% within the institutions. Notably, the PSTs in the Russian programs had the highest variation of MPCK scores among the institutions, whereas the PSTs in the United States generalist programs had the lowest MPCK variation among the institutions. Likewise, the CI of the MPCK ICCs at 95% and 99% are large enough for justifying a multi-level modeling approach for the relational analyses. However, ICC values at 95% and 99% CI for the United States specialist program MPCK score again indicate the multi-level analysis is not permissible. The OTL at level 2 is therefore not included in the analysis for the United States specialist programs.

In the following section, the findings concerning the relationships between the OTL and knowledge for teaching mathematics are discussed within and among the institutions. The selected items representing the OTL mathematics instruction for conceptual understanding are, as before, showing why procedures work, making distinctions between procedural and conceptual knowledge, and exploring multiple solution strategies. Tables 13.3, 13.4, 13.5, and 13.6 present summaries of the model relationships of OTL and PSTs' knowledge for teaching mathematics within and among the institutions. The three variables used are highly correlated and, therefore, each item is introduced in a separate model to show the unique relationships with the PSTs' knowledge for teaching mathematics.

OTL Mathematics Instruction for Conceptual Understanding and PSTs' MCK The relationships between the OTL mathematics instruction and PSTs' MCK had differing patterns within the institutions in the three countries. There were a number of significant positive relationships between OTL and PSTs' MCK within programs, as shown in Tables 13.3 and 13.4. The OTL *why procedures work* significantly predicted higher PSTs' MCK within the Russian program ($\beta = 5.97$, $p < .05$), the United States specialist program ($\beta = 17.79$, $p < .05$), and the Polish generalist

Table 13.3 Multi-level models of relationships between OTL mathematics instruction and MCK (United States and Russia)

	U.S. generalists			U.S. specialists			Russia		
Variables	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept	518.39***	520.40***	518.87***	524.57***	524.44***	524.43***	528.98***	525.54***	528.98***
	(3.77)	(4.20)	(4.23)	(7.33)	(7.36)	(7.36)	(10.11)	(9.68)	(10.11)
Level 1									
Gender	−23.43**	−22.47**	−22.76**	−3.66	2.77	2.55	−6.23**	−5.76**	−6.23**
	(7.71)	(7.87)	(7.68)	(16.17)	(18.44)	(18.17)	(2.60)	(2.64)	(2.60)
Parental years of schooling	3.02**	3.00**	2.92**	−2.34	−1.18	−1.25	0.42	0.46	0.42
	(1.29)	(1.27)	(1.28)	(3.05)	(3.31)	(3.22)	(0.27)	(0.27)	(0.27)
More than 100 books in the home	1.04	−0.17	0.91	41.81***	41.85**	42.04**	2.99	3.11**	2.99**
	(5.66)	(5.81)	(5.64)	(12.67)	(14.16)	(13.89)	(1.36)	(1.38)	(1.36)
Why procedures work	1.94			17.79**			5.97**		
	(4.32)			(6.00)			(2.66)		
Make distinctions between procedural and conceptual knowledge		−3.67			3.43			5.34***	
		(3.01)			(11.65)			(1.49)	
Multiple solution strategies			−1.58			0.93			5.97**
			(2.86)			(8.10)			(2.66)
Level 2									
Why procedures work	11.98**						6.90		
	(5.05)						(10.45)		
Make distinctions between procedural and conceptual knowledge		2.40						25.09	
		(4.34)						(15.45)	
Multiple solution strategies			6.30						6.90
			(4.42)						(10.45)
Variance components									
Level-2 between group *SD* ($\boldsymbol{u}_0$)	24.89	27.31	26.67	18.67	18.08	18.05	58.98	55.24	58.98
Level-1 within group *SD* ($\boldsymbol{r}$)	62.94	62.73	63.01	56.20	57.58	57.63	69.01	68.76	69.01

$^{**}p < .05$, $^{***}p < .001$: The models for the specialist programs in the United States do not have level 2 OTL because the ICC values do not allow for a multilevel analysis

Table 13.4 Multi-level models of relationships between OTL mathematics instruction and MCK (Poland)

	Poland generalists			Poland specialists		
Variables	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept	453.15***	452.81***	452.65***	614.84***	614.98***	616.54***
	(6.14)	(6.35)	(6.06)	(10.14)	(10.25)	(9.78)
Level 1						
Gender	−17.29	−16.63	−17.77	−10.41	−6.52	−9.58
	(10.84)	(10.51)	(11.00)	(10.86)	(11.19)	(11.10)
Parental years of schooling	0.01	−0.005	−0.02	0.91	0.86	0.87
	(0.30)	(0.29)	(0.30)	(0.57)	(0.55)	(0.53)
More than 100 books in the home	12.72***	14.02***	13.16***	20.90	21.98	20.85**
	(2.64)	(2.58)	(2.71)	(13.52)	(13.47)	(13.61)
Why procedures work	6.89**			5.47		
	(3.09)			(5.64)		
Make distinctions between procedural and conceptual knowledge		4.86			.77	
		(2.82)			(4.23)	
Multiple solution strategies			6.74**			4.60
			(3.10)			(5.58)
Level 2						
Why procedures work	−5.64			−4.69		
	(5.58)			(11.48)		
Make distinctions between procedural and conceptual knowledge		−1.21			−6.68	
		(2.17)			(10.20)	
Multiple solution strategies			−2.50			−15.13
			(6.99)			(13.25)
Variance components						
Level-2 between group *SD* (u_0)	33.76	33.60	33.81	48.79	48.36	46.74
Level-1 within group *SD* (*r*)	71.03	71.39	71.08	77.70	78.05	77.81

**$p < .05$, ***$p < .001$

program (β = 6.89, $p < .05$). PSTs' MCK also had a significant positive relationship with OTL *make distinctions* within the teacher preparation programs in Russia (β = 5.34, $p < .05$). Additionally, there was a significant positive relationship between OTL *multiple strategies* and PSTs' MCK within the Russian and Polish generalist programs.

Similarly, the program-level relationship patterns of OTL and PSTs' MCK differed among the institutions in the three countries. In the United States, the more frequently the programs provided OTL to show why a procedure works, the higher the average PSTs' MCK (γ_{01} = 11.98, $p < .05$). The other two OTL relationships in programs in the United States were positive (ns), as were all relationships in the Russian programs were positive. In the Polish programs, in contrast, the relationships among these variables were negative (ns).

Table 13.5 Multi-level models of relationships between OTL mathematics instruction and MPCK (United States and Russia)

	U.S. generalists			U.S. specialists			Russia		
Variables	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept	543.29***	545.04***	543.29***	546.03***	545.87***	545.86***	508.45***	504.42***	508.38***
	(3.25)	(3.21)	(3.25)	(6.54)	(6.52)	(6.53)	(7.21)	(6.92)	(7.31)
Level 1									
Gender	−6.01	−5.77	−6.01	6.63	2.53	4.63	−5.18	−4.02	−4.36
	(5.96)	(5.93)	(5.96)	(24.04)	(26.54)	(21.89)	(2.97)	(2.29)	(2.34)
Parental years of schooling	1.90**	2.09**	1.90**	1.02	0.04	0.53	0.27	0.32	0.29
	(0.90)	(0.87)	(0.90)	(1.54)	(1.83)	(0.94)	(0.27)	(0.25)	(0.26)
More than 100 books in the home	13.97**	12.23**	13.97**	53.09**	54.35**	53.81**	2.04	2.31	2.22
	(5.63)	(5.29)	(5.63)	(19.53)	(18.37)	(20.91)	(1.55)	(1.52)	(1.48)
Why procedures work	−0.79			18.50**			3.87**		
	(3.53)			(7.91)			(1.91)		
Make distinctions between procedural and conceptual knowledge		−1.87			11.14			3.41**	
		(3.53)			(12.71)			(1.54)	
Multiple solution strategies			−0.79			8.82			4.51
			(3.53)			(15.18)			(2.82)
Level 2									
Why procedures work	5.16						3.71		
	(3.50)						(3.73)		

(continued)

Table 13.5 (continued)

	U.S. generalists			U.S. specialists			Russia		
Variables	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Make distinctions between procedural and conceptual knowledge		1.28						27.33**	
		(3.52)						(8.71)	
Multiple solution strategies			5.16						5.97
			(3.50)						(7.45)
Variance components									
Level-2 between group *SD* (u_0)	14.21	15.22	14.21	5.84	4.52	4.26	47.38	41.60	47.44
Level-1 within group *SD* (r)	66.39	65.28	66.38	67.28	68.28	68.45	68.41	68.42	68.44

$^{**}p < .05$, $^{***}p < .001$: The models for the specialist programs in the United States do not have level 2 OTL because the ICC values do not allow for a multilevel analysis

Table 13.6 Multi-level models of relationships between OTL mathematics instruction and MPCK (Poland)

	Poland generalists			Poland specialists		
Variables	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Intercept	449.90***	449.55***	448.77***	573.95***	573.34***	574.96***
	(7.67)	(8.42)	(7.97)	(9.34)	(8.90)	(8.86)
Level 1						
Gender	−8.77	−8.71	−9.14	−0.19	2.12	3.16
	(5.39)	(5.12)	(5.53)	(7.16)	(7.45)	(6.88)
Parental years of schooling	0.04	0.06	0.02	0.64	0.61	0.61
	(0.35)	(0.34)	(0.35)	(0.32)	(0.31)	(0.31)
More than 100 books in the home	11.59***	11.14***	12.05***	5.47	6.18	6.19
	(2.88)	(2.82)	(2.92)	(9.18)	(9.26)	(9.14)
Why procedures work	5.87**			3.03		
	(3.00)			(2.96)		
Make distinctions between procedural and conceptual knowledge		7.31***			.12	
		(1.99)			(2.68)	
Multiple solution strategies			5.53			−.91
			(3.08)			(2.07)
Level 2						
Why procedures work	−11.59			−4.79		
	(7.01)			(8.87		
Make distinctions between procedural and conceptual knowledge		−1.35			6.70	
		(3.03)			(8.73)	
Multiple solution strategies			−7.64			−10.42
			(8.36)			(11.13)
Variance components						
Level-2 between group *SD* (u_0)	43.51	44.01	43.96	42.20	41.78	41.27
Level-1 within group *SD* (r)	88.21	87.69	88.29	64.90	65.04	65.01

$^{**}p < .05$, $^{***}p < .001$

The relationships between the OTL *make distinctions* and PSTs' MCK differed among institutions in the generalist programs in the three countries. Specifically, the higher the frequency of this OTL, the higher the average PSTs' MCK across the United States and Russian generalist programs (ns). In contrast, the relationships in the Polish programs suggest the higher the frequency of this OTL, the lower the average PSTs' MCK among the programs. The patterns of relationships between OTL how to make distinctions between procedural and conceptual knowledge and PSTs' MCK suggest that the more the programs have these opportunities in Polish teacher preparation programs, the lower the average PSTs' MCK.

Likewise, the patterns of the relationships between the OTL how to explore multiple solution strategies and PSTs' average MCK differed in the three countries among the institutions. In particular, the relationships between this OTL and PSTs' average MCK were positive in the United States and Russia (ns), but negative in the Polish programs (ns). The results indicate the frequency of having this OTL influences PSTs' MCK differently across the generalist programs.

OTL Mathematics Instruction for Conceptual Understanding and PSTs' MPCK Within the programs, the relationships between the OTL and PSTs' MPCK had different patterns across the selected contexts, as shown in Tables 13.5 and 13.6. Specifically, the more PSTs had the OTL how to show why procedures work, the higher their MPCK within the United States specialist program ($\beta = 18.50$, $p < .05$), the Russian program ($\beta = 3.87$, $p < .05$), and the Polish generalist program ($\beta = 5.87$, $p < .05$). Additionally, there was a significant positive relationship between the OTL how to make distinctions between procedural and conceptual knowledge and PSTs' MPCK in the Russian program ($\beta = 3.41$, $p < .05$) and the Polish generalists program ($\beta = 7.31$, $p < .05$). On the other hand, the relationships of OTL (a) how to make distinctions between procedural and conceptual knowledge and (b) exploring multiple solution strategies with PSTs' MPCK within the United States generalist programs were negative (ns).

The relationship patterns of OTL and the average PSTs' MPCK among the institutions were different across the three countries. Notably, the relationships between the OTL mathematics instructions for conceptual understanding and the average PSTs' MPCK were similar in the United States and Russia, but different in the Polish teacher preparation programs. Specifically, the OTL why procedures work was positively related to PSTs' average MPCK among the programs in the United States and Russia, but negative in the Polish programs (ns).

Additionally, there was a significant positive relationship between OTL how to make distinctions between procedural and conceptual knowledge and PSTs' average MPCK across the programs in Russia ($\gamma_{02} = 27.33$, $p < .05$). In contrast, the Polish programs had opposing patterns of relationships between this variable and average PSTs' MPCK (ns). Specifically, the relationships were positive for the Polish generalist programs, but negative for the Polish specialist programs. However, in the United States, the relationships between the OTL how to make distinctions between procedural and conceptual knowledge and PSTs' MPCK was positive, but not significant.

Finally, the relationships between the OTL how to explore multiple solution strategies and average PSTs' MPCK between the programs also differed across the three countries. Notably, the relationships between this OTL and the average PSTs' MPCK across the programs in the United States and Russia were positive (ns), whereas the relationships between this OTL and the average PSTs' MPCK were negative for the Polish programs. A summary of the relationships between the OTL mathematics instruction for conceptual understanding and PSTs' MPCK within and among the institutions is presented in Tables 13.5 and 13.6.

Discussion

The purpose of the larger TEDS-M study was to investigate how primary and lower secondary teachers are prepared to teach mathematics and assess the influence of the learning opportunities on future teachers' knowledge development. This study focused on opportunities that primary teachers have to learn mathematics instruction for conceptual understanding using three variables: learning to show why a procedure works, learning how to explore multiple solution strategies, and learning how to distinguish between procedural and conceptual knowledge. First, this study examined the differences between the intended curriculum from the teacher educators' perspectives and the experienced curriculum from the PSTs' perspectives. Further, this study investigated the relationships between the experienced OTL mathematics instruction for conceptual understanding and primary PSTs' knowledge of teaching mathematics.

There were marked differences between the frequency of the OTL mathematics instruction for conceptual understanding as reported by PSTs with respect the experienced curriculum and by teacher educators with respect to the intended curriculum. The PSTs' reported experience of the OTL was less frequent than what teacher educators reported at the program level in all programs across the three selected countries. Hsieh et al. (2011) found that teacher educators rated themselves higher than PSTs on the extent of the OTL. These researchers suggested that the relatively high estimates of OTL among teacher educators could be due to their confidence in the goals set for courses and the intended curriculum.

Comparing the teacher educators' and PSTs' views of OTL mathematics instruction for conceptual understanding across the three countries reveals some notable differences. Specifically, the Polish programs had the lowest percentage of PSTs and teacher educators who reported they experienced or provided the OTL mathematics instruction for conceptual understanding occasionally or often. Programs in the United States reported the highest frequency of OTL mathematics instruction for conceptual understanding among the three countries, with over 70% of PSTs and teacher educators indicating that they experienced or intended to offer each of the OTLs occasionally or often, respectively. Hsieh et al. (2011) found that the PSTs in programs in the United States reported that their instructors were effective and that the teaching was coherent. Although this finding did not focus on variables similar to the ones in this study, it parallels this study's findings in that most of the PSTs' perceptions of their experiences suggest they experienced what the programs or teacher educators intended, although to a lesser extent.

Does the OTL Mathematics Instruction for Conceptual Understanding Matter?

Some studies have shown that future teachers and in-service primary teachers have knowledge gaps in their mathematics content and pedagogical content knowledge (e.g., Ball, 1990; Ma, 1999). The knowledge teachers have for teaching influences their instruction and the lack of it can limit how they teach mathematics for understanding (Borko & Putnam, 1996). For this reason, providing opportunities for learning how to teach mathematics for conceptual understanding in teacher preparation programs and professional development forums can not only influence teachers knowledge and beliefs about teaching, but also support their teaching approaches in mathematics classrooms. Not only do these opportunities build on future teachers' mathematics content knowledge; they also introduce them to new ways of teaching mathematics that they may not have experienced in their earlier mathematics learning.

The findings from this study suggest frequent experiences in learning mathematics for conceptual understanding through particular strategies is related to an increase in PSTs' knowledge about teaching mathematics. In particular, more opportunities that allow PSTs to learn to show why a procedure works was significantly related to the average PSTs' increase in content knowledge among the teacher preparation programs in the United States. Similarly, within the Polish generalist, United States specialist, and Russian teacher preparation programs, the relationships between the OTL to show why procedures work and PSTs' knowledge for teaching were significant. If PSTs have opportunities to develop skills for providing explanations, the PSTs present appropriate explanations for given tasks (Charalambous et al., 2011). Similarly, PSTs with more opportunities to use representations were more successful in explaining mathematics concepts that involved place value (Chick, 2003). Indeed, the reasoning and proof process standards require students to explain why procedures work (NCTM, 2000). Also, the recent Common Core practice standards and principles to actions emphasize problem solving, reasoning and engaging in mathematical argumentation in learning mathematics across multiple levels in school (CCSS, 2010; NCTM, 2014). Students can only attain these expected competencies if their teachers have the knowledge and strategies to support students engaging in reasoning.

Furthermore, the frequency of OTL how to make distinctions between procedural and conceptual understanding was significantly related to an increase in average PSTs' MPCK in the Russian teacher preparation programs. Although the relationship patterns for the United States programs and Russia were all positive, they were not significant at the program level, but were significant within the Russian and the Polish generalist teacher preparation programs. The findings further support the previous work done by Bartell et al. (2012), in which they noted growth in PSTs' MPCK when the instructors provided opportunities for engaging in differentiating procedural and conceptual knowledge. Similarly, Crespo (2000) vfound that if PSTs have supportive contexts, they could identify conceptual understanding from students' responses.

All of these studies in mathematics teacher education also support the notion that providing opportunities that allow for PSTs to differentiate procedural and conceptual understanding of mathematics expands their knowledge of instruction. Hence, discussions on specific topics and opportunities that support PSTs to distinguish this procedural knowledge from conceptual knowledge is needed among the professional communities of practice in teacher preparation programs. This is an essential specialized knowledge that supports teaching for conceptual understanding of mathematics.

Although the OTL how to explore multiple solution strategies is part of the problem-solving process in teaching mathematics, at the program level it was not significantly related to the outcome variables of interest. However, within the programs in Russia and the Polish generalist programs, the relationships between this OTL and PSTs' MCK were significant. As shown in Figs. 13.2 and 13.3, across the countries, it was the form of mathematics instruction that PSTs reported learning about most frequently when compared to the other two. Less variation could explain the non-significant results among the teacher preparation programs. However, the patterns of the relationships in the United States and Russia show support for the previous findings of smaller-scale studies (e.g., Crespo, 2000; Grant & Lo, 2009; Ryken, 2009).

In the Polish programs, most of the OTL mathematics instruction for conceptual understanding examined in this study had negative correlations with the PSTs' average MCK and MPCK among the institutions. In contrast, for PSTs within the generalist programs, there were some significant relationships. Specifically, the OTL why procedures work (MCK and MPCK), multiple solution strategies (MCK), and the OTL how to make distinctions between conceptual and procedural knowledge (MPCK) had significant positive relationships. Although the negative relationships among the programs were not significant, some of the findings suggest there may be more emphasis on a particular form of knowledge within the programs by some instructors. Notably, the opportunities to distinguish between procedural and conceptual knowledge were perhaps experienced minimally, as presented in the descriptive findings. The generalist programs had the lowest scores for the knowledge for teaching mathematics scores, whereas the specialist programs had the highest scores.

An examination of the teacher educators' responses in the Polish programs indicates that the percentage of those who intended to provide OTL mathematics instruction for conceptual understanding more frequently were lower than for the teacher educators in the programs in Russia and the United States. Perhaps the documented significant improvements in the Polish students' mathematics test scores could be a result of the restructuring of the schools (OECD, 2011) rather than the teacher preparation programs. Although the findings showed mixed patterns across contexts, they can be generalizable to a wider population within the countries, and theories can be formulated to explain the relationships between approaches to learning mathematics instruction for conceptual understanding and future teachers' knowledge for teaching mathematics.

Limitations of the Study

The analysis provided in the study involved some limitations. First, the study researchers were not able to find comparable measures of previous knowledge that were common across countries. For this reason, future teachers' prior knowledge is not known; the study used as a proxy previous attainment as reported by future teachers. Second, the sampled teacher educators in the United States did not reach the minimum threshold for participation rate; therefore, the study is limited to information provided by the teacher educators who participated in the study. Third, the structure of teacher education programs is different from that of schools, for that reason PST's are not nested within classrooms with a single teacher educator as teacher educators teach across the program, thus future teachers are nested within programs. Both PSTs and educators are considered representatives of their institutions—the commonality is that they come from the same institution. Although there were some limitations in the study, the rigorous data-collection processes and the analytical procedures conducted in this study enable significant findings that are relevant to teacher education policy.

Significance of the Study

The use of a similar analytical process across the selected countries to examine the relationships between the mathematical practices for conceptual understanding and the PSTs' knowledge is useful for the generalization of findings and confirmation of theories. From this study, it is safe to say that some of the results from previous, small-scale studies on opportunities to learn mathematics instruction for conceptual understanding through teacher preparation were supported by this study on a larger scale. The findings can inform teacher preparation programs on the opportunities that can be provided to expand PSTs' knowledge for teaching mathematics for conceptual understanding.

The evaluation of the knowledge for teaching mathematics that future teachers have at the end of their teacher preparation programs informs teacher education and policy-makers about the opportunities that teacher preparation programs need to offer their future teachers to learn to teach primary school mathematics. This information is useful for professional development forums on the offerings that are needed to support in-service teachers to continue building on problem-solving strategies. Finally, countries can evaluate where there is a lack of emphasis in their programs and investigate ways of covering the knowledge deficiencies through the adoption of appropriate curriculum guides.

References

Ball, D., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special. *Journal of Teacher Education, 59*(5), 389–407.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The Elementary School Journal, 90*(4), 449–466.

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *The Elementary School Journal, 93*(4), 373–397.

Bartell, T., Webel, C., Bowen, B., & Dyson, N. (2012). Prospective teacher learning: Recognizing evidence of conceptual understanding. *Journal of Mathematics Teacher Education, 16*(1), 57–79. https://doi.org/10.1007/s10857-012-9205-4

Blömeke, S., & Kaiser, G. (2014). Homogeneity or heterogeneity? Profiles of opportunities to learn in primary teacher education and their relationship to cultural context and outcomes. In S. Blömeke, F.-J. Hsieh, G. Kaiser, & W. H. Schmidt (Eds.), *International perspectives on teacher knowledge, beliefs and opportunities to learn* (pp. 299–325). Dordrecht, The Netherlands: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-94-007-6437-8_14

Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning* (Revised and expanded edition). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Boaler, J. (2008). Promoting “relational equity” through a mathematics approach focused upon social justice. *British Educational Research Journal, 34*(2), 167–194.

Borko, H., & Putnam, R. T. (1996). Learning to teach. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 673–708). New York, NY: Macmillan.

Brese, F., & Tatto, M. (2012). *User guide for the TEDS-M international database*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Carnoy, M. (2015). *International test score comparison and educational policy: A review of the critiques.* National Education Policy Center. Retrieved from http://nepc.colorado.edu/publication/international-test-scores

Carroll, J. B. (1963). A model of school learning. *Teachers College Record, 64*, 723–733.

Carpenter, T. (1986). Conceptual knowledge as a foundation for procedural knowledge. In *Conceptual and procedural knowledge : The case of mathematics* (pp. 113–132). Hillsdale, NJ: Lawrence Erlbaum.

Charalambous, C., Hill, H., & Ball, D. (2011). Prospective teachers’ learning to provide instructional explanations: How does it look and what might it take? *Journal of Mathematics Teacher Education, 14*(6), 441–463. https://doi.org/10.1007/s10857-011-9182-z

Chick, H. L. (2003). Pre-service teachers’ explanations of two mathematical concepts. *Proceedings of the 2003 annual conference of the Australian association for research in education*, Auckland, New Zealand. Retrieved from http://www.aare.edu.au/03pap/chi03413.pdf

Chinnappan, M., & Forrester, T. (2014). Generating procedural and conceptual knowledge of fractions by pre-service teachers. *Mathematics Education Research Journal, 26*(4), 871–896.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics.* Retrieved from http://www.corestandards.org/

Conference Board of the Mathematical Sciences. (2000). *The mathematical education of teachers* (Vol. II). Providence, RI/Washington, DC: American Mathematical Society and Mathematical Association of America.

Conference Board of the Mathematical Sciences. (2012). *The mathematical education of teachers II* (MET-II). Providence, RI: American Mathematical Society and Mathematical Association of America.

Cowan, C. D., Hauser, R. M., Kominski, R. A., Levin, H. M., Lucas, S. R., Morgan, S. L., … Chapman, C. (2012). *Improving the measurement of socioeconomic status for the national assessment of educational progress: A theoretical foundation—Recommendations to the National Center for Education Statistics*. Washington, DC: National Center for Education Statistics.

Crespo, S. (2000). Seeing more than right and wrong answers: Prospective teachers' interpretations of students' mathematical work. *Journal of Mathematics Teacher Education, 3*(2), 155–181. https://doi.org/10.1023/A:1009999016764

Dewey, J. (1902). *The child and the curriculum*. Chicago, IL: The University of Chicago Press.

Fan, L., & Cheong, C. (2002). Investigating the sources of Singaporean mathematics teachers' pedagogical knowledge. In D. Edge & B. H. Yap (Eds.), *Mathematics education for a knowledge-based era. Proceedings of second East Asia regional conference on mathematics education & ninenth Southeast Asian conference on mathematics. Mathematics education for a knowledge-based era* (pp. 224–231). Derby, GB: Association of Mathematics Educators.

Floden, R. E. (2002). The measurement of opportunity to learn. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 231–266). Washington, DC: National Academy Press.

Grant, T. J., & Lo, J.-J. (2009). Reflecting on the process of task adaptation and extension: The case of computational starters. In B. Clarke, B. Grevholm, & R. Millman (Eds.), *Tasks in primary mathematics teacher education* (Vol. 4, pp. 25–36). Boston, MA: Springer US Retrieved from http://www.springerlink.com.proxy1.cl.msu.edu/content/r755848586375251/abstract/

Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York, NY: Teachers College Press.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Hsieh, F.-J., Law, C.-K., Shy, H.-Y., Wang, T.-Y., Hsieh, C.-J., & Wang, S.-J. (2011). Mathematics teacher education quality in TEDS-M: Globalizing the views of future teachers and teacher educators. *Journal of Teacher Education, 62*(2), 172–187. https://doi.org/10.1177/0022487110390819

Husén, T. (1967). International study of achievement in mathematics: A comparison of twelve countries. *International project for the evaluation of educational achievement, IEA* (Vols 1–2). New York, NY: Wiley.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education, 62*(3), 172–192.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics* (6th ed.). Reston, VA: Author.

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: Author.

Organization for Economic Cooperation and Development. (2011). *The impact of the 1999 education reform in Poland* (OECD Education Working Papers, No. 49). OECD Publishing. https://doi.org/10.1787/5kmbjgkm1m9x-en.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Raykov, T., & Marcoulides, G. A. (2012). *An introduction to applied multivariate analysis*. New York, NY: Routledge.

Ryken, A. (2009). Multiple representations as sites for teacher reflection about mathematics learning. *Journal of Mathematics Teacher Education, 12*(5), 347–364. https://doi.org/10.1007/s10857-009-9107-2

Schmidt, W. H., Burroughs, N., & Cogan, L. (2013). *World class standards for preparing teachers of mathematics*. (Working paper No. 37). East Lansing, MI: Education Policy Center, Michigan State University. Retrieved from http://education.msu.edu/csc/pdf/World-Class-Standards-for-Preparing-Teachers-of-Mathematics.pdf

Schmidt, W. H., McKnight, C. C., Wang, H. C., Cogan, L. S., Houang, R. T., Wiley, D. E., & Wolfe, R. G. (2001). *Why schools matter: A cross-national comparison of curriculum and learning*. San Francisco, CA: Jossey-Bass.

Schwab, J. J. (1978). *Science, curriculum, and liberal education: Selected essays*. Chicago, IL: University of Chicago Press.

Senk, S. L., Tatto, M. T., Reckase, M., Rowley, G., Peck, R., & Bankov, K. (2012). Knowledge of future primary teachers for teaching mathematics: An international comparative study. *ZDM, 44*(3), 307–324. https://doi.org/10.1007/s11858-012-0400-7

Shulman, L. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 3–36). New York, NY: Macmillan.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Tatto, M. T. (2013). *To teach primary and secondary mathematics in 17 countries. Technical report*. Amsterdam, The Netherlands: IEA Retrieved from http://www.iea.nl/teds-m.html

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Conceptual framework*. Amsterdam, The Netherlands: IEA Retrieved from http://www.iea.nl/teds-m.html

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA Teacher education and development study in mathematics (TEDS-M)*. Amsterdam, The Netherlands: IEA Retrieved from http://www.iea.nl/teds-m.html

Törnroos, J. (2005). Mathematics textbooks, opportunity to learn and student achievement. *Studies In Educational Evaluation, 31*(4), 315–327. https://doi.org/10.1016/j.stueduc.2005.11.005

Wong, K. Y., Boey, K. L., Lim-Teo, S. K., & Dindyal, J. (2012). The preparation of primary mathematics teachers in Singapore: Programs and outcomes from the TEDS-M study. *ZDM, 44*(3), 293–306. https://doi.org/10.1007/s11858-011-0370-1

Chapter 14
The Mathematical Education of Secondary Teachers

Maria Teresa Tatto

Abstract This chapter explores the influence of pre-service teacher education on future secondary teachers' mathematical knowledge for teaching across several of the countries/regions that participated in the Teacher Education and Development Study in Mathematics (TEDS-M) including Chile, Chinese Taipei, Germany, Malaysia, the Philippines, Poland, the Russian Federation, Singapore, Switzerland, and Thailand, and paying particular attention to the situation in the United States of America. This chapter uses survey and knowledge assessment data collected by TEDS-M from representative samples of teacher education programs and their future secondary teachers across these countries/regions. Multilevel analyses show wide variability in the knowledge for teaching mathematics future secondary teachers attain. Previous mathematics knowledge as a requirement for entry into teacher education and mathematics-rich opportunities to learn were associated with higher and deeper levels of mathematical and mathematical pedagogical knowledge, after controlling for individual characteristics. Beliefs espousing traditional orientations to learning mathematics were associated with lower levels of performance in the knowledge assessments. The discussion highlights the importance of self-study and self-regulation in teacher education.

Introduction

Since as far back as 1997, the quality of teachers has been a central policy concern in the United States. That year, President Clinton issued a call to action prioritizing "improving the quality of teachers in every American classroom" (Lewis, Basmat, Carey, Bartfai, Farris, & Smerdon, 1999, p. iii). The National Center for Education Statistics

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

M. T. Tatto (✉)
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: teresa.tatto@asu.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_14

undertook the task of producing a report on the preparation and qualifications of public school teachers as a follow-up to this call (*Teacher Quality: A Report on the Preparation and Qualifications of Public School Teachers*) and argued that while teacher quality

> …is a complex phenomenon, and there is little consensus on what it is or how to measure it […there] are, however, two broad elements that most observers agree characterize teacher quality: (1) teacher preparation, and qualifications (e.g., pre-service and continued learning), and (2) teaching practices (e.g., actual behaviors and practices that teachers exhibit in their classrooms). (Lewis et al., 1999, p. iii)

While the No Child Left Behind Act (NCLB) of 2001 identified providing highly qualified teachers to schools as one of its key goals and resulted in the development of teacher-evaluation systems based on students' test scores in every state, under the new law known as the Every Student Succeeds Act (ESSA) such a requirement is no longer in place. Instead, states must come up with ways to define quality in order to evaluate teachers. Under both NCLB and ESSA, teacher educators and teacher quality advocates have raised concerns about the lack of standards for teacher quality especially as little agreement exists regarding what constitutes quality teacher preparation and how it relates to quality teaching (Sawchuk, 2016; Zeichner & Conklin, 2005). Recent reviews of teacher education reveal the need for more systematic exploration of programs and their intended outcomes (Levine, 2006; National Council for Teacher Quality, 2013), and for rigorous research directed at producing system-level evidence of program effects (Darling-Hammond, 2013). Indeed, national-, state-, or even program-level evaluations of teacher education program effects are rare (Feuer, Floden, Chudowsky, & Ahn, 2013; Heafner, McIntyre, & Spooner, 2014). When they have been undertaken, evaluations have not shed much light on the acquisition of knowledge needed for teaching because they have not measured future teachers' knowledge outcomes and have, for the most part, relied on responses to satisfaction surveys (Thomas & Loadman, 2001). Attempts at applying value-added models to the evaluation of teacher education programs by measuring teachers' effects on their pupils' achievement as an indicator of teachers' program effectiveness (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Constantine et al., 2009; Goldhaber, Liddle, & Theobald, 2013) have lacked validity (Floden, 2012). Many studies have been descriptive, small in scale, not specific to subject-matter, and rarely comparative (Crowe, 2010; Tatto, 2011). The call by the U.S. Department of Education (DOE, 2014) to create accountability systems in teacher preparation had the potential to answer some of these questions; however, this proposal met with strong resistance and was rescinded in early 2017. Such criticism is understandable, because large-scale systemic efforts to evaluate teacher preparation are expensive and time-consuming, and the field lacks efficient models to guide such high-stakes measurement efforts.

One exception is the field of mathematics education. In response to calls from groups such as the National Commission on Mathematics and Science Teaching for the twenty-first century (2000) and the RAND Mathematics Study Panel (2003), the National Science Foundation funded a number of studies including, in 2005, the first international and comparative large-scale study of the outcomes of teacher education. The study, known as the Teacher Education and Development

Study in Mathematics (TEDS-M), explored the influence of teacher education on the knowledge that is considered important for future mathematics teachers to master at the end of their pre-service teacher education programs. The study is distinctive in that it was a collaborative effort of mathematics teacher educators to evaluate the knowledge outcomes of their own teacher education programs. The TEDS-M study included a comparative element bringing together the effort of teacher educators to study their own preparation systems for future secondary teachers in 14 countries including the U.S. The study was authoritative because it endeavored to use representative samples of teacher education programs within each country/region and assessed the knowledge that their graduates attained at the end of their programs.

Research Questions and Assumptions

While international studies of student achievement have helped to provide rigorous and comprehensive evaluations of schooling success within and across countries for several decades, the education of teachers has rarely been studied in such a manner. Thus, while much is known about the factors that seem to make a difference in improving students' learning, including the quality of their teachers, we lack similar valid and generalizable outcome data for teacher education programs concerning the factors that make a difference on future teachers' knowledge acquisition. Such data would help inform programs' efforts to improve future teachers' learning and the overall quality of teacher education.

This chapter uses the TEDS-M database to investigate the outcomes of secondary mathematics teacher education as indicated by future teachers' knowledge at the end of their programs, using samples of pre-service programs and their future secondary mathematics teachers across a variety of countries/regions. The research questions are

1. What is the level of the mathematical and mathematical pedagogical content knowledge future teachers attain?
2. How are specific characteristics of future teachers (such as socioeconomic status, age, gender, prior attainment, and beliefs) associated with their attained levels of knowledge?
3. What are some of the key learning opportunities available to future teachers in their teacher education programs, and how are these associated with their attained levels of knowledge?

This chapter's central assumption is that mathematics teacher education is designed to enable future teachers to be ready to teach once they finish their teacher education. In order to do this, programs plan opportunities to learn (OTL) according to the individuals they recruit into teaching and according to the curricular demands of the school systems where their teachers will be hired. The degree to which programs succeed in preparing knowledgeable teachers may depend on the degree to which they are able to recruit individuals with adequate content knowledge into teaching and to regulate the quality of their programs.

In the past, studies have included indicators of program effectiveness using proxies such as teachers' certification status, or performance on state certification tests (Clotfelter, Ladd, & Vigdor, 2007; Croninger, Rice, Rathbun, & Nishio 2007; Goldhaber & Brewer 2000). However, these studies do not tell us what teachers know, how they are able to use this knowledge in teaching situations, and how this knowledge is in turn, related to programs' practices and OTL. To investigate these questions and assumptions, this chapter reports on the findings from Chile, Chinese Taipei, Germany, Malaysia, the Philippines, Poland, the Russian Federation, Singapore, Switzerland (German-speaking cantons only), Thailand, and the United States (public institutions only).[1]

The Education of Future Secondary Teachers

The education of secondary mathematics teachers (i.e., individuals who may teach grades ranging from the upper grades of primary[2] schooling to the upper grades of secondary schooling or children aged 11–18) has become an area of concern in the U.S. and elsewhere, in part as a result of findings from international assessments such as the TIMSS studies.

The latest TIMSS 2015[3] results show that while the performance level of U.S. eighth-graders in mathematics improved from 1995 to 2011 to 2015 (with mean scores of 492, 509, and 518 respectively), pupils in eight education systems scored significantly higher in 2015 than those in the U.S. system. These systems include

[1] TEDS-M was funded by the US National Science Foundation NSF REC 0514431. TEDS-M collected data from 14 countries who volunteered to participate in the study of secondary teacher education. Three countries were excluded from the analysis for this chapter because the data collected did not meet the coverage requirement to ascertain representativeness or because as it is often the case with surveys, there was a significant number of missing values in the variables included in the analysis. In addition, while the TEDS-M study endeavored to draw nationally representative samples of teacher education programs in each participating country, the National Research Center (NRC) in the United States decided to collect nationally representative data from public institutions only as these provide most of the teachers in the country (about 80 percent). The NRC in Switzerland decided to collect representative data from German speaking areas to draw a comparison with the German system (France and Italy did not participate in TEDS-M thus, the NRC did not find a comparative group to justify inclusion). The samples are representative of the programs and populations the United States and Switzerland NRCs decided to survey.

[2] Note that in the U.S. 'primary' usually refers to grade K-3, while 'elementary' is used for grades K-5 or K-6. In this chapter the term elementary is used as used in TEDS-M (see Tatto et al. 2012, pp. 29–32 for specific definitions within countries as to what grades are included as primary or secondary).

[3] The TIMSS results are reported in simple descriptive statistics; the reference point is the international mean which is centered at 500, with a standard deviation of 100.

Canada (527), the Russian Federation (538), Quebec-CAN (543), Japan (586), Hong Kong-CHN (594), Chinese Taipei-CHN (599), the Republic of Korea (606), and Singapore (621). When looking at the domains measured (number, data, algebra, and geometry), significant gains in scores occurred only in the domains of algebra (507, 512, and 525) and geometry (480, 485, and 500).

Further, only 10% of eighth-grade students in the United States reached the TIMSS 2015 advanced benchmark (mean score of 625), compared with 13% in Israel, 14% in the Russian Federation, 15% in Kazakistan, 34% in Japan, 37% in Hong Kong, 43% in the Republic of Korea, 44% in Chinese Taipei, and 54% in Singapore. Where are eighth-grade U.S. students failing? According to the description of the performance levels developed by TIMSS (Provasnik, Kastberg, Ferraro, Lemanski, Roey, and Jenkins, 2012), eighth-grade pupils at advanced levels should demonstrate the following abilities: reason with information, draw conclusions, make generalizations, and solve linear equations … solve a variety of fraction, proportion, and percent problems and justify their conclusions … express generalizations algebraically and model situations … solve a variety of problems involving equations, formulas, and functions … reason with geometric figures to solve problems … reason with data from several sources or unfamiliar representations to solve multi-step problems (Provasnik et al. 2012, p. 19).

In the TIMSS 2015 Advanced, a study that measures the performance of students at the *advanced* level (12th-graders) implemented in a reduced number of countries, United States and Russian students had an average score of 485, scoring significantly higher than those in Sweden, France, Italy, and Slovenia (whose scores ranged from 431 to 460), but significantly lower than students in Lebanon (with an average score of 532) and advanced students in the Russian Federation (with an average score of 540). According to the TIMSS Advanced international benchmark, at the advanced level students should be able to "demonstrate thorough understanding of concepts, mastery of procedures and mathematical reasoning skills [and should be able to] solve problems in complex contexts in algebra, calculus, geometry and trigonometry" (Carr 2016).

The TEDS-M study, which could be seen as the TIMSS equivalent for teachers, assessed the mathematical knowledge attained by future secondary teachers after completion of their pre-service program requirements, close to the time they would receive their teaching credential. The international comparison of the TEDS-M assessments parallel those reported by TIMSS for secondary pupils. TEDS-M described two international performance levels or benchmarks (anchor points) to give substantial meaning to the scale scores attained by future secondary teachers in the mathematical content assessment; one defines a basic level of performance, and the other an advanced level. For instance in the U.S. among lower secondary teachers expected to teach up to Grade 10 (i.e., middle school mathematics), only 34% reached a basic level of performance, while 2% reached an advanced level (corresponding to a scale score of 559, where the international mean is equal to 500 and the standard deviation is equal to 100). Future secondary teachers expected to teach up to Grade 11 and above (i.e., high school mathematics) did better, with 87% reaching the basic, and 45% the advanced performance levels. These levels of performance, especially at the advanced levels, are significantly lower than those of other systems, such as Chinese Taipei, Germany, Poland, the Russian Federation, Singapore, and Switzerland. More detail on these performance levels will be presented later in this chapter; for now, it is important to point out that the areas where U.S. future secondary teachers had more challenges in the TEDS-M knowledge assessment was in demonstrating acceptable levels of conceptual understanding, problem-solving and reasoning capacity in essential mathematics domains, and, more specifically, in geometry. These low levels of performance suggest that it will be challenging to implement the ambitious standards for learning to teach mathematics, such as the newly-issued Association of Mathematics Teacher Educators (AMTE, 2017) Standards for the Preparation of Teachers of Mathematics, the Principles to Actions guidelines (National Council of Teachers of Mathematics, 2014), and the Common Core State Standards for Mathematics (CCSS-M, Common Core State Standards Initiative, 2010).

The TEDS-M assessments, designed to parallel the cognitive domains measured by the TIMSS assessments for pupils, were developed independently of TIMSS and in collaboration with groups of mathematicians and mathematics teacher educators representing the countries that took part in the study. The similarity in the patterns of knowledge and cognitive performance between students and future teachers in the TIMSS and TEDS-M assessments in the countries/regions that participated in both studies, however, suggests the need to examine the preparation that future secondary teachers obtain in their pre-service teacher education, and how it may contribute to future teachers' knowledge.

Previous Research

What teachers know and how they teach is seen as the most important school factor affecting student achievement (Goe, 2007; Hill, Rowan, & Ball, 2005; Kaplan & Owings, 2001). Because teachers are expected to teach the school curriculum, subject matter knowledge and subject matter pedagogy knowledge are seen as essential in their preparation. This has long been the consensus regarding mathematics (Baumert et al., 2010; Hill et al., 2005). But while teacher education is increasingly considered to be a key factor in teacher quality (Darling-Hammond, 2000; Darling-Hammond & Bransford, 2005; Levine, 2006), existing research is mixed regarding how mathematics teacher education may result in the kinds of high-quality teachers envisioned (Adler, 2017; Kilpatrick, Swafford, & Findell, 2001; Ball, 2003). While a number of studies have looked at the relationship between teachers' mathematics knowledge and pupil achievement (see Hawkins, Stancavage, & Dossey, 1998; Monk, 1994; Monk & King, 1994), prior to TEDS-M, there have been no large-scale, representative international and comparative research studies linking teacher education programs' practices (such as recruitment and selection strategies) and characteristics (such as coherent OTL) with teachers' knowledge at program completion (for early single country studies, see Tatto & Kularatna, 1993; Tatto, Nielsen, Cummings, Kularatna, & Dharmadasa, 1993).

Characteristics of Programs

A number of studies suggest that preparing effective mathematics teachers depends on rigorous recruitment and selection strategies, as well as programs' ability to provide OTL that integrate theoretical components and practical experiences successfully, and, in designing these opportunities, to account for individual characteristics of aspiring teachers (Grossman, Hammerness, & McDonald, 2009). Yet these studies have proven inconclusive (Casey & Childs, 2011; Levine, 2006; Mikitovicsa & Crehanb, 2002).

Scholars also have studied programs in a more holistic manner and found that the degree to which a program's OTL are aligned with an inquiry based theory about learning to teach contributes to its effectiveness as measured by cognitive changes (e.g., Tatto, 1996, 1998, 1999). Recent work on teacher education quality assurance reveals that, in addition to rigorous entry and exit requirements, program alignment with accreditation demands may be better able to produce highly knowledgeable teachers (Tatto et al., 2012; Ingvarson et al., 2013).

Characteristics of Future Teachers

Understanding how key individual characteristics such as socioeconomic status (SES), gender, age, and previous ability are related to future teachers' learning may help to create more successful teacher preparation programs. As noted above, studies that relate teacher education outcomes to the individual characteristics of future teachers are rare. Most studies explore the relationship of individual characteristics and teaching effectiveness as measured by gains in student achievement, with older studies also including measures of teachers' classroom performance. Darling-Hammond (2000) cites research studies as far back as the 1940s attempting to find relationships between teachers' ability (IQ or other measures) and teaching performance; she remarks that while the relationships were positive, they were rarely significant, likely due to the lack of variability among teachers in this measure. More informative studies have found that teachers' verbal ability is related to student achievement (Monk, 1994; Mullens, Murnane, & Willett, 1996) and that "this relationship may be differentially strong for teachers of different types of students [… and] may be a more sensitive measure of teachers' abilities to convey ideas in clear and convincing ways" (Darling-Hammond, 2000, p. 3). A more recent review (Wayne & Youngs, 2003) examined the literature spanning more than 25 years (from 1975 to 2002) in search of studies exploring the relationship between teacher characteristics (defined as ratings of colleges teachers attended, teachers' test scores, teachers' degrees and coursework, and teachers' certification status) and student achievement gains. While the search yielded only 21 studies, some with mixed results, the authors conclude that in the case of high school mathematics "students learn more from teachers with more mathematics-related coursework" (p. 103). Past research has established that knowledge of the subject before teacher education (e.g., previous attainment) and after teacher education (e.g., knowledge and beliefs about teaching and learning) in mathematics seem to be importantly related to how much attention teachers give to the subject once in the classroom (Rowland, 2012; Schmidt & Buchmann, 1983) and to the beliefs they hold as they teach diverse learners (Tatto, 1996). The field has for the most part assumed that there is a link between future teachers' characteristics and teacher education outcomes and has left unexplored the important question of what key features of teacher education seem to make a difference in the development of more knowledgeable teachers. Exploring OTL helps to answer this question.

Opportunities to Learn

There are significant disagreements about what content knowledge is important for teachers to acquire, how it is to be acquired, and whether acquiring that knowledge enables future teachers to tackle challenging mathematics problems and enact that knowledge in pedagogical situations (Rowland & Turner, 2008). One position is that if teachers are intelligent and well prepared in mathematics, they can generally learn most of what they need to know about teaching informally on the job (e.g., through formal mentoring and apprenticeship relationships). Another position is that in addition to content knowledge, mathematics teachers need opportunities to learn pedagogical content knowledge. Yet another position emphasizes the importance of general pedagogical knowledge (almost exclusively or in addition to content knowledge and pedagogical content knowledge). Teacher education programs seem to vary in the degrees to which knowledge of mathematics, knowledge of the mathematics school curriculum, and knowledge of mathematics pedagogy are emphasized, and this very question has been the subject of much research (e.g., Ball & Bass, 2003; Ball, Thames, & Phelps, 2008; Speer & King, 2009). In general, the preparation of secondary teachers has been characterized by an emphasis on the disciplines that future teachers are expected to teach. In the case of future secondary mathematics teachers, the knowledge that is considered essential is the knowledge of mathematics needed at the upper levels of schooling.

A related focus of research is the degree of program coherence across opportunities to learn. Scholars have found that more effective mathematics [and writing] programs are those able to deliver more coherent experiences (Tatto, 1996, 1998, 1999). Yet studies exploring the distinct impact of coherent course offerings are rare.

Mathematical Content Knowledge A number of research studies have explored the degree to which exposure to more mathematics courses contributes to teacher knowledge. While many research studies have also made important advances in exploring the knowledge needed for teaching primary mathematics (see Ball & Bass, 2000; Hill, Ball, & Schilling, 2009; Venkat & Spaull, 2015), equivalent work concerning secondary mathematics is still limited (Goldhaber & Brewer, 2000;

Krauss et al., 2008; Rowland, 2012; Speer & King, 2009). Overall, the evidence concerning the relationship between mathematics coursework and teacher knowledge is mixed. One line of research has tended to indicate that completing the upper-division college mathematics courses required for the mathematics major does little to improve the conceptual understanding of prospective elementary and secondary mathematics teachers (Ball, 1991; Rowan, Correnti, & Miller, 2002).

U.S.-based scholars who have explored these issues in more depth find that the content of courses that prospective secondary school teachers take does not prepare them to understand key underlying concepts of basic mathematics, let alone more advanced mathematics such as algebra or geometry (Ferrini-Mundy & Findell, 2000; Hill, Sleep, Lewis, & Ball, 2007). A general conclusion from a long chain of studies is that elementary and middle school teachers possess a limited knowledge of mathematics: "Teachers may know the facts and procedures that they teach but often have relatively weak understandings of the conceptual basis for that knowledge [and…] have difficulty clarifying mathematical ideas or solving problems that involve more than routine calculations" (Ball, 1991). Researchers also have remarked that this knowledge is not only limited; it is also not the kind of mathematics knowledge that is likely to be useful in and for teaching (Ball & Bass, 2000; Graham, Portnoy, & Grundmeier, 2002; Usiskin, Peressini, Marchisotto, & Stanley, 2003; Van Dooren, Verschaffel, & Onghena, 2002).

Mathematical Pedagogical Content Knowledge Scholars also have explored whether exposure to courses on the pedagogy of mathematics (commonly known as mathematical pedagogical content knowledge, or MPCK), have resulted in more knowledgeable teachers. Here again, the results are mixed. Ball, Thames, and Phelps (2008) have found that the knowledge that teachers need to effectively teach mathematics is "a kind of mathematical reasoning that most adults do not need to do on a regular basis" (p. 299); they found that this special kind of teaching knowledge cannot be acquired in mathematics courses alone and that adults, even if well-educated, do not necessarily possess this knowledge.

Unfolding what this special knowledge is has been a slow process. Building from the work of several scholars in the field, Ball and her colleagues suggest the helpful concept of mathematical knowledge for teaching an amalgam of mathematical content knowledge and mathematical pedagogical knowledge, which includes (a) knowing mathematics and being able to act on that knowledge (e.g., knowing concepts and procedures in decompressed or unpacked form, being able to recast or organize mathematical ideas); (b) knowing and being able to use mathematical concepts, processes and procedures of the school curriculum (e.g., knowing the mathematical precursors and trajectories of concepts and procedures, relating school mathematics to mathematicians' mathematics); and (c) knowing and working with mathematical arguments (e.g., collectively finding the mathematical rationale of an algorithm, judging equivalence of arguments, and making a positive reading of student thinking) (Ball & Bass, 2000). This framework has emerged from years of concentrated empirical study; however, it still requires validation at a large scale to be generalizable.

While teacher knowledge—both knowledge of content and knowledge of pedagogy for specific content—was and still is seen as critical to teachers developing deeper understandings for teaching, studies have found that in some cases mathematics courses taken by pre-service teachers in the United States contain knowledge that has little relevance for teaching, and that methods courses often teach little mathematics content, spending time instead on giving teachers a generally applicable "bag of tricks" (Floden, McDiarmid, & Jennings, 1996).

Other kinds of knowledge also are considered important for future teachers, including knowledge of students and knowledge of the curriculum. Studies have found that courses that teach this knowledge in the context of content-specific pedagogy seem to be more effective (Clift & Brady, 2005; Floden & Meniketti, 2005).

Espoused Beliefs

Teacher education programs dedicate a significant amount of time to addressing concerns related to future teachers' beliefs. Empirical literature is well-established on teachers' and student-teachers' beliefs (see, e.g., De Corte, Op't Eynde, & Verschaffel, 2002; Handal, 2003; McCleod, 1992; Staub & Stern, 2002) and how these are related to OTLs in methods courses. A number of studies have found that methods courses and field experiences seem to affect candidates' beliefs about teaching mathematics and their ability to demonstrate knowledge of inquiry-oriented principles when planning instruction (Beswick, 2007, 2009; Langrall & Mooney, 2002; Mewborn, 2000).

Framework

Teacher education's theory of action entails a complex combination of inputs and implementation processes. The most important input in teacher education is the characteristics of those who enroll (e.g., prospective teachers' academic ability, and background such as SES, gender, and age). Teacher education's most relevant indicator of implementation processes is the provision of key coherent and relevant OTLs that are offered to future teachers, all of which is shaped by programs' structure and accreditation guidelines and by curricular standards. While the ultimate outcome of teacher education's theory of action is to positively impact pupils' learning, its most immediate outcome is to positively influence future teachers' learning (i.e., knowledge and beliefs). The argument in this chapter is that to understand teachers' influence on pupil learning it is necessary to understand first how teachers' knowledge is affected by their learning experiences in teacher education programs.

This chapter explores whether there are key common factors (future teachers' characteristics, programs' characteristics, etc.) associated with the successful preparation of future secondary teachers (with respect to their MCK and MPCK) at the country/regional level. The focus is on the United States in comparison with several countries/regions.

Methods

The cross-national data were collected via surveys administered from 2008 to 2009 to representative samples of programs and their future teachers who were in their last year of their teacher preparation. The data from programs were collected using a questionnaire, and the data from future secondary teachers were collected using background questionnaires and assessments of MCK and MPCK.[4] This chapter, uses the data from the future secondary teacher survey.

Participants

Future Secondary Teachers Future teachers in their last year of their teacher education program were surveyed. A teacher education program was defined as a pathway that exists within an institution that requires students to undertake a set of subjects and experiences, and leads to the award of a common credential on completion. Within the TEDS-M study, all teacher education programs were located within higher education institutions (HEI) or the equivalent. A future teacher was defined as a person enrolled in a teacher education program that is explicitly intended to prepare teachers qualified to teach mathematics at the secondary school level.

Sampling The study used a stratified multi-stage probability sampling design drawing representative samples of teacher education programs in the participating countries/regions. Programs were randomly selected from a national list of teacher education programs, and future teachers were randomly selected from a list of in-scope future teachers within each of the teacher preparation programs. In countries with few programs, all teacher preparation programs were selected to participate in the study.[5] The data analysis takes into account the sampling weights for programs and future teachers, which also provide for a non-response adjustment factor for all the estimates (for a detailed description of the estimation of weights, participation rates, and sampling error, consult Tatto, 2013, chapter 10).

[4] Human subject protection procedures were implemented in every country and monitored by Michigan State University, the principal investigator's institution at the time of the study. The methods are described in detail in Tatto (2013).

[5] The minimum sample size was set at 50 institutions per level and an effective sample size of 400 future teachers per level in a given country. "Effective sample size" means that the sample design must be as efficient (i.e., precise) as a simple random sample of 400 teachers from a (hypothetical) list of all eligible future teachers.

Table 14.1 Number of participating programs offering secondary teaching credential and future secondary teachers within countries in the TEDS-M Study

Country/ Region	Number of participating programs offering preparation to teach at the secondary level per country/region	Number of participating future secondary teachers (sample size)	Number of participating future secondary teachers (valid N)	Percent missing
Chile	37	746	648	13.1
Chinese Taipei	19	365	355	2.7
Germany	28	771	620	19.6
Malaysia	6	389	357	8.2
Philippines	48	733	668	8.9
Poland	35	298	247	17.1
Russian Federation	48	2141	1951	8.8
Singapore	4	393	371	5.6
Switzerland (German)	7	141	137	2.8
Thailand	53	652	614	5.8
United States (public only)	72	607	461	24.1

Source. Tatto (2013, pp. 214–251)

The number of programs participating in the study by country/region and the number of participating future secondary teachers are shown in Table 14.1. All participating countries with response rates of 85% or more on the assessments were included. Country/region data were analyzed across the same variables.

Procedures

Instruments The data reported in this chapter come from a one-time survey of future secondary teachers administered at program completion and before graduation. The survey included a questionnaire and an assessment of knowledge. Researchers from the different participating countries/regions were asked to contribute questions for the questionnaire and items for the assessment. Several items were also provided by other studies, including the Study of Instructional Improvement (SII) Learning Mathematics for Teaching,[6] the Developing Subject Matter Knowledge in Math Middle School Teachers,[7] and Knowing Mathematics

[6] Consortium for Policy Research in Education (CPRE), University of Michigan, School of Education, Ann Arbor, MI (measures development supported by NSF grants REC-9979873, REC-0207649, EHR-0233456 & EHR 0335411).

[7] Measures development for lower secondary teachers was supported by NSF Grant REC-0231886. Michigan State University, East Lansing, MI.

Table 14.2 Future teachers' overall booklet structure and allocated times for administration

Future teacher questionnaire sections	Time (min)
Part A: General background	5
Part B: Opportunity to learn	15
Part C: Mathematics for teaching (MCK/MPCK)	60
Part D: Beliefs about mathematics and teaching	10

Source. Tatto et al., (2008, p. 33)

for Teacher Algebra (KAT).[8] The questions and items were piloted, and the instruments were field-tested before the instruments were declared ready for the main study. The final TEDS-M instruments were rigorously developed in collaboration with psychometricians and translated from the English to the local languages and back-translated to confirm accuracy and consistency. Factor analysis provided construct-related validity evidence and curriculum analysis and expert review provided content-related validity evidence. Further details on the design of the study and the methods used, including the creation of the scales and fit and reliability indices, can be found in the TEDS-M Technical Report (Tatto, 2013).

The future secondary teacher survey was administered in one and a half hours and consisted of four parts (shown in Table 14.2), with questions about background characteristics (Part A), questions about the program's OTLs (Part B), an assessment of MCK and MPCK (Part C), and questions about beliefs about the nature of teaching and learning mathematics (Part D). While recognizing that there are many aspects of preparation of future teachers that could be considered, the areas explored in the survey were necessarily limited by the time available for instrument administration.

Measures

While there is much information available for each of the countries studied, a number of factors limited the explanatory variables that could be considered in the analysis for this chapter. The goal of including as many countries as possible in the analysis limited the number of common variables. To avoid inflation of results, some variables were excluded based on results of collinearity tests[9] (e.g., applying the variance inflation factor, or VIF, results in a variable reduction strategy). In addition and given VIF results, some variables were excluded because one single vari-

[8] Measures development was supported by NSF Grant REC-0337595. Michigan State University, East Lansing, MI.

[9] In statistics, collinearity is a phenomenon in which two or more predictor variables in a model are highly correlated meaning that one can be predicted from the others with a substantial degree of accuracy.

able represented the best indicator of a domain (such as geometry, as an indicator of having the opportunity to learn university-level mathematics).

A detailed description of the variables is in the TEDS-M User Guide for the International Database (Brese & Tatto, 2012).

Future Teachers' Knowledge for Teaching Secondary Level Mathematics Two assessments were developed to measure future teachers' knowledge as indicators of teacher education outcomes. The Mathematical Content Knowledge (MCK) assessment comprised two-thirds of the assessment and measured four domains: number and operations, algebra and functions, geometry and measurement, and data and chance. The Mathematical Pedagogical Content Knowledge (MPCK) assessment comprised one-third of the assessment and measured three domains: curricular knowledge, knowledge of planning for teaching, and knowledge for enacting teaching. Three blocks of items were assembled for the secondary assessment, each with 12–15 questions. Each future teacher received a booklet with two of the blocks of items about knowledge for teaching mathematics. The assessment was designed to be answered in no more than 60 min under a controlled administration by trained researchers. To sample all the domains, the study used a matrix sampling design for the assessments (Mazzeo, Lazer, & Zieky, 2006). To obtain comparable estimates of performance, item response theory (IRT) was used (see, e.g., De Ayala, 2009). The final results were calibrated and used to estimate the location of the examinees on a common IRT scale with the international mean set at 500 on each of the MCK and MPCK scales, and the international standard deviation set at 100 (Wu, Adams, Wilson, & Haldane, 2007). For the international sample, the reliability for the mathematical content knowledge assessment was .91, and for the mathematical pedagogical content knowledge assessment, it was .72. The assessment results are reported in score scales and performance levels.

Future Teachers' Characteristics These included data on the background of respondents—specifically, information about individuals' SES,[10] age, gender, and prior attainment (in Tables 14.3 and 14.4). Future secondary teachers' background reflects recruitment and selection policies as well as the social and economic level of those who are attracted into teaching. With the exceptions of the Russian Federation, the United States, Germany, and Switzerland, future secondary teachers come from low SES and are in their early to mid-twenties. Across countries, most are female. Their self-reported level of attainment as indicated by average grades in high school was above average.[11]

[10] Using principal components analysis, a scale was created to obtain a proxy measure of socioeconomic status, including variables such as the possessions in the parents or guardians home, and others such as father's highest level of education and mother's highest level of education. Its aggregate within a program constitutes the variable program's SES. A positive value signals greater SES.

[11] 5 While desirable, the TEDS-M research team was not able to administer a "pre-test" to future teachers or to obtain comparable measures of knowledge across countries. Instead, the study asked future teachers to report their levels of previous attainment. Given the results of the assessments and the high positive correlations with previous attainment, it is possible to confirm the reliability of their report.

Table 14.3 Means and standard deviations for future secondary teachers in countries with large program samples

	Chile		Philippines		Poland	
	M	*SD*	*M*	*SD*	*M*	*SD*
Future teachers secondary (level 1)						
MCK score [AP_{basic} = 490; $AP_{advanced}$ = 559]	356.35	84.60	449.36	48.29	536.05	88.91
MPCK score [$AP_{proficient}$ = 509]	394.58	87.70	451.50	60.75	525.98	95.04
SES*[SES]	−0.29	0.80	−0.63	0.86	−0.11	0.73
Age [MFA001]	23.85	2.80	20.96	2.00	23.13	5.33
Proportion female [MFA002_]1 = F; 0 = M	.84	.36	.65	.48	.81	.22
Prior attainment: Average grades in secondary school (1 = below average for year level; 5 = always at top of year level) [MFA009_]	3.28	1.14	3.07	0.95	3.28	0.84
Program characteristics (level 2)						
Average number of university level mathematics topics in geometry ever studied (range 0–4) [MFB1GEOM]	1.87	0.54	2.78	0.45	3.23	0.47
Average number of school level mathematics topics in function, probability and calculus studied as part of the TE program (range 0–4) [MFB2SLMF]	1.53	0.45	2.74	0.50	3.82	0.25
Average frequency with which future teachers engaged in reading research on teaching and mathematics (scales centered at 10 representing neutral) [MFB5READ]	9.22	0.96	10.68	0.74	8.15	1.34
Average level of program coherence (scales centered at 10 representing neutral)[MFB15COH]	11.90	1.21	13.59	0.88	11.53	1.14
Average SES for each program (aggregated from future teachers SES) [SES]	−0.21	0.47	−0.64	0.49	−0.10	0.28
Beliefs						
Average agreement with the belief that mathematics is a collection of rules and procedures (scales centered at 10 representing neutral) [MFD1RULE]	10.94	0.55	12.67	0.63	10.39	0.51
Average agreement with the belief that mathematics is better learned through active learning (scales centered at 10 representing neutral) [MFD2ACTV]	12.73	0.48	11.80	0.50	12.29	0.79

(continued)

Table 14.3 (continued)

	Russian Fed.		Thailand		United States	
	M	*SD*	*M*	*SD*	*M*	*SD*
Future teachers secondary (level 1)						
MCK score [AP_{basic} = 490; $AP_{advanced}$ = 559]	593.26	90.84	478.46	57.91	536.13	65.36
MPCK score [$AP_{proficient}$ = 509]	569.01	94.67	477.41	64.37	529.15	80.55
SES*[SES]	0.60	0.64	−0.90	1.06	0.46	0.84
Age [MFA001]	22.01	1.59	22.34	0.81	25.26	6.45
Proportion female [MFA002_]1 = F; 0 = M	0.72	0.45	0.75	0.43	0.69	0.28
Prior attainment: Average grades in secondary school (1 = below average for year level; 5 = always at top of year level) [MFA009_]	3.80	0.89	3.29	0.84	3.88	1.00
Program characteristics (level 2)						
Average number of university level mathematics topics in geometry ever studied (range 0–4) [MFB1GEOM]	3.81	0.21	3.41	0.42	2.59	0.74
Average number of school level mathematics topics in function, probability and calculus studied as part of the TE program (range 0–4) [MFB2SLMF]	3.46	0.32	3.51	0.61	2.81	0.79
Average frequency with which future teachers engaged in reading research on teaching and mathematics (scales centered at 10 representing neutral) [MFB5READ]	10.28	0.75	10.31	0.75	10.61	1.34
Average level of program coherence (scales centered at 10 representing neutral)[MFB15COH]	12.93	0.75	13.01	0.99	12.78	1.63
Average SES for each program (aggregated from future teachers SES) [SES]	0.60	0.17	−0.83	0.54	0.47	0.48
Beliefs						
Average agreement with the belief that mathematics is a collection of rules and procedures (scales centered at 10 representing neutral) [MFD1RULE]	10.52	0.28	11.83	0.56	10.71	0.59
Average agreement with the belief that mathematics is better learned through active learning (scales centered at 10 representing neutral) [MFD2ACTV]	11.89	0.47	11.96	0.55	12.11	0.90

Table 14.4 Means and standard deviations for future secondary teachers in countries/regions with small program samples

	Chinese Taipei		Germany		Malaysia	
	M	*SD*	*M*	*SD*	*M*	*SD*
Future teachers (level 1)						
MCK score [AP_{basic} = 490; $AP_{advanced}$ = 559]	666.58	75.37	541.91	84.33	495.06	51.09
MPCK score [$AP_{proficient}$ = 509]	647.67	94.46	553.12	98.16	472.95	62.55
SES*[Zscore: REGR factor score…]	−0.50	0.88	0.41	0.94	−0.69	0.79
Age [MFA001]	24.06	2.28	28.98	4.91	22.70	2.32
Proportion female [MFA002_]1 = F; 0 = M	0.38	0.49	0.62	0.48	0.82	0.38
Prior attainment: Average grades in secondary school (1 = below average for year level; 5 = always at top of year level) [MFA009_]	3.66	1.07	3.32	0.88	3.82	0.96
Program characteristics (level 2)						
Average number of university level mathematics topics in geometry ever studied (range 0–4) [MFB1GEOM]	3.23	0.33	2.20	0.47	2.77	0.45
Average number of school level mathematics topics in function, probability and calculus studied as part of the TE program (range 0–4) [MFB2SLMF]	3.45	0.35	2.48	0.57	3.45	0.17
Average frequency with which future teachers engaged in reading research on teaching and mathematics (scales centered at 10 representing neutral) [MFB5READ]	9.69	0.86	8.01	0.49	10.37	0.30
Average level of program coherence (scales centered at 10 representing neutral)[MFB15COH]	11.97	0.57	9.17	0.48	12.73	0.50
Average SES for each program (aggregated from future teachers SES) [SES]	−0.50	0.20	0.41	0.30	−0.70	0.15
Beliefs						
Average agreement with the belief that mathematics is a collection of rules and procedures (scales centered at 10 representing neutral) [MFD1RULE]	10.81	0.20	9.66	0.15	11.63	0.19
Average agreement with the belief that mathematics is better learned through active learning (scales centered at 10 representing neutral) [MFD2ACTV]	12.35	0.26	12.43	0.33	11.38	0.21

(continued)

Table 14.4 (continued)

	Singapore		Switzerland	
	M	*SD*	*M*	*SD*
Future teachers (level 1)				
MCK score [AP_{basic} = 490; $AP_{advanced}$ = 559	573.92	60.72	530.64	48.79
MPCK score [$AP_{proficient}$ = 509	554.84	84.69	546.48	73.03
SES*[Zscore: REGR factor score…]	−0.55	0.79	0.12	0.91
Age [MFA001]	26.73	4.00	26.20	4.30
Proportion female [MFA002_]1 = F; 0 = M	0.48	0.50	0.42	0.49
Prior attainment: Average grades in secondary school (1 = below average for year level; 5 = always at top of year level) [MFA009_]	3.52	0.95	3.41	0.91
Program characteristics (level 2)				
Average number of university level mathematics topics in geometry ever studied (range 0–4) [MFB1GEOM]	1.49	0.40	2.78	0.43
Average number of school level mathematics topics in function, probability and calculus studied as part of the TE program (range 0–4) [MFB2SLMF]	2.63	0.30	2.90	0.40
Average frequency with which future teachers engaged in reading research on teaching and mathematics (scales centered at 10 representing neutral) [MFB5READ]	9.12	0.15	8.75	0.80
Average level of program coherence (scales centered at 10 representing neutral)[MFB15COH]	12.03	0.17	10.45	0.87
Average SES for each program (aggregated from future teachers SES) [SES]	−0.55	0.11	0.12	0.20
Beliefs				
Average agreement with the belief that mathematics is a collection of rules and procedures (scales centered at 10 representing neutral) [MFD1RULE]	10.91	0.07	9.85	0.28
Average agreement with the belief that mathematics is better learned through active learning (scales centered at 10 representing neutral) [MFD2ACTV]	11.53	0.15	12.47	0.43

Program Characteristics These variables were aggregated from the future secondary teachers' answers to program-related questions. The future teacher questionnaire was designed so that it could be completed in 30 minutes and substantial information about the program was collected, yet, as explained previously, because of collinearity issues, not all program variables were fit to include in the analysis. For instance, among the OTL variables, and based on VIF analysis geometry was considered the best indicator of the university-level mathematics domain for future secondary teachers. The same rationale holds true for the other variables in the model.

Future Teachers' Opportunities to Learn A number of indices were developed that were based on counts of topics studied. These included (a) university-level mathematics, using geometry as an indicator, including topics such as foundations of geometry or axiomatic geometry, analytic/coordinate geometry, non-Euclidean geometry, and differential geometry; and (b) upper-school level mathematics (within school-level mathematics), including functions, relations, equations, data representation, probability, statistics, calculus, and validation, structuring, and abstracting. Fit indices for these indicators are robust for university-level mathematics (CFI = .969, TLI = .986, RMSEA = .032) and acceptable for school-level mathematics (CFI .892, TLI .846, RMSEA .085).[12] Other OTL indices were based on a 4-point scale (e.g., expressing frequency such as *never* to *often*) and included questions on whether future teachers had the opportunity to learn topics in mathematical pedagogy, general education and pedagogy, accommodations to classroom diversity and reflections on practice, and from school experience and the practicum. The reliabilities for these 4-point scale indices ranged from .83 to .97.[13] These were aggregated to constitute the different learning opportunities offered to future secondary teachers, such as university- and school-level mathematics, and the average frequency with which future teachers engaged in reading research on teaching and mathematics as shown across countries in Tables 14.3 and 14.4.

Program Coherence An index was developed from questions asking whether future secondary teachers had a coherent experience in their teacher education program (e.g., whether each of the courses was clearly designed to prepare future teachers to meet a common set of explicit standard expectations for beginning

[12] Comparative Fit Index (CFI): The CFI depends in large part on the average size of the correlations in the data. If the average correlation between variables is not high, then the CFI will not be very high. An acceptable model is indicated by a CFI larger than .93, but .85 is acceptable (Bollen, 1989).The Tucker Lewis index (TLI) is relatively independent of sample size (Marsh, Balla, & McDonald, 1988). Values over .90 or .95 are considered acceptable (e.g., Hu & Bentler, 1999). Root Mean Square Error of Approximation (RMSEA): Another test of model fit, good models are considered to have a RMSEA of .05 or less. Models whose RMSEA is .1 or more have a poor fit.

[13] The reliabilities for the OTL and beliefs scales are unweighted and were estimated using jMetrik 2.1 (Meyer, 2011). The reliability estimates are based on the congeneric measurement model, which allows each item to load on the common factor at different levels and allows item error variances to vary freely (each item can be measured with a different level of precision). This is the most flexible measurement model and most appropriate for measures with few items.

teachers, and whether there were clear links between most of the courses in the teacher education program and practicum experiences). Based on a series of confirmatory factor analyses, these indices were scaled using the Rasch model and are based on a score scale where 10 is located at the neutral position where values lower than 10 indicate less coherence and values larger than 10 more coherence. The reliability for this 4-point rating scale index is .97. Tables 14.3 and 14.4 show that, in general, future teachers judged their programs to be coherent with all means higher than the neutral point.

Programs' Socioeconomic Status Socioeconomic status is an aggregated scale created by measures of future teachers' home possessions, including number of books at home and parents' levels of education.

Beliefs Future teachers were asked a number of questions to explore their beliefs about teaching and learning mathematics using 6-point rating scales (*strongly agree* to *strongly disagree*) in two key areas: (a) the "nature of mathematics" questions explored how future teachers perceive mathematics as a subject (e.g., mathematics as formal, structural, procedural, or applied); and (b) the "learning mathematics" questions explored ideas about the appropriateness of particular instructional activities, including questions about students' cognition processes, and questions about the purposes of mathematics as a school subject. After factor analysis, these belief items were scaled using the Rasch model (based on a score of 10 located at the neutral position) to form two scales: the "mathematics as a set of rules and procedures" scale (with a reliability of .93) and the "learning mathematics through active involvement" scale (with a reliability of .92).

On average, future secondary teachers show a tendency to agree with widely accepted beliefs on learner-centered teaching (e.g., "teachers must focus on what the learner is thinking when learning—and not solely on the subject/lesson to be taught"), and less with the notion that mathematics can be learned by mastering a collection of rules and procedures, a belief that, if upheld, would imply a more procedural view of mathematics and, if rejected, may indicate a philosophy more attuned to current and more progressive thinking in education (see Tables 14.3 and 14.4). The next section presents the results of the analysis.

Knowledge for Teaching Secondary Level Mathematics and Teacher Education Characteristics: Analysis and Findings[14]

Exploring Mathematical and Mathematical Pedagogical Content Knowledge

The mean scores and standard deviation of the MCK and MPCK assessments are reported in Tables 14.3 and 14.4. When looking at the tables, and taking into account that the international mean score was set at 500 with a standard deviation of 100, one can see the wide variation in the level of knowledge attained by future secondary teachers in the different participating countries. Table 14.3, for instance, shows the MCK score for the lowest scoring country (Chile) and the highest scoring country (Russia), and a simple comparison indicates more than two standard deviations difference in the scores (356 versus 593). To describe more concretely individuals' levels of performance, anchor points (AP) or benchmarks, were identified based on future secondary teachers' scores at specific points reached on the MCK and MPCK scales. The goal was to find out not only the scores the future teachers reached, but also the areas where they were successful (more knowledgeable) and the areas where they had difficulties (less knowledgeable). A panel of mathematicians and mathematics educators were asked to analyze the items classified at these APs and to formulate empirically based descriptions of the knowledge that future teachers demonstrated at each AP. Items used to describe performance at the APs were selected based on the probability that a future teacher with a score at that point would get the relevant items right. Based on the number of items for each measure, two APs for MCK and one AP for MPCK were identified.

Mathematics Content Knowledge For MCK, the basic knowledge level anchor point (AP_{basic}) corresponded to a scale score of 490 (just below the international mean of 500) and included items representing a .70 probability of answering the items correctly, while the higher knowledge level anchor point ($AP_{advanced}$) corresponded to a scale score of 559 and represented a .50 or less probability of answering the items correctly. For instance, future secondary teachers reaching on average a scale score of 490 or above were successful at demonstrating knowledge of concepts related to whole numbers, integers, and rational numbers, and the associated computations; evaluating algebraic expressions correctly; solving simple linear and quadratic equations, particularly those that can be solved by substitution or trial and error; demonstrating knowledge of standard geometric figures in the plane and

[14] In this chapter, the concern is with exploring the relationships between knowledge for teaching (MCK and MPCK) and program characteristics for the United States and for the other countries included in the analysis. While noting between-country differences especially with respect to the knowledge assessments, the concern is not with statistically testing differences between countries. Care has been taken when discussing how variables play a role within a country and how this differs across countries. These kinds of observations do not require statistical hypothesis tests.

space; identifying and applying simple relations in plane geometry; and interpreting and solving more complex problems about numbers, algebra, and geometry if the context or problem type was commonly taught in lower-secondary schools. However, these teachers had difficulty answering items that asked them to describe general patterns; to solve multi-step problems with complex linguistic or mathematical relations; to relate equivalent representations of concepts (with a tendency to overgeneralize concepts); and to reason mathematically, including recognizing faulty arguments and justifying or proving conclusions.

In contrast, future secondary teachers reaching on average a scale score of 559 were likely to correctly do all the mathematics items reached by future secondary teachers at the basic level AP and, in addition, they were successful at demonstrating knowledge of functions (particularly linear, quadratic, and exponential); reading, analyzing, and applying abstract definitions and notation; making and recognizing simple arguments; and demonstrating knowledge of some definitions and theorems typically taught in tertiary-level courses, such as calculus, abstract algebra, and college geometry, and then applying these in straightforward situations. But they still had difficulty solving problems stated in purely abstract terms; working competently on foundational material, such as axiomatic systems; reasoning logically (e.g., they failed to attend to all conditions of definitions or theorems and confused the truth of a statement with the validity of an argument); recognizing valid proofs of more complex statements; and constructing and completing mathematical proofs.

Tables 14.3 and 14.4 show that the average MCK scores of future secondary teachers in the United States reached and surpassed the basic level AP of 490, but only future secondary teachers in Russia, Chinese Taipei, and Singapore reached and surpassed the higher-level AP (559).

Figure 14.1 below includes an example of a constructed response item measuring MCK basic and advanced levels, and the MPCK proficient AP. The first two items are designed to measure applied MCK in algebra and asked future teachers to solve two different word problems about linear relations. For item 1 (playing with marbles), the international average percent correct was 72% indicating that close to three-fourths of the future secondary teachers in the international sample were able to answer that item correctly. For item 2 (about money) the international percent correct was 50% indicating that this was a more difficult item for half of the future secondary teachers in the international sample. The MPCK also in the algebra domain was designed to measure MPCK as enacted. The item asked future secondary teachers to analyze why one word problem is more difficult than another. In this case the international average percent correct was 39% revealing a higher level of difficulty as close to 60% of the international sample of future secondary teachers were unable to answer this item correctly.

Mathematical Pedagogical Content Knowledge For MPCK, only one anchor point ($AP_{proficient}$) was identified, representing a score of 509 on the scale. Overall, the MPCK items were more challenging for all future secondary teachers. For instance, those future secondary teachers at the AP were able to demonstrate knowledge to support planning for instructional purposes (e.g., identifying prerequisites

The following problems appear in a mathematics textbook for <lower secondary school>.

1. [Peter], [David], and [James] play a game with marbles. They have 198 marbles altogether. [Peter] has 6 times as many marbles as [David], and [James] has 2 times as many marbles as [David]. How many marbles does each boy have?

2. Three children [Wendy], [Joyce] and [Gabriella] have 198 zeds altogether. [Wendy] has 6 times as much money as [Joyce], and 3 times as much as [Gabriela]. How many zeds does each child have?

(a) Solve each problem.

Solution to Problem 1:
Solution to Problem 2:

Typically, Problem 2 is more difficult than Problem 1 for <lower secondary> students. Give one reason that might account for the difference in difficulty level.

Source: Brese and Tatto (2012, Supplement 4)

Fig. 14.1 Example of MCK and MPCK anchor point item

for teaching a derivation of the quadratic formula, and determining consequences of moving the concept of square root from the lower-secondary to the upper-secondary school mathematics curriculum). They were able to demonstrate knowledge for enacting school mathematics teaching and to evaluate students' mathematical work correctly in some situations (for instance, determining whether a student's diagram satisfied certain given conditions in geometry, and recognizing a student's correct argument about divisibility of whole numbers). They were able to analyze students' errors when the students' work involved a single step or short explanations (for example, identifying an error in a histogram). Yet they had difficulty with identifying or analyzing errors in more complex mathematical situations (for instance, they could not consistently apply a rubric with descriptions of three performance levels to evaluate students' solutions to a problem about linear and non-linear growth); understanding and interpreting students' thinking or determining appropriate responses to students; and understanding the concept and meaning of a valid mathematical argument (for example, they were unable to evaluate an argument as invalid by recognizing that examples are not sufficient to constitute a proof).

Future secondary teachers in Poland, Russia, Singapore, Switzerland, and the United States reached and surpassed the MPCK anchor point (see Tables 14.3 and 14.4).

Exploring Relationships Between Knowledge for Teaching Secondary-Level Mathematics and Teacher Education Characteristics

To explore relationships between MCK and MPCK results and teacher education program characteristics, hierarchical linear modeling (HLM) and Ordinary Least Squares (OLS) were used for the analyses. Because of hypothesizing a multilevel structure, in which the influence of teachers' background may differ according to the programs they are in, a multilevel statistical analysis such as HLM (Raudenbush Stephen, Bryk, Cheong, & Congdon, 2004) is the optimal approach to investigating the relationship between teacher education programs' characteristics and mathematical knowledge for teaching (MCK and MPCK). However, the number of programs preparing future secondary teachers varied widely across countries—with the extreme cases being Singapore, with four programs, and the United States, with 72—with a consequent impact on future teachers' sample sizes (see Table 14.1). Thus, while HLM was generally used for countries with large program samples, it was necessary to use Ordinary Least Squares (OLS) for some countries/regions. In some cases, OLS was necessary because a country/region had a small sample of programs with a very similar structure, including Singapore, with four programs, Malaysia with six, Switzerland with seven, Chinese Taipei with 19, and Germany with 28. In other cases, OLS was used because the country/region had a strict centralized system or norms. Because of the smaller numbers of programs and/or because of institutional isomorphism within these countries, variance between institutions is not reliably estimated, so that variance (if it exists systematically in the population) is part of the individual future-teacher variability and not partitioned separately. In order to proceed with the analysis of the association between teacher background and knowledge, program-level characteristics were added as additional explanatory variables in OLS models.

The analyses used standardized coefficients to explore individual- and program-level features associated with MCK and MPCK as teacher education outcomes, using the same HLM or OLS model within country/region groups, but analyzing each country/region separately.

All participating countries/regions with acceptable response rates on the assessments were included. Each country/region's data were analyzed across the same variables. Before inclusion in the model, variables across countries/regions were examined for missing values across variables of interest, and for collinearity; this resulted in a number of variables (and programs) being excluded from the analysis. Other variables were excluded because they lacked variability (e.g., all programs reported offering field experience).

Analysis for Countries with Large Samples HLM was used to investigate the association between future secondary teachers' mathematical knowledge for teaching (defined for this study as MCK and MPCK) and teacher education features. As noted, the same model was used across all (equally weighted) countries, controlling for future teachers' background variables. The analysis in this study uses a random intercept model in which future teachers are considered as nested within their teacher education programs (modeling the teachers' response) within their countries. The model is detailed below.

STEP 1
Full Unconditional Model (Level 1 & 2)

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}, \text{where } u_{0j} \sim \mathrm{N}\left(0, \tau_{00}\right) \text{and } r_{ij} \sim \mathrm{N}\left(0, \sigma^2\right)$$
$$\mathrm{ICC} = \left(\tau_{00}\right) / \left(\tau_{00} + \sigma^2\right)$$

STEP 2
Level-1 Model

$$Y_{ij} = \beta_{0j} + \beta_{1j}\left(\mathrm{SES}\right)_{ij} + \beta_{2j}\left(\mathrm{MFA001}\right)_{ij} + \beta_{3j}\left(\mathrm{MFA002}\right)_{ij} + \beta_{4j}\left(\mathrm{MFA009}\right)_{ij} + r_{ij}, \text{ where } r_{ij} \sim \mathrm{N}\left(0, \sigma^2\right)$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\left(\mathrm{MFB1GEOM}\right)_j + \gamma_{02}\left(\mathrm{MFB2SLMF}\right)_j + \gamma_{03}\left(\mathrm{MFB5READ}\right)_j + \gamma_{04}\left(\mathrm{MFB15COH}\right)_j + \gamma_{05}\left(\mathrm{MeanSES}\right)_j + \gamma_{06}\left(\mathrm{MFD1RULE}\right)_j + \gamma_{07}\left(\mathrm{MFD2ACTV}\right)_j + u_{0j}, \text{ where } u_{0j} \sim \mathrm{N}\left(0, \tau_{00}\right)$$

$$\beta_{pj} = \gamma_{p0}$$

for $p = 1$ to 4 (the four level-1 predictors, which are fixed).

The coefficients in the model (which have been standardized within each country) help explain variation in the outcome. For the United States, for example, the resulting coefficient, called a standardized beta or β, is 0.28 (see Table 14.5, United States MCK column and the row indicating "opportunity to learn university-level mathematics," in this case geometry), meaning that, after controlling for background characteristics, a change of one SD in the opportunity to learn university-level geometry is associated with a 0.28 SD increase in the mathematics content knowledge score.

Table 14.5 shows the results for the HLM model for secondary teachers and programs. The ICCs from the unconditional model at the bottom of the table indicate the association between program features and the knowledge that future teachers attain at the end of their pre-service teacher education across countries. For the mathematics content assessment (MCK), the higher ICCs are in Russia (57%), the United States (43%), Poland (31%), and Thailand (35%), indicating that, with the

Table 14.5 Coefficients from HLM regression of future secondary teachers' mathematics and mathematics pedagogy scores on future secondary teachers' background and program characteristics

Variable	Chile				Philippines				Poland			
	$N_{Level1 = 648}$; $N_{Level2 = 37}$				$N_{Level1 = 668}$; $N_{Level2 = 46}$				$N_{Level1 = 247}$; $N_{Level2 = 34}$			
	Coefficient (SE)				Coefficient (SE)				Coefficient (SE)			
	MCK	SE	MPCK	SE	MCK	SE	MPCK	SE	MCK	SE	MPCK	SE
Future teacher characteristics												
SES	0.07*	0.03	0.08	0.05	−0.05	0.05	0.11**	0.04	0.05	0.05	0.05	0.04
Age	−0.07	0.05	−0.05	0.05	−0.08	0.08	−0.09***	0.02	−0.14*	0.06	−0.11	0.06
Gender [1 = F; 0 = M]	−0.09**	0.03	−0.01	0.04	−0.10*	0.04	0.05	0.04	−0.17**	0.06	−0.15**	0.05
Prior attainment	0.07*	0.04	0.15**	0.05	0.07	0.04	−0.02	0.04	0.18*	0.07	0.09	0.06
Program level characteristics												
OTL: University level mathematics: Geometry	0.11	0.05	0.05	0.07	0.05	0.04	0.08	0.05	0.14	0.08	0.08	0.09
OTL: School level mathematics: Function, probability and calculus	0.02	0.04	0.02	0.05	0.02	0.05	0.05	0.05	−0.03	0.08	−0.10	0.08
OTL: Reading research on teaching and mathematics	0.04	0.05	−0.02	0.06	0.00	0.06	0.09	0.05	−0.10	0.07	−0.12	0.07
Program coherence	0.04	0.05	0.10*	0.04	0.10*	0.04	−0.09*	0.04	0.04	0.08	0.06	0.08
Average SES for each program (aggregated from future teachers SES)	−0.06	0.04	−0.09	0.05	0.20***	0.06	0.11*	0.05	0.13	0.11	−0.01	0.11
Beliefs												
Mathematics is a collection of rules and procedures	−0.10*	0.04	−0.14**	0.04	−0.04	0.05	0.08	0.04	−0.24	0.09	−0.26*	0.10
Mathematics is better learned through active learning	0.08*	0.04	0.00	0.03	0.08	0.06	0.07	0.06	−0.03	0.08	−0.01	0.08
ICC (from the unconditional model)	10%		8%		16%		12%		31%		24%	
% of variance explained within programs (level 1)	2%		2%		2%		1%		5%		1%	
% of variance explained between programs (level 2)	30%		39%		64%		74%		50%		35%	

Variable	Russian Federation				Thailand				United States			
	$N_{Level1 = 1951}$; $N_{Level2 = 48}$				$N_{Level1 = 614}$; $N_{Level2 = 52}$				$N_{Level1 = 461}$; $N_{Level2 = 68}$			
	Coefficient (SE)				Coefficient (SE)				Coefficient (SE)			
	MCK	SE	MPCK	SE	MCK	SE	MPCK	SE	MCK	SE	MPCK	SE
Future teacher characteristics												
SES	–0.01	0.01	0.01	0.02	0.10**	0.03	0.01	0.04	0.05	0.04	0.06	0.05
Age	–0.01	0.02	–0.02	0.02	0.02	0.04	0.03	0.03	–0.10*	0.04	–0.08	0.06
Gender [1 = F; 0 = M]	0.05*	0.02	0.06*	0.02	–0.06	0.03	0.05	0.04	–0.26***	0.05	–0.14*	0.06
Prior attainment	0.06**	0.02	0.06*	0.02	0.18***	0.04	0.07*	0.03	0.19***	0.05	0.11*	0.06
Program level characteristics												
OTL: University level mathematics: Geometry	0.13*	0.06	0.17***	0.05	–0.09	0.08	–0.10	0.05	0.28***	0.07	0.15*	0.07
OTL: School level mathematics: Function, probability and calculus	0.03	0.06	–0.03	0.04	0.08	0.05	0.14***	0.03	0.10	0.07	0.13*	0.07
OTL: Reading research on teaching and mathematics	0.21	0.12	0.26**	0.09	0.10	0.06	0.07	0.04	–0.07	0.06	–0.05	0.05
Program coherence	0.18	0.11	0.08	0.09	–0.06	0.06	0.05	0.05	–0.09*	0.04	–0.03	0.03
Average SES for each program (aggregated from future teachers SES)	0.03	0.09	0.05	0.06	0.21*	0.08	0.21***	0.06	0.04	0.05	0.03	0.05
Beliefs												
Mathematics is a collection of rules and procedures	–0.34***	0.08	–0.12	0.07	–0.14	0.09	–0.15**	0.05	–0.20***	0.05	–0.24***	0.04
Mathematics is better learned through active learning	–0.02	0.08	–0.01	0.06	0.02	0.04	–0.07	0.06	0.00	0.06	–0.04	0.04
ICC (from the unconditional model)	57%		32%		35%		18%		43%		21%	
% of variance explained within programs (level 1)	0%		1%		4%		0%		12%		5%	
% of variance explained between programs (level 2)	31%		31%		46%		58%		84%		86%	

Note. The coefficients have been standardized. Method of estimation: restricted maximum likelihood. The numbers under the SE columns are robust standard errors. Weighting specification: level 1 and level 2 estimates have been weighted and normalized

HLM = hierarchical linear regression models

*$p < .05$; **$p < .01$; ***$p < .001$

exception of Russia (where programs seem more differentiated), a larger proportion of the variance in the MCK assessment results occurs within programs. For the mathematical pedagogical knowledge assessment (MPCK), the higher ICCs are in Russia (32%), Poland (24%), the United States (21%), and Thailand (18%) indicating that a larger proportion of the variance in the assessment results occurs within programs as well.[15] In all cases, the proportion of variance explained by the model is higher between programs than within programs, indicating the influence of program standards or norms (e.g., selection criteria), as well as the influence of programs' OTLs.

Analysis for Countries/Regions with Small Samples Regression analysis was done using the IEA's IDB Analyzer, and it was used to explore the relationships between mathematical knowledge for teaching and program features in the countries with small samples. In regression analysis, the main objective is to explore the relationship between a dependent variable, in this case mathematical knowledge for teaching (defined in this study as MCK and MPCK), and one or more explanatory variables, controlling for the characteristics of individuals who enroll in these programs. The model is described below.

The OLS estimation considers the weights repeated replicates and thus accounts for the sample mean complex design. Unlike HLM, OLS models do not include random coefficients between programs.

Partial Model

The model specification is similar from the one presented above (without random terms for the intercept). The partial model is expressed on the following single equation.

$$Y_i = \beta_0 + \beta_1 (\text{SES})_i + \beta_2 (\text{MFA001})_i + \beta_3 (\text{MFA002})_i + \beta_4 (\text{MFA009})_i + r_i$$

[15] The ICC reports the percent of variance between programs, where 100-ICC is the percent of variance within programs. This is always reported simply as the ICC for percent of variance between groups. The % of variance explained in the last two rows then simply states how much variance each model explains. In Table 14.5, in Chile, 10% of the variance in MCK performance is between programs, 90% is within programs. The model including all variables explains 2% of the variance within programs (i.e., student characteristics do not explain much of the variance in their performance within program); that is, the student characteristics explain 2% of the 90% within programs. The model also explains 30% of the variance between programs; that is, the program characteristics explain 30% of the 10% between programs.

Full Model

The full model is expressed by the following single equation (again, random terms for the intercept are not considered; the beliefs variables are aggregated to the program level).

$$Y_i = \beta_0 + \beta_1 (\text{SES})_i + \beta_2 (\text{MFA001})_i + \beta_3 (\text{MFA002})_i + \beta_4 (\text{MFA009})_i + \beta_5 (\text{MFB1GEOM})_i + \beta_6 (\text{MFB2SLMF})_i + \beta_7 (\text{MFB5READ})_i + \beta_8 (\text{MFB15COH})_i + \beta_9 (\text{MeanSES})_i + \beta_{10} (\text{MFD1RULE})_i + \beta_{11} (\text{MFD2ACTV})_i + \beta_{12} (\text{TARGETs\_g10})_i + r_i$$

The results are given in regression coefficients or unstandardized betas (B), as well as in standardized betas or β in Table 14.6. The unstandardized betas are useful in comparing each independent variable between the regressions. For instance, the prior attainment of future secondary teachers is positively and significantly related to performance in MCK across Chinese Taipei, Germany, Malaysia, Singapore, and Switzerland. The table also presents the standardized beta coefficients or β in parenthesis, which can be used to compare the importance of the independent variables included in the analysis within a country/region (Cohen, Cohen, West, & Aiken, 2002; Nardi, 2006). For instance, within Germany a number of variables are positively associated with the MCK assessment score; the most important association after controlling for future teacher characteristics is having the opportunity to learn (OTL) "school-level mathematics—function, probability and calculus" with a β of 0.37 showing a moderate to strong positive and significant association with high scores in the MCK assessment. In Chinese Taipei, however, the most important positive correlation with MCK and MPCK scores is the opportunity to learn "university-level mathematics" and more specifically "Geometry".

Relationships Between Future Teachers' Background and Mathematical Knowledge for Teaching

Analysis for Countries with Large Samples In the countries with large samples in Table 14.5, higher levels of knowledge are related to higher socioeconomic levels (SES) in the case of MCK in Chile and in Thailand and in the case of MPCK in Chile and the Philippines. The age of future teachers across all countries (with the exception of Thailand) had a negative correlation with MCK and MPCK, meaning that younger future teachers scored higher in the study's assessments; this relationship was significant in the Philippines, Poland, and the United States. Higher levels of knowledge were more common among future secondary male teachers, and this relationship was significant for the MCK assessment in Chile, the Philippines,

Table 14.6 Unstandardized B and standardized B (in parenthesis) OLS regression coefficients for correlations between knowledge measures in each of the mathematics and mathematics pedagogy assessments (across top) and independent variables (left side) for future secondary teachers in the participating countries

	C. Taipei		Germany		Malaysia	
	N = 355		N = 620		N = 357	
Variable	MCK	MPCK	MCK	MPCK	MCK	MPCK
Future teacher characteristics						
SES	1.75 (0.02)	2.44 (0.02)	10.14* (0.10)	5.88 (0.06)	5.20 (0.08)	4.23 (0.05)
Age	−1.82 (−0.05)	−0.34 (−0.01)	2.09* (0.13)	0.72 (0.04)	0.14 (0.01)	−0.39 (−0.01)
Gender [1 = F; 0 = M]	−27.65** (−0.18)	−4.27 (−0.02)	−30.95** (−0.15)	−15.20 (−0.07)	−11.75** (−0.09)	−1.92 (−0.01)
Prior attainment	7.42* (0.11)	3.17 (0.04)	23.49** (0.24)	17.35** (0.17)	7.76** (0.15)	9.47** (0.15)
Program characteristics						
OTL: University level mathematics: Geometry	62.51** (0.27)	87.23** (0.30)	−7.03 (−0.04)	−4.28 (0.02)	79.66 (0.68)	46.20 (0.33)
OTL: School level mathematics: Function, probability and calculus	19.44* (0.09)	−5.71 (−0.02)	54.29** (0.37)	40.45* (0.27)	−110.88 (−0.36)	−128.14 (−0.35)
OTL: Reading research on teaching and mathematics	−6.86* (−0.08)	2.37 (0.02)	−12.23* (−0.06)	−8.20 (−0.04)	182.40 (1.10)	146.67 (0.74)
Program coherence	1.66 (0.01)	−7.05 (−0.04)	−2.87 (−0.01)	−10.59 (−0.05)	−8.98 (−0.09)	11.55 (0.09)
Average SES for each program (aggregated from future teachers SES)	11.93 (0.03)	71.91** (0.15)	0.23 (0.00)	2.75 (0.01)	−236.60 (−0.66)	−205.40 (−0.47)
Beliefs						
Mathematics is a collection of rules and procedures	−21.25 (−0.06)	−33.30 (−0.07)	−107.75** (−0.15)	−111.17** (−0.15)	−82.69 (−0.33)	−112.54 (−0.37)
Mathematics is better learned through active learning	10.00 (0.03)	28.08 (0.08)	−5.46 (−0.03)	−18.24 (−0.09)	21.51 (0.09)	13.23 (0.05)
R^2	.16	.12	.39	.18	.24	.13
F	5.93***	4.25***	35.33***	12.13***	9.90***	4.68***

(continued)

Table 14.6 (continued)

	Singapore		Switzerland	
	N = 371		N = 137	
Variable	MCK	MPCK	MCK	MPCK
Future teacher characteristics				
SES	16.44** (0.22)	11.63** (0.11)	−6.89 (−0.13)	−22.51** (−0.12)
Age	−3.98** (−0.26)	−3.74** (−0.18)	−1.78 (−0.16)	−0.66 −0.04
Gender [1 = F; 0 = M]	−22.86** (−0.19)	−30.92** (−0.18)	−9.12 (−0.09)	−21.70 (−0.15)
Prior attainment	8.70** (0.14)	7.56* (0.09)	13.00** (0.24)	10.09** (0.13)
Program characteristics				
OTL: University level mathematics: Geometry	77.97 (0.52)	55.21 (0.27)	−32.85 (−0.28)	−38.40 (−0.23)
OTL: School level mathematics: Function, probability and calculus	−165.49 (−0.82)	−155.65 (−0.56)	22.03 (0.18)	44.13 (0.25)
OTL: Reading research on teaching and mathematics	183.05 (0.48)	168.27 (0.32)	−8.64 (−0.15)	−21.99 (−0.27)
Program coherence	−16.63 (−0.05)	−11.07 (−0.02)	8.03 (0.15)	13.05 (0.17)
Average SES for each program (aggregated from future teachers SES)	−122.28 (−0.22)	−91.04 (−0.12)	8.19 (0.03)	66.49 (0.19)
Beliefs				
Mathematics is a collection of rules and procedures	−17.82 (−0.02)	−31.12 (−0.02)	−96.74 (−0.53)	−78.83 (−0.30)
Mathematics is better learned through active learning	−55.01 (−0.13)	−34.15 (−0.06)	−35.09 (−0.29)	−12.26 (−0.07)
R^2	.30	.10	.19	.16
F	13.98***	3.62***	2.66**	2.16*

OLS = ordinary least squares linear regression
**p* < .05; ***p* < .01

Poland, and the United States, and, for the MPCK assessment, for Poland and the United States. Only in the Russian Federation did females have significantly higher scores than males in the assessments. As expected, prior attainment levels were positively and strongly related to higher results in both assessments in all countries, with the exception of the Philippines and Poland in the MPCK assessment.

Analysis for Countries/Regions with Small Samples For the countries/regions with small samples in Table 14.6, the relationship of socioeconomic status with the level of mathematical knowledge for teaching demonstrated by future teachers was positive overall, as shown for MCK in Germany and Malaysia (β = 0.10, $p < .05$, β = 0.08, $p < .10$), and in Singapore in both assessments (with β = 0.22, $p < .01$; β = 0.11, $p < .05$). In Chinese Taipei and in Singapore, younger future teachers scored higher in the study's assessments, but this was not the case in Germany. Male future teachers did significantly better in the mathematics portion of the assessment in Chinese Taipei, Germany, and Malaysia, while in Singapore and in Switzerland, they did well in both MCK and MPCK. Only in the Russian Federation did females perform better on both MCK and MPCK assessments. Prior attainment seems to play a significant role in mathematical knowledge across most countries (with βs ranging from $\beta = 0.24$ to $\beta = 0.09$).

The next sections describe the relationship between program characteristics and teacher education outcomes controlling for background characteristics.

Relationships Between Learning Opportunities Available to Future Teachers in Their Teacher Education Programs and Mathematics Knowledge for Teaching

Analysis for Countries with Large Samples The type and number of topics studied in university-level geometry had a positive and significant correlation with the study's MCK assessment scores in Chile, Poland, the Russian Federation, and the United States ($\beta = 0.28$, $p < .001$); and with the MPCK assessment in the Philippines, the Russian Federation, and the United States ($\beta = 0.15$, $p < .05$). The number of topics covered in school-level mathematics—specifically function, probability and calculus—had an overall positive and significant correlation with the MPCK scores demonstrated in the study's assessments of future teachers' in Thailand and in the United States ($\beta = 0.13$, $p < .05$). Future secondary teachers in programs that emphasized reading research connected with mathematics teaching and learning scored higher in the MCK and MPCK assessments in Russia, in the Philippines (in MPCK), and in Thailand (in MCK). Program coherence had a positive correlation with the assessment results in Chile (specifically in the MPCK assessment), and in the Philippines (in the MCK assessment).

Analysis for Countries/Regions with Small Samples Among the countries/regions with small samples (in Table 14.6), in Chinese Taipei having the opportunity to learn geometry at the university level was positively and significantly correlated with high scores in the MCK and MPCK assessments ($\beta = 0.27$, $p < .01$; $\beta = 0.30$, $p < .01$). Future secondary teachers who had higher scores in the assessments were in programs that emphasized OTLs school mathematics topics such as functions, probability, and calculus; this is the case in Chinese Taipei (MCK with $\beta = 0.09$, $p < .05$) and in Germany (in both assessments, with $\beta = 0.37$, $p < .01$; and $\beta = 0.27$, $p < .05$). However, unlike the analysis for countries with large samples, higher scorers were not seen in programs that provided opportunities to read research on mathematics teaching and learning (showing small and negative correlations, with β ranging from $\beta = -0.06$ to $\beta = -0.08$, $p < .05$).

Program's Socioeconomic Status

Overall, and across all countries/regions with large and small samples, a program's socioeconomic status had a positive correlation with the level of performance; these correlations were significant only in the Philippines, Thailand, and in Chinese Taipei.

Beliefs

Analysis for Countries with Large Samples There was a general negative correlation between performance in both assessments and the view that learning mathematics consists in mastering rules and procedures. Table 14.5 shows that future secondary teachers who strongly believed that mathematics can be seen as a collection of rules and procedures scored significantly lower in the assessments in Chile (in MCK and MPCK), in Poland (in MPCK), in Russia (in MCK), in Thailand (in MPCK), and in the United States (in MCK $\beta = -0.20$, $p < .001$, and in MPCK with $\beta = -0.24$, $p < .001$). The view that mathematics is better learned through inquiry-oriented learning received a weak endorsement among future teachers in these secondary-level programs, with the exception of Chile, where future secondary teachers in programs espousing such a view scored close to 10 points higher in the MCK assessment.

Analysis for Countries with Small Samples Among the countries with small samples in Table 14.6, Germany shows a negative relationship between performance in both assessments and the view that learning mathematics consists of mastering rules and procedures in MCK and MPCK (both with $\beta = -0.15$, $p < .01$).

In sum, while variable, a number of program features and individual beliefs show an association with teacher education outcomes after controlling for future teachers' background. These findings are discussed in the section below.

Discussion

While much attention in the current policy environment is given to value-added models to evaluate teachers' effectiveness using as key indicators their pupil's achievement in tests, little attention has been given to the study of teacher education as a key factor in the development of knowledgeable and competent teachers. The process of learning to teach has been for the most part assumed, and the immediate outcomes of the programs on teachers' knowledge ignored. Those who enroll in teacher education programs do so expecting to learn what they need to know to become effective teachers. Those who teach in these programs assume that their future teachers will be prepared to teach after a series of courses and field experiences. Yet we have lacked strong evidence to confirm or challenge programs' "learning to teach theory". This lack of evidence has made it difficult for teacher educators to defend their programs from their critics and makes clear that more needs to be learned about whether teacher education programs are accomplishing their goals and what factors seem to contribute to their success or failure.

The data presented in this chapter provides such rigorous evidence using an international and comparative framework. The most important finding is related to the levels of knowledge for teaching secondary-level mathematics that future teachers have attained close to program completion (and are presumably considered ready to teach). The results show that while future secondary teachers do have (in different degrees) MCK and MPCK, many do not have the kind of knowledge that was expected in the participating countries/regions when the assessments were developed. The key question that guided the development of the TEDS-M items and the overall assessments for the mathematicians and mathematics teacher educators in the countries that participated in the study was a normative one: What kind of mathematical knowledge for teaching should future secondary teachers know and be able to demonstrate after completion of their teacher education program's requirements and when they are ready to enter the profession? To answer this question, the participants engaged in a rigorous analysis of the standards for teacher education programs and for the school curriculum in their countries and used those as the frame to develop an assessment of future teachers' knowledge across different countries (see Chap. 4 of this book for an example of the analysis in the U.S., and Tatto & Hordern, 2017). Future teachers were given the assessment prior to graduation and as close as possible to the time they would earn their teaching credential. Consequently, the assessments could be seen as a measurement of the knowledge that is expected future teachers would have attained after undergoing a teacher education program. As explained earlier, some, but not all future secondary teachers were able to reach the basic level of mathematical knowledge measured by the assessments and many failed to reach the more advanced level, with the exceptions being Russia, Chinese Taipei, and Singapore. U.S. teachers' attained knowledge lagged behind the advanced mathematical knowledge of future teachers in countries/regions such as Russia, Chinese Taipei, and Singapore, by 40–130 points, depending on the country/region of comparison.

Mathematics knowledge seems to go hand to hand with the demonstrated levels of mathematics pedagogical content knowledge, which is the exclusive domain of teacher education programs. For instance, U.S. future secondary teachers were able to successfully demonstrate a proficient level of mathematical pedagogical knowledge that was similar to that demonstrated by Poland, yet close to 25 points below Germany and Singapore, and 118 and 40 points below Chinese Taipei and Russian future secondary teachers, respectively. Overall, U.S. future secondary teachers are well prepared to implement the school curriculum, to plan, to evaluate students' work correctly, and to analyze simple students' errors. However, some are challenged when addressing more complex mathematics learning, when attempting to understand or interpret students' thinking, and when asked to engage with more complex concepts and abstract reasoning of mathematical arguments. In sum, as concerns the United States, there is room for improvement in helping future teachers reach advanced levels of mathematics and mathematical pedagogical knowledge (similar conclusions have been reached by Rowland, 2012).

The findings show that across countries/regions the background characteristics of future secondary teachers have an important influence on how knowledgeable they are at the end of their programs (as per the TEDS-M assessments). Two factors associated with high levels of mathematics attainment (as per the TEDS-M assessments) are socioeconomic status and gender. In countries such as Chile, the Philippines, Thailand, Germany, and Singapore, for the most part wealthier individuals do better in the assessments and while not significant in all countries, and with the exceptions of future teachers in the Philippines, the Russian Federation, and Thailand, male future secondary teachers consistently outperform female future secondary teachers in one or both assessments. These findings continue to highlight a concern with social justice and equity frequently voiced by the mathematics education community (see for instance Association of Mathematics Teacher Educators (AMTE), 2015; Bartell, 2013; Boaler, 2002; Lerman, 2000; NCSM and TODOS, n.d.; Turner et al., 2012).

It is clear that much work is needed in providing more equitable opportunities, with respect to gender and socioeconomic status, not only at the program level, but also throughout schooling. Recognizing that teachers are a product of the systems they are in, future work may explore how programs in countries where this pattern is not observed, such as Russia, prepare their future teachers, and how these in turn teach secondary-level pupils.

Not surprisingly, one of the most important factors correlated with high levels of knowledge is previous attainment in mathematics. This finding has important policy implications. In countries where there is a great supply of knowledgeable candidates, it is possible to be highly selective, yet in other countries, where the teaching profession is not as highly regarded and the supply of high-quality candidates is limited, selectivity may not be possible. In situations such as this, teacher education opportunities to learn become more important, as these institutions and programs may be the only venue that can provide mathematics courses—in some cases, remedial—to future teachers.

Data from TEDS-M (Tatto et al., 2012) allows us to document that quality assurance in some countries/regions, such as Chinese Taipei, Russia, and Singapore, has resulted in rigorous selection and exit criteria based on proven knowledge of mathematics. However, in some countries where teacher education programs are less selective, programs play a key role in improving future teachers' knowledge before they begin to teach. Shorter or more experiential and field-based modalities may not be able to compensate for the gaps in mathematics knowledge accumulated in previous schooling (including academic mathematics courses in university) as longer programs might. With respect to this study, it can be argued that in a number of the participating countries, pre-service teacher education programs seem uniquely positioned to provide the mathematical pedagogical content knowledge that can complement teachers' mathematical knowledge by nurturing the ability to develop deeper understandings of mathematics and of how to teach it. This is a particularly important finding that underscores the need for more, not less, teacher education before future secondary teachers are declared as ready to teach and given a credential.

Understanding that teacher education quality is highly dependent on program's selectivity provides information for future policy. An important question for teacher educators, however, is whether their programs contribute knowledge, skills, and dispositions beyond those that individuals bring with them when they enter a program. After controlling for background characteristics, the analysis shows that, among those who did well in the assessments, some program features seemed to make a difference. Findings show that teachers benefit more from programs offering opportunities to learn mathematics (e.g., geometry) at the university level and opportunities to learn school-level mathematics such as functions, probability and calculus and, in the case of Russia, to study research on mathematics teaching and learning. These findings indicate the contributions of teacher education after controlling for background characteristics.

In addition, strong program norms may be able to challenge naïve norms about how mathematics is better learned. Lower levels of knowledge in the assessments were observed among future secondary teachers who hold the general belief that mathematics learning consists of mastering a series of rules and procedures. Future teachers whose philosophy was critical of such beliefs did significantly better in the assessments. The finding that those who seem less knowledgeable espouse the beliefs that mathematics can be taught and learned as a set of rules and procedures represents a challenge to the teacher education community and requires critically considering whether secondary programs may be reinforcing or simply not challenging these beliefs. It is worth asking whether teachers with these beliefs may reproduce the same detrimental association among their students when (and if) they enact these beliefs in their teaching. A long period of study allowing for in-depth examination of alternative beliefs about teaching and learning mathematics beyond those acquired through the apprenticeship of observation is needed to provide future teachers with a framework that will enable them to adapt instruction to the various needs of a diverse student population.

This chapter based on the TEDS-M study data highlights the importance of developing rigorous evaluation studies to understand and improve the degree to which teacher education succeeds in helping teachers acquire important knowledge and abilities to teach secondary mathematics. The study described in this chapter presents an assessment model developed collaboratively by teacher educators, which can potentially be used to evaluate the outcomes of teacher education programs in systematic and useful ways.

The TEDS-M study was designed to measure the knowledge outcomes of teacher education and represents an important step in advancing our understanding of the association between high-quality teacher education and the knowledge needed for initial teaching practice.

References

Adler, J. (2017). Mathematics in mathematics education. *South African Journal of Science, 113*(3/4). Retrieved from https://doi.org/10.17159/sajs.2017/a0201

Association of Mathematics Teacher Educators (AMTE). (2015). *Position: Equity in mathematics teacher education*. Retrieved from https://amte.net/sites/default/files/amte_equitypositionstatement_sept2015.pdf

Association of Mathematics Teacher Educators (AMTE). (2017). *Standards for preparing teachers of mathematics*. Retrieved from http://amte.net/standards

Ball, D. L. (1991). Research on teaching mathematics: Making subject matter knowledge part of the equation. In J. Brophy (Ed.), *Advances in research on teaching, Volume 2: Teachers knowledge of subject matter as it relates to their teaching practice* (pp. 1–48). Greenwich, CT: JAI Press.

Ball, D.L. (Chair). (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education, RAND mathematics study panel*. Washington, DC.: U.S. Department of Education's Office of Educational Research and Improvement Retrieved from http://www.rand.org/pubs/research_briefs/RB8023/index1.html

Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 83–104). Westport, CT: Ablex.

Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 annual meeting of the Canadian Mathematics Education Study Group* (pp. 3–14). Edmonton, AB: CMESG/GDEDM.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*, 389–407.

Bartell, T. G. (2013). Learning to teach mathematics for social justice: Negotiating social justice and mathematical goals. *Journal for Research in Mathematics Education, 44*(1), 129–163.

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180. https://doi.org/10.3102/0002831209345157

Beswick, K. (2007). Teachers' beliefs that matter in secondary mathematics classrooms. *Educational Studies in Mathematics, 65*(1), 95–120.

Beswick, K. (2009). School mathematics and mathematicians' mathematics: Teachers' beliefs about the nature of mathematics. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proceedings of the 33 rd annual conference of the international group for the psychology of mathematics education* (Vol. 2, pp. 153–160). Thessaloniki, Greece: IGPME.

Boaler, J. (2002). *Experiencing school mathematics: Traditional and reform approaches to teaching and their impact on student learning*. Mahwah, NJ: Erlbaum.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods Research, 17*(3), 303–316.

Boyd, D., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 416–440.

Brese, F., & Tatto, M. T. (Eds.) (2012). *User guide for the TEDS-M international database*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Carr, P. G. (2016). *Highlights from TIMSS and TIMSS advanced 2015* (the Commissioner's presentation). National Center for Education Statistics. Washinton, DC: Institute of Education Sciences. Retrieved from https://nces.ed.gov/timss/timss2015/

Casey, C., & Childs, R. (2011). Teacher education admission criteria as measure of preparedness for teaching. *Canadian Journal of Education, 34*(2), 3–20.

Clift, R. T., & Brady, P. (2005). Research on methods courses and field experiences. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher education: the report of the AERA panel on* *research and teacher education* (pp. 309–424). Washington, DC: American Educational Research Association.

Clotfelter, C. T., Ladd, H., & Vigdor, J. (2007). *Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects* (CALDER Working Paper 11). Washington, DC: The Urban Institute.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). New York, NY: Routledge.

Common Core State Standards Initiative (CCSS-M). (2010). *Common core state standards for mathematics*. Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., & Deke, J. (2009). *An evaluation of teachers trained through different routes to certification, final report. (NCEE 2009–4043)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education..

Croninger, R. G., Rice, J. K., Rathbun, A., & Nishio, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement. *Economics of Education Review, 26*(3), 312–324.

Crowe, E. (2010). *Measuring what matters: A stronger accountability model for teachers*. Washington, DC: Center for American Progress.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives, 8*(1).

Darling-Hammond, L. (2013, June 18). Why the NCTQ teacher prep ratings are nonsense. *Washington Post*. Retrieved from http://www.washingtonpost.com/blogs/answer-sheet/wp/2013/06/18/why-the-nctq-teacher-prep-ratings-are-nonsense/

Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco, CA: Jossey-Bass.

De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: The Guilford Press.

De Corte, E., Op't Eynde, P., & Verschaffel, L. (2002). Knowing what to believe. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, NJ: Erlbaum.

DOE (Department of Education). (2014). Teacher education issues. Proposed rule. *Federal Register, 79*, 232. Retrieved from http://www.gpo.gov/fdsys/pkg/FR-2014-12-03/pdf/2014-28218.pdf

Ferrini-Mundy, J., & Findell, B. (2000, October). *The mathematical education of prospective teachers of secondary school mathematics: Old assumptions, new challenges*. Paper prepared for the Mathematical Association of America Committee on the Undergraduate Program in Mathematics. Retrieved from http://www.math.purdue.edu/~cowen/BrBag/MAAmetDraft.pdf

Feuer, M. J., Floden, R. E., Chudowsky, N., & Ahn, J. (2013). *Evaluation of teacher preparation programs: Purposes, methods, and policy options*. Washington, DC: National Academy of Education Retrieved from http://naeducation.org/cs/groups/naedsite/documents/webpage/naed_085581.pdf

Floden, R. E. (2012). Teacher value added as a measure of program quality: Interpret with caution. *Journal of Teacher Education, 63*(5), 356–360.

Floden, R., & Meniketti, M. (2005). Research on the effects of coursework in the arts and sciences and in the foundations of education. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher education: the report of the AERA panel on research and teacher education* (pp. 261–308). Washington, DC: American Educational Research Association.

Floden, R. E., McDiarmid, G. W., & Jennings, N. (1996). Learning about mathematics in elementary methods courses. In D. J. McIntyre & D. M. Byrd (Eds.), *Preparing tomorrow's teachers: The field experience. Teacher education yearbook IV* (pp. 225–241). Thousand Oaks, CA: Corwin.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality Retrieved from http://eric.ed.gov/?id=ED521219

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*, 129–146.

Goldhaber, D., Liddle, S., & Theobald, R. (2013). The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review, 34*, 29–44.

Graham, K. J., Portnoy, N., & Grundmeier, T. (2002). Making mathematical connections in programs for prospective teachers. *Proceedings of the twenty-fourth annual meeting of the North American chapter of the international group for the psychology of mathematics education* (Vol. 4), 1930–1932.

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teacher: Re-imagining teacher education. *Teachers and Teaching: Theory and Practice, 15*(2), 273–290.

Handal, B. (2003). Teachers' mathematical beliefs: A review. *The Mathematics Educator, 13*(2), 47–57.

Hawkins, E. F., Stancavage, F. B., & Dossey, J. A. (1998). *School policies and practices affecting instruction in mathematics (NCES 98–495)*. Washington, DC: National Center for Education Statistics Retrieved from http://nces.ed.gov/spider/webspider/98495.shtml

Heafner, T., McIntyre, E., & Spooner, M. (2014). The CAEP standards and research on educator preparation programs: Linking clinical partnerships with program impact. *Peabody Journal of Education, 89*(4), 516–532.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Hill, H. C., Sleep, L., Lewis, J. M., & Ball, D. L. (2007). Assessing teachers' mathematical knowledge: What matters and what evidence counts? In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*. Charlotte, NC: Information Age.

Hill, H., Ball, D. L., & Schilling, S. G. (2009). Unpacking pedagogical content knowledge: Conceptualizing and measuring Teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*(4), 372–400.

Hu, L.-t., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

Ingvarson, L., Schwille, J., Tatto, M. T., Rowley, G., Peck, R., & Senk, S. (2013). *An analysis of teacher education context, structure, and quality assurance arrangements in TEDS-M countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Kaplan, L., & Owings, W. A. (2001). Teacher quality and student achievement: Recommendations for principals. *NASSP Bulletin, 85*(628), 64–73.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. National Research Council, Mathematics Learning Study Committee. Washington, DC: National Academy Press. Retrieved from http://www.nap.edu

Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology, 100*(3), 716–725. https://doi.org/10.1037/0022-0663.100.3.716

Langrall, C. W., & Mooney, E. S. (2002). *The development of a framework characterizing middle school students' statistical thinking*. Retrieved from http://icots6.haifa.ac.il/PAPERS/6B3_LANG.PDF

Lerman, S. (2000). The social turn in mathematics education research. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 19–44). Westport, CT: Ablex Publishing.

Levine, A. (2006). *Educating school teachers*. Retrieved from http://www.edschools.org/pdf/Educating_Teachers_Report.pdf

Lewis, L., Basmat, P., Carey, N., Bartfai, N., Farris, E., & Smerdon, B. (1999). *Teacher quality: A report on the preparation and qualifications of public school teachers*. Washington, DC: U.S. Department of Education, National Center for Education Statistics (NCES 1999-080).
Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 102*, 391–410.
Mazzeo, J., Lazer, S., & Zieky, M. J. (2006). Monitoring educational progress with group-score assessments. In R. L. Brennan (Ed.), *Educational measurement* (4th ed.). Westport, CT: Praeger.
McCleod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*. Macmillan: New York, NY.
Mewborn, D. S. (2000, April). *Changing actions vs. changing beliefs: What is the goal of mathematics teacher education*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
Meyer, J. P. (2011). *jMetrik (Version 2.1) Computer software*. Charlottesville, VA: University of Virginia Retrieved from http://www.itemanalysis.com
Mikitovicsa, A., & Crehanb, K. D. (2002). Pre-professional skills test scores as college of education admission criteria. *Journal of Educational Research, 95*(4), 215–223.
Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review, 13*, 125–145.
Monk, D. H., & King, J. (1994). Multi-level teacher resource effects on pupil performance in secondary mathematics and science: The role of teacher subject matter preparation. In R. Ehrenberg (Ed.), *Contemporary policy issues: Choices and consequences in education* (pp. 29–58). Ithaca, NY: ILR Press.
Mullens, J. E., Murnane, R. J., & Willett, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from belize. *Comparative Education Review, 40*(2), 139–157.
Nardi, P. M. (2006). *Interpreting data*. Boston, MA: Pearson.
National Commission on Mathematics and Science Teaching for the 21st Century. (2000). *Before it's too late: A report to the nation*. Retrieved from http://www.ed.gov/americacounts/glenn/toc.html
National Council for Teacher Quality. (2013). *Teacher prep review: A review of the nation's teacher preparation programs*. Retrieved from http://www.nctq.org/dmsView/Teacher_Prep_Review_2013_Report
National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions*. Reston, VA: NCTM.
NCSM & TODOS. (n.d.). *Mathematics education through the lens of social justice*. Retrieved from http://www.todosmath.org/assets/docs2016/2016Enews/3.pospaper16_wtodos_8pp.pdf
Provasnik, S., Kastberg, D., Ferraro, D., Lemanski, N., Roey, S., & Jenkins, F. (2012). *Highlights from TIMSS 2011: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context (NCES 2013-009)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education Retrieved from http://nces.ed.gov/pubs2013/2013009_1.pdf
RAND Mathematics Study Panel. (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education*. RAND Mathematics Study Panel, Deborah Loewenberg Ball, Chair. Retrieved from http://www.rand.org/content/dam/rand/pubs/monograph_reports/MR1643/MR1643.pref.pdf
Raudenbush Stephen, W., Bryk, A. S., Cheong, Y. F., & Congdon, R. (2004). *HLM 6: Hierarchical Linear Modeling*. Chicago, IL: Scientific Software International.
Rowan, B., Correnti, R., & Miller, R. J. (2002, November). *What large-scale, survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools*. Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania.

Rowland, T. (2012). Contrasting knowledge for elementary and secondary mathematics teaching. *For the Learning of Mathematics, 32*(1), 16–21.

Rowland, T. & Turner, F. (2008). How shall we talk about 'subject knowledge' for mathematics teaching? In M. Joubert (Ed.) *Proceedings of the British Society for Research into Learning Mathematics,* 28(2), 91–96.

Sawchuk, S. (2016, January 6). Law could spur changes in teacher requirements. *Education Week, 35* (15), pp. 14–15.

Schmidt, W. H., & Buchmann, M. (1983). Six teachers' beliefs and attitudes and their curricular time allocations. *The Elementary School Journal, 84*(2), 162–172.

Speer, N., & King, K. (2009). *Examining mathematical knowledge for teaching in secondary and post-secondary contexts*. Retrieved from http://sigmaa.maa.org/rume/crume2009/Speer_LONG.pdf

Staub, F. C., & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology, 94*(2), 344–355.

Tatto, M. T. (1996). Examining values and beliefs about teaching diverse students: Understanding the challenges for teacher education. *Educational Evaluation and Policy Analysis, 18*, 155–180.

Tatto, M. T. (1998). The influence of teacher education on teachers' beliefs about purposes of education, roles and practice. *Journal of Teacher Education, 49*, 66–77.

Tatto, M. T. (1999). The socializing influence of normative cohesive teacher education on teachers' beliefs about instructional choice. *Teachers and Teaching, 5*, 111–134.

Tatto, M. T. (2011). Reimagining the education of teachers: The role of comparative and international research. *Comparative Education Review, 55*, 495–516.

Tatto, M. T. (Ed.). (2013). *The Teacher Education and Development Study In Mathematics (TEDS-M). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries* (Technical report). Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Tatto, M. T. & Hordern, J. (2017). The configuration of teacher education as a professional field of practice: A comparative study of mathematics education. In J. Furlong & G. Whitty (eds.), *Knowledge and the study of education: An international exploration* (pp. 255–274). Oxford, UK: Oxford Comparative Education Series, Symposium Books.

Tatto, M. T., & Kularatna, N. G. (1993). The interpersonal dimension of teacher education: Comparing distance education with two other programs in Sri Lanka. *International Journal of Educational Research, 19*, 755–778.

Tatto, M. T., Nielsen, H. D., Cummings, W. C., Kularatna, N. G., & Dharmadasa, D. H. (1993). Comparing the effectiveness and costs of different approaches for educating primary school teachers in Sri Lanka. *Teaching and Teacher Education, 9*, 41–64.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Conceptual framework*. Teacher Education and Development International Study Center, Michigan State University, East Lansing, MI, and IEA.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Thomas, A. M., & Loadman, W. E. (2001). Evaluating teacher education programs using a national survey. *The Journal of Educational Research, 94*(4), 195–206.

Turner, E. E., Drake, C., Roth McDuffie, A., Aguirre, J. M., Bartell, T. G., & Foote, M. Q. (2012). Promoting equity in mathematics teacher preparation: A framework for advancing teacher learning of children's multiple mathematics knowledge bases. *Journal of Mathematics Teacher Education, 15*(1), 67–82.

Usiskin, Z., Peressini, A., Marchisotto, E. A., & Stanley, D. (2003). *Mathematics for high school teachers: An advanced perspective*. Upper Saddle River, NJ: Prentice Hall.

Van Dooren, W., Verschaffel, L., & Onghena, P. (2002). The impact of pre-service teachers' content knowledge on their evaluation of students' strategies for solving arithmetic and algebra word problems. *Journal for Research in Mathematics Education, 33*(5), 319–351.

Venkat, H., & Spaull, N. (2015). What do we know about primary teachers' mathematical content knowledge in South Africa? An analysis of SACMEQ 2007. *International Journal of Educational Development, 41*, 121–130. https://doi.org/10.1016/j.ijedudev.2015.02.002

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89–122.

Wu, M., Ray, A., Wilson, M., & Haldane, S. (2007). *ACER conquest: Generalised item response modelling software (Version 2.0)*. Melbourne, VIC: ACER.

Zeichner, K., & Conklin, H. (2005). Teacher education programs. In M. Cochran-Smith & K. Zeichner (Eds.), *Studying teacher preparation: Report of the AERA Panel on Research and Teacher Education* (pp. 645–735). Mahwah, NJ: Lawrence Erlbaum.

Chapter 15
An International Study of the Relationship Between Learning to Teach Students from Diverse Backgrounds and Mathematical Knowledge for Teaching in Future Secondary Mathematics Teachers

Elizabeth B. Dyer

Abstract This study examines the role teacher education plays in developing highly knowledgeable secondary mathematics teachers prepared to work with students from diverse backgrounds. Hierarchical linear modeling is used to investigate the relationship between opportunities to learn to teach students from diverse backgrounds during teacher preparation and teachers' mathematical knowledge for teaching using the TEDS-M international dataset. In some countries, a negative relationship within teacher preparation programs was found: Teachers with more opportunities to address the learning needs of students from diverse backgrounds have lower levels of mathematical knowledge for teaching. These results suggest that teachers with tools for addressing the learning needs of students from diverse backgrounds may lack adequate mathematical preparation.

Introduction

The achievement gap in mathematics between students of differing socioeconomic status is well documented internationally and has been persistent over time (Organisation for Economic Co-operation and Development, 2013, 2014; Simon, Malgorzata, & Beatriz, 2007). While some countries see greater differences between these groups, this inequity presents a worldwide challenge to development and

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

E. B. Dyer (✉)
STEM Program, WestEd, Redwood City, CA, USA
e-mail: edyer@wested.org

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_15

quality of life. While there are numerous policy recommendations to help reduce inequities in student outcomes at multiple levels of the educational system, a main focus has been to provide equity in access to high-quality learning environments (Simon et al., 2007). As teachers have a strong influence on student learning (Rivkin, Hanushek, & Kain, 2005), equitable access to high-quality teachers is likely essential in reaching goals of equity in student achievement. Unfortunately, gaps in access to high-quality teachers are also prevalent in many countries (Akiba, LeTendre, & Scribner, 2007; Kang & Hong, 2008; Little & Bartlett, 2010; Luschei et al., 2013).

While the inequities are clear, it is unclear what role teacher preparation programs play in providing equitable access to high-quality teachers. This study begins to explore the role of teacher education in supporting equitable access to mathematics teachers. It examines how opportunities for future secondary mathematics teachers to learn about teaching students from diverse backgrounds are related to teachers' mathematical knowledge for teaching at the end of their teacher preparation programs. Therefore, this study takes a first step at identifying whether teachers are prepared as well mathematically as they are for working with diverse student populations.

Reasons can be found for why teachers both could and could not be expected to be equally well-prepared in the mathematical aspects of teaching and in teachers' preparedness to teach students from diverse backgrounds. First, if programs are of high quality in general, these programs would prepare teachers well in all aspects of teaching. However, because disparities in mathematics achievement between different groups of students are long-standing, they would likely persist in teacher preparation programs, particularly if teachers with diverse backgrounds are more likely to go on to teach students from diverse backgrounds. This analysis of the relationship between teachers' mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds is primarily exploratory in nature.

This study uses the TEDS-M international dataset of mathematics teacher preparation to investigate this relationship. As these data include teachers nested in preparation programs, the analysis uses hierarchical linear modeling, separating the relationship among teachers in the same program and between different programs, in order to reduce bias. Additionally, teacher background characteristics and opportunities to learn mathematics will be accounted for to investigate whether these factors drive any relationship found. The results of this analysis could help identify ways in which teacher preparation programs may wish to focus their efforts in attempting to prepare high-quality mathematics teachers for students from diverse backgrounds.

Theoretical Background

High-Quality Mathematics Teachers for Students with Diverse Backgrounds

While many domains of expertise in teaching have been proposed for high-quality mathematics teaching, mathematical knowledge for teaching (MKT) is perhaps the best-researched domain. A variety of knowledge related to mathematics, or mathematical knowledge for teaching, is theorized to be used by mathematics teachers in their everyday practice (Ball, Thames, & Phelps, 2008; Hill, Sleep, Lewis, & Ball, 2007). Two domains of knowledge have been shown to be related to student learning: common content knowledge and specialized content knowledge (Hill et al., 2008; Hill, Rowan, & Ball, 2005; Hill, Umland, Litke, & Kapitula, 2012). Common content knowledge includes knowledge that would normally be considered part of mathematics, and includes the mathematics that teachers might expect to teach to students. On the other hand, specialized content knowledge includes the mathematical knowledge that is particular to the domain of teaching. Knowledge of different solution strategies students use and their accuracy or of different ways of representing mathematical concepts (e.g., different representations of ¼) are examples of this type of knowledge.

While research on mathematical knowledge for teaching has assumed its importance for teaching all types of students, research on effective teaching for students from diverse backgrounds has assumed teachers have special pedagogical knowledge and skills for teaching different students. This idea is supported by research that begins to show that, depending on their backgrounds—including cultural, racial, socioeconomic, and linguistic background—students benefit more or less from teachers with particular backgrounds or particular styles of teaching (Dee, 2005; Villegas & Irvine, 2010). Generally, teaching that is considered more equitable incorporates understanding of students' backgrounds, often through non-deficit orientations, and works to understand and influence the role of culture, race, and power in the classroom (Delpit, 1995; Hand, 2012; Irvine, 2003). This work suggests that students from diverse backgrounds, particularly backgrounds seen as disadvantaged, have rich sources of knowledge that can be leveraged to support their interest, motivation, and learning in the classroom (González, Andrade, Civil, & Moll, 2001; González, Moll, & Amanti, 2005; Moll, Amanti, Neff, & González, 1992).

Researchers have typically treated these two domains of teacher competence, mathematical knowledge for teaching and pedagogical knowledge for teaching students from diverse backgrounds, as separate. Recently, Turner and Drake (2016) have suggested that teaching mathematics to students from diverse backgrounds requires knowledge at the intersection of these two domains, particularly students' mathematical thinking and cultural funds of knowledge. In this case, teachers need a deep understanding of children's multiple mathematical knowledge bases. Turner et al. (2012) suggest that effective teaching requires teachers to attend to and make use of the broad and often cultural resources that students bring with them into the classroom to support students' mathematical learning in particular. However, in practice, teachers may be more skilled at noticing and leveraging the resources students from more privileged backgrounds bring into the classroom (Battey & Franke, 2013; Valencia, 1997). Therefore, effective mathematics teaching for students from more disadvantaged or non-dominant backgrounds is supported by knowledge and skills around understanding the resources those particular students bring for learning mathematics.

This study examines all three of these domains of teaching based on the assumption that they are all important for equitable mathematics instruction. In particular, this vision of equitable mathematics teaching suggests that teacher effectiveness is specific to the subject being taught (i.e., mathematics) and the students in the classroom (i.e., students from diverse backgrounds), as well as the intersection of the two. Therefore, equal access to teachers who are either well prepared mathematically or well prepared for working with students from diverse backgrounds, but not both, may be insufficient for promoting equity. Instead, students need equal access to teachers who are effective mathematics teachers for students of their particular background. Furthermore, this idea of effective teaching suggests teachers need more comprehensive preparation for mathematics teaching.

Teacher Preparation for Teaching Mathematics to Students from Diverse Backgrounds

While countries have different systems for preparing and developing teachers on different time frames, most countries emphasize formal teacher preparation (Schwille, Ingvarson, & Holdgreve-Resendez, 2013). In many cases, teacher preparation programs have limited time, but numerous goals, meaning programs must decide which aspects of teacher preparation are most important to develop before teachers are put in classrooms.

Mathematical development of teachers is a generally accepted focus for teacher preparation programs (Conference Board of the Mathematical Sciences [CBMS], 2001, 2012; Tatto et al., 2008). This area is critically important for secondary mathematics teachers, as teachers entering into teacher preparation programs may have just learned the mathematics they plan to teach. Some experts have suggested teachers need to learn an additional five years of mathematics beyond what they will be teaching (CBMS, 2012). Therefore, many secondary teacher preparation programs have a central goal of developing teachers' mathematical content knowledge. Programs often accomplish this goal through mathematics content courses in university-level mathematics, most of which are not designed specifically for teachers. Programs have started to recognize the importance of developing teachers' specialized content knowledge in addition to more advanced common content knowledge (CBMS, 2001, 2012).

In addition to teachers' mathematical preparation, teacher preparation programs focus on teachers' pedagogical preparation. Several countries have recently been placing more emphasis on pedagogical preparation in learning to teach students from diverse backgrounds (Schwille, Ingvarson, & Holdgreve-Resendez, 2013). However, it is unclear whether this emphasis is widespread, particularly at the secondary level. In the area of learning to teach students from diverse backgrounds, programs have often focused on shifting teachers' attitudes and beliefs with respect to students of diverse backgrounds, as many future teachers hold deficit-oriented views that are seen as unproductive for equitable teaching approaches (Castro, 2010; Foote et al., 2013). Research has identified several types of experiences for future teachers that support developing productive beliefs and pedagogical skills, including cross-cultural experiences (Adams, Bondy, & Kuhel, 2005; Garmon, 2004, 2005; Whipp, 2013) and courses focused on culturally responsive teaching or teaching for social justice (Freedman & Appleman, 2009).

Although programs often include general pedagogical preparation for teaching students from diverse backgrounds, little research has looked at whether mathematics-specific approaches are widespread. Several promising new experiences for future teachers have been developed and researched at the primary level that integrate mathematics with preparation for students from diverse backgrounds. For example, identifying mathematical practices and funds of knowledge through community engagement (Bartell et al., 2010), analyzing videos of students engaging in mathematics with equity-oriented analytical lenses (Aguirre et al., 2012; McDuffie, Foote, Bolson, et al., 2014; McDuffie, Foote, Drake, et al., 2014), and developing mathematics lesson plans (Aguirre et al., 2013) can all shift teachers' focus to consider the broad resources students bring into the classroom in ways that inform mathematics teaching.

Relationship Between Preparation in Mathematics and Preparation for Diversity

Given these different approaches to preparing teachers to teach mathematics to students from diverse backgrounds, this study examines whether teachers are prepared equally in mathematical knowledge for teaching and teaching students from diverse backgrounds. The analysis accomplishes this aim by examining the relationship between future teachers' mathematical knowledge for teaching at the end of their teacher preparation programs and their self-reported opportunities to learn to teach students from diverse backgrounds. This study answers the following research questions:

How are opportunities for secondary future mathematics teachers to learn to teach students from diverse backgrounds associated with mathematical knowledge for teaching?

1. Is this association found within teachers in the same program and between different teacher preparation programs?
2. Does this association vary by teacher preparation program and by country?
3. Do teacher background characteristics or opportunities to learn mathematics partly account for any association found?

There are several reasons why teachers with more opportunities to learn to teach students from diverse backgrounds might have lower levels of mathematical knowledge for teaching. Teachers who hope to teach students from diverse backgrounds may begin teacher preparation with less strong mathematical backgrounds, inasmuch as these teachers may be drawn to teaching to bring about educational equity rather than for their love of mathematics. Teachers with this motivation to enter teaching may have a less strong mathematics background and may seek out fewer opportunities to strengthen their mathematical background. Additionally, teachers often take jobs in the area where they grew up (Boyd, Lankford, Loeb, & Wyckoff, 2005; Reininger, 2012), so teachers with more opportunities to learn to teach students from diverse backgrounds may have diverse or non-dominant backgrounds themselves. In this way, the differences seen in student achievement would be reinforced by the teachers who take jobs in disadvantaged schools. Finally, teachers who have more opportunities to learn to teach students from diverse backgrounds may have fewer opportunities to learn mathematics simply because they use their time to pursue these pedagogical opportunities rather than opportunities to learn mathematics. These teachers would then end up learning less mathematical knowledge for teaching by the end of their teacher preparation programs, leading to a negative relationship of mathematical knowledge for teaching with opportunities to learn to teach students from diverse backgrounds.

However, there are also reasons to expect a positive relationship between opportunities to learn to teach students from diverse backgrounds and mathematical knowledge for teaching. Programs may be higher quality overall, meaning that teachers generally have similar quality opportunities to learn in all areas of teacher preparation. Additionally, if teacher preparation programs develop teachers' abilities to teach students from diverse backgrounds in a way that is specific to mathematics teaching, teachers with more of those opportunities may increase their mathematical knowledge for teaching at the same time. In particular, developing an understanding of and skills associated with students' multiple mathematical knowledge bases is likely to develop teachers' pedagogical content knowledge, as well as make teachers more effective when teaching students from diverse backgrounds.

The analysis in this study tests these varied explanations for the existence of both positive and negative relationships using different analytical models. Teacher background likely plays a role in determining teachers' mathematical knowledge for teaching. These background characteristics include gender, age, native language, socioeconomic status, and racial or ethnic background. Other characteristics, such as reasons for entering teaching and previous educational achievement, could have an impact on any relationship between mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds. In addition to teacher characteristics, teachers' experiences during teacher preparation likely will have an influence on any relationship seen. In particular, opportunities to learn mathematics could help explain any relationship found. For example, teachers could spend less time learning mathematics in order to have more opportunity to learn to teach students from diverse backgrounds, leading to lower mathematical knowledge for teaching. Because the theory presented suggests these teacher characteristics and opportunity to learn factors may account for the relationship in question, they are the variables tested.

Methods

Data

The TEDS-M international dataset is used as the basis for this study. The TEDS-M project was a large-scale international data collection effort around mathematics teacher preparation programs in 15 countries. All countries had a response rate of at least 76 percent for future teachers. The study was conducted to investigate how policies and practices in teacher preparation programs influence what beginning mathematics teachers know and can do.

Populations and Samples The TEDS-M dataset consists of nationally representative probability samples of future teachers.[1] All sampling was completed in accordance with IEA's quality standards (see IEA, 2007; Dumais, Meinck, Tatto, Schwille, & Ingvarson, 2013; Dumais & Meinck, 2013, for sampling details). It used a stratified multistage probability sampling design for secondary future teachers in their last year of training. Institutions were first sampled from the list of all institutions that prepare future mathematics teachers, which was provided by each country. Within each institution selected, all programs that prepared teachers, or teacher preparation units, were included in the sample. Teachers within each of the programs were sampled to reach the TEDS-M precision requirements of at least 30 teachers. In cases of programs with fewer than 30 teachers, all of the teachers were surveyed. Because of this variability in the number of teachers in each program in the data, some programs have fewer than 10 teachers, while others have up to 30, creating an unbalanced panel. In total, the sample includes 8,207 future secondary teachers and 381 teacher preparation programs. Individual country sample sizes are shown in Table 15.1. Dumais and Meinck (2013) used balanced repeated replication based on the sampling design and created the estimation weight. This study uses weights in all the analyses allowing for population estimates.

Measurement The TEDS-M study collected data about the future teachers' general background, opportunity to learn during their program, mathematics and mathematical pedagogical content knowledge for teaching, and beliefs about mathematics and teaching. In this study, three main variables are investigated: *mathematical content knowledge (MCK)*, *mathematics pedagogy content knowledge* (*MPCK*), and *opportunity to learn teaching for diversity* (referred to as *OTL DIVERSITY* in this study). Additional variables, including teacher background characteristics and opportunity to learn mathematics, are used to test alternate hypotheses in the three sub-research questions. Reliabilities for the scaled scores used in this study can be found in Table 15.2. Details about item development, assessment frameworks, and scaling can be found in Tatto et al. (2008, 2013).

[1] In the United States, the dataset contains a nationally representative sample of future teachers in public institutions only.

Table 15.1 Country by country program and future teacher sample size

Country	Teacher preparation programs	Teacher participants
Botswana	3	53
Chile	37	746
Chinese Taipei	19	365
Georgia	7	78
Germany	28[a]	771
Malaysia	6	389
Oman	8	268
Philippines	48	733
Poland	35	298
Russia	48	2,141
Singapore	4	393
Switzerland	8	141
Thailand	53	652
United States	72	607
Total	381	8,207

Note: [a]Although 28 programs were sampled in Germany, programs were not identified for teachers in the sample from Germany. Programs were taken into account in determining sampling weights

Table 15.2 Reliabilities for scaled scores using congeneric measurement model

Scale	Reliability
Mathematics Content Knowledge (MCK)[a]	.91
Mathematics Pedagogical Content Knowledge (MPCK)[a]	.72
Opportunity to learn to teach students from diverse backgrounds[a]	.90
Teaching for impact and change	.77

[a]Tatto et al. (2013)

Mathematical Knowledge for Teaching Teachers' mathematical knowledge for teaching (MKT) was measured through assessments in two domains: mathematical content knowledge (MCK) and mathematical pedagogical content knowledge (MPCK). These instruments were based on previous instruments developed for MKT, including the instrument developed for the Mathematics Teaching in the 21st Century study, Knowing Mathematics for Teaching Algebra Project, and the Learning Mathematics through Teaching Project. The MCK instrument is focused on common and horizon content knowledge (i.e., the mathematics taught in schools at or slightly above secondary grades). The MPCK instrument includes specialized content knowledge (i.e., mathematical knowledge specific to the profession of teaching), as well as knowledge of content and students (i.e., knowledge of how students think about mathematics). While these two domains do not provide a

Table 15.3 Items in the opportunity to learn to teach students from diverse backgrounds scaled score

In your current teacher preparation program, how frequently did you engage in activities that gave you the opportunity to learn how to do the following?
Develop specific strategies for teaching students with behavioral and emotional problems
Develop specific strategies and curriculum for teaching pupils with learning disabilities
Develop specific strategies and curriculum for teaching gifted pupils
Develop specific strategies and curriculum for teaching pupils from diverse cultural backgrounds
Accommodate the needs of pupils with physical disabilities in your classroom
Work with children from poor or disadvantaged backgrounds

comprehensive measurement of teachers' mathematical knowledge for teaching, they are two domains that have been shown to be associated with teaching quality, and both are main areas of focus for many teacher preparation programs (CBMS, 2001, 2012). These two domains were measured through two separate assessments designed specifically for secondary mathematics teachers, which used a balanced-incomplete block design to reduce time required to complete the assessments. Tatto et al. (2013) developed scaled scores for each assessment in consultation with ACER using the standard Rasch model for dichotomous items and the partial credit model for polytomous items. The scores were standardized to a mean score of 500 and standard deviation of 100. Tatto et al. (2013) calculated the reliability estimate of .91 for MCK and .72 for MPCK using a congeneric measurement model.

Opportunity to Learn The TEDS-M study measured teachers' opportunity to learn many topics in four broad areas: mathematics content, mathematics education pedagogy, general education pedagogy, and school-based experiences. The 24 opportunity-to-learn subscales across the four domains were developed based on logical organization of the topics, according to concordance with curricular areas. Three opportunity-to-learn scales are used in this study.

Opportunities to learn to teach different populations of students with potentially different needs in the classroom is used to measure teachers' opportunities to learn to teach students from diverse backgrounds. Tatto et al. (2013) created the composite score for teachers' opportunities to learn to teach students from diverse backgrounds that is used in this study based on six items (listed in Table 15.3) using Rasch modeling. These scales were centered on a value of 10, which corresponded to the middle of the response scale. These items had a four-option response scale of *never*, *rarely*, *occasionally*, and *often*. Tatto et al. (2013) calculated the reliability estimate of .90 for this scale using a congeneric measurement model.

Two composite scores for school-level mathematics (i.e., secondary mathematics) are used in this study: (a) functions, probability and calculus and (b) numbers, measurement, and geometry. The functions composite contains three items, while the numbers composite contains four items (a list of items can be found in Table 15.4). These items had two response choices: *studied* or *not studied*. Tatto

Table 15.4 Items in the opportunity to learn school-level mathematics composites

Consider the following list of mathematics topics that are often taught at the secondary school level. Please indicate whether you have studied each topic as part of your current teacher preparation program. (options: studied, not studied)
Numbers, measurement, and geometry composite
Numbers (e.g., whole numbers, fractions, decimals, integer, rational, and real numbers; number concepts; number theory; estimation; ratio and proportionality)
Measurement (e.g., measurement units; computations and properties of length, perimeter, area, and volume; estimation and error)
Geometry (e.g., 1-D and 2-D coordinate geometry, Euclidean geometry, transformational geometry, congruence and similarity, constructions with straightedge and compass, 3-D geometry, vector geometry)
Functions, probability and calculus composite
Functions, Relations, and Equations (e.g., algebra, trigonometry, analytic geometry)
Data Representation, Probability, and Statistics
Calculus (e.g., infinite processes, change, differentiation, integration)
Validation, Structuring, and Abstracting (e.g., Boolean algebra, mathematical induction, logical connectives, sets, groups, fields, linear space, isomorphism, homomorphism)

et al. (2013) created the composite scores used in this study by summing the number of topics marked as studied in each domain. The reliability of these two scales is not calculated because they are simple sums.

Teacher Background A variety of teacher background characteristics were measured through the TEDS-M study. The background characteristics that have been shown to have an influence on mathematics achievement are used in the analysis as control variables (see Barton & Coley, 2009; Reardon, 2011; Reardon & Galindo, 2009; Reardon, Robinson-Cimpian, & Weathers, 2015; Robinson & Lubienski, 2011). These variables include age, gender, self-reported typical level of grades obtained in secondary school, whether the language of the test was typically spoken in their home, and mother's education:

Age. Age was measured through an item asking the teacher to write in her or his age, and the age written was used directly in the model.

Gender. Gender was measured through a question asking teachers to report being male or female, and an indicator variable for female was created for analysis.

Grades. An indicator variable was also created with an item asking about teachers' grades in secondary school. Teachers who reported that their grades were always or usually near the top of their year level were coded as 1, while the other responses of generally above, about, or below average for their year level were coded as 0.

Non-native language speaker. An indicator variable for being a non-native language speaker was created for the item asking about the frequency with which the teacher spoke the language of the test at home. Teachers reporting they spoke the language of the test *sometimes* or *never* were coded as 1, while responses of *always* or *almost always* were coded as 0.

Table 15.5 Items in the teaching for impact scaled score

To what extent does each of the following identify your reasons for becoming a teacher?(options: not a reason, a minor reason, a significant reason, a major reason)
I believe that I have a talent for teaching
I like working with young people
I want to have an influence on the next generation

Mother's education. Although there is no measure of socioeconomic status in the TEDS-M data, mother's education was measured, which can be used as a proxy for SES. Mother's education was reported at a variety of levels, so indicator variables were created for obtaining at most secondary-level and at least tertiary-level education.

Teaching for impact. In addition to the variables found in the TEDS-M dataset, a composite variable about the reasons for becoming a teacher was created to help control for selection in the types of teachers who may seek out more opportunities to learn to teach students from diverse backgrounds. This score was based on three items asking teachers to rate three reasons for becoming a teacher—influencing students, having talent for teaching, and working with young people—on a scale that included *not a reason*, *a minor reason*, *a significant reason*, and *a major reason*. The items can be found in Table 15.5. This composite score (*TEACHING FOR IMPACT*) was not originally created in the released TEDS-M dataset, but was calculated for this study using Rasch modeling with a partial credit structure. All teachers in the sample were scaled with the same model. Reliability was estimated to be .77 using a congeneric measurement model with the jMetrik software.

Analysis

Hierarchical linear modeling was the primary analysis technique used. All analysis was completed using the HLM7 software with the full maximum likelihood estimation method. To investigate relationships internationally, two-level HLM was used with teachers nested within teacher preparation programs. Separate models were run for each country because of the relatively small number of countries in the sample and to refrain from generating average effects internationally, which are difficult to interpret given the wide variability between countries. The relationship between mathematical knowledge for teaching and opportunity to learn to teach students from diverse backgrounds was the main focus. This was modeled with *MKT* as the outcome or dependent variable and *OTL DIVERSITY* as the independent variable. The relationship was modeled as a linear relationship because adding nonlinear terms did not increase model fit and would make the interpretation of the results less straightforward. As two parts of MKT were measured in the data, *MCK* and *MPCK*, separate models were constructed for each outcome. In this study, the models are represented with *MKT* as the outcome variable. Sampling weights were used in all analyses.

Four different model specifications were tested. The first specification is the unconditional specification, which only includes the outcome variable in the model. This specification serves as a basis for the three primary specifications by showing the variation in the outcome variable at each level of the model (the between-program level and the within-program level). The additional three specifications serve two purposes: They allow us to examine the robustness to alternate specification and to test hypotheses about factors that might account for any associations identified. The basic specification (two) only includes the relationship between *OTL DIVERSITY* and *MKT*. The third specification added teacher background characteristics to the model, which allows us to examine whether teacher background characteristics account for any relationship seen between opportunity to learn to teach students from diverse backgrounds and MKT. The fourth specification added opportunity to learn mathematics to the teacher background characteristics, which allows us to investigate whether any relationship seen between opportunity to learn to teach students from diverse backgrounds and MKT continues to exist when controlling for opportunity to learn mathematics. The additional variables in the third and fourth specifications primarily serve as control variables, although they also slightly increase the model fit.

In the second through fourth specifications, the between-program and within-program variation was separated using the centered within context, with reintroduction of the mean at level 2 approach (i.e., a CWC(M) approach). To isolate the within-program variation at level 1, the variables were all group mean centered. This centering reduces exogenous variation in the estimates of the coefficients at the teacher level by removing variation seen due to selection or sorting into different teacher preparation programs. As teachers are certainly not randomly assigned to different programs, and instead are admitted and/or choose particular programs, separating the within and between variation reduces bias in the level-1 coefficient estimates (Enders & Tofighi, 2007; Preacher, Zyphur, & Zhang, 2010; Zhang, Zyphur, & Preacher, 2009). In addition to isolating the within-program variation, the between-program variation was explored within countries. Variables for the mean of each of the teacher-level variables for all the teachers in the same program were included in the equation for the intercept at level 2. The same relationships examined within teachers can now be investigated between different programs in the same county with the addition of these mean values for the teachers in each program. In all of the analyses, the relationship within teacher preparation programs is always separated from the relationship between teacher preparation programs, and is always reported separately.

The parameters estimated for the opportunity to learn to teach students from diverse backgrounds were modeled as random effects in order to investigate the variation of this relationship in different teacher preparation programs. All parameters estimating the slope for the teacher background characteristics and opportunity to learn mathematics were modeled as fixed effects. This final (or most comprehensive) estimated model is shown below for level 1, and for the equation for the intercept at level 2. The subscripts *ij* denote the value for the *i*th teacher in the *j*th program. To simplify the presentation of the model, T_{ij} is used as a vector of variables of teacher characteristics and M_{ij} is used as a vector of variables of opportunity to learn school-level mathematics at the teacher level, and $\overline{T_j}$ and $\overline{M_j}$ are used at the program level.

$$\text{Level 1}: MKT_{ij} = \beta_{0j} + \beta_{1j} OTL\ DIVERSITY_{ij} + \beta_{2j} T_{ij} + \beta_{3j} M_{ij} + r_{ij}$$

$$\text{Level 2}: \beta_{0j} = \gamma_{00} + \gamma_{01} \overline{OTL\ DIVERSITY_j} + \gamma_{02} \overline{T_j} + \gamma_{03} \overline{M_j} + u_{0j}$$

The single equation mixed model representation is given below:

$$MKT_{ij} = \gamma_{00} + \gamma_{01} \overline{OTL\ DIVERSITY_j} + \gamma_{02} \overline{T_j} + \gamma_{03} \overline{M_j} + \gamma_{10} OTL\ DIVERSITY_{ij} + \gamma_{20} T_{ij} + \gamma_{30} M_{ij} + u_{0j} + u_{1j} OTL\ DIVERSITY_{ij} + r_{ij}$$

To test the model fit of each specification, three-level HLM was used in order to determine the model fit across all countries. Teachers were nested within programs, which were nested within countries. The model was equivalent to the two-level models for each country, and the slope of the coefficient on *OTL DIVERSITY* was modeled as a random effect at level 3 as well. In each specification, the model fit was calculated with the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Results in Table 15.6 show that the full model, or fourth specification, generally has the lowest values for both criteria, and, therefore, has the best fit overall. In this study, the analysis is aimed at investigating the relationship between mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds. As such, both the second (basic) and fourth (full) specifications are used in the results section and compared, not just the model with best fit. This comparison of specifications helps to examine whether controlling for additional covariates accounts for the main relationship seen.

Table 15.6 Model fit statistics for 3-level models

Model specification	Mathematics content knowledge	Mathematics pedagogical content knowledge
Akaike information criterion (AIC)	90,558.05	92,850.44
Unconditional		
Basic	87,298.67	89,571.05
Teacher Background	77,540.89	79,809.10
Math Opportunity	77,235.68	79,606.96
Bayesian information criterion (BIC)	90,586.02	92,878.41
Unconditional		
Basic	87,389.12	89,661.50
Teacher Background	77,725.73	79,993.93
Math Opportunity	77,447.84	79,819.11

Note: Decreases in value between models indicate better fit

In the third (background) and fourth (full) model, including teacher background characteristics and opportunity to learn mathematics pose additional challenges at the program level (level 2). In many of the countries, the number of programs is smaller than the number of variables added to the model as controls. In these cases, the level 2 intercept equation does not include additional variables at the program level. The intercept equation still includes the variable for mean opportunity to learn to teach students from diverse backgrounds, such that the within and between relationships for this variable are separated. However, the between-program relationship is not reported in the full model results because of the lack of control variables, making those results not comparable with the models including these control variables. Results for the within-program relationship are still reported as the control variables are included at this level.

Finally, although future teachers from Germany were sampled within programs, and this information was used to determine sampling weights, these program assignments were not included in the data. Therefore, all results presented from Germany combine the relationship within and between programs. These results are shown in the within-program columns in the tables, and no results are shown for between-program relationships. Because of this missing data, the results from Germany should be interpreted with caution, as any bias introduced from selection into programs is not accounted for in the estimates calculated.

Results

Across the different specifications, there is some evidence of a negative relationship between opportunities to learn to teach students from diverse backgrounds and mathematical knowledge for teaching. This relationship is found both within teachers from the same program and across programs. A negative relationship suggests that teachers or programs with more opportunities to learn to teach students from diverse backgrounds tend to have lower levels of mathematical knowledge for teaching. Evidence of a positive relationship was only found between different programs in one country. Teacher background characteristics and opportunity to learn mathematics seem to account for part of this relationship between programs, but do not play a significant role in the relationship within programs. Finally, countries tend to show different results, and many countries show variance in the relationship within different programs. The results for the basic specification are presented first followed by the full specification.

In the basic model results in Table 15.7, some countries show a negative relationship between mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds. For mathematical content knowledge, there is a statistically significant negative relationship between programs in Thailand and the United States (−15.43 and −24.08, respectively). Germany is the only country to show a statistically significant relationship for mathematical content knowledge but this result should be treated with caution because the within and between relationship

Table 15.7 Two-level model results by country for the estimates of the association of mathematical knowledge for teaching with opportunity to learn to teach students from diverse backgrounds in the basic specification (β_{1j} and γ_{01})

Country	N (programs)	N (teachers)	Mathematics content knowledge		Mathematics pedagogical content knowledge	
			Within program	Between programs	Within program	Between programs
Botswana	3	53	−5.80 (16.96)	−19.83 (3.67)	−1.55 (6.33)	36.61 (29.20)
Chile	37	746	1.53 (2.20)	8.67 (5.89)	3.90⁺ (1.96)	−4.82 (5.35)
Chinese Taipei	19	365	1.01 (2.46)	2.81 (19.54)	−3.12 (3.08)	4.41 (19.80)
Georgia	7	78	−16.58⁺ (6.73)	3.64 (9.93)	−7.28 (5.85)	−10.28 (8.97)
Germany	1	771	−4.30* (2.07)		−4.13 (2.77)	
Malaysia	6	389	−2.13 (1.44)	14.87 (34.37)	−3.26 (3.18)	0.28 (26.33)
Oman	8	268	1.68 (1.72)	6.31 (10.35)	−0.42 (2.17)	−3.70 (10.93)
Philippines	48	733	−2.27 (2.27)	6.47 (6.40)	0.51 (2.31)	3.35 (6.26)
Poland	35	298	−6.21⁺ (3.40)	−1.54 (6.37)	−8.95⁺ (4.98)	1.62 (6.40)
Russia	48	2141	0.36 (0.94)	20.06 (13.61)	0.62 (1.33)	17.38⁺ (9.92)
Singapore	4	393	0.82 (2.15)	86.38⁺ (27.94)	−2.48 (3.14)	55.97 (25.86)
Switzerland	8	141	−2.96 (3.04)	14.95 (9.13)	−1.63 (4.45)	−7.15 (13.07)
Thailand	53	652	−0.80 (1.15)	−15.43** (5.45)	−1.19 (1.06)	−8.21⁺ (4.86)
United States	71	606	−5.66 (3.53)	−24.08*** (6.06)	−4.36* (1.84)	−22.87*** (4.87)

Note: Robust standard errors are reported in parentheses and clustered at the program level. Within and between relationships are separated in all countries except Germany where program assignment data was not available.

⁺$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

cannot be separated because of the data missing about program assignment. For MPCK, there is a statistically significant negative relationship for the United States both within and between programs (−4.36 and −22.87, respectively). No countries show statistically significant evidence of a positive relationship for MPCK.

The full model, which accounts for teacher background characteristics and opportunity to learn school-level mathematics, yielded more statistically significant results, as shown in Table 15.8, but both positive and negative results are found. Looking at the results for MCK, there is a statistically significant negative relation-

Table 15.8 Two-level model results by country for the estimates of the association of mathematical knowledge for teaching with opportunity to learn to teach students from diverse backgrounds in the full specification (β_{1j} and γ_{01})

Country	N (Programs)	N (Teachers)	Mathematics content knowledge		Mathematics pedagogical content knowledge	
			Within program	Between programs	Within program	Between programs
Botswana	3	53	−4.59 (4.62)		−21.44 (24.27)	
Chile	37	746	0.00 (2.79)	9.88 (6.22)	2.44 (2.26)	−0.85 (4.35)
Chinese Taipei	19	365	0.51 (2.61)	17.41 (15.07)	−2.54 (3.48)	42.69⁺ (21.04)
Georgia	7	78	−24.05* (6.03)		−9.43 (5.17)	
Germany	1	771	−5.49** (2.12)		−5.37⁺ (2.91)	
Malaysia	6	389	−1.55 (1.50)		−5.19 (4.35)	
Oman	8	268	1.25 (1.88)		−2.39 (2.99)	
Philippines	48	733	−0.70 (2.14)	2.83 (3.08)	1.32 (1.52)	4.82 (4.74)
Poland	35	298	−6.86⁺ (3.79)	−0.96 (8.54)	−7.19 (5.73)	−2.78 (7.27)
Russia	48	2141	−0.59 (1.17)	33.25*** (8.76)	0.31 (1.61)	24.16*** (6.30)
Singapore	4	393	−0.59 (2.12)		−2.88 (3.78)	
Switzerland	8	141	0.56 (3.45)		−3.15 (5.07)	
Thailand	53	652	−1.25 (1.16)	−7.82* (3.64)	−1.31 (1.14)	−6.65 (5.74)
United States	71	606	−4.54 (3.09)	−6.75* (3.23)	−4.37* (1.89)	−13.97*** (4.02)

Note: Robust standard errors are reported in parentheses and clustered at the program level. All estimates are from regressions that control for variables at the teacher level and mean values of teacher variables at the program level. Countries without reported between relationship estimates did not include control variables at the program level due to small program sample size, but still separated the within and between relationships in the model. Controls include age; gender; self-reported typical level of grades obtained in secondary school; whether the language of the test was typically spoken in the respondent's home; mother's education; opportunity to learn functions, probability, and calculus; and opportunity to learn numbers, measurement, and geometry. Within and between relationships are separated in all countries except Germany, where program assignment data was not available.

$^{+}p < .10$; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

ship within programs in Germany (similar to the Basic model), and in Georgia. There is evidence of a negative relationship between programs in Thailand and the United States, similar to the Basic model. Additionally, there is evidence of a positive relationship between programs in Russia. For MPCK, there is still a statistically significant relationship within and between programs in the United States. Finally, there is evidence of a statistically significant positive relationship in Russia between programs.

Across the different specifications, there is evidence for robustness of the results within programs, but less evidence for the results between programs. All results that were statistically significant in the basic model stay statistically significant in the full model. In fact, new statistically significant relationships emerge when adding control variables. However, the estimates or the strength of the relationships between programs generally tend to decrease, suggesting the relationships at this level may suffer from omitted variables bias. Selection bias into teacher preparation programs could account for a large part of any association of MKT with opportunities to learn to teach students from diverse backgrounds between teacher preparation programs. If this were the case, programs with more opportunities for teachers to learn to teach students from diverse backgrounds on average would tend to have teachers with lower levels of MKT simply because different types of teachers attend these programs, rather than because of the program itself. However, the results within programs are relatively stable between the two models, suggesting that the results found are quite robust.

The variation of the association of mathematical knowledge for teaching with opportunities to learn to teach students from diverse backgrounds within teachers from the same program is examined by modeling this relationship as a random effect at level 2. This random effect allows the estimate of this within relationship to vary across different programs in each country, rather than requiring all programs' estimates to be the same. In the random effects results shown in Table 15.9, there is evidence of variation in the estimate of the within relationship in many of the countries, particularly in countries with large numbers of programs. The variance in the relationship is statistically significant (shown in the level 2 *OTL DIVERSITY* columns) in Chile, Poland, and the United States for both outcomes. The variance is statistically significant for MCK in the Philippines and for MPCK in Botswana, Malaysia, and Russia.

Discussion

In the preceding section, a negative relationship was found between teachers' mathematical knowledge for teaching and opportunities to learn to teach students from diverse backgrounds in some countries. This relationship suggests that teachers in some countries are not as prepared in the mathematical aspects of teaching as they are in learning to teach students from diverse backgrounds. Additionally, the relationship varied considerably between countries, with the United States being a

Table 15.9 Two-level model results by country for the variance components and standard deviation of random effects in the full specification

Country	Mathematics content knowledge			Mathematics pedagogical content knowledge		
	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*	Level 1 Intercept	Level 2 Intercept	Level 2 *OTL DIVERSITY*
Botswana	1029.40*** (32.08)	0.10 (0.32)	0.04 (0.20)	1655.39*** (40.69)	15.04 (3.88)	1552.98*** (39.41)
Chile	5635.02*** (75.07)	408.84*** (20.22)	89.87*** (9.48)	6180.76*** (78.62)	12.94* (3.60)	38.42* (6.20)
Chinese Taipei	4356.88*** (66.01)	7.95** (2.82)	8.84 (2.97)	7618.29*** (87.28)	92.66* (9.63)	16.79 (4.10)
Georgia	3845.86*** (62.02)	152.03$^+$ (12.33)	0.33 (0.58)	2833.10*** (53.23)	1.60 (1.26)	0.12 (0.35)
Germany						
Malaysia	1978.04*** (44.48)	755.12*** (27.48)	0.20 (0.45)	3147.18*** (56.10)	397.78*** (19.94)	76.97** (8.77)
Oman	1509.29*** (38.85)	96.07* (9.80)	0.13 (0.36)	3850.55*** (62.05)	0.36 (0.60)	0.12 (0.34)
Philippines	1848.38*** (42.99)	34.62* (5.88)	22.94* (4.79)	3182.55*** (56.41)	0.58 (0.76)	0.23 (0.48)
Poland	2407.08*** (49.06)	401.58*** (20.04)	232.87*** (15.26)	3887.22*** (62.35)	413.11*** (20.33)	484.39*** (22.01)
Russia	4027.46*** (63.46)	2570.37*** (50.70)	2.35 (1.53)	6171.71*** (78.56)	1284.36*** (35.84)	13.06* (3.61)
Singapore	2679.30*** (51.76)	163.62*** (12.79)	0.58 (0.76)	6199.71*** (78.74)	57.32* (7.57)	15.29 (3.91)
Switzerland	2196.70*** (46.87)	6.27 (2.50)	0.44 (0.66)	4672.03*** (68.35)	27.78 (5.27)	2.98 (1.73)
Thailand	2167.76*** (46.56)	363.01*** (19.05)	1.30 (1.14)	3300.75*** (57.45)	288.54*** (16.99)	1.51 (1.23)
United States	1827.29*** (42.75)	104.59*** (10.23)	55.15** (7.43)	3624.40*** (60.20)	52.47 (7.24)	5.81** (33.70)

Note: Standard deviations are reported in parentheses. All estimates are from regressions that control for variables at the teacher level and mean values of teacher variables at the program level. Countries with eight or less programs sampled did not include program level control variables. Controls include: age, gender, self-reported typical level of grades obtained in secondary school, whether the language of the test was typically spoken in their home, mother's education, opportunity to learn functions, probability and calculus, and opportunity to learn numbers, measurement, and geometry. Within and between relationships are separated in all countries except Germany where program assignment data was not available.

$^+p < .10$; $^*p < .05$; $^{**}p < .01$; $^{***}p < .001$

consistent outlier showing the strongest negative relationship both within teachers from the same program and between programs. Finally, the relationship found within teachers from the same preparation program varied across programs in different countries, which suggests that countries that did not show an average negative relationship are still likely to have programs with negative relationships, as well as programs with positive relationships.

Explanations of the Relationship Found

This negative relationship between opportunities to learn to teach students from diverse backgrounds and mathematical knowledge for teaching could exist for several reasons. Selection likely drives the relationship found rather than direct or indirect causal mechanisms. This study examined one indirect causal mechanism, less opportunity to learn mathematics, and did not find it accounted for the relationship. Additionally, no causal explanations were theorized for a negative relationship, so a selection explanation is more plausible.

Explanations based on selection mechanisms suggest that teachers who report increased opportunities to learn to teach students from diverse backgrounds are different in ways that contribute to lower levels of mathematical knowledge for teaching at the end of teacher preparation. The previous analysis controls for some aspects of what might make these teachers different. However, the TEDS-M data does not include some background characteristics of teachers that likely contribute to the differences. For example, TEDS-M did not collect data on racial and ethnic background of teachers, which is likely associated with lower levels of mathematical content knowledge as well as teachers seeking more opportunities to learn to teach students from diverse backgrounds. Additionally, race may help explain why this study finds stronger relationships in the United States—that is, because of its high level of racial diversity and the history of differences in educational achievement by race. More exploratory work could identify additional factors that may account for the relationship seen.

In addition, more data on teacher background characteristics and longitudinal data collection on future teachers would help identify possible explanations for the relationship found. These data could help to distinguish between differences at the end of teacher preparation and changes in differences from the beginning to end of teacher preparation using a difference in difference approach. While these data would still not account for differences in teachers' ability to learn, they would provide a first step toward identifying the unique influence of teacher preparation programs without random assignment.

The TEDS-M dataset only includes data about teachers at the very end of their program, meaning analysis cannot examine changes over time in any relationship found. In fact, the negative relationship identified may have diminished, widened, or stayed the same during teacher preparation. This type of pre-post data on teachers' mathematical knowledge for teaching is essential in order to tease out the impact teacher preparation has on the differences seen. The relationship with mathematical pedagogical content knowledge may better measure the influence of teacher preparation programs than mathematical content knowledge, as teachers have little opportunity to develop knowledge about learning and teaching mathematics before teacher preparation. Given that this study did not find a clear trend of stronger or more frequent significant results for MCK than for MPCK, the relationship found may arise during teacher preparation programs. However, longitudinal data on teachers would provide much stronger evidence of the unique contribution of teacher preparation.

Unfortunately, for several reasons, this type of pre-post data collection is particularly complex in the case of international teacher preparation. First, programs in different countries have quite different lengths and identify future teachers are different times. Therefore, researchers have difficulty selecting comparable times to collect pre-data across different countries, and efforts of individual countries to design this type of data collection may be more successful than international efforts. Researchers could also collect data on future teachers longitudinally, perhaps at the end of each year of their teacher preparation programs. In this case, countries could have varying numbers of data collection points depending on the length of their programs. This design would allow for appropriate comparisons based on the number of years before the end of teacher preparation programs. With even larger data collection efforts, researchers could track future teachers starting in secondary school, similar to the longitudinal studies following secondary students through post-secondary education (e.g. the HSLS dataset). This type of longitudinal data would provide strong evidence for reproduction of differences in achievement in schooling and teacher preparation. Finally, study designs that collect data at the beginning and end of future teachers' final year of teacher preparation programs may be most feasible. While teachers do not complete all relevant coursework in the final year, this data could still contribute to understanding the unique impact of teacher education programs.

Implications for Research, Teacher Preparation, and Equitable Access to Math Teachers

The differences in hypothesized explanations for the relationship found suggest that further research in this area is greatly needed. Future research would benefit from additional measures of teachers' opportunities to learn to teach students from diverse backgrounds. The TEDS-M dataset includes six items asking about opportunities with different populations of students in general. Future data collection could focus on different practices or activities involved in teaching students from diverse backgrounds, such as incorporating students' cultural funds of knowledge into teaching decisions, designing curriculum, or developing better understanding of student communities. Additionally, questionnaires could ask about experiences teachers have with students with different backgrounds in their teacher preparation programs, both in school settings and out of school settings. Research on developing equitable mathematics teachers suggests these experiences can be particularly helpful for future teachers (Adams et al., 2005; Garmon, 2004, 2005; Whipp, 2013).

In addition to more detailed data collection on teachers' experiences in teacher preparation with respect to teaching students from diverse backgrounds, future research would benefit from continuing to follow teachers as they start their teaching careers. These data would allow researchers to see whether the differences seen at the end of teacher preparation programs contribute to differences in teacher quality in teachers' first year of teaching. Additionally, it would allow for an empirical investigation of how different teacher preparation is associated with teachers working in different school and community contexts. Specifically, research could examine whether teachers with increased opportunities to learn to teach students from diverse backgrounds are more likely to teach successfully students from diverse backgrounds in nationally representative samples.

This study has important implications regarding the overall quality of teachers' preparation. Specifically, teachers are not equally prepared in mathematics and in learning to teach students from diverse backgrounds in some countries. Teachers tend to be better prepared in one area or the other, rather than being well prepared in both. In particular, teachers with strong mathematical preparation are less prepared for the unique challenges of teaching students from diverse backgrounds, which could result in the larger proportion of teachers seen leaving the profession in schools that serve diverse student populations (Boyd et al., 2011; Boyd et al., 2005; Guarino, Santibañez, & Daley, 2006; Hanushek, Kain, & Rivkin, 2004; Ingersoll, 2001; Smith & Ingersoll, 2004). Conversely, teachers who are better prepared to teach students from diverse backgrounds may be less effective teachers of mathematics because of their lower levels of content knowledge. Combining the two tendencies, the majority of teachers may not be adequately prepared to effectively teach mathematics to students from diverse backgrounds. However, preparation in some areas of teaching may be more important than others when working in specific contexts. Specifically, for teachers working in schools with students from diverse backgrounds, it may be more important to consider the unique aspects of teaching students from these populations than to better understand the mathematics they will be teaching. There is important research to be done in this area to better inform teacher preparation programs' allocation of time to studying different topics (see Tatto & Hordern, 2017, and chapter 4 of this book).

Conclusion

This study shows that future secondary mathematics teachers in some countries may not be adequately prepared for the mathematical aspects of teaching as well as for the pedagogical aspects of teaching students from diverse backgrounds. Specifically, the analysis finds a negative relationship between opportunities to learn to teach students from diverse backgrounds and mathematical knowledge for teaching in the United States, Thailand, Georgia, and Germany. Some of these results exist among teachers from the same preparation programs, suggesting selection into particular programs does not fully account for the relationships found.

Although the discussion of results has erred on the side of caution in the interpretation, there are some clear implications for teacher preparation programs. First, programs would benefit from considering whether their future teachers have ample opportunities to learn both to teach students from diverse backgrounds and to learn mathematics, particularly if many teachers go on to teach in settings with diverse student populations. Additionally, teacher preparation programs may wish to consider developing teachers' mathematical knowledge for teaching and pedagogical skills with students from diverse backgrounds at the same time. Such an approach may develop mathematical knowledge for teaching that is more applicable to contexts with students from diverse backgrounds. Current work on the preparation of elementary teachers in equitable mathematics teaching suggests these types of integrated courses are particularly effective (Aguirre et al., 2013; Foote et al., 2013; McDuffie et al., 2014), but more research to understand the design of these courses and their effectiveness at the secondary level is needed.

Acknowledgments This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-0824162, and the NAEd/Spencer Dissertation Fellowship Program. Its contents are solely the responsibility of the author and do not necessarily represent the official views of the previous organizations.

References

Adams, A., Bondy, E., & Kuhel, K. (2005). Preservice teacher learning in an unfamiliar setting. *Teacher Education Quarterly, 32*(2), 41–62.

Aguirre, J. M., Turner, E. E., Bartell, T. G., Drake, C., Foote, M. Q., & McDuffie, A. R. (2012). Analyzing effective mathematics lessons for English learners: A multiple mathematical lens approach. In N. Ramirez & S. Celedón-Pattichis (Eds.), *Beyond good teaching: Advancing mathematics education for ELLs* (pp. 207–222). Reston, VA: National Council of Teachers of Mathematics.

Aguirre, J. M., Turner, E. E., Bartell, T. G., Kalinec-Craig, C., Foote, M. Q., McDuffie, A. R., & Drake, C. (2013). Making connections in practice: How prospective elementary teachers connect to children's mathematical thinking and community funds of knowledge in mathematics instruction. *Journal of Teacher Education, 64*(2), 178–192. https://doi.org/10.1177/0022487112466900

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher, 36*(7), 369–387. https://doi.org/10.3102/0013189X07308739

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.

Bartell, T. G., Foote, M. Q., Aguirre, J. M., Roth McDuffie, A., Drake, C., & Turner, E. E. (2010). Preparing preK-8 teachers to connect children's mathematical thinking and community based funds of knowledge. In P. Brosnan, D. B. Erchick, & L. Flevares (Eds.), *Proceedings of the 32nd annual meeting of the North American chapter of the international group for the psychology of mathematics education* (pp. 1183–1191). Columbus, OH: The Ohio State University.

Barton, P. E., & Coley, R. J. (2009). *Parsing the achievement gap II. Policy information report*. Princeton, NJ: Educational Testing Service Retrieved from http://eric.ed.gov/?id=ED505163

Battey, D., & Franke, M. (2013). Integrating professional development on mathematics and equity: Countering deficit views of students of color. *Education and Urban Society, 47*, 433. https://doi.org/10.1177/0013124513497788

Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). The influence of school administrators on teacher retention decisions. *American Educational Research Journal, 48*(2), 303–333. https://doi.org/10.3102/0002831210380788

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management, 24*(1), 113–132. https://doi.org/10.1002/pam.20072

Castro, A. J. (2010). Themes in the research on preservice teachers' views of cultural diversity implications for researching millennial preservice teachers. *Educational Researcher, 39*(3), 198–210. https://doi.org/10.3102/0013189X10363819

Conference Board of the Mathematical Sciences. (2001). *The mathematical education of teachers* (Vol. 11). Providence, RI/Washington, DC: American Mathematical Society/Mathematical Association of America. Retrieved from http://www.cbmsweb.org/MET_Document/

Conference Board of the Mathematical Sciences. (Ed.). (2012). *The mathematical education of teachers II* (Vol. 17). Providence, RI/Washington, DC: American Mathematical Society/Mathematical Association of America. Retrieved from http://cbmsweb.org/MET2/.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *The American Economic Review, 95*(2), 158–165.

Delpit, L. D. (1995). *Other people's children: Cultural conflict in the classroom*. New York, NY: The New Press.

Dumais, J., & Meinck, S. (2013). Chapter 10: Estimation of weights, participation rates, and sampling error. In M. T. Tatto (Ed.), *The teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Dumais, J., Meinck, S., Tatto, M. T., Schwille, J., & Ingvarson, L. (2013). Chapter 6: Sampling design. In M. T. Tatto (Ed.), *The teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Foote, M. Q., Roth McDuffie, A., Turner, E. E., Aguirre, J. M., Bartell, T. G., & Drake, C. (2013). Orientations of prospective teachers toward students' family and community. *Teaching and Teacher Education, 35*, 126–136. https://doi.org/10.1016/j.tate.2013.06.003

Freedman, S. W., & Appleman, D. (2009). "In it for the long haul"— How teacher education can contribute to teacher retention in high-poverty, urban schools. *Journal of Teacher Education, 60*(3), 323–337. https://doi.org/10.1177/0022487109336181

Garmon, M. A. (2004). Changing preservice teachers' attitudes/beliefs about diversity: What are the critical factors? *Journal of Teacher Education, 55*(3), 201–213. https://doi.org/10.1177/0022487104263080

Garmon, M. A. (2005). Six key factors for changing preservice teachers' attitudes/beliefs about diversity. *Educational Studies, 38*(3), 275–286. https://doi.org/10.1207/s15326993es3803_7

González, N., Andrade, R., Civil, M., & Moll, L. (2001). Bridging funds of distributed knowledge: Creating zones of practices in mathematics. *Journal of Education for Students Placed at Risk, 6*(1–2), 115–132. https://doi.org/10.1207/S15327671ESPR0601-2_7

González, N., Moll, L. C., & Amanti, C. (Eds.). (2005). *Funds of knowledge: Theorizing practices in households and classrooms*. New York, NY: Routledge.

Guarino, C. M., Santibañez, L., & Daley, G. A. (2006). Teacher recruitment and retention: A review of the recent empirical literature. *Review of Educational Research, 76*(2), 173–208. https://doi.org/10.3102/00346543076002173

Hand, V. (2012). Seeing culture and power in mathematical learning: Toward a model of equitable instruction. *Educational Studies in Mathematics, 80*(1–2), 233–247. https://doi.org/10.1007/s10649-012-9387-9

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources, 39*(2), 326–354.

Hill, H. C., Blunk, M., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430–511.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Hill, H. C., Sleep, L., Lewis, J. M., & Ball, D. L. (2007). Assessing teachers' mathematical knowledge: What knowledge matters and what evidence counts. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, 2nd ed., pp. 111–155). Charlotte, NC: Information Age Publishing Inc.

Hill, H. C., Umland, K., Litke, & Kapitula, L. R. (2012). Teacher quality and quality teaching: Examining the relationship of a teacher assessment to practice. *American Journal of Education, 118*(4), 489–519. https://doi.org/10.1086/666380

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*(3), 499–534. https://doi.org/10.3102/00028312038003499

International Association for the Evaluation of Educational Achievement. (2007). *TEDS-M sample preparation manual*. Hamburg, Germany: International Education Association.

Irvine, J. J. (2003). *Educating teachers for diversity: Seeing with a cultural eye*. New York, NY: Teachers College Press.

Kang, N.-H., & Hong, M. (2008). Achieving excellence in teacher workforce and equity in learning opportunities in South Korea. *Educational Researcher, 37*(4), 200–207. https://doi.org/10.3102/0013189X08319571

Little, J. W., & Bartlett, L. (2010). The teacher workforce and problems of educational equity. *Review of Research in Education, 34*(1), 285–328. https://doi.org/10.3102/0091732X09356099

Luschei, T. F., Chudgar, A., Rew, W., Klugman, J., Francois, C., & Mehta, J. (2013). Exploring differences in the distribution of teacher qualifications across Mexico and South Korea: Evidence from the teaching and learning international survey. *Teachers College Record, 115*(5).

McDuffie, A. R., Foote, M. Q., Bolson, C., Turner, E. E., Aguirre, J. M., Bartell, T. G., … Land, T. (2014). Using video analysis to support prospective K-8 teachers' noticing of students' multiple mathematical knowledge bases. *Journal of Mathematics Teacher Education, 17*(3), 245–270. https://doi.org/10.1007/s10857-013-9257-0

McDuffie, A. R., Foote, M. Q., Drake, C., Turner, E., Aguirre, J., Bartell, T. G., & Bolson, C. (2014). Use of video analysis to support prospective K-8 teachers' noticing of equitable practices. *Mathematics Teacher Educator, 2*(2), 108–140.

Moll, L. C., Amanti, C., Neff, D., & González, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory Into Practice, 31*(2), 132–141.

Organisation for Economic Co-operation and Development. (2013). *PISA 2012 results: Excellence through equity* (Vol. II). Paris, France: OECD Publishing Retrieved from http://www.oecd-ilibrary.org/education/pisa-2012-results-excellence-through-equity-volume-ii_9789264201132-en

Organisation for Economic Co-operation and Development. (2014). *Education at a glance 2014*. Paris, France: OECD Publishing Retrieved from http://www.oecd-ilibrary.org/education/education-at-a-glance-2014_eag-2014-en

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233. https://doi.org/10.1037/a0020141

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity?: Rising inequality, schools, and children's life chances* (pp. 91–115). New York, NY: Russell Sage Foundation.

Reardon, S. F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal, 46*(3), 853–891. https://doi.org/10.3102/0002831209333184

Reardon, S. F., Robinson-Cimpian, J. P., & Weathers, E. S. (2015). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In H. F. Ladd & M. G. Goertz (Eds.), *Handbook of research in education finance and policy* (2nd ed., pp. 491–509). Mahwah, NJ: Lawrence Erlbaum.

Reininger, M. (2012). Hometown disadvantage? It depends on where you're from: Teachers' location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis, 34*(2), 127–145. https://doi.org/10.3102/0162373711420864

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school examining direct cognitive assessments and teacher ratings. *American Educational Research Journal, 48*(2), 268–302.

Schwille, J., Ingvarson, L., & Holdgreve-Resendez, R. (2013). *TEDS-M Encyclopedia: A guide to teacher education context, structure, and quality assurance in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Simon, F., Malgorzata, K., & Beatriz, P. (2007). *Education and training policy. No more failures ten steps to equity in education: Ten steps to equity in education*. Paris, France: OECD Publishing.

Smith, T. M., & Ingersoll, R. M. (2004). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal, 41*(3), 681–714. https://doi.org/10.3102/00028312041003681

Tatto, M. T., & Hordern, J. (2017). The configuration of teacher education as a professional field of practice: A comparative study of mathematics education. In J. Furlong & G. Whitty (Eds.), *Knowledge and the study of education: An international exploration* (pp. 255–274). Oxford, UK: Oxford Comparative Education Series, Symposium Books.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics: Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.

Tatto, M. T., Rodríguez, M. C., Reckase, M., Rowley, G., & Lu, Y. (2013). Chapter 11: Scale development and reporting: Opportunities to learn, beliefs, and mathematics knowledge for teaching. In M. T. Tatto (Ed.), *The teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Turner, E. E., & Drake, C. (2016). A review of research on prospective teachers' learning about children's mathematical thinking and cultural funds of knowledge. *Journal of Teacher Education, 67*(1), 32–46. https://doi.org/10.1177/0022487115597476

Turner, E. E., Drake, C., McDuffie, A. R., Aguirre, J., Bartell, T. G., & Foote, M. Q. (2012). Promoting equity in mathematics teacher preparation: A framework for advancing teacher learning of children's multiple mathematics knowledge bases. *Journal of Mathematics Teacher Education, 15*(1), 67–82. https://doi.org/10.1007/s10857-011-9196-6

Valencia, R. R. (1997). *The evolution of deficit thinking: Educational thought and practice*. London, UK: Falmer Press.

Villegas, A. M., & Irvine, J. J. (2010). Diversifying the teaching force: An examination of major arguments. *The Urban Review, 42*(3), 175–192. https://doi.org/10.1007/s11256-010-0150-1

Whipp, J. L. (2013). Developing socially just teachers: The interaction of experiences before, during, and after teacher preparation in beginning urban teachers. *Journal of Teacher Education, 64*(5), 454–467. https://doi.org/10.1177/0022487113494845

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models problems and solutions. *Organizational Research Methods, 12*(4), 695–719. https://doi.org/10.1177/1094428108327450

Part III
Methodological Challenges and Strategies

Chapter 16
Introduction: Methodological Challenges and Strategies in the TEDS-M Study

Maria Teresa Tatto and Wendy M. Smith

Abstract This chapter provides an introduction to Part III of the book which highlights specific methodological challenges and strategies in the Teacher Education and Development Study in Mathematics (TEDS-M), a cross-national study of teacher education programs that prepare future primary and secondary mathematics teachers. The TEDS-M Technical Report (Tatto MT, The Teacher Education and Development Study in Mathematics (TEDS-M). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: technical report. International Association for the Evaluation of Student Achievement, Amsterdam, 2013) provides the complete methodology and procedures. Chapters in this part discuss challenges with sampling, interpretation, and validation. The chapter on sampling introduces non-statisticians to the basic concepts of statistical sampling and its application in TEDS-M using examples and graphical illustrations. Instruments to capture the demographic and outcome variables of interest in

The text in this introduction to Part III contains shortened and slightly edited versions of text that has appeared in the following publications: Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M., & Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M).* Amsterdam, the Netherlands: International Association for the Evaluation of Student Achievement (IEA); Brese, F., & Tatto, M.T. (Eds.) (2012). *User guide for the TEDS-M international database*. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA); Tatto, M.T. (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report*. Amsterdam, the Netherlands: International Association for the Evaluation of Student Achievement (IEA). Text cited directly or indirectly from those sources will not be made recognizable. An extensive report on the methods and strategies used in TEDS-M can be found in Tatto (2013). We summarize the key aspects here to orient the reader; the chapters in Part III however are original contributions written exclusively for this book.

M. T. Tatto (✉)
Mary Lou Fulton Teachers College, Arizona State University, Tempe, AZ, USA
e-mail: teresa.tatto@asu.edu

W. M. Smith
University of Nebraska, Lincoln, NE, USA
e-mail: wsmith5@unl.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_16

TEDS-M did not exist, much less for use internationally, so TEDS-M researchers had to create, validate, and interpret appropriate measurements for the target populations. Importantly, the TEDS-M study, at the request of the participating countries, included assessments of knowledge, an approach that had not been attempted at the time in any country, much less at an international scale. The chapter on anchor points explains the development of a method to help interpret the assessment results. While the TEDS-M Technical Report includes information on overall instrument validation, the final two chapters in this Part III, look more closely at differential item functioning, and at the alignment of the teacher knowledge assessments to national expectations concerning mathematical knowledge for teaching using the United States case as an example.

Key Methodological Issues

The work that went into developing, implementing and interpreting the TEDS-M study is very extensive, and we cannot do it justice in this short introduction or in this book. The reader is encouraged to consult the TEDS-M Technical Report (Tatto, 2013) to gain a comprehensive understanding of the TEDS-M methods. Here we only highlight sampling, interpretation and validation challenges.

Sampling, Instrument Development, and Validation

Conducting international studies is exceedingly difficult, particularly when one seeks to study nationally representative samples. The Teacher Education and Development Study in Mathematics (TEDS-M) was and is still the first large-scale assessment of the outcomes of teacher education that uses statistical sampling. Thus, no precedent existed when researchers began to design the sampling frame. The sampling design that was developed for the study allowed the computation of correct, precise, and unbiased population estimates for all characteristics investigated in the survey. One goal of Chap. 17 is to introduce non-statisticians to the basic concepts of statistical sampling and its application in TEDS-M using examples and graphical illustrations.

Another challenge was the development of instruments. While TEDS-M was primarily a survey research project, no other available instruments both aligned with the research questions and were internationally validated. Thus, the TEDS-M researchers faced a number of key methodological issues in addition to the construction of a framework to select nationally representative samples: those of creating, validating, and interpreting appropriate measurements for the target populations. Importantly, the TEDS-M study, at the request of the participating countries, included assessments of knowledge, an approach that had not been attempted at the time in any country, much less at an international scale. Not only did the instruments have to be developed and field tested, but benchmarks for the knowledge levels that should be reached by future primary and secondary teachers were nonexistent.[1] Chapter 18 discusses issues of interpretation and reporting for the knowledge assessments.

The development of instruments in a rigorous study such as TEDS-M required multiple piloting and psychometric analyses to test the fit of items and questions to the study participants. These efforts are reported in detail in the technical report (Tatto, 2013). Further analyses, once the final data have been collected, are always recommended. The last two chapters in this part present post hoc analyses of TEDS-M results. Chapter 19 examines variability in item difficulty in mathematics performance that is accounted for by gender, and the extent to which gender differential item functioning (DIF) is explained by person predictors (e.g., opportunity to learn) and item characteristics (e.g., item format). The last chapter (Chap. 20) reports the results of a post hoc validation study in which publicly released TEDS-M knowledge assessment items were validated against teacher knowledge expectations in the United States.

In this last section, we highlight key approaches that made the study both scientific and rigorous, in the hopes that these reflections may inspire future rigorous, empirical work on teacher education. We first present an overview of definitions and an outline of the sampling procedure, which is followed by a brief summary of the development of the instruments to measure knowledge and opportunities to learn. References are made to other important publications for researchers and stakeholders to examine when undertaking similar approaches in the future.

Sampling

The population of interest in each country included "all institutions where future primary and secondary teachers were receiving their preparation to teach mathematics, the teacher educators who were preparing them in mathematics and mathematics pedagogy as well as in general pedagogy, and the future teachers in

---

[1] An effort of this magnitude was possible due to previous work in the area by the principal investigators and collaborators across the globe. For instance, questionnaire items were received from several sources, including study investigators, national research coordinators, and mathematics consultants. Several items were adapted from other studies, including the Study of Instructional Improvement (SII)/Learning Mathematics for Teaching/Consortium for Policy Research in Education (CPRE), University of Michigan, School of Education (measures development supported by NSF grants REC-9979873, REC- 0207649, EHR-0233456, & EHR 0335411); Developing Subject Matter Knowledge in Math Middle School Teachers (P-TEDS/MT-21), supported by NSF Grant to Michigan State University REC-0231886; and Knowing Mathematics for Teacher Algebra (KAT), supported by NSF Grant REC-0337595. The instruments and assessments developed by TEDS-M were the result of the collaborative efforts of the international centers at Michigan State University, the Australian Council for Educational Research, and the IEA under the direction of Professor Maria Teresa Tatto at Michigan State University with support from a grant from the National Science Foundation (Award No. REC-0514432).

their last year of training" (Tatto et al., 2012, p. 259). In most countries, TEDS-M implemented a two-stage random sampling design. First, the sampling unit of the IEA Data Processing and Research Center (DPC) worked with each participating country's national research center to select samples representative of the national population of teacher preparation (TP) institutions offering education to future teachers intending to teach mathematics at the primary and/or lower-secondary levels. Second, once an institution had been selected, all programs within that institution offering mathematics preparation were identified including both concurrent and consecutive programs. Programs were sampled within countries, and then individuals were sampled from the programs. The international target population of TP institutions was defined as follows: The set of secondary or post-secondary schools, colleges, or universities that offer structured opportunities to learn (i.e., a program or programs) on a regular and frequent basis to future teachers within a route of teacher preparation.

The national research coordinators (NRCs) for each participating country were asked to list all routes where TP programs could be found and to indicate which were of principal interest to TEDS-M (i.e., a major route), and which were of marginal interest. Each NRC and the sampling team sought agreement as to which routes would constitute the national desired target population for the country of interest. Countries could also opt to exclude routes or institutions of very small size. The remaining populations are referred to, within the context of TEDS-M, as the national defined target populations, a customization of the defined international target population to recognize country-specific conditions. A TP institution did not have to be teaching mathematics content to be part of the target population. However, it was necessary for the institution to be teaching mathematics pedagogy (Tatto et al., 2012, p. 260).

If the findings from the TEDS-M study were to be meaningful to country stakeholders, estimates needed to reflect inferences about the population rather than merely the sample, in which particular groups within each strata were guaranteed to be sampled, regardless of the size of the group. Thus, TEDS-M employed a complex weighting strategy (see Tatto et al., 2012, p. 267). In the countries that administered a census, it was sufficient to adjust the collected data for non-response to obtain unbiased estimates of the population parameters. When the sample design is complex and involves stratification and unequal probabilities of selection, estimation weights are required to achieve unbiased estimates (Lohr, 1999). Estimation weights are the product of one or many design or base weights and one of many adjustment factors; the former are the inverse of the selection probability at each selection stage, and the latter compensate for non-response, again at each selection stage. These design weights and adjustment factors are specific to each stage of the sample design and to each strata. Because each country participating in TEDS-M had to adapt the general TEDS-M sample design to its own conditions, the estimation weights had to conform to the national adaptations. Usually, one set of estimation weights is produced for each participating country in a study. However, in the case of TEDS-M, four sets of estimation weights were required according to each of the surveys used by TEDS-M: the institutions, the teacher educators, the future teachers of primary school mathematics, and the future teachers of lower-secondary school mathematics. All estimates computed for any one of the four TEDS-M surveys were produced using the appropriate estimation weight, as developed by Horwitz-Thompson (Lohr, 1999). Chapter 11 of the IEA technical report (Tatto, 2013) provides a detailed description of how TEDS-M calculated the different weight components and the resulting estimation weights for the four populations.

According to Tatto et al. (2012, p. 267), surveys with complex designs such as TEDS-M require special attention to estimation, especially estimation of the sampling error. Both the survey design and the unequal weights need to be considered to obtain (approximately) design-unbiased estimates of sampling error. TEDS-M adopted the balanced repeated replication (BRR) technique (McCarthy, 1966) to estimate sampling error. More specifically, TEDS-M used the variant of this technique known as Fay's method (Fay, 1989). BRR is a well-established and well-documented technique that is used in other international educational studies—notably, the Programme for International Student Assessment (PISA) and the Teaching and Learning International Survey (TALIS), both conducted by the Organisation for Economic Co-operation and Development (OECD). Chapter 11 of the TEDS-M technical report (Tatto, 2013) describes how the replicates were created and how the BRR estimates of sampling error were computed for TEDS-M. These estimates of the sampling error are another key element of the statistical quality of survey outcomes.

Instruments

MCK and MPCK Assessments The TEDS-M tests of future teachers' mathematics content knowledge (MCK) and mathematics pedagogical content knowledge (MPCK) used a balanced-incomplete block design so that the desired content would be appropriately covered while simultaneously allowing the test to be completed within a reasonable administration time. This meant that each future teacher was given only a portion of the full set of items. Because the set of items taken by each teacher was not comparable, summing the scores on the items taken by that person would not have yielded meaningful results. If summed scores were to be comparable, all of the test booklets would have to be constructed to be equivalent in content and difficulty. This was not possible because of the complexity of the content domains and the short time allowed for the assessment. To obtain comparable estimates of performance, TEDS-M used item response theory (IRT; De Ayala, 2009). IRT allows estimates of performance to be obtained on the same scale even when the set of items taken by each individual is different (Tatto et al., 2012, p. 273).

TEDS-M used item response models from the Rasch family to carry out calibration. The standard Rasch (1980) model was used for the dichotomous items, and the partial credit model (Masters, 1982) was used to fit the matrix of item scores for the polytomous items. Both item types were analyzed simultaneously using ACER Conquest software (Wu, Adams, Wilson, & Haldane, 2007).

The calibration data were used to carry out standardization. TEDS-M standardized the achievement estimates (in logits) to a mean of 500 and a standard deviation of 100, in line with the procedure followed in TIMSS, wherein all countries are weighted so that they contribute equally to the standardization sample. This process was repeated for each of the four key measures: MCK (primary), MCK (lower-secondary), MPCK (primary), and MPCK (lower-secondary). Once standardization was completed, scores were computed for all participants for whom MCK and MPCK estimates could be obtained. The mean of 500 and the standard deviation of 100 thus apply to the calibration sample rather than to the complete set of scores (Tatto et al., 2012, p. 274).

The calibration results were also used to identify anchor points for the score scale. Anchor points are specific values on the score scale, each of which pertains to a description of what examinees at this point know and can do. TEDS-M identified two sets of test items to support development of the descriptions of the skills and knowledge at each anchor point. The first set of test items contained those items that a person at that anchor point on the scale score would, according to the IRT model, be able to answer correctly with a probability of .70 or greater. The second set of test items included those items that a person at that anchor point on the scale score would, based on the IRT model, have a probability of .50 or less of answering correctly. The anchor points identified were points at which there would be sufficient items of each type (between 10 and 12 items) to develop a description of the skills and knowledge that a person at that point would have. Given these requirements, two anchor points were identified for the MCK primary scale and two for the MCK lower-secondary scale: Anchor Point 1 represented a lower level of performance, and Anchor Point 2 represented a higher level. Only one anchor point, representing an acceptable level of performance, was selected for the MPCK scales because the MPCK assessment had fewer items than MCK assessment did (Tatto et al., 2012, p. 274–275).

To develop descriptions of the capabilities of persons near each anchor point on the scales, committees of mathematicians and mathematics educators conducted detailed analyses of the sets of items for the respective anchor points. They did this work in workshops specifically set up for this purpose at the international research center at Michigan State University (MSU). The resulting anchor point descriptions give tangible meaning to points on the reporting score scales. They can be found in Tatto et al. (2012, Chap. 6); a more detailed description is included in the TEDS-M technical report (Tatto, 2013).

Opportunity to Learn Opportunity to learn (OTL) measures were based on scales and items developed in a variety of ways. Several were based on previous research conducted at MSU and elsewhere. Some were based on previous research conducted at the Australian Council for Educational Research (ACER), and some were developed specifically for TEDS-M, in collaborative workshops and meetings that included the researchers in the international research centers at MSU and ACER, and in the national research centers in the participating countries. After completing an extensive pilot of a larger set of items, TEDS-M researchers selected items that provided information on program, institution, and country variation. Items that survived initial exploratory factor analyses were used in the operational forms for the main study.

The researchers then conducted a confirmatory factor analysis that was based on a conceptualization of OTL based on previous analysis of teacher education program curriculum, encompassing four broad categories relating to mathematics content areas: tertiary and school-level mathematics, mathematics education pedagogy, general education pedagogy, and school-based experiences. The aim of the analysis was to assess the fit of each OTL index (measure) to the data and the index interrelations. Each of the four broad categories contained several indices, which taken together across the categories resulted in 24 distinct OTL indices.

Using as their reference the best-fitting models, the researchers then created OTL index scores. The OTL indices for the areas of mathematics, mathematics pedagogy, and general pedagogy were derived from summing the number of topics studied in each area. Rasch logit scores were estimated for the OTL indices using rating scales (e.g., activities in which future teachers participated from *never* to *often*). These scores were centered at the point on the OTL scale that is associated with the middle of the rating scale (essentially, neutral). More precisely, this step involved using the test characteristic curve to identify the point on the θ-scale associated with the mid-point on the summed score scale. The θ-value was then used to center the OTL scale so that it would be located at a scaled value of 10.

All OTL scales consisting of *number of topics* are interpretable using the mean proportions to report outcomes in terms of number of topics studied for each OTL index (for instance, a mean proportion of .52 would indicate that about half of the future teachers reported studying a given topic). All OTL scales based on Rasch logit scores can be interpreted given the location of the mid-point, where 10 is associated with the neutral position with a standard deviation determined by the Rasch model (Tatto et al., 2012, p. 281).

Chapters in This Part

In Chap. 17, "Sampling for TEDS-M," Meinck and Dumais explain the sampling design of TEDS-M, the first large-scale assessment in teacher education using statistical sampling. The sampling design applied in the study allows the computation of correct, precise, and unbiased population estimates for all characteristics investigated in the survey. This chapter introduces non-statisticians to the basic concepts of statistical sampling and its application in TEDS-M using examples and graphical illustrations. Readers will be made familiar with the features of complex samples in general and with the TEDS-M survey in particular. Implications for the conduct of statistical analysis of TEDS-M data and the interpretation of the results are presented. TEDS-M comprises four target populations, namely, teacher preparation institutions, future primary/lower secondary mathematics teachers, and their educators, for which reliable estimates of main characteristics were required. The demanding study goals, accounting for the complexity and differences of the teacher education systems in the 17 participating countries, on the other hand, posed particular challenges in the design of a multi-purpose international sampling plan. Experiences gained throughout the implementation of this study provide a valuable contribution to the specification of sampling designs for future studies in higher education.

In Chap. 18, "Developing Anchor Points to Enhance the Meaning of the Mathematical and Mathematical Pedagogy Score Scales from the TEDS-M Study," Reckase explains the development of benchmarks or anchor points that were used to report the information about prospective teachers' knowledge and skills in a way that is easy to understand by the numerous audiences for the results of the study. He explains that the numerical values that are typically used to report the results of an achievement test do not have any intrinsic meaning. Thus, to help with the interpretation of the results of the TEDS-M study, work was done to give substantive meaning to the scale used for reporting results so that those interested in the results could have a reasonable sense of what the numerical values meant in relation to teachers' knowledge. Meaning was added to the reporting scales by attaching detailed substantive descriptors to selected points on the scale called anchor points. The purpose of this chapter is to share the logic and methodology used to add descriptions to the TEDS-M anchor points, and to discuss how the interpretations derived from these descriptions differ from those used with other assessments.

In Chap. 19, "Examining Sources of Gender DIF in Mathematics Knowledge of Future Teachers Using Cross-Classified IRT Models," Cai and Albano use TEDS-M data to study differential item functioning (DIF). This chapter serves as a proof-of-concept for methodology to examine variability in item difficulty in mathematics performance that is accounted for by gender, referred to as "gender DIF," and the extent to which gender DIF is explained by both person predictors (e.g., opportunity to learn) and item characteristics (e.g., item format). Cai and Albano use cross-classification multilevel IRT models to examine the relationships among item difficulty, gender, OTL, and item format. The cross-classified multilevel model is a flexible tool to explain potential DIF sources related to item and person characteristics. This approach results in more economical models where DIF can be detected within an omnibus test. This approach can be helpful in creating and adapting appropriate measurement tools when constructing or translating items. Moreover, in terms of person characteristics, researchers can take variables such as OTL into account, and thus improve DIF detection and estimation. Doing so can improve the validity of group comparisons.

In Chap. 20, "Standing the Test of Time: Validating the TEDS-M Knowledge Assessment Against MET II Expectations," Silver and Mortimer present the results of a post hoc content validation study of the TEDS-M knowledge assessment. They used the publicly released TEDS-M knowledge assessment items, originally validated prior to the 2008 study, and compared them to the specifications for teacher knowledge found in the United States' Mathematical Education of Teachers II report (MET II; CBMS, 2012). The validation was based on the expert judgments of two authors of the MET II report. The findings suggest strong content validity for

TEDS-M items. Validating the TEDS-M knowledge assessment against more recent standards serves to strengthen claims about future teacher knowledge based on this knowledge assessment.

References

Brese, F., & Tatto, M. T. (Eds.). (2012). *User guide for the TEDS-M international database*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Conference Board of the Mathematical Sciences (CBMS). (2012). *The mathematical education of teachers II (MET II)*. Providence, RI/Washington, DC: American Mathematical Society/ Mathematical Association of America.

De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: The Guilford Press.

Fay, R. E. (1989). Theoretical application of weighting for variance calculation. In *Proceedings of the section on survey research methods of the American Statistical Association* (pp. 212–217). Alexandria, VA: American Statistical Association.

Lohr, L. S. (1999). *Sampling: Design and analysis*. Pacific Grove, CA: Duxbury Press.

Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149–174.

McCarthy, P. (1966). Replication: An approach to the analysis of data from complex surveys. In *Vital and health statistics* (Series 2, No. 14). Hyatsville, MD: National Center for Health Statistics.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (originally published 1960). Chicago, IL: University of Chicago Press.

Tatto, M. T. (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA teacher education and development study in mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Student Achievement.

Wu, M., Adams, R., Wilson, M., & Haldane, S. (2007). *ACER conquest: Generalised item response modelling software (Version 2.0)* [Computer software]. Melbourne, Australia: Australian Council for Educational Research (ACER).

Chapter 17
Sampling for TEDS-M

Sabine Meinck and Jean Dumais

Abstract The Teacher Education and Development Study in Mathematics (TEDS-M) is the first large-scale assessment in teacher education using statistical sampling. The sampling design applied in the study allows the computation of correct, precise, and unbiased population estimates for all characteristics investigated in the survey. This chapter introduces non-statisticians to the basic concepts of statistical sampling and its application in TEDS-M using examples and graphical illustrations. Readers will be made familiar with the features of complex samples in general and with the TEDS-M survey in particular. Implications for the conduct of statistical analysis of TEDS-M data and the interpretation of the results will be presented. TEDS-M comprises four target populations, namely, teacher preparation institutions, future primary/lower secondary math teachers, and their educators, for whom reliable estimates of main characteristics were required. The demanding study goals—both combining four target populations into one survey and capturing the complexity and differences of the teacher education systems in the 17 participating countries—posed particular challenges in the design of a multi-purpose international sampling plan. Experiences gained throughout the implementation of this study provide a valuable contribution to the specification of sampling designs for future studies in higher education.

Important note to reader: This chapter is in major parts a re-print of the following sources: Dumais, Meinck, Tatto, Schwille, and Ingvarson (2013), and Dumais and Meinck (2013). Text cited directly or indirectly from those sources will not be made recognizable.

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

S. Meinck (✉)
IEA, Hamburg, Germany
e-mail: sabine.meinck@iea-hamburg.de

J. Dumais
Statistics Canada, Ottawa, ON, Canada
e-mail: gagnedumais@sympatico.ca

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_17

Introduction

TEDS-M was the first large-scale international comparative study in higher education, and is still the only one conducted so far. The study was supported by the U.S. National Science Foundation (NSF REC 0514431) and the International Association for the Evaluation of Educational Achievement (IEA), and applied IEA's rigorous standards for technical quality, which ensure the highest validity of the results (Gregory & Martin, 2001). The implementation of statistical sampling procedures is one of the measures ensuring international comparability of the studies' outcomes. More specifically, the sampling design applied in TEDS-M allows the computation of correct, precise, and unbiased population estimates for all characteristics investigated in the survey.

This chapter draws in broad terms from two sections of the technical report for the TEDS-M study (Dumais & Meinck, 2013; Dumais, Meinck, Tatto, Schwille, & Ingvarson, 2013), but aims to address a broader audience. It will introduce the basic concepts of statistical sampling and how they were applied in TEDS-M, using illuminating examples and graphical illustration accessible to non-statisticians. The consequences of the sampling design on the analysis of TEDS-M data and on the interpretation of the results arising from the complex sampling design will also be presented.

This chapter will start with a review of the target populations of the study, will continue with an illustration of the sampling strategies and weighting procedures, and, finally, will introduce the strategies for estimating population characteristics and their standard errors.

Sampling Principles in International Large-Scale Assessments

All international large-scale assessments (ILSA) conducted by IEA or the Organisation for Economic Co-operation and Development (OECD)[1], such as TIMSS, PIRLS, PISA, and TALIS, rely on statistical sampling of schools and people.

In those international studies, as well as in national assessments of educational achievement, statistical sampling is used because the number of schools, principals, teachers, and students is far too large to allow for enumerating and interviewing all members of the population; statistical sampling helps in reducing the workload while providing estimates that are close enough to values from a complete enumeration to meet most purposes. The smaller countries, e.g., Iceland, Bahrein, Cyprus, Iceland, Malta, or Singapore, can afford to do a complete enumeration, but mid-size and large countries, e.g., Australia, Brazil, Denmark, England, or Japan, must resort to more economical means of accessing the information. Hence, a statistical

[1] https://www.oecd.org/

sampling strategy must be developed that will (a) be cheaper than a complete enumeration, (b) yield estimates that are close to what a complete enumeration would have given and precise enough for the intended purposes, (c) be useable in a variety of educational systems, and (d) allow for international comparisons of the national results. The expressions "design-unbiased" or "design-unbiasedness" and "sampling precision" are often used to summarize those characteristics of a statistical sampling design.

Common features of samples for ILSAs are stratification, multiple sampling stages, cluster sampling, and unequal sampling probabilities. In most ILSAs, these methods are part of the international sampling design, are applied in most but not all countries, and can be mixed. Those design features are reviewed in the following sections. Typically, the international sampling design is optimized with respect to the specific circumstances of a particular country, while complying with the international objectives of design-unbiasedness and sampling precision.

Like any other statistical survey, ILSAs measure two values for each characteristic of interest: an estimate of its level (e.g., the average age of teachers, the proportion of students who have access to a home computer, or the average mathematics score of Grade 8 students) and an estimate of its precision (e.g., a margin of error for the estimated mathematics score of Grade 8 students).

Comprehensive introductions to the topic that are also accessible to a non-technical readership are given, for example, in "Survey Methods and Practices" (Statistics Canada, 2003) and Rust (2014).

Stratification

Stratification is the division of the population into homogeneous groups from which samples are drawn. Stratification can be used to achieve higher-precision estimates (i.e., estimates with a narrower margin of error) with a fixed sample size. Stratification can also be used to guarantee a specific sample size for units from certain population groups in order to obtain reliable estimates for each of the groups.

Strata are exhaustive and mutually exclusive groups of individuals or units. That is, each individual or unit is in one and only one stratum. The total sample is allocated into the various strata. Strata must be created before the sample can be selected from the sampling frame[2].

Stratification variables should be selected with care. The following guidelines help in the identification of useful stratification variables:

[2] A sampling frame is a list that covers the units of interest to the survey that gives access to those units of interest; in ILSAs, the sampling frame is often the list of schools available from the ministry of education or statistical offices. If multiple stages of sampling are required, sampling frames are needed at each stage, e.g., after school sampling a list of students within each sampled school is needed.

- Before it can be used for stratification, a variable must be available for every unit on the frame.
- A stratification variable should be related to the key characteristics of the study. Otherwise, no gains in precision can be expected.
- When several variables are used for stratification, they can be crossed (e.g. age group by gender), nested (e.g. school districts within provinces), or mixed (e.g. school size in urban areas, post code in rural areas) in all sorts of ways as long as each unit on the sampling frame is assigned to only one level or category of each stratification variable.

In the TEDS-M study, stratification was employed in several countries. For details on the specific sampling designs of each country, refer to Meinck and Dumais (2012).

Cluster Sampling and Multiple Sampling Stages

In most ILSAs, the individuals of interest can be found within a hierarchical clustered structure. For example, students are nested in classes, classes are nested in schools, and schools in regions within countries. In most studies, the relationships between and among the levels of the hierarchy are of interest to the researchers. This is one reason why whole classes or many students within a school are selected, rather than selecting individuals from one comprehensive list of individuals. This sampling procedure is referred to as cluster sampling.

However, there is another important reason for cluster sampling: Comprehensive lists of individuals are usually not available, at least not at reasonable costs. A list of schools (or of teacher preparation institutions, as in the case of TEDS-M), however, may be more readily available, and a list of classes or individuals within the sampled institutions can usually be obtained at minimal additional cost.

Cluster sampling typically comes along with multiple-stage sampling in ILSAs. In a first stage, the selection of clusters is conducted. In a second stage, full classes, students, teachers, or other units are selected. More sampling stages may become necessary under specific circumstances, but may decrease the precision of the estimates.

In many countries participating in TEDS-M, all teacher preparation (TP) institutions were asked to participate in the survey. Only a few of the participating countries had to implement multiple-stage cluster sampling. Depending on the number of TP institutions, courses and future teachers or educators, the following sampling stages were possible:

1. TP institution
2. Course
3. Individual (educator, staff or future teacher)

Unequal Sampling Probabilities

ILSAs customarily aim for a so-called multi-stage self-weighted design. This is a sampling design that will have each respondent represent the same number of persons in the population, regardless of the size of the institution where they are selected, or the number of classes available or the number of students comprising those classes. Self-weighted designs can be constructed by sampling institutions and courses with probabilities proportional to their size (PPS)—for instance, by first sampling institutions, with larger institutions more likely to be selected, and then sampling a fixed number classes within the selected institutions, again with probability proportional to their size, and finally by sampling a fixed number of students within each selected class. In this type of design, larger institutions and classes are more likely to be selected than the smaller institutions and classes; and so the individuals found in them should be more likely to be in the final sample. Conversely, smaller institutions and classes have a lesser chance of being selected and so the individuals they comprise should be less likely to be in the final sample. Through the interplay of the various selection probabilities and sample sizes at each stage, the varying selection probabilities even out.

However, selection probabilities vary among units at each stage, a fact that has to be accounted for when analyzing data arising from such survey designs. In the TEDS-M survey, because of various factors, the selection probabilities varied drastically among participating individuals. For example, in countries with a census of all TP institutions, individuals in smaller TP institutions were selected with certainty, while individuals in large universities had rather small selection probabilities. The principal way of accounting for this in analysis is by using sampling weights, a topic discussed in detail later in this article.

International Sampling Plan for TEDS-M

Seventeen countries participated in the TEDS-M study. It is not difficult to imagine that very diverse systems of future teacher training can be found within and across all these countries. In order to help countries to identify the correct targets when preparing for and conducting sample selection, national research coordinators had to agree on common terms for key survey units. Furthermore, an international template sampling design had to be developed that could be implemented in all countries, serving the international study goals while at the same time being flexible enough to fit national conditions.

The international sampling plan implemented in the TEDS-M survey is a stratified multi-stage probability sampling design. This means that the targeted individuals (educators and future teachers) were randomly selected from a list of in-scope educators and future teachers for each of the randomly selected TP institutions.

The universes of interest comprised institutions where future primary and secondary teachers receive mathematics preparation, the teacher educators teaching mathematics, mathematics pedagogy, and general pedagogy, and finally the future teachers in their last year of training, prepared to teach at primary school or lower secondary school.

Random samples were required for each population.

Programs and Routes

Two concepts play a key role in the organization of teacher preparation: the program and the route. A program is a specific pathway that exists within an institution that requires students to undertake a set of subjects and experiences, and leads to the award of a common credential or credentials on completion. A route is a set of teacher education programs available in a given country. TP programs within a given route share a number of common features that distinguish them from TP programs in other routes and can be identified in similar ways in different participating countries. For the purposes of TEDS-M, two kinds of routes were defined (Tatto et al., 2008)[3]:

- C*oncurrent* routes consist of a single program that includes studies in the subjects future teachers will be teaching (academic studies), studies of pedagogy and education (professional studies) and practical experience in the classroom;
- C*onsecutive* routes consist of a first phase for academic studies (leading to a degree or diploma), followed by a second phase of professional studies and practical experience (leading to a separate credential/qualification); the first and second phases need not have been completed in the same institution. No route can be considered consecutive if the institution or the government authorities do not award a degree, diploma, or official certificate at the end of the first phase. Moreover, it may be customary or required for future teachers to complete the first and second phases in different institutions.

Moreover, sets of programs identified within a country sharing further common features (e.g., leading to a certain degree), are referred to as different program types. Table 17.1 lists the identified program types in participating countries and gives an overview of the structure of the teacher preparation systems. Program types and their sizes (estimated from the sample) are displayed along with the number of institutions that offer a specific program type. Note that the numbers of institutions do not necessarily add up to the total number of institutions in a country since some institutions offer more than one program type. This becomes most obvious when looking at the situation in Singapore: There is only one a university preparing future

[3] One participating country (USA) identified in addition an apprenticeship route (a route predominantly consisting of school-based experience with other institutions playing only a minor, marginal, supporting role), but this particular apprenticeship route was not included in the survey.

mathematics teacher; however, it offers ten different program types to master this education. To give another example, Botswana defined the program type “Diploma in Primary Education” as a concurrent route that is offered in four teacher education institutions.

Table 17.1 Structure of mathematics teacher preparation by participating country

Country	Level[1]	Route	Program type	No. of Institutions (sample estimate)	No. of future teachers(final year, sample estimate)
Botswana	1	Concurrent	Diploma in Primary Education	4	100
	2	Concurrent	Bachelor of Secondary Education (Science), University of Botswana	1	25
	2	Concurrent	Diploma in Secondary Education, Colleges of Education	2	35
Chile	3	Concurrent	Generalist	36[2]	2018[2]
	2	Concurrent	Generalist with further mathematics education	8	181
Chinese Taipei	1	Concurrent	Elementary Teacher Education	18	3595
	2	Concurrent	Secondary Mathematics Teacher Education	19	375
Georgia	1	Concurrent	Bachelor in Pedagogy *(4 years)*	9	636
	1	Concurrent	Bachelor in Pedagogy *(5 years)*	1	23
	2	Concurrent	Bachelor of Arts in Mathematics	5	99
	2	Concurrent[6]	Master of Science in Mathematics	2	17
Germany[3]	1	Consecutive	Teachers for Grades 1–4 with Mathematics as Teaching Subject (Type 1A)	7	1286
	1	Consecutive	Teachers for Grades 1–4 without Mathematics as Teaching Subject (Type 1B)	4	1430
	3	Consecutive	Teachers of Grades 1–9/10 with Mathematics as Teaching Subject (Type 2A)	7	1093[2]
	1	Consecutive	Teachers for Grades 1–10 without Mathematics as Teaching Subject (Type 2B)	7	2433
	2	Consecutive	Teachers for Grades 5/7–9/10 with Mathematics as Teaching Subject (Type 3)	9	1162
	2	Consecutive	Teachers for Grades 5/7–12/13 with Mathematics as a Teaching Subject (Type 4)	12	1200

(continued)

Table 17.1 (continued)

Country	Level[1]	Route	Program type	No. of Institutions (sample estimate)	No. of future teachers(final year, sample estimate)
Malaysia	1	Concurrent	Malaysian Diploma of Teaching (Mathematics)	22	558
	1	Concurrent	Bachelor of Education, Primary	1	19
	1	Concurrent	Diploma of Education (Mathematics)	2	50
	2	Concurrent	Bachelor of Education (Mathematics), Secondary	1	82
	2	Concurrent	Bachelor of Science in Education (Mathematics), Secondary	6	521
	1	Concurrent	Bachelor of Education in Teaching of English as Second Language with minor in Mathematics	1[4]	No estimation possible due to low participation
	2	Consecutive	Post graduate Diploma of Education (Mathematics)	5[4]	No eligible future teachers at the time of testing
Norway	3	Concurrent	General Teacher Education (ALU) without Mathematics Option[5]	16[2]	1429[2]
	3	Concurrent	General Teacher Education (ALU) with Mathematics Option	16[2]	433[2]
	2	Consecutive	Teacher Education Program (PPU)	7	78
	2	Concurrent	Master of Science [5]	6	28
Oman	2	Concurrent	Bachelor of Education, University	1	36
	2	Consecutive	Educational Diploma after Bachelor of Science	1	17
	2	Concurrent	Bachelor of Education, Colleges of Education	6	235

(continued)

Table 17.1 (continued)

Country	Level[1]	Route	Program type	No. of Institutions (sample estimate)	No. of future teachers(final year, sample estimate)
Philippines	1	Concurrent	Bachelor in Elementary Education	171	2921
	2	Concurrent	Bachelor in Secondary Education	252	3135
Poland	3	Concurrent	Bachelor of Arts in Mathematics, First Cycle *(full-time teacher education programs); Years: 3*	16[2]	459[2]
	3	Concurrent	Master of Arts in Mathematics, Long Cycle *(full-time teacher education programs); Years: 5*	15[2]	696[2]
	3	Concurrent	Bachelor of Arts in Mathematics, First Cycle *(part-time teacher education programs); Years: 3*	4[2]	67[2]
	3	Concurrent	Master of Arts in Mathematics, Long Cycle *(part-time teacher education programs); Years: 5*	4[2]	91[2]
	1	Concurrent	Bachelor of Pedagogy Integrated Teaching, First cycle *(full-time programs); Years: 3*	27	1206
	1	Concurrent	Master of Arts Integrated Teaching, Long Cycle *(full-time programs); Years: 5*	14	864
	1	Concurrent	Bachelor of Pedagogy Integrated Teaching, First cycle *(part-time programs); Years: 3*	37	2195
	1	Concurrent	Master of Arts Integrated Teaching, Long Cycle *(part-time programs); Years: 5*	10	566
Russian Federation	1	Concurrent	Primary Teacher Education	161	8563
	2	Concurrent	Teacher of Mathematics	116	5915

(continued)

Table 17.1 (continued)

Country	Level[1]	Route	Program type	No. of Institutions (sample estimate)	No. of future teachers(final year, sample estimate)
Singapore	1	Concurrent	Diploma of Education, Primary Option A	1	53
	1	Concurrent	Diploma of Education, Primary Option C	1	119
	1	Concurrent	Bachelor of Arts in Education, Primary	1	33
	1	Concurrent	Bachelor of Science in Education, Primary	1	42
	1	Consecutive	Post-Graduate Diploma in Education, Primary Option A	1	75
	1	Consecutive	Post-Graduate Diploma in Education, Primary Option C	1	102
	2	Consecutive	Post-Graduate Diploma in Education, Secondary*, January 2007 intake*	1	111
	2	Consecutive	Post-Graduate Diploma in Education, Lower Secondary, *January 2007 intake*	1	67
	2	Consecutive	Post-Graduate Diploma in Education, Secondary*, July 2007 intake*	1	153
	2	Consecutive	Post-Graduate Diploma in Education, Lower Secondary, *July 2007 intake*	1	100
Spain (primary education only)	1	Concurrent	Teacher of Primary Education	72	3845
Switzerland (German speaking parts)	1	Concurrent	Teachers for Grades 1–2/3 *(Kindergarten and 1.-2. Grade)*	5	106
	1	Concurrent	Teachers for Grades 1–2/3 *(Kindergarten and 1.-3. Grade)*	2	54
	1	Concurrent	Teachers for Primary School (Grades 1–6) *(Kindergarten and 1.-6. Grade)*	2	304
	1	Concurrent	Teachers for Primary School (Grades 1–6)	12	745
	1	Concurrent	Teachers for Primary School (Grades 3–6)	2	43
	2	Concurrent	Teachers for Secondary School (Grades 7–9)	6	177

(continued)

Table 17.1 (continued)

Country	Level[1]	Route	Program type	No. of Institutions (sample estimate)	No. of future teachers(final year, sample estimate)
Thailand	3	Concurrent	Bachelor of Education	45	1240[2]
	3	Consecutive	Graduate Diploma in Teaching Profession	9	124[2]
United States (public institutions)	1	Concurrent	Primary Concurrent	382	20597
	2	Concurrent	Secondary Concurrent	303	2246
	3	Concurrent	Primary + Secondary Concurrent	74[2]	3472[2]
	1	Consecutive	Primary Consecutive	81	2031
	2	Consecutive	Secondary Consecutive	85	620
	3	Consecutive	Primary + Secondary Consecutive	20[2]	172[2]

[1]1 = primary; 2 = lower secondary; 3 = primary and lower secondary
[2]Estimate from sample of future teachers who took the primary test
[3]The administrative units of the 16 federal states are considered as being the institutions in the sense of the TEDS definition
[4]Estimate from sampling frame; could not be estimated from sample data
[5]Program was not considered as being part of the TEDS-M core target population. Further information is given in the appendix
[6]According to information given by the national research coordinator after the survey administration, this program-type takes a consecutive structure within one of the two institutions. Note that both programs are labeled *concurrent* in the International Database

Target Populations: International Requirements and National Implementation

TEDS-M covered all program types preparing future teachers for the teaching of mathematics at primary and lower secondary school levels. Both concurrent and consecutive program types were of interest.

The international target population of *TP institutions* was defined as the set of secondary or post-secondary schools, colleges, or universities that offer structured opportunities to learn (i.e., a program or programs) on a regular and frequent basis to future teachers within a route of teacher preparation (see the TEDS-M Conceptual Framework [Tatto et al., 2008] for key definitions).

National research coordinators were asked to list all TP program types within the defined routes and to indicate which were of principal interest to TEDS-M (Table 17.1). The institutions providing those program types are referred to as the national defined target population.

Programs surveyed in TEDS-M did not necessarily offer mathematics classes, but always provided courses on the pedagogy of mathematics (IEA, 2007b).

The target population of *educators* was defined as all persons with regular, repeated responsibility for teaching future teachers of mathematics one of the compulsory courses of their program in any year of the program. That target population could comprise up to three subpopulations:

- **Educators in mathematics or mathematics pedagogy**: persons responsible for teaching one or more of the program's required courses in mathematics or mathematics pedagogy during the study's data collection year at any stage of the institution's TP program;
- **Educators in general pedagogy**: persons responsible for teaching one or more of the program's required courses in foundations or general pedagogy (other than a mathematics or mathematics pedagogy course) during the study's data collection year at any stage of the institution's teacher preparation program;
- **Educators belonging to both groups as described above**: persons responsible for teaching one or more of the program's required courses in mathematics or mathematics pedagogy and required courses of general pedagogy during the study's data collection year at any stage of the institution's teacher preparation program.

Finally, the target population of *future teachers* comprises all members of a route in their last year of training enrolled in an institution offering formal opportunities to learn to teach mathematics, with the explicit goal of preparing individuals to teach mathematics in any of Grades 1 to 8. TEDS-M distinguishes between two different groups of future teachers: future teachers who would be certified to teach to primary students and future teachers who would be certified to teach to lower secondary students. These two groups are referred to as two distinct levels. In some countries, the distinction between primary and lower secondary levels is not feasible within a program. For example, teachers may be prepared for both levels because they will be expected to teach at any level from Grade 1 to Grade 8 in the school where they will work.

Table 17.2 presents the size of the national target populations of TEDS-M. It should be noted that not all institutions listed offer education for both the primary and the secondary levels. Also, as can be seen in Table 17.2, the population sizes estimated by the national teams before sampling and data collection (in columns labeled "Sampling Frame") deviate, sometimes considerably, from those estimated from the surveyed sample (columns labeled "Sample Estimate" [sum of weights]). These deviations reflect the fact that, for some participating countries, the compilation of a reliable sampling frame with proper measures of the size of the institution was a task that could not be fulfilled. One of the typical consequences of working with imperfect sampling frames (i.e. incomplete or out-of-date information) is increased sampling errors (i.e., wider margins of error) for the survey estimates.

Table 17.2 Size of national defined target populations by participating country

Country	Institutions		Future primary teachers		Future lower secondary teachers		Educators
	Sampling frame[a]	Sample estimate (sum of weights)	Sampling frame[a]	Sample estimate (sum of weights)	Sampling frame[a]	Sample estimate (sum of weights)	Sample estimate (sum of weights)[b]
Botswana	7	7	91	100	56	60	44
Canada (4 provinces)	30	30	Not available	728	Not available	686	282
Chile	50	40	2378	2018	2511	2242	729
Chinese Taipei	34	39	3589	3595	444	375	339
Georgia	10	10	697	659	113	116	64
Germany	16	16	8145	6242	3789	3383	3944
Malaysia	34	30	3110	627	845	603	457
Norway	45	45	1589	1862	1689	2092	Data not processed
Oman	7	7	No primary education at present		287	288	103
Philippines	417	289	4593	2921	3266	3135	2847
Poland	92	91	5800	6144	1308	1344	1181
Russian Federation	182	177	15618	8563	6872	5915	3135
Singapore	1	1	433	424	462	431	91
Spain (primary education only)	72	72	7028	3845	Not covered		770
Switzerland (German speaking parts)	16	16	1230	1252	175	177	416
Thailand	46	46	1354	1364	1354	1368	354
USA (Public institutions)	498	408	45482	26272	15160	7098	9500

[a]After institution level exclusions
[b]Population figures for educators were not available on the sampling frames

Sample Size Requirements and Implementation

Minimum sample sizes within countries were set to allow for reliable estimation and modelling, while also allowing for some amount of non-response: 50 institutions per route and level, 30 mathematics and mathematics pedagogy educators and 30 educators of general pedagogy per selected institution, and an effective sample size of 400 future teachers per route and level in a given country.

The expression *effective sample size* means that the sample design must be as efficient (i.e., precise) as a simple random sample of 400 future teachers from a (hypothetical) list of all eligible future teachers that can be found in a level and route.

When a two-stage sample design was implemented, the sample size required for each level and route was larger than the nominal 400 because such designs are typically less precise than a simple random sample due to the clustering effect. The actual number of future teachers required for each level and route within the selected TP institutions and overall was dictated mainly by (a) the total number of institutions in the country; (b) the various sizes of the institutions in the country; and (c) the sample selection method (e.g., simple random, cluster random sampling) used in the institutions.

TP institutions that offered education both to future primary and to lower secondary school teachers could be part of both samples. Similarly, TP institutions that offered more than one route to students could be part of more than one sample.

Among the 17 countries participating in TEDS-M, 12 identified fewer than the minimum of 50 (or only slightly more than 50) eligible institutions. These countries conducted a census of institutions. Therefore, in these countries, the sample design can no longer be described as a two-stage cluster design; the design has become a stratified simple random sample, which is more efficient than an un-stratified simple random sample. The high precision of the estimates for these samples illustrates this effect.

For operational purposes, each institution in the sample was divided into subgroups defined by the level × route × program type combinations. These subgroups were called Teacher Preparation Units (TPUs, see IEA, 2007a) and comprise the actual programs offered in a given institution. All programs within selected institutions were automatically part of the sample.

For example, at the time of TEDS-M, the Philippines was offering only one teacher education route (concurrent) per education level (see Table 17.1), with one program type associated with each level: the Bachelor in Elementary Education and the Bachelor in Secondary Education. Hence, teacher preparation institutions in the Philippines had either one TPU, for either the primary or secondary level (not both), or two TPUs, for both the primary and secondary levels.

In Malaysia, at the time of TEDS-M, teacher preparation institutions were offering, collectively, four different program-types for future primary teachers and three different program-types for future lower-secondary teachers (see Table 17.1). Hence, in theory, there could be up to seven TPUs in one institution. However, in practice, institutions were usually offering, if not just one program-type, only a few of the possible program-types.

Every future teacher in-scope for TEDS-M had to be allocated to exactly one and only one TPU. The minimum sample size of future teachers within institutions was set to 30 future teachers per TPU, they were selected using stratified random sampling with equal probabilities TPUs that had fewer than 30 future teachers in their final year, or where more than half of the future teachers would have been selected, were to be surveyed in full.

In countries where the number of TP institutions in a participating country was small, or where the institutions themselves were small on average, it was necessary to select all eligible future teachers for the survey to reach the TEDS-M precision requirements.

National Sampling Strategies

Participating countries could suggest variations or adaptations of the international sampling plan to better suit their national needs. All changes to the international sampling plan had to be reviewed and approved by the sampling team, the sampling referee and the international study center. Details of the national sampling plans are given in Meinck and Dumais (2012).

One important modification was the reduction of scope of the national implementation. Countries could choose to reduce their target populations for political, organizational, or operational reasons that would otherwise have made it extremely difficult for some national research coordinators to conduct the survey. This reduced coverage means that the survey results cannot be deemed representative of the entire national teacher education system in target of TEDS-M. The international reports reflect the reduced coverage of the national desired target populations using appropriate annotations. The national target population could be further reduced to avoid surveying very small institutions or programs of marginal importance. Where those exclusions have amounted to more than 5% of the national target population, tables and charts were annotated accordingly. The remaining population to be surveyed is called the *national target population*. Table 17.3 gives an overview of population coverage and exclusions.

Sample Selection

Since TEDS-M targeted four different populations (institutions, educators, primary school and secondary school future teachers), four different sampling plans were designed and implemented. Exhibit 4 illustrates sampling units and stages for the different populations (Fig. 17.1).

Sampling of Institutions

Participating countries were asked to provide the sampling team with a current and complete list of institutions, organized by route, level, and any classification variable deemed relevant to national interests.

Table 17.3 Coverage and exclusions

Country	Exclusions[a]	Coverage
Botswana	None	100% in all target populations
Chile	2% of institutions; 2% of educators; 3.8% of future primary teachers; 3.6% of future lower secondary teachers	100% in all target populations
Chinese Taipei	26.1% of institutions; <4% of educators; 4.5% of future primary teachers; 4.7% of future lower secondary teachers	100% in all target populations
Georgia	1.4% of future primary teachers; 1.7% of future lower secondary teachers	100% in all target populations
Germany	6% of institutions offering primary education and 3.7% of future primary teachers; 7% of institutions offering lower secondary education and 5.6% of future lower secondary teachers	100% in all target populations
	22% of institutions participating in the educator survey; <5% of educators	
Malaysia	None	Due to low participation, program type 'Bachelor of Education in Teaching of English as Second Language with minor in mathematics' not covered. (<5% of future primary teachers)
Norway	None	100% in all target populations
Oman[b]	None	100% in all target populations
Philippines	7.4% of institutions; <5% of educators; 2.1% of future primary teachers; 1.7% of future lower secondary teachers	100% in all target populations
Poland	3.8% of institutions; <5% of educators; 3.0% of future primary teachers; 0.4% of future lower secondary teachers	Institutions with consecutives programs only were not covered (8.5% of institutions; percentage of not covered educators unknown; 23.6% of future primary teachers; 29.0% of future lower secondary teachers)
Russian Federation	None	Secondary Pedagogical Institutions (amount unknown)
Singapore	None	100% in all target populations
Spain (primary education only)	None	Only institutions offering education to future primary teachers covered
Switzerland (German speaking parts only)	None	German speaking parts covered only
Thailand	None	100% in all target populations
USA (public institutions)	None	public institutions covered only[a]

[a]Refer to appendix for reasons for exclusions and further information
[b]Oman did not have any future primary education teachers during the data collection period

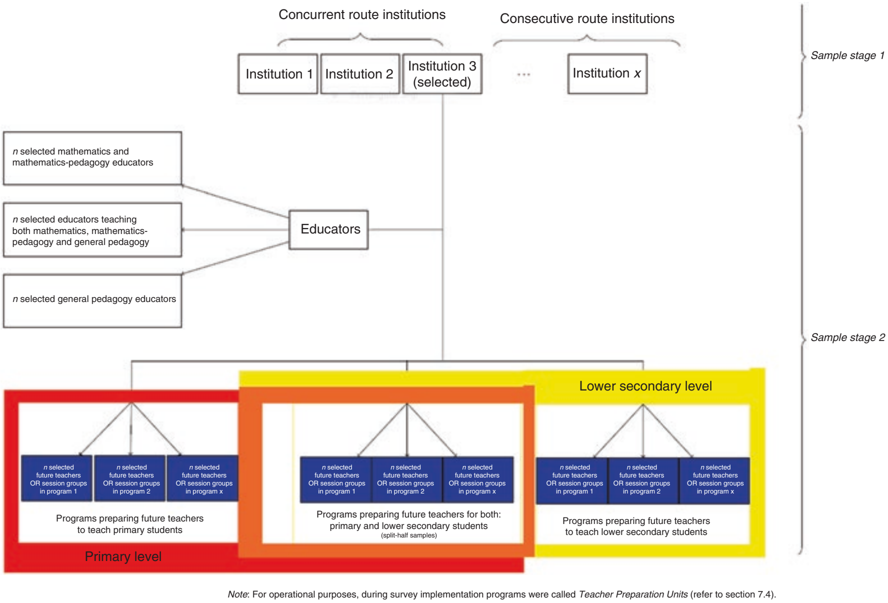


Fig. 17.1 Sampling stages and units

Where required, samples of institutions were selected by systematic random sampling within strata, according to the national sampling plans. If reliable measures of size for the institutions were available, institutions were sampled with probability proportional to size (PPS). Otherwise, or if the institutions were so small that censuses of individuals within the institutions were expected, institutions were sampled with equal probabilities. In some circumstances, institutions were sorted by additional variables and a measure of size prior to sampling. This process is often referred to as implicit stratification and ensures an approximate proportional allocation of the sample to the implicit strata.

Figure 17.2 represents the process of systematic PPS sampling within a stratum. In this diagram, the units in the sampling frame are sorted in descending order by measures of size (here, number of students), and the height of the cells reflects this size measure. A random start selects the second unit in the list, and a constant sampling interval determines the next sampled unit. In the figure, the sampling interval is represented by a fixed distance, where larger units are more likely to be selected because they span more vertical space. Sampled units are shaded.

Whenever possible, up to two replacement units were designated for each unit selected for the sample of the main survey; this was applicable solely for the sample of institutions. Institutions could be replaced if they refused to participate in the survey. Non-responding individuals, educators or future teachers could not be replaced.

The sample of institutions was also used as the first-stage sample for the educator and future teacher populations.

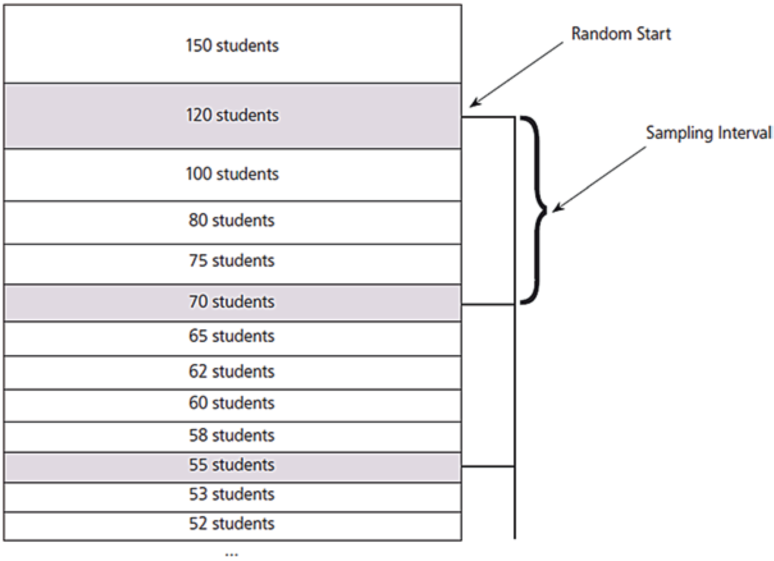


Fig. 17.2 Visualization of PPS systematic sampling. (Source: Zuehlke, 2011)

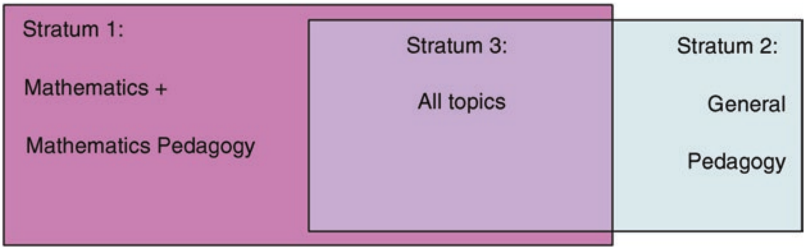


Fig. 17.3 Three strata of educators

Sampling Within Institutions – Educators

For each selected institution, a comprehensive list of eligible educators was compiled. Each educator had to be allocated to one of the educator groups described in Fig. 17.3 above. Then, using software provided by IEA DPC (WinW3S-Within Institution Sampling Software), a systematic random sample of at least 30 mathematics or mathematics-pedagogy educators and a systematic random sample of 30 general pedagogy educators were selected. In all participating countries, a census of educators was conducted in institutions where fewer than 30 educators were found in a given group, which was the case in a vast majority of institutions in all countries.

Sampling Within Institutions—Future Teachers

For the selection of future teachers within TPUs, two different procedures were implemented, using WinW3S:

1. Selection of whole session groups
   In some participating countries (e.g., Germany, Chinese Taipei, Russia), or in some selected institutions, future teachers were grouped together for organizational purposes (e.g., into tutorial groups or classroom-like groups). Such groups were called session groups in the TEDS-M survey. Particularly in very large institutions, it was sometimes operationally desirable and more convenient to select whole session groups instead of individual future teachers. When this approach was chosen, a comprehensive list of session groups was compiled. Each eligible future teacher in a TPU had to be allocated to one, and only one, session group. Then, predetermined numbers of session groups were randomly selected with equal probability. All future teachers within the selected session groups were asked to participate in the survey. The downside of this sampling approach is that the sampling design is expected to be less efficient because of clustering effects that might affect such groups. This was countered by appraising each situation and possibly increasing the within-institution sample sizes.
2. Selection of individual future teachers
   For each TPU, a comprehensive list of eligible future teachers was compiled. Then, at least 30 future teachers for that TPU were randomly selected (or all, in institutions where the random selection would be more than half the total and in institutions with fewer than 30 future teachers).

Again, because of small populations within the country or within the institution, all eligible future teachers were asked to participate in the survey in many instances.

Selected future teachers being prepared to teach to both levels were randomly divided into two groups, with one group being asked to answer for primary education and the other group being asked to answer for lower secondary education.

All sampling procedures and processes were extensively documented either by the sampling team (institution samples) or automatically by WinW3S, so that every selection step remains reproducible at any time.

Sampling for Field Trial

Prior to the main data collection, a field trial was conducted between January and April 2007 in Botswana, Chile, Chinese Taipei, Georgia, Germany, Oman, Philippines, Poland, Singapore, Spain, Switzerland, and Thailand. The other participants joined the study too late to participate in the field trial.

Convenience samples were selected for the field trial. Since, in almost all countries participating in the field trial, overlap of the field trial and main survey samples

Table 17.4 Response rates within participating institutions—field trial and main survey

Countries	Response rate future primary teachers (%, over all participating institutions) Field trial	Main survey	Increase (%)
Botswana	95	86	−9
Chile	66	79	13
Chinese Taipei	46	90	44
Georgia	55	77	22
Germany	77	82	5
Philippines	87	91	4
Poland	63	79	16
Singapore	No calculation possible	90	n.a.
Spain	32	87	55
Switzerland (German speaking parts)	41	76	35
Thailand	91	99	8

could not be avoided, convenience selection for the field trial gave countries the possibility to purposively select institutions that would be willing to participate in both parts of the survey. In almost every country, a convenience sample of five institutions for each level and route was selected by the national study center.

The field trial helped to identify one big challenge of this survey: achieving high participation rates. Many national research coordinators reported difficulties with picking a convenience sample of institutions; also, the response rates within institutions were often very low. Experiences from the field trial contributed in significant ways to developing strategies to enhance the willingness of all targeted populations to participate in this survey. In fact, out of ten countries participating in the field trial, nine were able to significantly increase the response rates of future primary teachers in the main survey (Table 17.4).

Computing Sampling Weights for TEDS-M

Most of the statistics produced for TEDS-M are derived from data obtained through samples of institutions, educators, and future primary and secondary school teachers. For these statistics to be meaningful for a country, they need to reflect the whole population from which they were drawn and not merely the sample used to collect them.

Each national sampling plan is unique, ranging from a stratified multi-stage probability sampling plan with unequal probabilities of selection to a simple and complete census of all units of interest. The unequal chances of institutions and individuals being part of the sample need to be considered for any analysis in order to obtain unbiased estimates of the population features (Lohr, 1999). This can be done by using so-called *estimation* (or *final*) weights. The TEDS-M databases provide estimation weights for each population. This chapter gives an overview of how these weights were computed. Works illustrating the principles of weighting in ILSAs include Statistics Canada (2003), Rust (2014), and Meinck (2015).

The estimation weight in complex sample surveys is computed as the product of several weighting and non-response adjustment factors reflecting the sampling probabilities and non-response patterns at each sampling stage and in each stratum. The estimation weight indicates how many population units a participating unit represents, taking all of these factors into consideration.

Clearly, since each country had to adapt the general sample design of TEDS-M to its own conditions, the estimation weights had to conform to the national adaptations. In countries where censuses were conducted, it was sufficient to adjust the collected data for non-response[4] in order to obtain unbiased estimates of the population parameters.

In many ILSAs, one set of country-level estimation weights is produced for each participating country. However, in the case of TEDS-M, four sets of estimation weights are required to reflect the various surveys that comprise TEDS-M: the institutions, the educators, and the future teachers of mathematics at primary and at lower secondary levels. While this chapter will only give a simplified introduction to the process of weighting, details can be obtained from Dumais and Meinck (2013).

Institution Base Weight (Institution Design Weight)

The first stage of sampling in TEDS-M is the sampling of institutions. In many countries, or strata within a country, the sample of institutions is a census; in other countries, or strata within a country, the sample of institutions was drawn according to a systematic random sampling scheme with selection probabilities proportional to size. When a census sample of institutions was implemented in a country or in a stratum of a country, then the institution base weight is set to 1. The institution base weight is given by

$$WGTFAC1_{hi} = \begin{cases} 1 & \text{for censuses} \\ \dfrac{M_h}{n_h \times M_{hi}} & \text{for PPS random samples} \end{cases}$$

for each institution *i,* and each stratum *h* with M_{hi} denoting the measure of size of a specific selected institution and M_h the cumulated measure of size in stratum *h*. For example, a country provided the total number of future teachers as size measure. Then, the institution base weight of an institution with 50 future teachers in a

[4] Under the hypothesis of non-informative response model, or that items are missing completely at random.

stratum with a total of 1000 future teachers and the sample size n in that stratum being 2 institutions, would be 1000/(2 × 50) = 100.

The institution base weight was computed once and then fixed, irrespective of which of the subsequent four different target populations of TEDS-M was concerned.

It should be noted that the computation of any base weight follows the same rationale: It is computed within a given stratum, respecting the number of sampled units and their measure of size, if PPS sampling was adopted as shown in the formula above. Design weights for simple random sampling (SRS) samples are even simpler to compute: The total number of units is divided by the number of units to sample. For example, when selecting four units out of ten, each sampled unit gets a design weight of 2.5 (=10/4), because the set of four sampled units represents the ten units on the sampling frame.

Institution Non-response Adjustment Factor

In spite of all efforts to secure the full participation of all selected institutions and of their members, some were unable or unwilling to participate. Those institutions where the participation of individuals was below 50% are deemed to be non–participating for the respective population of interest. The institutions that the non-participating institutions would have represented must still be represented. Therefore, a non-response adjustment factor is required within each stratum. Because of the multiplicity of types of respondents in TEDS-M, multiple institution non-response adjustment factors are required: for the educators and the future teachers of mathematics at primary and lower secondary levels.

For each stratum h if r_h out of the n_h selected institutions participated in TEDS-M, then the non-response adjustment factor is given by

$$WGTADJ1_h = \begin{cases} \dfrac{n_h}{r_h}, & \text{for participating institutions} \\ 0, & \text{for non-participating institutions} \end{cases}$$

acknowledging that, if the form is identical, the value of the adjustment factor may change with the population of interest.

Again, it should be noted that non-response adjustment factors are computed always applying the same rationale. In a first step, it is assumed that, within a stratum, the non-respondents (or non-responding units) do not differ systematically from the responding ones. Then, the adjustment factor is computed such that the weight carried by the non-respondents gets evenly re-distributed among those participating. Continuing with the example given in 4.1, if only 2 of the four sampled units agree to participate, the adjustment factor for each participating unit is 2 (=4/2).

Final Institution Weight

The final institution weight is the product of the institution base weight and the institution non-response adjustment factor.

Because there are several populations of interest, the institution final weight was computed separately for each population of interest. This is reflected in the population identifier (I = institution, E = educator, P = future primary teacher, S = future secondary teacher) attached to the name of each of the final institution weights in the respective file of the international database (INSWGT**I**, INSWGT**E**, INSWGT**P**, INSWGT**S**)[5]. All estimates pertaining to institution-specific features should use the appropriate final institution weight.

TPU (Teacher Preparation Unit) Non-response Adjustment Factor

For operational purposes, each institution in the sample was divided into subgroups of future teachers that are defined by the combination of the level (primary, secondary), the route (concurrent, consecutive) and the specific program type. These subgroups are called Teacher Preparation Units (TPUs) or programs (refer also to section "Sample selection" in this chapter). Within each selected institution, all TPUs were automatically selected to participate in the survey. Hence, it was not necessary to apply a TPU base weight (it would always be equal to 1).

A selected institution was asked to complete one Institutional Program Questionnaire for each TPU. The data coming from these questionnaires are stored in the institution files (DIG files) of the International Data Base (IDB).[6] Despite all efforts to gather all requested questionnaires in the participating institutions, the questionnaire was not completed for one or more TPUs in some institutions, in which case, a TPU non-response adjustment factor had to be calculated. This adjustment was done within strata, across institutions but within the level-route combination. Thus, the estimation weight for, say, all concurrent primary TPUs within one stratum that responded to the institutional program questionnaire was adjusted to account for those that did not respond.

Final TPU (Teacher Preparation Unit) Weight

The final TPU weight is the product of the institution base weight, the institution non-response adjustment factor, and the TPU non-response adjustment factor.

[5] Refer to the TEDS-M IDB User Guide for further information (Brese & Tatto, 2012).

[6] The TEDS-M IDB can be accessed through http://www.iea.nl/our-data.

The final TPU weight can be found in the institution file (DIG) only. Any analysis producing estimates based on data from the Institutional Program Questionnaires should use the final TPU weight.

Session Group Base Weight

As explained earlier, within each TPU, it was possible to further divide future teachers into subgroups, called *session groups*, for organizational purposes and (in rare instances) to select only some session groups from a list of session groups according to the national sampling plan. This selection step had to be taken into account when calculating the final future teacher weight. In many participating countries, however, it was decided not to select (some out of many) session groups but rather individual future teachers from an exhaustive list of all future teachers within one TPU. In this case, one single session group is created, and its base weight is set to 1.

Because there are two populations of interest (future teachers of primary and lower secondary schools), the session group base weight is calculated separately for each target population. This is reflected in the population identifier attached to the name of the session group base weight in the respective file of the international database (WGTFAC2**P** and WGTFAC2**S**).

Admitting that session groups were often artificial groups of students brought together for test administration, no session group non-response adjustment factor was calculated. Instead, the non-response adjustment was calculated at the future teacher level (see section "Future teacher non-response adjustment factor" below).

Future Teacher Base Weight

If no session group sampling was performed, systematic random samples of future teachers with equal probabilities were selected from each TPU (at least 30 future teachers by design). The future teacher base weight is computed as the inverse of the selection probability of a future teacher within a TPU.

In institutions where session group sampling was performed, all future teachers within a selected session group were automatically selected for the survey. In this case, the future teacher base weight is 1.

Because there are two populations of interest (future teachers of primary school and lower secondary school), the future teacher base weight is calculated separately for each target population. This is reflected in the population identifier attached to the name of future teacher base weight in the respective file of the international database (WGTFAC3**P** and WGTFAC3**S**).

Future Teacher Non-response Adjustment Factor

Unfortunately, not all selected future teachers were able or willing to participate in TEDS-M. The future teachers that were represented by the non-participating future teachers still need to be represented by the sample. This is why a non-response adjustment factor is introduced. The non-response adjustment was done within each TPU but across session groups. You may refer to Fig. 17.1 to visualize this process. The adjustment was done within the blue boxes. Let us imagine 30 future teachers were sampled within the unit represented by one of the blue boxes. Ten of them, however, did not participate. In this case, the remaining 20 received an adjustment factor of 30/20 = 1.5.

As already pointed out, because there are two populations of interest (future teachers of primary and lower secondary), the future teacher non-response adjustment factor is calculated separately for both target populations. This fact will be reflected in the population identifier attached to the name of the future teacher non-response adjustment factor in the respective file of the international database (WGTADJ3**P** and WGTADJ3**S**).

Future Teacher Level Weight

In some participating countries, future teachers would be certified to teach to primary *and* lower secondary students. Those future teachers are eligible for both target populations of the TEDS-M future teacher survey. However, it would have been very difficult to convince those future teachers to participate in both surveys, i.e., to complete both a primary and a lower secondary questionnaire. Thus, those future teachers were randomly assigned to one of the two surveys. The future teacher-level weight will adjust for this procedure, making sure the future teachers assigned to the respective other group are still represented by the sample.

Final Future Teacher Weight

The final future teacher weight (estimation weight) is the product of the final institution weight, the session group base weight, the future teacher base weight, the future teacher non-response adjustment factor, and the future teacher level weight, calculated for the respective future teacher population (INSWGT**P** or INSWGT**S**). All estimates pertaining to the populations of future teachers should be computed by using the final future teacher weight.

Educator Base Weight

In each participating institution, up to three strata of educators could be created: that of mathematics and mathematics pedagogy educators (stratum 1), that of general pedagogy educators (stratum 2) and that of those educators teaching all topics (stratum 3, Fig. 17.3).

Samples of 30 educators were required for each of the two groups of educators (mathematics and mathematics pedagogy, and general pedagogy). All educators were asked to complete specific parts in the (unique) educator questionnaire. While educators belonging to strata 1 and 2 had to complete only the parts that concerned their specific teaching responsibilities, educators belonging to stratum 3 were asked to complete the whole questionnaire.

For educators, systematic random samples with equal probabilities were selected from each stratum. The educator base weight is used to bring the individual educator's information to the level of his or her institution. Generally, all TEDS-M-eligible educators in an institution were selected. For those educators, the base weight is equal to 1, and each educator represents only him or herself within his or her institution.

Educator Non-response Adjustment Factor

Not all selected educators were able or willing to participate in TEDS-M. The educators who were represented by the non-participating educators still need to be represented by the sample. Hence, an educator non-response adjustment factor is introduced within each institution. Again, one can refer to Fig. 17.1 for getting a clearer idea of this process: the adjustment was carried out within the unshaded boxes on the left-hand side, with the effect that, e.g., only general pedagogy educators who participated in the study would represent general pedagogy educators who refused participation.

In some cases, none of the selected educators in an educator group within an institution responded at all. Consequently, the non-response adjustment could not be calculated according to the standard procedures, since there were no respondents who could carry the weight of the non-respondents. In these situations, the non-response adjustment for educators in this educator group in the affected institution was done across institutions, but within stratum and within the educator group (i.e., mathematics and mathematics pedagogy, and general pedagogy).

Final Educator Weight

The final educator weight is the product of the educator base weight, the educator non-response adjustment factor, and the final institution weight calculated for the educator population (INSWGTE). All estimates pertaining to the populations of educators should be computed using the final educator weight.

Table 17.5 Example of incorrect (unweighted) and correct (weighted) analyses

Variable		Unweighted, % (Biased estimate)	Weighted, % (Unbiased estimate)
MFA012A	Yes	29.8	18.7
Had family responsibilities that made it difficult to do my best	No	70.2	81.3
MFA012B	Yes	18.3	23.3
Had to borrow money	No	81.7	76.7
MFA012C	Yes	51.1	45.5
Had to work at a job	No	48.9	54.5

Importance of Using Sampling Weights for Data Analysis

Although the international sampling plan was prepared as a self-weighting design (i.e., aiming for equal final weights of respondents), the actual conditions in the field made that ideal plan impossible to be executed. Neglecting that the final weights are different among the responding elements of a given sample during data analysis will lead to biased results.

The following example, taken from the TEDS-M user guide for the international database (Brese & Tatto, 2012), illustrates the risk of bias if not using weights in research with TEDS-M data.

A researcher may be interested in finding out whether there were any circumstances that prevented future teachers from focusing on their teacher preparation program. Consider, for example, primary mathematics specialists (DPG file, variable TARGETP = 4) in Germany. The corresponding analysis variables are MFA012A, MFA012B, and MFA012C. The results of this analysis, using unweighted versus weighted data, are displayed in Table 17.5. Focusing on the variable MFA012A, if one omitted the final weights, one would conclude that almost one in three future teachers in the population of interest had family responsibilities (29.8%), while, in actual fact, this is only the case for about one in five students (18.7%) once the data are correctly weighted.

Participation Rates

Non-response adjustments were done in TEDS-M assuming that non-response within an adjustment cell occurred completely at random. Of course, this is a strong assumption, given that very little or even nothing at all is known about the non-responding units and individuals. This is the reason why all ILSAs, and therefore also TEDS-M, employ strict standards for participation rates.

The following example illustrates the risk related to non-response and thereby shows the importance of high participation rates. Consider a country with an average achievement score of 530 points, as shown in Fig. 17.4. If, in such country,

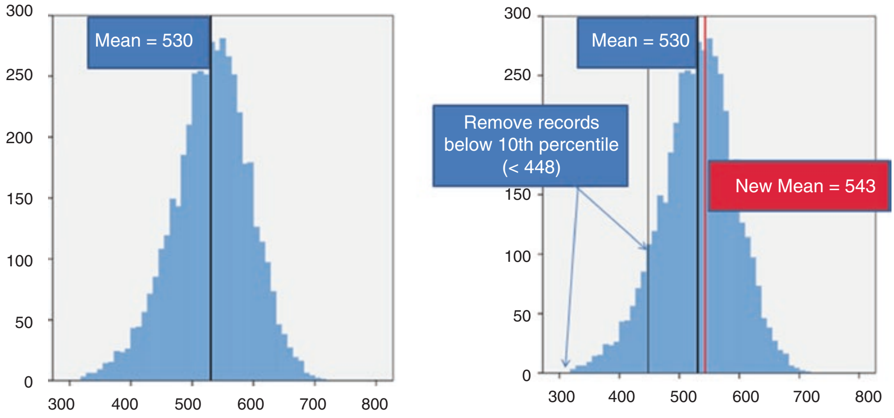


Fig. 17.4 Risks of bias due to non-response

all of, and only, the 10% lowest achievers refuse to participate in the survey, the average score rises by 13 points—a significant increase leading to misleading conclusions.

The TEDS-M quality standards required minimum participation rates for all target populations of the survey if statistics are to be reported purporting to describe characteristics of those populations. The aim of these standards is to ensure that bias resulting from non-response is kept within acceptable limits.

For each country, the participation rates are calculated and reported separately for each of the four different target populations of TEDS-M. Reports describing the results for each target population considered the participation rate for that target population only.

The minimum requirement for TEDS-M to publish statistical key data for international comparisons for each population is

- that the overall (combined) participation rate (weighted or un-weighted) of that population is at least 75%

  OR

- that the participation rate (weighted or un-weighted) of institutions for the considered population and the participation rate for individuals within the participating institutions are both at least 85%.

The intention of the following section is to report the participation rates achieved in TEDS-M. Details on the mathematical computations can be found in Dumais and Meinck (2013).

Sampling Adjudication Outcomes

The adjudication of the data was done for each participating country and each of the four TEDS-M survey populations separately, following the recommendations of the sampling referees[7] and in agreement with all participants of the sampling adjudication meetings.[8]

If a country did not meet the required participation rates for one (or more) of the populations, statistics were still reported for that country, but the fact of failing the requirements is annotated in the international report. This is to point any reader to the reduced reliability of the data.

The following adjudication comments were observed in reporting:

1. Reporting without any annotation—No annotation was made if all participation rate requirements were met, the exclusion rate was below 5%, and full coverage of the target population was observed.
2. Annotation because of low participation rates—This comment was added if the participation rate was below the requirement but the combined participation rate was still above 60%.[9]
3. Participation rates clearly below standards, reporting together with other countries not advisable—This comment was added if the combined participation rate dropped below 60% but was still above 30%.
4. Unacceptable (move to appendix)—This comment was added if the combined participation rate dropped below 30%.

A summary of the adjudication outcomes is shown in Table 17.6.

The achieved participation rates are displayed in Table 17.7. Compared to other social surveys, the rates are notably high, with most countries meeting the high standards in their target populations, underlining the high quality of TEDS-M survey results. Tables 17.8, 17.9, 17.10, and 17.11 show the expected and the achieved sample sizes for each population.

Estimating Sampling Error in TEDS-M

Surveys with complex designs like TEDS-M, i.e. stratified, unequal probability cluster sampling, require special attention when it comes to estimation, especially estimation of the sampling error. Both the survey design and the unequal weights need to be taken into account to obtain (approximately) unbiased estimates of

[7] Jean Dumais and Marc Joncas, Statistics Canada.

[8] Maria Teresa Tatto, Principal Investigator (PI) andTEDS-M Executive Director, Inese Berzina-Pitcher Coordinator, John Schwille Co-PI, and Sharon Senk Co-PI) as representatives of the international study center; Sabine Meinck as the representative of the IEA DPC Sampling Team.

[9] Annotations were also advised if the exclusion rate exceeded 5% or reduced coverage of the target population was observed.

Table 17.6 Summary of adjudication results

Countries	Institutions	Teacher educators	Future primary teachers	Future lower-secondary teachers
Botswana	None	None	None	None
Canada (4 provinces)	Unacceptably low participation rates; data remains un-weighted, and is not reported.			
Chile	None	Low participation rates, data is highlighted to make readers aware of increased likelihood of bias.	Combined participation rate between 60 and 75%.	Combined participation rate between 60 and 75%.
Chinese Taipei	Exclusion rate >5% (very small institutions were excluded).	None	None	None
Georgia	None	None	None	Combined participation rate between 60 and 75%; An exception was made to accept data from two institutions because, in each case, one additional participant would have brought the response rate to above the 50% threshold.
Germany	None	Low participation rates, data is highlighted to make readers aware of increased likelihood of bias. Survey of institutions and future teachers are not connected with survey of educators.	None	None

(continued)

Table 17.6 (continued)

Countries	Institutions	Teacher educators	Future primary teachers	Future lower-secondary teachers
Malaysia	Low participation rates, data is highlighted to make readers aware of increased likelihood of bias.	Low participation rates, data is highlighted to make readers aware of increased likelihood of bias.	None	None
Norway	None	Participation rates could not be calculated; data remains un-weighted, and is not reported.	Combined participation rate between 60 and 75%. An exception was made to accept data from one institution because one additional participant would have brought the response rate to above the 50% threshold. Program types ‘ALU’ and ‘ALU plus Math’ are partly overlapping populations; analysis across program types is inappropriate due to this overlap.	Participation rates low, data is highlighted to make readers aware of increased likelihood of bias. Program types “ALU”, “ALU plus Math”, and “Master’s” are partly overlapping populations, results derived from analysis across program types should be conducted with care to avoid undue overlap of populations.
Oman	Oman provided education for future secondary teachers only at the time of testing.	Oman provided education for future secondary teachers only at the time of testing.	Not applicable	None
Philippines	Exclusion rate >5% (very small institutions were excluded).	None	None	None

(continued)

Table 17.6 (continued)

Countries	Institutions	Teacher educators	Future primary teachers	Future lower-secondary teachers
Poland	Institutions with consecutive programs only were not covered.	Combined participation rate between 60 and 75%; institutions with consecutive programs only were not covered.	Combined participation rate between 60 and 75%; institutions with consecutive programs only were not covered.	Combined participation rate between 60 and 75%; institutions with consecutive programs only were not covered.
Russian Federation	Secondary pedagogical institutions were not covered.	Secondary pedagogical institutions were not covered.	Secondary pedagogical institutions were not covered.	An unknown percentage of surveyed future teachers were already certificated primary teachers.
Singapore	None	None	None	None
Spain (Primary education only)	None	None	None	Not applicable
Switzerland (German speaking parts)	None	Low participation rates, data is highlighted to make readers aware of increased likelihood of bias.	None	None
Thailand	None	None	None	None
United States (Public Institutions)	None	Unacceptably low participation rates; data remains un-weighted, and is not reported.	An exception was made to accept data from two institutions because, in each case, one additional participant would have brought the response rate to above the 50% threshold. Items with low responses are clearly marked.	Combined participation rate between 60 and 75% only; An exception was made to accept data from one institution because one additional participant would have brought the response rate to above the 50% threshold. Items with low responses are clearly marked.

Table 17.7 Un-weighted participation rates for institutions, future primary and lower secondary teachers and educators

	Institutions (completion of IPQs)	Future primary teachers			Future lower secondary teachers			Educators		
	IPR_I	IPR_P	WPR_P	CPR_P	IPR_S	WPR_S	CPR_S	IPR_E	WPR_E	CPR_E
Country	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
Botswana	100	100	86	86	100	88	88	100	98	98
Canada (4 provinces)	37	7	69	5	29	72	21	33	79	26
Chile	88	86	79	68	83	76	63	70	77	54
Chinese Taipei	100	100	90	90	100	97	97	100	95	95
Georgia	100	100	77	77	100	67	67	100	97	97
Germany	100	93	82	76	100	81	81	92	61	56
Malaysia	57	96	97	93	86	84	72	73	77	57
Norway	96	81	78	63	73	79	58	Data not processed		
Oman	100	Not applicable			100	93	93	100	85	85
Philippines	85	80	91	75[a]	91	92	83	85	94	80
Poland	86	86	79	68	82	84	69	79	86	68
Russian Federation	91	96	94	91	98	94	92	98	92	91
Singapore	100	100	90	90	100	91	91	100	85	85
Spain (primary education only)	96	90	87	78	Not applicable			92	93	85
Switzerland (German speaking parts)	94	100	76	76	100	81	81	75	69	52
Thailand	96	98	99	97	98	98	96	93	94	88
USA (public institutions)	83	85	85[a]	71	82	84	69	23	59	14

[a]Weighted participation rate

Table 17.8 Institutions—expected and achieved sample sizes

Countries	Number of institutions in original sample	Ineligible institutions	Total number of institutions providing response to the IPQ	Number of expected IPQs within participating institutions	Number of returned IPQs within participating institutions
Botswana	7	0	7	7	7
Canada (4 provinces)	30	0	11	32	23
Chile	50	10	35	42	38
Chinese Taipei	19	0	19	19	19
Georgia	10	0	10	17	17
Germany	16	0	16	51	51
Malaysia	34	4	17	33	20
Norway	47	2	43	43	43
Oman	7	0	7	8	8
Philippines	80	20	51	83	82
Poland	92	1	78	130	125
Russian Federation	58	1	52	98	88
Singapore	1	0	1	10	10
Spain (primary education only)	50	0	48	48	48
Switzerland (German speaking parts)	16	0	15	32	28
Thailand	46	0	44	53	51
USA (public institutions)	60	0	50	136	117

sampling error. Failing to do so can lead to severe underestimation of the sampling error. While exact formulae exist in theory for such designs, the required computations become practically impossible as soon as the number of primary units selected per stratum exceeds two. Approximate solutions have been proposed over the years for handling such cases. An important class of solutions is that of *resampling* or *replication*. *Interpenetrating sub-samples*, *Balanced Repeated Replication*, the *jackknife*, and the *bootstrap* are the best known examples of replication methods (see, for example, Lohr 1999, Rust and Rao 1996, or Wolter 2007 for a review of these methods).

The Balanced Repeated Replication (BRR) (McCarthy, 1966), with Fay's modification (Fay, 1989; Judkins, 1990), was adopted for the estimation of the sampling

Table 17.9 Future primary teachers—expected and achieved sample sizes

Countries	Number of institutions in original sample	Ineligible institutions	Total number of institutions participated	Number of sampled future primary teachers in participating institutions	Number of participating future primary teachers
Botswana	4	0	4	100	86
Canada (4 provinces)	28	0	2	52	36
Chile	50	14	31	836	657
Chinese Taipei	11	0	11	1023	923
Georgia	9	0	9	659	506
Germany	15	0	14	1261	1032
Malaysia	28	4	23	595	576
Norway	32	0	26	709	551
Oman	Not applicable				
Philippines	60	19	33	653	592
Poland	91	0	78	2673	2112
Russian Federation	52	1	49	2403	2266
Singapore	1	0	1	424	380
Spain (primary education only)	50	0	45	1259	1093
Switzerland (German speaking parts)	14	0	14	1230	936
Thailand	46	0	45	666	660
USA (public institutions)	60	0	51	1807	1501

error of the estimates produced for TEDS-M. A detailed description of the method and its application in TEDS-M can be obtained from Dumais and Meinck (2013).

Each of the four TEDS data files, comprising data from the four different populations, contains two sets of BRR variables. One set refers to the respective final institution weight. The second set of BRR variables refers to the respective final population weight (final TPU weight, final future teacher weight—primary/secondary, final educator weight). These variables have to be used when estimating sampling error and confidence intervals and when performing significance tests. See Brese and Tatto (2012) for further details on the handling of variance estimation variables.

Table 17.10 Future lower secondary teachers—expected and achieved sample sizes

Countries	Number of institutions in original sample	Ineligible institutions	Total number of institutions participated	Number of sampled future lower secondary teachers in participating institutions	Number of participating future lower secondary teachers
Botswana	3	0	3	60	53
Canada (4 provinces)	28	0	8	174	125
Chile	50	10	33	977	746
Chinese Taipei	21	2	19	375	365
Georgia	6	0	6	116	78
Germany	13	0	13	952	771
Malaysia	7	0	6	462	389
Norway	47	2	33	724	572
Oman	7	0	7	288	268
Philippines	60	7	48	800	733
Poland	28	0	23	355	298
Russian Federation	50	1	48	2275	2141
Singapore	1	0	1	431	393
Spain (primary education only)	Not applicable				
Switzerland (German speaking parts)	6	0	6	174	141
Thailand	46	0	45	667	652
USA (public institutions)	59	3	46	726	607

Conclusions

The conduct of the TEDS-M study taught us that the implementation of a sound random sampling design covering higher education is particularly challenging, but not impossible. Key elements where the careful inspection of the structure of all participating education systems, and determining common grounds and procedures how to address national specifics. This chapter can give valuable guidance to researchers wishing to enter similar endeavors.

Table 17.11 Educators—expected and achieved sample sizes

Countries	Number of institutions in original sample	Ineligible institutions	Total number of institutions participated	Number of sampled educators in participating institutions	Number of participating educators
Botswana	7	0	7	44	43
Canada (4 provinces)	30	0	10	94	74
Chile	50	10	28	510	392
Chinese Taipei	19	0	19	205	195
Georgia	10	0	10	64	62
Germany	50	0	46	792	482
Malaysia	34	4	22	330	255
Norway	Data not processed				
Oman	7	0	7	99	84
Philippines	80	20	51	626	589
Poland	92	1	72	857	734
Russian Federation	58	1	56	1311	1212
Singapore	1	0	1	91	77
Spain (primary education only)	50	0	46	574	533
Switzerland (German speaking parts)	16	0	12	318	220
Thailand	46	0	43	331	312
USA (public institutions)	60	0	14	407	241

References

Brese, F., & Tatto, M. T. (Eds.). (2012). *TEDS-M 2008 user guide for the international database*. Amsterdam, The Netherlands: International Association for Educational Achievement (IEA).

Dumais, J., & Meinck, S. (2013). Estimation weights, participation rates, and sampling error. In M. T. Tatto (Ed.), *Teacher Education and Development Study in Mathematics (TEDS-M): Technical report*. Amsterdam, The Netherlands: International Association for Educational Achievement (IEA).

Dumais, J., Meinck, S., Tatto, M. T., Schwille, J., & Ingvarson, L. (2013). Sampling design. In M. T. Tatto (Ed.), *Teacher Education and Development Study in Mathematics (TEDS-M): Technical report*. Amsterdam, The Netherlands: International Association for Educational Achievement (IEA).

Fay, R. E. (1989). Theoretical application of weighting for variance calculation. In *Proceedings of the section on survey research methods of the ASA* (pp. 212–217).
Gregory, K. D., & Martin, M. O. (2001). *Technical standards for IEA studies. An annotated bibiography*. Amsterdam, The Netherlands: International Association for Educational Achievement (IEA).
IEA. (2007a). *TEDS-M survey operations procedures*. Hamburg, Germany: IEA-DPC.
IEA. (2007b). *TEDS-M sample preparation manual*. Hamburg, Germany: IEA-DPC.
Judkins, D. (1990). Fay's method for variance estimation. *Journal of Official Statistics, 6*, 223–240.
Lohr, S. (1999). *Sampling: Design and analysis*. New York, NY: Duxbury Press.
McCarthy, P. J. (1966). Replication: An approach to the analysis of data from complex surveys. In *Vital and health statistics* (Series 2, No. 14). Hyattsville, MD: National Center for Health Statistics.
Meinck, S. (2015). Computing sampling weights in large-scale assessments in education. *Survey insights: Methods from the field*, *weighting: Practical issues and 'how to' approach*. Retrieved from http://surveyinsights.org/?p=5353
Meinck, S., & Dumais, J. (2012). Characteristics of national samples: Implementations of the international sampling design in participating countries. In M. T. Tatto (Ed.), *Teacher Education and Development Study in Mathematics (TEDS-M): Technical report*. Amsterdam, The Netherlands: International Association for Educational Achievement (IEA).
Rust, K. (2014). Sampling, weighting and variance estimation in international large-scale assessments. In L. Rutkowski, M. Von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment*. Boca Raton, FL: CRC Press.
Rust, K., & Rao, J. N. K. (1996). Variance estimation for complex estimators in sample surveys. *Statistics in Medical Research, 5*, 381–397.
Statistics Canada. (2003). *Survey methods and practices*. Retrieved from http://www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf
Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.
Wolter, K. (2007). *Introduction to variance estimation* (2nd ed.). New York, NY: Springer.
Zuehlke, O. (2011). Sample design and implementation. In W. Schulz, J. Ainley, & J. Fraillon (Eds.), *ICCS technical report* (pp. 59–68). Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Chapter 18
Developing Anchor Points to Enhance the Meaning of the Mathematical and Mathematical Pedagogy Score Scales from the TEDS-M Study

Mark D. Reckase

Abstract The Teacher Education and Development Study (TEDS-M), a cross-national study of teacher education programs that prepare future primary and secondary mathematics teachers, included a series of measures of mathematics achievement designed to determine what prospective teachers knew and could do concerning the mathematics that they would likely teach. One of the goals of the study was to report the information about prospective teachers' knowledge and skills in a way that is easy to understand by the numerous audiences for the results of the study. The values that are reported are often obtained using an item response theory (IRT) model that gives estimates of a location on the latent scale for a hypothetical construct. For any of these numerical values, it is difficult to interpret what a person knows or can do. At best, the numbers can be used to order persons or groups according to the magnitude of what they know or can do, but not whether they have particular capabilities such as being able to solve systems of linear equations. Meaning was added to the reporting scales by attaching detailed substantive descriptors to selected points on the scale called anchor points. The logic and methodology used to add descriptions to anchor points is similar in many ways to methodologies used elsewhere, such as item mapping or scale anchoring. However, there are some important differences between the procedure used for TEDS-M and those used with other testing programs such as NAEP, TIMSS, and PISA. The purpose of this chapter is to describe how the descriptive information was added to the scales using the TEDS-M anchor points, and to discuss how the interpretations derived from these descriptions differ from those used with other assessment programs.

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

M. D. Reckase (✉)
Michigan State University, East Lansing, MI, USA
e-mail: reckase@msu.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_18

Introduction

The TEDS-M study included a series of measures of mathematics achievement designed to determine what prospective teachers knew and could do in the subject matter that they would likely teach (Tatto et al., 2012). One of the goals of the study was to report the information about prospective teachers' knowledge and skills in a way that is easy to understand by the numerous audiences for the results of the study. Unfortunately, the numerical values that are typically used to report the results of an achievement test do not have any intrinsic meaning. For simple testing programs, such as classroom achievement tests, the numerical values are often simple sums of the scores on all the items. The simplest interpretation is that the score indicates how much of what a teacher (or test developer) chooses to put on the test is known by the examinee. Interpretations beyond that are challenging without information about test specifications, test difficulty, average scores, etc. If parents are told that their child obtained a value of 46 on the test, it has little interpretive meaning without much additional information and explanation.

For more elaborate testing programs, such as large-scale state assessments and international comparison studies, the numerical values that are used to report the results of the assessment are even more abstract. Examinees may respond to different sets of items, and the test specifications need to be appropriate for multiple schools and/or multiple countries (e.g., Watermann & Klieme, 2002). The values that are reported are often obtained using an item response theory (IRT) model that gives estimates of a location on the latent scale for a hypothetical construct. This scale has an arbitrary zero point and an arbitrary unit of measurement (see DeMars, 2010, for an introduction to IRT). For example, the scale may be set to have a mean value for an analysis sample of 0.0 and a standard deviation of 1.0. In this case, the value reported for a person might be 1.13, indicating that the person is well above average, and the mean for a group might be −0.088, indicating that the group mean is slightly below the analysis sample mean. Further, the initial values from the data analysis are often transformed to whole numbers to avoid negative values and decimal places. This is done using a linear transformation such as $200x + 500$ and rounding to the nearest whole number. The score for the person would be converted to 726 ($200 \times 1.13 + 500$) on this scale, and the mean for the group mean to 482 ($200 \times -0.88 + 500$). On this reporting scale, values are compared to 500 to determine whether they are above or below average. But, for any of these numerical values, it is difficult to interpret what a person knows or can do. At best, the numbers can be used to order persons or groups according to the magnitude of what they know or can do, but not whether they have particular capabilities such as being able to solve systems of linear equations.

The TEDS-M international comparison study (Tatto et al., 2012) assessed the preparation of individuals to teach mathematics using the same kind of IRT-based, abstract scale as other large-scale assessment programs. To help with the interpretation of the results of this study, work was done to give substantive meaning to the scale so that, when results were reported, those interested in the results could have a reasonable sense of what the numerical values meant in practice. Meaning was added to the reporting scales by attaching detailed substantive descriptors to selected points on the scale called "anchor points." The logic and methodology used to add descriptions to anchor points is similar in many ways to methodologies used elsewhere, such as item mapping or scale anchoring (Beaton & Allen, 1992; Zwick, Senturk, Wang, & Loomis, 2001). However, there are some important differences between the procedure used for TEDS-M and those used with other testing programs such as NAEP, TIMSS, and PISA. The purpose of this chapter is to describe how the descriptive information was added to the scales using the TEDS-M anchor points, and to discuss how the interpretations derived from these descriptions differ from those used with other assessment programs.

Conceptual Framework

A very useful property of IRT models is that once the items used in a testing program are calibrated (i.e., the item parameters are estimated), the probability that a person at a specified point on the scale defined by the model obtains a particular score on a particular item can be calculated. For items that are scored as correct (1 point) or incorrect (0 points), the location of a person on the scale and the parameter estimates for the item can be used to determine the probability of a correct response or an incorrect response to the item. If the item has more than two score categories, the probability of each score category can be computed. These computed probabilities are considered to be accurate if the IRT model has good fit to the item response data for the testing program and the sample size for calibration is large enough to yield small errors of estimation for the item parameters.

Once a point on the scale for the IRT model is selected, it is possible to compute the probabilities of a correct response for persons at the selected point for each of the calibrated items. These probabilities can then be used to identify the items that a person at that point on the scale would likely answer correctly, indicating that they know the concepts behind the item. The items that the person would most likely answer incorrectly can also be identified. For the items with likely incorrect responses, the person at the specified point on the scale has some deficiencies related to the skills and knowledge needed to respond correctly.

The goal of attaching descriptions to anchor points on the scales is to provide verbal descriptions of the skills and knowledge that are needed to answer those items that are correct while being careful not to include skills and knowledge that are shown to be lacking by the items that are answered incorrectly. Put differently, the descriptions are an attempt to provide substantive detail about what a person at the point on the IRT scale most likely knows and can do, and what they most likely do not know and cannot do.

To produce meaningful descriptions related to a point on the IRT scale, there must be sufficient test items in the likely correct (know) and the likely incorrect (do not know) categories to avoid the instabilities and inaccuracies that result from basing a description on limited information. As the descriptions of skills and knowledge from the test items are developed, the persons writing the descriptions need to constantly test the credibility of the description against the full set of items that fall into the *know* and *do not know* sets.

Implementation of the Anchor Point Procedure

The Anchor Point Procedure was implemented for the TEDS-M study for tests of prospective teachers' mathematical content knowledge (MCK) and mathematical pedagogical content knowledge (MPCK) for teaching at the primary level and secondary level. Thus, four sets of anchor points and descriptions were produced. The test items for the different domains and levels were administered to prospective teachers using an incomplete block design, with each prospective teacher taking two blocks of items. The overlapping sets of items in the administration design were calibrated together using the Rasch model (see DeMars, 2010, for a description of the model), so all the items in a domain were on the same scale. Table 18.1 lists the number of items that were calibrated for each domain and each level.

When developing the descriptions of anchor points on the score scale used for reporting, it was important to select probabilities that give operational definitions for what a person at a certain point on the scale *can do* and *cannot do*. The literature on scale anchoring uses a wide range of probabilities to define "can do," from .5 to as high as .9. Generally, there is no discussion in the literature about a probability for "cannot do," so there was no guidance on the value to be selected for that purpose. One consideration in selecting the values to operationally define these terms is to be sure there are enough items that meet these probability requirements to be used to develop stable descriptions of the capabilities of a person at an anchor point.

After a careful review of the items related to each of the scales, it was determined that using .70 or more as the definition of "can do" would work well. This is also the value that is often used for standard setting with the bookmark procedure (Mitzel, Lewis, Patz, & Green, 2001). The value selected for the "cannot do" probability was .5 or less. This value gave enough items for creating the description and it also took into account guessing probabilities for the multiple-choice items.

Table 18.1 Number of items for each level and each content domain

	Content domain	
Level	MCK	MPCK
Primary	74	32
Secondary	76	27

MCK = Mathematics Content Knowledge, *MPCK* = Mathematics Pedagogical Content Knowledge

After selecting the operational definitions for "can do" and "cannot do," the next step was to select the points along the reporting score scale to use for developing anchor point descriptions. To do that, points along the Rasch scale at 0.1 intervals from −3 to 3 were used to determine the probability of correct response for each item. For primary MCK, this resulted in a 61 × 74 matrix of probabilities with each row representing a scale point and each column representing an item. The columns were ordered from the easiest item to the hardest item. From this matrix, the number of items that met the "can do" and "cannot do" criteria for each row could be determined. From a review of this matrix for the MCK items, it was determined that the full set of items would have sufficient "can do" and "cannot do" items to support accurate descriptions at two anchor points. These points were at −0.8 and 0.2 on the Rasch scale for the primary level and−0.5 and 0.1 for the Rasch scale at the secondary level. Because there are fewer MPCK items, they full set of items would only support the development of one anchor point for each scale. The anchor point was at 0.4 for the Rasch scale for the primary level and 0.0 for the Rasch scale for the secondary level.

Note that the anchor points are not based on some predetermined amounts of skills and knowledge, but on the criterion of having sufficient items for a well-supported description. Therefore, they should not be interpreted as standards of performance but rather as descriptions developed to give substantive meaning to points on the scale. Figure 18.1 summarizes these initial steps in the procedure for selecting anchor points and developing the descriptions.

For the MPCK items at both the primary and secondary levels, the sets of items that matched the "can do" requirements and the "cannot do" requirements were determined. Each of the "can do" and "cannot do" sets were further divided into three subsets for the development of the anchor point description. The three subsets of items were (a) example items for use in describing the task to persons working on the descriptions, (b) a definition set for developing the initial descriptions of the anchor points, and (c) a validity set for use in checking and refining the descriptions. Each of the sets of items contained both "can do" and "cannot do" items. The items were collected into separate booklets for use in the anchor point description workshop.

For the MCK items, the process was more complex because items were needed for each of the two anchor points. First the "can do" items were identified for the upper anchor point and the "cannot do" items were identified for the lower anchor

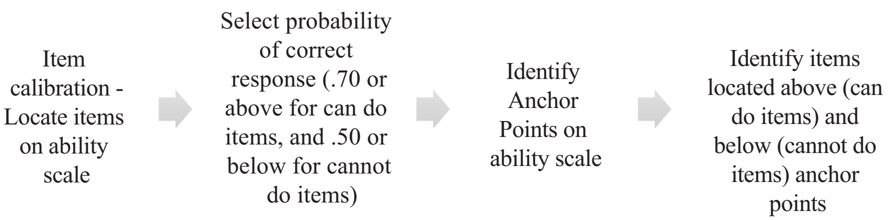


Fig. 18.1 Summary of the initial steps in the anchor point procedure

point. These items were then set aside to use for the descriptions. Then, from the items that were left, the “can do” items were identified for the lower anchor point, and the “cannot do” items were identified for the upper anchor point. This process gave non-overlapping sets of items for developing the anchor point descriptions. As with the MPCK items, the sets for each anchor point were divided into three subsets to use as (a) examples, (b) definition sets for initial descriptions of the anchor points, and (c) validation sets for refinement of the descriptions.

Once the anchor points and items were identified, the task shifted to developing the descriptions of what a person at each anchor point could and could not do. The descriptions were developed by mathematicians and mathematics educators from a number of universities and organizations from across the United States. Ten persons were recruited to work on the primary level anchor point descriptions and 11 were recruited to work on the secondary level anchor points. For each level, the participants were divided into small groups of three or four who initially worked independently, but then shared their results at a later point in the process. The primary and secondary workshops each took two days.

At the beginning of the workshop sessions, the participants were told to review the sets of “can do” and “cannot do” items that they were given and to develop a description of a person who responded correctly to the “can do” items and incorrectly to the “cannot do” items. They were not to describe the items, but rather they were to describe the skills and knowledge of a person who responded to the items in the way indicated. They were to think about the person’s deficiencies in skills and knowledge that would cause them to provide incorrect responses to the “cannot do” items.

As the groups worked on developing the descriptions, they were also told that they would later be given new sets of items and would be asked to use their descriptions to predict which items persons at the anchor point would answer correctly or incorrectly. Their descriptions would need to be sufficiently detailed to support making the predictions.

After each small group completed the first draft of their descriptions, they shared their work with the other small groups to determine similarities and differences. They were encouraged to revise their descriptions after the discussions to better reflect the skills needed to respond to the test items.

Once the groups were satisfied with the descriptions developed with the first set of items, they were given a second set of test items called a “validity” set. They were not told which items were “can do” or “cannot do” items. Their initial task was to predict based on the descriptions that they had written the items that belonged in each category. After they made their predictions, they were informed of the actual classification of the items and were asked how their descriptions could be revised so that they would improve their rate of correct classification. After the groups had made revisions based on the information from the validity item sets, there was a full-group discussion of the resulting descriptions, with the goal of merging the small-group results into a single description. Figure 18.2 summarizes the final steps in the procedure.

Subsets of items selected:
1. Training set
2. Definition set
3. Validity set

→

Anchor point description workshop (after training, the definition set is used to develop anchor point descriptions)

→

Validation process (using descriptions to predict performance on the validity set and refine descriptions)

Fig. 18.2 Summary of the final steps in the anchor point procedure

Table 18.2 Anchor points on the MCK and MPCK scales

	MCK	MPCK
Primary	431	544
	516	
Secondary	490	509
	559	

Results

The results of the anchor point process were (a) the selection of two anchor points for each of the MCK score scales and one anchor point for each of the MPCK scales; (b) the selection of "can do" and "cannot do" sets of items for each anchor point; and (c) the development of verbal descriptions of the skills and knowledge possessed by persons who were at the anchor points on each scale. The numerical values for the anchor points for Primary and Secondary and the MCK and MPCK tests are presented in Table 18.2. These values are on the reporting scales for the TEDS-M study, which is set to have a mean of 500 and standard deviation of 100 for the full set of prospective teachers in the study.

The anchor points are more related to the characteristics of the items on the test than the capabilities of the prospective teachers because they were selected such that there would be sufficient items to support the development of descriptions. For the MCK scales, the anchor points are on either side of the overall mean of 500. This means that the center of the score distribution, with the highest density of prospective teachers, is in the region between the two anchor points. For the MPCK scales, the single anchor point that was selected is slightly above the mean for the sample of prospective teachers.

These anchor points are not standards of performance. There was no attempt to determine whether they define what is good enough for new teachers. They are intended to give information about the meaning of points on the scales so there would be content meaning associated with the abstract numerical values on the scale.

Verbal descriptions were developed for each of the anchor points—six descriptions in all were produced—two for each MCK scale and one for each MPCK scale. The full descriptions along with short excerpts of the items that were used to gener-

The following problems appear in a mathematics textbook for <lower secondary school>.

1. [Peter], [David], and [James] play a game with marbles. They have 198 marbles altogether. [Peter] has 6 times as many marbles as [David], and [James] has 2 times as many marbles as [David]. How many marbles does each boy have?

2. Three children [Wendy], [Joyce] and [Gabriela] have 198 zeds altogether. [Wendy] has 6 times as much money as [Joyce], and 3 times as much as [Gabriela]. How many zeds does each child have?

Solve each problem.

Fig. 18.3 Algebra Items MCF604A1 and MCF604A2

ate the descriptions are provided in the Appendix to this chapter. These descriptions refer to performance of persons who are located at specific points on the MCK and MPCK scales. They do not describe performance over the range of the scale between two points or above the higher anchor point. Prospective teachers who scored higher than an anchor point would have greater competence than the description indicates. Those scoring below a point would have more deficiencies. The descriptions associated with the anchor points on the scales allow richer interpretations of the scores than the numerical values alone.

An example of the kinds of information that the participants in the workshops would use to develop the descriptions is the following item pair (see Fig. 18.3). Prospective teachers could answer part 1 of this item with greater than a .70 probability, but they had less than a .50 probability of responding correctly to part 2. The anchor point descriptions that resulted from this were that prospective teachers were likely to “Solve word problems involving ratios of whole numbers (See Released Item MFC604A1) or sums of consecutive integers”, but were not likely to “Solve a word problem with a more complex linguistic or logical structure or one in which the choice of variable is not obvious. (See Released Item MFC604A2)”. Information from many pairs of items were used to develop the descriptions in the Appendix.

Discussion

The descriptions provided in the Appendix give much more detail than is typically seen in descriptions attached points on IRT-score scales. The reason for the difference is the need to include nuances in the descriptions to account for the “cannot do” aspect of performance as well as the “can do” aspect. Further, typical descriptions related to score scales use simple declarative sentences without any qualifying words. For example, the statement might say “can solve problems with fractions.”

This type of statement is unqualified, suggesting that a person at a point on the scale can solve all possible problems with fractions. In contrast to the typical statement, the description for Anchor Point 1 at the primary level is "able to solve some problems with fractions." This subtle difference suggests that the more difficult problems with fractions do not have a high probability of being solved by a person at this point. In other words, there are limits to that person's capabilities with fractions. Similar limitations are frequently included in the anchor point descriptions. For example, people at the primary Anchor Point 2 can find areas and perimeters of *simple figures* instead of all figures. These nuances were included because of the need to make predictions of performance on the validity sets of items. When the statements are left without qualifiers, they imply a higher level of performance than is actually observed.

For the primary and secondary MCK item sets, it was possible to give descriptions to two anchor points that were 0.6 to 1.0 unit apart, respectively, on the Rasch scale. The result is that detailed information is provided about persons whose scores fall between the two anchor points: They know more than people at the lower point description do, and less than those at the upper point description. For example, at .6 on the Primary MCK scale, teachers can solve some problems with fractions, but at 1.0 on that scale they are competent using fractions. As they acquire the skills and knowledge needed to move the .4 points on that scale, they are refining their capabilities to use fractions to solve problems.

It would be helpful to have more anchor points spaced along the IRT scale, but describing the capabilities at more points would require more test items that are well spread in difficulty. Further, those items would need to be calibrated using the IRT model, requiring more data acquired through either more testing time or more examinees. For the TEDS-M project, obtaining more data was not possible because of practical constraints present in the countries at the time of the study. Testing time could not be extended beyond one hour, and acquiring larger samples was either impossible or very expensive.

Deciding to select anchor points and develop the descriptions came fairly late in the TEDS-M study. This was because of the innovative nature of the full project and the complexities of working with multiple countries. Until the full set of data was collected, it was not possible to determine how many anchor points could be supported. Future studies can take advantage of what was learned during the TEDS-M study to plan for the use of anchor points and descriptions as part of the reporting process. Early planning would allow for more detailed reporting based on more anchor points and refined descriptions.

The anchor point description approach described here helps to give enhanced meaning to the score scale, so that the particular skills and knowledge acquired by a prospective teacher can be described. The approach is more nuanced than previous scale anchoring procedures because it considers both what a person at a point on the scale is likely to be able to do and what they are unlikely to be able to do. This should improve the interpretation of results of studies like TEDS-M.

Appendix[1]

Anchor Point Descriptions for the Mathematics Content Knowledge Assessment of Future Primary Teachers

Anchor Point 1 Future teachers of primary school mathematics at Anchor Point 1 are successful at performing basic computations with whole numbers, understand properties of operations with whole numbers, and are able to reason about related concepts such as odd or even numbers. They are able to solve some problems with fractions. Future teachers at this Anchor point are successful at visualizing and interpreting 2-dimensional and 3-dimensional geometric figures, and can solve simple problems about perimeter. They can also understand straightforward uses of variables and the concept of equivalence, and can solve problems involving simple expressions and equations.

Future teachers at Anchor Point 1 are able to apply whole number arithmetic in simple problem-solving situations, however they tend to over-generalize and have difficulty solving abstract problems and those requiring multiple steps. They have limited understanding of the concept of least common multiple, the number line, and the density of the real numbers. Their knowledge of proportionality and multiplicative reasoning is weak. They have difficulty solving problems that involve coordinates and problems about relations between geometric figures. Future teachers at this Anchor point can make simple deductions, but they have difficulty reasoning about multiple statements and relationships among several mathematical concepts.

Anchor Point 2 Future teachers at Anchor Point 2 are successful at the mathematical tasks at Anchor Point 1. In addition, future teachers at Anchor Point 2 are more successful than future teachers at Anchor Point 1 at using fractions to solve story problems, and recognize examples of rational and irrational numbers. They know how to find the least common multiple of two numbers in a familiar context, and can recognize that some arguments about whole numbers are logically weak. They are able to determine areas and perimeters of simple figures, and have some notion of class inclusion among polygons. Future teachers at Anchor Point 2 also have some familiarity with linear expressions and functions.

However, while future primary teachers at Anchor Point 2 can solve some problems involving proportional reasoning, they have trouble reasoning about factors, multiples, and percentages. They are unable to solve problems about area of obtuse-angled triangles involving coordinate geometry. They do not recognize applications of quadratic or exponential functions, and have limited skills in algebraic reasoning.

[1]The anchor point descriptions were taken from Appendix Q of the TEDS-M Technical Report, Tatto (2013).

Overall, future teachers at Anchor Point 2 do well on items testing "knowing," and on standard problems about numbers, geometry, and algebra classified as "applying," but they are not able to answer problems that require more complex reasoning in applied or non-routine situations.

Some specific examples of items at Anchor Point 1 and Anchor Point 2 follow.

Anchor Point 1 Following are examples of items on which future primary teachers at Anchor Point 1 answered successfully at least 70% of the time.

- Reason about fractions to interpret simple numerical statements relating to a word problem.
- Identify the least likely outcome for a simple random experiment involving fractions with different denominators.
- Determine whether subtraction and division are commutative and addition is associative. (See released items MFC202A, B & C)
- Interpret a diagram of a pan balance to determine the mass of an unknown quantity. (See released item MFC303) Determine whether the results of particular operations with even or odd numbers are odd or even. Recognize a net for a triangular based prism. (See released item MFC501)
- Interpret a bar chart and some verbal clues to solve a problem about the number of items sold. (See released item MFC502A)
- Identify common rational numbers. (See released item MFC503B)

The following are examples of items on which future teachers at Anchor Point 1 answered successfully less than 50% of the time.

- Determine whether subtraction of whole numbers is associative. (See released item MFC202D.)
- Identify the correct Venn diagram to illustrate the relation between four types of quadrilateral. (See released item MFC204)
- Understand that there are an infinite number of decimal numbers between two given numbers. (See released item MFC304)
- Find a linear algebraic rule to describe a general situation illustrated by a diagram. (see released item MFC308)
- Find the area of a triangle drawn on a grid. (See released item MFC408)
- Identify an algebraic representation of a numerical relationship between three consecutive even numbers.

Anchor Point 2 Following are examples of items on which future teachers at Anchor Point 2 perform successfully at least 70% of the time.

- Identify the truth of a statement about the solvability of a word problem involving proportional reasoning.
- Determine whether subtraction of whole numbers is associative. (See released item MFC202D.)
- Determine the area of a walkway around a rectangular pool. (See released item MFC203)

- Interpret Venn diagrams representing relationships between quadrilaterals. (See released item 204)
- Identify the solution to a word problem involving a rate and requiring some proportional reasoning. (See released item 206A)
- Recognize whether some story problems correctly model the subtraction of two fractions.
- Identify the difference between the perimeter and area of a rectangle drawn on dot paper.
- Indicate whether π and $\sqrt{49}$ are rational or irrational. (See released items MFC503A & C.)
- Identify a future term in a linear rule represented visually. (See released item MFC508)

Following are examples of items on which future teachers at Anchor Point 2 answered successfully less than 50% of the time.

- Use proportional reasoning to interpret numerical statements involving percentage relating to a word problem.
- Identify the true probability statement relating to a game involving two dice. (See released item MFC106)
- Write a correct statement about the reflection image of the point with coordinates (a, b) over the x-axis.
- Identify a set of geometric statements that uniquely define a square.
- Describe properties of the function defined by the ratio of the area and circumference of a circle
- Identify whether $-\frac{3}{2}$ is rational or irrational. (See released item MFC503D)
- Determine the conditions for which one linear algebraic expression is greater than or equal to another. (See released item MFC509)
- Compare lengths on a cube and a cylinder with common dimensions. (See released item MFC 513)

Anchor Point Descriptions for the Mathematics Pedagogical Content Knowledge Assessment of Future Primary Teachers

Anchor Point Future primary teachers at this Anchor Point are able to recognize the correctness of a teaching strategy for a particular concrete example, and are able to evaluate students' work when the mathematics content is conventional or typical of primary grades. They are able to identify the arithmetic elements of single-step story problems that influence their difficulty. (See released item MFC505).

While future primary teachers at the primary MPCK Anchor point have some ability to interpret student solution methods, identify the skills inherent in a task and identify student difficulties, they may not be able to articulate them as clearly and concisely as more able future teachers. (See released item MFC502B). Similarly, future teachers at this Anchor Point can partially identify and compare the attributes of the graphical representations of young children but not as well as their more able counterparts. (See released item MFC410).[2]

However future teachers at this Anchor Point may not know how to use concrete representations to support students' learning (See released item MFC312), and may not recognize how a student's thinking is related to a particular algebraic representation (See released item MFC108). They may not sufficiently understand some measurement or probability concepts in order to reword or design a task. (See released item MFC307B).

Future teacher at this Anchor Point may not know why a particular teaching strategy would make sense (See released item MFC513), whether a strategy can be generalized to a larger class of problems, or if it will always work. They may be unaware of common misconceptions and unable to conceive useful representations of numerical concepts. (See released items MFC208A & B)

Anchor Point Descriptions for the Mathematics Content Knowledge Assessment of Future Secondary Teachers

Anchor Point 1 Future teachers of lower secondary school mathematics who perform at Anchor Point 1 know concepts related to whole numbers, integers, and rational numbers, and can compute with them. They also can evaluate algebraic expressions and solve simple linear and quadratic equations, particularly those that are solvable by substitution or trial and error. They are familiar with standard geometric figures in the plane and space, and can identify and apply simple relations in plane geometry. They are also able to interpret and solve more complex problems in number, algebra, and geometry if the context or the problem type is a commonly taught topic in lower secondary schools.

However, future teachers at anchor point 1 have difficulty describing general patterns, solving multi-step problems if they have complex linguistic or mathematical relations, and relating equivalent representations of concepts. They tend to over-generalize concepts, and do not have a good grasp of mathematical reasoning. In particular, they do not consistently recognize faulty arguments or are able to justify or prove conclusions.

Anchor Point 2 Future teachers who perform at Anchor Point 2 perform successfully at all the mathematics problems in Anchor Point 1. In addition, they seem to have a more robust notion of function, especially of linear, quadratic, and exponential functions, are better able to read analyze and apply abstract definitions and

[2] MFC410 and MFC502B are examples of where future teachers at the Anchor point have been awarded partial credit for their responses thereby indicating some proficiency.

notation, and have greater ability to make and recognize simple arguments, than a future lower secondary teacher at Anchor Point 1. They also know some definitions and theorems from university level courses such as calculus, abstract algebra, and college geometry, and can apply them in straightforward situations.

However, future teachers at Anchor Point 2 usually are not consistently successful in solving problems stated in purely abstract terms, or with problems containing foundational material such as axiomatic systems in geometry. Additionally, they make errors in logical reasoning, such as not attending to all conditions of definitions or theorems and confusing the truth of a statement with the validity of an argument, and are unable to recognize valid proofs of more complex statements. Even though they may be able to make some progress in constructing a mathematical proof, future teachers performing at anchor point 2 are not generally successful at completing mathematical proofs.

Some specific examples of items that future lower secondary teachers were successful at solving at Anchor Points 1 and 2 follow.

Anchor Point 1 Following are examples of items on which future lower secondary teachers at this Anchor Point **are successful at least 70% of the time**.

- Solve a simple linear or quadratic equation and identify the smallest set of numbers to which the solution belongs.
- Solve word problems involving ratios of whole numbers (See Released Item MFC604A1) or sums of consecutive integers.
- Determine if angles in a triangle are congruent using given information.
- Determine the number of lines of symmetry in a regular polygon (See Released Items MFC808A1, A2, B1, and B2)
- Determine whether a given translation or reflection maps one figure to another.

The following are examples of items on which future teachers at Anchor Point 1 **are successful less than 50% of the time**.

- Solve a word problem with a more complex linguistic or logical structure or one in which the choice of variable is not obvious. (See Released Item MFC604A2)
- Generalize patterns involving linear and non-linear growth.
- Determine whether a given composite of transformations maps one figure to another.
- Solve equations in one variable and describe the solution set in the coordinate plane or space. (See Released Item MFC705A & B)
- Write a proof of a statement about the sum of two functions. (See released item MFC711)
- Identify an appropriate definition for a function that is continuous at a point.
- Identify consequences of replacing a particular axiom in geometry.

Anchor Point 2 Following are examples of items on which future teachers at Anchor Point 2 **are successful at least 70% of the time**.

- Solve problems about properties of angles or triangles.
- Determine if the relation "is similar to" satisfies the reflexive, symmetric, and transitive properties.
- Identify a situation that is modeled by an exponential function. (See released Item MFC710 A, B & C)
- Identify consequences of replacing a particular axiom in geometry.
- Make some progress toward solving a problem about conditional probability.
- Write part of a proof related to the sum of two functions. (See released item MFC711)
- Recognize that a particular algebraic argument about the divisibility of a square of any natural number is a valid proof. (See released item MFC802B)

Following are examples of items on which future teachers at Anchor Point 2 are successful **less than 50% of the time**.

- Determine properties of absolute value.
- Find solutions to equations in the set of complex numbers or integers modulo 6.
- Interpret standard deviation when distributions are presented visually.
- Determine whether statements about abstract concepts are equivalent.
- Work with foundational materials such as axiomatic systems in geometry.
- Write a complete proof about the sum of two functions. (See Released Item MFC711)
- Solve problems about combinations. (See Released Item MFC804)

Anchor Point Descriptions for the Mathematics Pedagogical Content Knowledge Assessment of Future Secondary Teachers

Anchor Point Future teachers who are at the Anchor Point on the Mathematics Pedagogical Content Knowledge scale have a variable range of knowledge of the lower secondary curriculum and of planning for instruction. For instance, they know prerequisite knowledge and steps for teaching a derivation of the quadratic formula (See Released Item MFC 712A, B, C & D) and can determine consequences of moving the concept of square root from the lower secondary to the upper secondary school mathematics curriculum. However, they have difficulty deciding what would be a helpful mathematics concept to use in a proof about isosceles triangles.

They also have some skill in enacting school mathematics. Future teachers at this Anchor Point can sometimes correctly evaluate students' mathematical work. For example, they can determine whether a student's diagram satisfies certain given conditions in geometry, and they can recognize a student's correct prose argument about divisibility of whole numbers. (See Released Item MFC709A).

However, they cannot identify the correct solution to a trigonometry problem, and cannot consistently apply a rubric with descriptions of three performance levels to evaluate students' solutions to a problem about linear and non-linear growth.

Future teachers at this Anchor Point are successful at analyzing students' errors when the students' work involves single step or short explanations, but they are less successful at identifying or analyzing errors in more complex mathematical situations. For instance, future teachers at this level can identify an error in misreading a histogram (See Released Item MFC806B), but cannot explain why one word problem is more difficult for students than another (See Released Item MFC604B).

In general, future teachers' own depth of mathematical understanding seems to influence their ability to interpret students' thinking or to determine appropriate responses to students. Because future teachers at this level lack a well-developed concept of the meaning of a valid mathematical argument, they have difficulty evaluating some invalid arguments. In particular, they do not recognize that examples are not sufficient to constitute a proof (See Released Item MFC709B). They also are not able to recognize whether certain word problems correctly exemplify expressions involving the division of fractions.

References

Beaton, A. E., & Allen, N. L. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics, 17*(2), 191–204.

DeMars, C. (2010). *Item response theory*. New York, NY: Oxford University Press.

Mitzel, H. C., Lewis, D. M., Patz, R. J., & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Ed.), *Standard setting: Concepts, methods, and perspectives* (pp. 249–281). Mahwah, NJ: Erlbaum.

Tatto, M. T. (Ed.). (2013). *The Teacher Education and Development Study in Mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries – Technical report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Tatto, M. T., Schwille, J., Senk, S. L., Ingvarson, L., Rowley, G., Peck, R., … Reckase, M. (2012). *Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries: Findings from the IEA Teacher Education and Development Study in Mathematics (TEDS-M)*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Watermann, R., & Klieme, E. (2002). Reporting results of large-scale assessment in psychologically and educationally meaningful terms: Construct validation and proficiency scaling in TIMSS. *European Journal of Psychological Assessment, 18*(3), 190–203.

Zwick, R., Senturk, D., Wang, J., & Loomis, S. C. (2001). An investigation of alternative methods for item mapping in the National Assessment of Educational Progress. *Educational Measurement: Issues and Practice, 20*(2), 15–25.

Chapter 19
Examining Sources of Gender DIF in Mathematics Knowledge of Future Teachers Using Cross-Classified IRT Models

Liuhan Sophie Cai and Anthony D. Albano

Abstract Research on differential item functioning (DIF) has focused traditionally on the detection of effects. However, recent studies have investigated potential sources of DIF, in an attempt to determine how or why it may occur. This study examines variability in item difficulty in math performance that is accounted for by gender, referred to as gender DIF, and the extent to which gender DIF is explained by both person predictors (opportunity to learn [OTL]) and item characteristics (item format). Cross-classified multilevel IRT models are used to examine the relationships among item difficulty, gender, OTL, and item format. Data come from the U.S. cohort of an international study of future math teachers, the Teacher Education and Development Study in Mathematics.

Introduction

Gender differences in math performance have been studied widely on many different populations. Results of large-scale assessments such as the Programme for the International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS) indicate consistently higher average math scores for male students across countries (Else-Quest, Hyde, & Linn, 2010). The Teacher Education and Development Study in Mathematics (TEDS-M) is an international study of teacher preparation programs and candidates that utilized different instruments than did PISA or TIMSS. TEDS-M tested the math

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

L. S. Cai (✉) · A. D. Albano
University of Nebraska-Lincoln, Lincoln, NE, USA
e-mail: albano@unl.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_19

proficiency of future primary and lower-secondary school teachers. We were interested to see whether the same gendered results from PISA and TIMSS would also manifest in TEDS-M. Gender difference in math performance is of interest to many countries. An examination of measurement invariance of math assessments is necessary before any comparison of the test-level performance is carried out. Policy reforms and teaching practices may be informed by results from these assessments. Score comparability and measurement invariance over person subgroups is essential (Meredith, 1993). Measurement invariance is a statistical property that is achieved for an assessment when item parameters do not vary meaningfully across person groups and person parameters do not vary meaningfully across time points or measurement conditions (Rupp & Zumbo, 2006). Differential item functioning (DIF), a special case of measurement invariance, is the result of lack of item parameter invariance (Hambleton, Swaminathan, & Rogers, 1991). Significant amounts of DIF in an assessment can invalidate score interpretations and compromise test fairness.

DIF results when variables other than the construct of interest have an influence on performance (Ackerman, 1992). An item is said to be free of DIF if all individuals with the same underlying ability or construct have equal probability of answering the item correctly, regardless of group membership (Hambleton et al., 1991). DIF detection involves the identification of items impacted by these extraneous variables. Numerous DIF analysis and detection techniques have been developed to examine invariance in both person and item parameters, including ones based on contingency tables (e.g., Holland & Thayer, 1988), regression models (e.g., Swaminathan & Rogers, 1990), item response theory models (IRT, e.g., Thissen, Steinberg, & Wainer, 1993), multidimensional models (e.g., Roussos & Stout, 1996), structural equation models (e.g., Muthén, Kao, & Burstein, 1991), and multilevel models (e.g., Cheong, 2006; Kamata, 2001).

Traditionally, research on DIF has focused on the detection of biased assessment items (Kim, Cohen, Alagoz, & Kim, 2007). However, studies recently have investigated how and why DIF may occur. Improvements in statistical modeling techniques have made it possible to explore additional covariates as potential sources of DIF by measuring the extent to which these covariates account for variability in item difficulty parameters.

The present study uses data from TEDS-M to examine the overall variability of item difficulty parameters in math performance that is accounted for by gender. This variability in item difficulty parameters by gender is referred to as lack of measurement invariance by gender or gender DIF. This study then investigates the extent to which gender DIF can be explained by both person predictors (i.e., opportunity to learn [OTL]) and item characteristics (i.e., item format). A cross-classified multilevel IRT model framework is used to examine the relationships among item difficulty, gender, OTL, and item format.

The following section first reviews previous work on the implementation of a multilevel item response model for testing item difficulty parameter invariance and how the model can be extended to explore sources of DIF. Previous research on gender performance in math at both the test level and the item level is reviewed. The section ends with a summary of research on OTL, and the research questions that guided this study.

Literature Review

Modeling Parameter Variance

A variety of methods have been developed to investigate DIF. These methods typically involve multiple statistical tests for individual items or each pair of focal and reference groups. A limitation of using multiple statistical tests is the expected increase in false positives (Longford, Holland, & Thayer, 1993). Compared to the traditional DIF detection procedures, the logistic mixed model is more economical as it can detect DIF within an omnibus test, rather than by targeting individual items. DIF can be interpreted via significant interactions between item difficulty (at the item level) and group membership (at the person level). Sources of DIF can be explained by modeling item or person covariates through exploratory mixture model analysis (Cohen & Bolt, 2005; Van den Noortgate & De Boeck, 2005).

Previous research has demonstrated the formulation of traditional IRT models as multilevel logistic models (e.g., Adams, Wilson, & Wu, 1997; Kamata, 2001). In the basic one-parameter IRT or Rasch model (Rasch, 1960), the log-odds of correct response to item i for person j are modeled as

$$Logit\,P\left(Y_{ij}=1\right)=\ln\left(\frac{P}{1-P}\right)=\eta_{ij}=\theta_j-b_i. \quad (19.1)$$

Here, Y_{ij} represents the scored response of person j to item i (1 = correct, 0 = incorrect). Item responses are modeled as a logistic function of the difference between person ability θ_j and item difficulty b_i. This model can also be reformulated as a cross-classified multilevel model with random person and random item effects (Van den Noortgate, De Boeck, & Meulders, 2003):

$$\eta_{ij}=\beta_{0ij} \quad (19.2)$$

$$\beta_{0ij}=\gamma_0+u_{0i}+u_{0j} \quad (19.3)$$

with $u_{0i}\sim N(0,\sigma^2_{u_i}), u_{0j}\sim N(0,\sigma^2_{u_j})$.

In this two-level model, item responses are the level-one unit and persons and items are both level-two units. Thus, item responses are considered to be nested within persons and items, which are assumed to be random samples from populations of items and persons. The log-odds of correct response are modeled as a summation of the random item and person parameters u_{0i} and u_{0j} with γ_0 representing the estimated log-odds of correct response for a person of average ability on an item of average difficulty.

The baseline model in Eq. 19.2 is only a descriptive model. Item and person covariates can be incorporated into the model to examine DIF and its potential sources (Van den Noortgate & De Boeck, 2005). DIF can first be tested by allowing group effects to vary over items at level two (see examples below). When group main effect and item-by-group interaction effects are included in the model, the random effects of group over items represent the residual DIF. The model can be further extended by adding item predictors or person characteristics to explain the DIF. If the item or person covariates explain the DIF effects, one would expect that the group main effect on the additional item or person covariates would differ from zero and the variance of the random group effects over items would decrease.

Gender

Gender effects in math achievement have been studied for decades. Researchers are especially interested in understanding and mitigating the underrepresentation of women in science, technology, engineering, and mathematics (STEM) disciplines (Hyde, Lindberg, Linn, Ellis, & Williams, 2008). Men are reported to achieve higher math scores than women in national and international large-scale assessments (e.g., Baker & Jones, 1993; Beller & Gafni, 1996; Gallagher & Kaufman, 2005; Gamer & Engelhard Jr., 1999). Meanwhile, research has consistently reported math and reading achievement parity between genders in early grades, with increasing male advantages in math and female advantages in reading achievement as students move up through the grades (e.g., Willingham & Cole, 1997).

A considerable amount of research has documented the gender gap in mathematics performance at the overall test level. Hyde, Fennema, and Lamon (1990) conducted a meta-analysis on gender effects on math performance. A weighted mean effect size of 0.15 was found from over 100 studies. This small effect size indicated that, overall, males outperformed females by a small but not negligible amount. The study also reported that starting from high school to college, the gender discrepancy favoring male students emerged in the area of complex problem solving and geometry, but no gender differences existed in arithmetic or algebra performance.

A more recent meta-analysis reported similar findings. Lindberg, Hyde, Petersen, and Linn (2010) examined 242 studies of gender math performance from 1990 to 2007 and found small gender variations in mean math achievement. However, the gender gap was not found to decline from 1990 to 2007. Performance differences favoring males peaked during high school, with an effect size of 0.23, and declined among college students. Furthermore, no gender variations in performance on different math content domains or depth of knowledge were found. In the same paper, Lindberg et al. (2010) conducted another meta-analysis using large national data sets collected after 1990 in the United States. The data sets yielded an average weighted effect size of 0.07, indicating a small male advantage in mean math performance in the United States.

Gender effects at the item level have been researched on different dimensions such as item difficulty, item format, and math content domain. Penner (2003) examined the relationship between gender differences and item difficulty in math items using the 1995 TIMSS data set. The study showed a general pattern of a male advantage on easy math items and an increasing male advantage on more difficult math items. Other studies also have found significant gender-by-item difficulty interactions. Bielinski and Davison (1998) studied the minimum competency math test outcomes and found that easy items tended to be easier for female students than male students, while harder items tended to be harder for female students than male students. The significant negative correlations (−.47 and −.43) between gender differences in item difficulty, and in item difficulty estimated on the overall samples over the two studies, indicated that as item difficulty increased, the male advantage also increased. To extend their previous study, Bielinski and Davison (2001) used three national data sets and reported a similar phenomenon: Math tests with harder items generally favored men, and this gender variability grew in late adolescence.

Additional research has revealed that item format, most often expressed as either multiple choice (MC) or constructed-response (CR), is related to gender DIF; however, findings are inconsistent. Taylor and Lee (2012) analyzed state math tests by using the POLYSIBTEST DIF procedure and a Rasch procedure to explore gender DIF based on item format. Both procedures showed the same direction of DIF effects, where MC items favored male students and CR items favored female students. Other studies have found similar results (e.g., Becker, 1990; DeMars, 1998; Gamer & Engelhard Jr., 1999). In contrast, Liu and Wilson (2009) examined item format and gender effects in math assessments using a multidimensional Rasch model. The results suggested no measurable gender differences on traditional MC items. However, a male advantage was found on CR items, though the effect sizes were small. The largest gender gap was found for complex MC items (an unconventional item format) where male students significantly outperformed female students with an effect size of 0.19.

Mendes-Barnett and Ercikan (2006) used the data of 12th-grade students' math exams to investigate the relationship between gender DIF and math content domain. Using differential bundle functioning analyses, they found that individual geometry items exhibited high DIF, especially those that used visuals. In the content area of computation, items with no equations were found to favor female students, while algebra items favored male students. Becker (1990) found that algebra items were more difficult for women than men, but suggested no significant difference between mean difficulties of women and men for arithmetic and geometric items. Conversely, in examining the math section of the SAT, Harris and Carlton (1993) found that after controlling for mean ability, men performed better on geometry items, while women performed better on algebra items. Among items showing DIF, eight out of 15 items came from geometry and measurement in favor of men, but none were from algebra. On the other hand, nine out of 16 items came from algebra in favor of women. Men also were found to have significant advantages in number and computation, data analysis, and proportional reasoning.

Differential course taking by gender is a potential explanation for male advantages in math performance (Meece, Parsons, Kaczala, & Goff, 1982). Beginning in high school, female students tend to take fewer advanced math and science courses in which students are trained intensively in problem-solving skills. However, in the United States, the gender gap in course enrollment is disappearing gradually. Gender differences in patterns of interest could be a factor that explains course choice variations (Su, Rounds, & Armstrong, 2009). In addition, parents' and teachers' expectation discrepancy in math ability among males and females can play an important role in students' course choices (Jacobs, DavisKean, Bleeker, Eccles, & Malanchuk, 2005; Eccles, 1994). Wiseman (2008) suggested that gender parity was achieved only when there was equity in enrollment, access to resources, and OTL for both males and females. Likewise, Else-Quest et al. (2010) concluded that cross-national variability in differential math performance by gender was associated with country-level disparity in opportunity structures for females. Gender equity in school enrollment, women's share of research jobs, and women's parliamentary representation were found to contribute to variability in gender distinction in math performance.

Opportunity to Learn

The concept of OTL was first introduced by the International Association for the Evaluation of Educational Achievement (IEA) in the 1960s in relation to differential math performance across countries (McDonnell, 1995). Husén (1967) described OTL in the context of testing as the opportunity to study a particular topic or learn how to solve a particular type of problem presented by the test. He argued that the likelihood of answering test items correctly would decrease subsequently if students have not had opportunities to learn the pertinent topics. The concept of OTL has evolved since this early work. OTL has been subsequently defined as the content coverage of knowledge, specifically the topics being taught, the relative emphasis on different aspects of a subject, and students' achievements on the relatively important aspects of the subject (Travers & Westbury, 1989). Differential OTL is reflected in the deviation from the opportunities a student is supposed to have, as established by curriculum at national, state, or district levels, from the educational opportunities a student actually is provided in class (Floden, 2002). The distinctions in OTL definitions suggest that OTL can be measured in various ways.

OTL has been used primarily to make cross-national comparisons. McDonnell (1995) suggested that OTL should be considered to ensure fairness when making performance comparisons. Through the examination of math textbooks and their use in lower-secondary classrooms, Haggarty and Pepin (2002) found that learners from different countries were provided with different types of math knowledge and were offered different levels of OTL in math. Research has also shown that, in international contexts, countries with higher levels of OTL outperform those with lower levels of OTL (e.g., Mullis, Martin, & Foy, 2008).

OTL also has been researched at the individual level. Boscardin et al. (2005) used hierarchical linear modeling to investigate the impacts of various OTL variables on student outcomes in English and algebra. The first level in the model was the student level, where individual students were the unit of analysis. Students were nested within teachers at level two. Findings suggested that teacher expertise in these two content areas was positively correlated with student performance. Moreover, content coverage, as an indicator of OTL, also was found to have a consistently positive relationship with outcomes from the algebra and English assessments. Specifically, with one more week spent on relevant content, there was an expected increase of 0.85 in algebra test scores. One additional week covering English resulted in an increase of 1.59 points on the English test. Additionally, Blömeke, Suhl, Kaiser, and Döhrmann (2012) found that among future primary teachers, OTL in math not only had a strong positive effect on math performance, but also a significant effect on math pedagogical content knowledge, presumably by mediating the effects of OTL in math pedagogy.

Relatively fewer studies have addressed the relationship between OTL and person grouping variables with respect to item performance. Albano and Rodriguez (2013) used hierarchical generalized linear modeling to investigate parameter invariance over covariates at the student level. Item responses were nested within students. Gender and OTL were both examined as potential sources of variability in item difficulty parameters. A two-level model was used, where gender was the person group covariate at level two and OTL was the person covariate at the same level. Future secondary teachers from three TEDS-M countries (Germany, Singapore, and the United States) were examined. For the Singapore cohort, item difficulty did not differ significantly by gender. In Germany, controlling for mean ability, items functioned in favor of men. The inclusion of OTL impact effects and item-by-OTL interaction effects did not reduce the number of items showing gender DIF, though some items did function differentially by OTL; thus, OTL was not found to be a source of DIF in Germany. For the U.S. cohort, the best-fitting model included main effects for items, gender, and OTL, and the two-way interaction effects of item-by-gender and item-by-OTL. Difficulty estimates for eight out of 22 items were found to vary by gender when OTL was not included in the model. These items were initially identified as exhibiting gender DIF. When OTL main effect and item-by-OTL interaction effects were introduced to the model, the mean proportion correct was estimated to increase by 0.15 logits for a one-unit increase in OTL. When these effects were introduced, three out of eight items were no longer found to display gender DIF. These results were taken to indicate that person-level OTL can mediate the relationship between item difficulty and gender. Thus, differential OTL may partly contribute to differential math performance by gender.

Burkes (2009) used multilevel-DIF methodology to examine item performance differences across two socioeconomic status (SES) groups. Results of the study revealed that eight out of 71 items exhibited DIF, all of which favored students with higher SES. However, when item difficulties and DIF effects for SES were modeled at the classroom level as a function of OTL, only one item still exhibited DIF. Because of unequal OTL, the seven items were systematically more difficult for students with lower SES. This study indicated that classroom-level OTL differences were the source of SES-based DIF.

Summary

Previous research has demonstrated the use of multilevel cross-classified models for examining covariates such as item and person characteristics as sources of variation in item difficulty parameters across multiple levels of nested data. Previous research also indicates that gender effects in mathematics are related to variables such as course-taking (e.g., Meece et al., 1982), educational resources (e.g., Else-Quest et al., 2010), and item characteristics (e.g., Taylor & Lee, 2012). The purpose of this study is to use cross-classified models to examine the extent to which item difficulty on a math test varies by gender, item format, and OTL. This is realized by addressing three main research questions: (a) How does item difficulty on a math test differ by gender conditional on overall ability in a sample of future teachers? (b) To what extent does item format explain variability in item difficulty by gender? (c) To what extent does OTL explain variability in item difficulty by gender? These questions were examined using data from TEDS-M an international and comparative study of future mathematics teachers.

Method

Data

Data for this study came from TEDS-M future lower-secondary teachers. The target population was defined as future teachers in the final year of their teacher education programs who would be eligible to teach mathematics in secondary schools (Tatto et al., 2008). Future secondary teachers from 15 countries participated in the TEDS-M study. The analyses in this study were conducted using the U.S. cohort, which contained data from 475 students (69% female, 31% male).

The TEDS-M study measured future teachers' mathematics content knowledge (MCK) and mathematics pedagogical content knowledge (MPCK) as the outcomes at the end of secondary teacher education. The assessment was administered in a standardized and monitored test session with a 60-minute completion time. The present study used scored item responses from the MCK assessment. The MCK assessment contained a total of 76 items. Item formats were MC (multiple-choice and complex multiple-choice) and CR (constructed-response). There were 58 MC and 18 CR items. Each item fell into one of the four domains: number (27 items), geometry (23), algebra (22), and data (four).

In TEDS-M, OTL was measured at both the individual and program levels. This study used individual OTL, which was defined as future teachers' occasion to learn about particular topics during the course of their teacher education. OTL in tertiary math topics was considered to be most relevant to secondary education. Tertiary OTL was based on the future secondary teachers' responses to whether they had the opportunity to learn 19 topics in four key areas: (a) geometry (e.g., axiomatic

Table 19.1 Descriptive statistics by gender

		Prop correct		OTL		
Gender	*N*	*M*	*SD*	*M*	*SD*	*r*
F	325	.56	.14	11.33	4.06	.52
M	149	.66	.13	13.03	3.25	.39

Note: Prop correct is the proportion correct score across the set of items administered to a student. *r* is the correlation between proportion correct and OTL

Table 19.2 Descriptive statistics by gender and format

		Prop correct (MC)		Prop correct (CR)	
Gender	*N*	*M*	*SD*	*M*	*SD*
F	325	.60	.13	.43	.22
M	149	.68	.13	.60	.21

Note: Prop correct is the proportion correct score across items with different formats

geometry or analytic geometry), (b) discrete structures and logic (e.g., linear algebra or number theory), (c) continuity and functions (e.g., multivariate or advanced calculus), and (d) probability and statistics (e.g., distributions).

Descriptive statistics for proportion correct scores on the MCK items and OTL by gender are provided in Table 19.1. The proportion correct was .10 higher among men. Also, in terms of OTL, men (M = 13.03, SD = 3.25) had a higher mean than women (M = 11.33, SD = 4.06), $t(340.29) = -4.80$, $p < .05$. The difference in OTL indicated that, overall, men had studied 1.6 more math topics than women. Women were found to have higher correlations between proportion correct and OTL than men. These descriptive statistics suggest that item level performance may be a function of gender, and that OTL may moderate the relationship between gender and performance.

Table 19.2 contains descriptive statistics for average proportion correct by gender and item format. The mean proportion correct on MC and CR items were .08 and .17 higher for men than women. Overall, both men and women performed better on MC than CR items. The preliminary findings indicate that item format may influence the item-level performance for men and women in different ways.

Models

In this study, model fit was compared for each model, with one model considered to be a reduced form of the subsequent model. Chi-squared likelihood ratio (χ^2) tests were conducted to test the appropriateness of more complex models. AIC (Akaike information criterion) and BIC (Bayesian information criterion) were also used to understand model fit. If the χ^2 was statistically significant and AIC was reduced for a more complex model in comparison to another, then the additional terms in that model were considered statistically significant. BIC provided supplemental fit information. In some cases, this model fit comparison approach was used to determine statistical significance for sets of parameters. Thus, individual effects were tested as needed using two-sided Wald tests with an alpha level of .05.

The baseline Model M0 in Eq. 19.2 includes random effects for both items and people. Since the means of both residual terms are set to 0, the intercept represents the mean difficulty for a person with mean ability, or the estimated log-odds of a correct response for an average person on an average item. The larger the estimated value, the easier items would be for an average person. Model M1 includes a main effect for gender, where $Gender_j$ equals 0 if person j belongs to the reference group women, and 1 if person j belongs to the focal group men:

$$\eta_{ij} = \gamma_0 + \gamma_1 Gender_j + u_{0i} + u_{0j}. \tag{19.4}$$

In this model, γ_0 estimates the mean performance for women, and γ_1 estimates the difference in mean performance for men compared to women. Thus, mean performance for men is expressed as ($\gamma_0 + \gamma_1$). The residual terms u_{0i} and u_{0j} still represent the random item and person effects.

Model M2 examines gender impact and item-by-gender interaction effects:

$$\eta_{ij} = \gamma_0 + \gamma_1 Gender_j + u_{1i} Gender_j + u_{0i} + u_{0j}. \tag{19.5}$$

The residual term u_{0i} now represents the random item effects for women, and u_{1i} estimates the overall differential item effects for men compared to women. This model is used to examine gender DIF. Variance in u_{1i} is taken as evidence of uniform DIF effects over gender groups. The random effects u_{0i} and u_{1i} can also be correlated. A positive correlation indicates that after controlling for the overall performance of all the people, the items with higher difficulty are harder for men. Item or person covariates would be included to explore DIF sources only if overall gender DIF is detected in Model M2.

Model M3 examines format impact and gender-by-format interaction effects. $Format_i$ is 0 if item i is MC, and is 1 if item i is CR. Format is added to the model as an item covariate to determine whether it contributes to DIF:

$$\begin{aligned}\eta_{ij} = \gamma_0 + \gamma_1 Gender_j + \gamma_2 Format_i + \gamma_3 Gender_j Format_i + \\ u_{1i} Gender_j + u_{0i} + u_{0j}.\end{aligned} \tag{19.6}$$

γ_0 now estimates the mean MC performance for women, and γ_1 estimates the difference in MC performance for men. γ_2 estimates the difference in CR performance compared to MC performance for women, and the interaction term γ_3 then estimates the difference in CR performance for men. This model can be used to examine how the inclusion of item format as an item covariate influences the gender DIF effects (u_{1i}). A statistically significant gender-by-item format interaction and/or a reduction in the variance of u_{1i} would provide evidence that item format contributes to the explanation of gender DIF.

Model M4 includes an OTL main effect, where the grand mean-centered OTL (OTL_j) is added to the model as a person covariate:

$$\begin{aligned}\eta_{ij} = \gamma_0 + \gamma_1 Gender_j + \gamma_2 Format_i + \gamma_3 Gender_j Format_i + \\ \gamma_4 OTL_j + u_{1i} Gender_j + u_{0i} + u_{0j}.\end{aligned} \tag{19.7}$$

The terms from Eq. 19.6 are now estimated while controlling for the impact of OTL. The additional term γ_4 estimates the effect of OTL on performance, controlling for gender and item format. Model M5 then investigates OTL impact and item-by-OTL interaction effects:

$$\begin{aligned}\eta_{ij} = \gamma_0 + \gamma_1 Gender_j + \gamma_2 Format_i + \gamma_3 Gender_j Format_i + \\ \gamma_4 OTL_j + u_{1i} Gender_j + u_{2i} OTL_j + u_{0i} + u_{0j},\end{aligned} \tag{19.8}$$

where u_{2i} estimates OTL effects at the item level. If the item-by-OTL interaction effects are significant and variance in the gender DIF effects $\sigma^2_{u_{0i}}$ is reduced, there is evidence that OTL contributes to the explanation of DIF.

Model M6 examines two-way interaction effects between gender and OTL:

$$\begin{aligned}\eta_{ij} = \gamma_0 + \gamma_1 Gender_j + \gamma_2 Format_i + \gamma_3 Gender_j Format_i + \\ \gamma_4 OTL_j + \gamma_5 Gender_j OTL_j + u_{1i} Gender_j + u_{2i} OTL_j + u_{0i} + u_{0j},\end{aligned} \tag{19.9}$$

where γ_5 estimates the extent to which the overall impact of OTL differs between men and women, or the extent to which the overall gender effect differs by OTL. Finally, Model M7 adds the three-way interaction effects between items, gender and OTL:

$$\begin{aligned}\eta_{ij} = \gamma_0 + \gamma_1 Gender_j + \gamma_2 Format_i + \gamma_3 Gender_j Format_i + \\ \gamma_4 OTL_j + \gamma_5 Gender_j OTL_j + u_{1i} Gender_j + u_{2i} OTL_j + \\ u_{3i} Gender_j OTL_j + u_{0i} + u_{0j},\end{aligned} \tag{19.10}$$

where u_{3i} estimates whether gender DIF for all items depends on OTL, or whether the impact of OTL at the item level differs by gender.

Models were fit sequentially based on significance. First, M2 was compared to M1, providing evidence of gender DIF. Starting from M3, if the inclusion of a covariate significantly improved model fit and reduced DIF, this covariate remained in subsequent models; if the covariate did not improve fit or contribute to the explanation of DIF, it was omitted from subsequent models.

Results

Models were fit using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in the statistical environment R (R Development Core Team, 2015). Table 19.3 contains results for fixed and random effects. Table 19.4 contains model fit results.

Gender

In the baseline Model M0, the intercept was estimated at 0.44 logits (see Table 19.3); thus, the probability for an average student to give a correct response on an average item was .61. The student variance term indicated that, for a student with an ability of one standard deviation lower and one standard deviation higher than the average ability, the expected probabilities of giving a correct answer to an item with average difficulty were .45 and .75, respectively, as calculated from the anti-logs of $\left(0.44-\sqrt{0.43}\right)$ and $\left(0.44+\sqrt{0.43}\right)$. The size of the item variance indicated that for a student with an average ability, the probabilities of answering an item correctly with a difficulty of one standard deviation lower or one standard deviation higher than the average difficulty were .36 and .81, which were the anti-logs of $\left(0.44-\sqrt{1.01}\right)$ and $\left(0.44+\sqrt{1.01}\right)$.

Both M1 and M2 were found to have significantly better model fit over the previous models. As shown in Table 19.3 for M2, the gender effect, which represented the difference between women and men in mean performance, was 0.52 logits

Table 19.3 Estimates of the parameters

Parameter	Notation	M0	M1	M2	M3	M4	M5	M6
Fixed								
Intercept	γ_0	0.44	0.28	0.27	0.50	0.55	0.55	0.55
Gender	γ_1		0.52	0.52	0.40	0.25	0.30	0.30
Format	γ_2				−0.98	−0.98	−1.00	−1.00
Gender*Format	γ_3				0.50	0.09	0.34	0.34
OTL	γ_4					0.52	0.09	0.09
Gender*OTL	γ_5							−0.003
Random								
Student	$\sigma^2_{u_{0j}}$	0.43	0.38	0.38	0.38	0.27	0.28	0.28
Item	$\sigma^2_{u_{0i}}$	1.01	1.01	1.14	0.97	0.97	1.02	1.02
Gender*Item	$\sigma^2_{u_{1i}}$			0.17	0.12	0.13	0.09	0.09
OTL*Item	$\sigma^2_{u_{2i}}$						0.005	0.005

Table 19.4 Model fit results

Model	*df*	AIC	BIC	Log Likelihood	χ^2	χ^2 *df*	*p*
M0	3	26627	26651	−13310			
M1	4	26528	26560	−13260	100.62	1	<.001
M2	6	26463	26512	−13226	68.92	2	<.001
M3	8	26447	26511	−13216	20.13	2	<.001
M4	9	25640	25712	−12811	808.96	1	<.001
M5	12	25458	25554	−12717	188.49	3	<.001
M6	13	25460	25564	−12717	0.02	1	.880

($z = 6.11$, $p < .001$). The log-odds of correct response for women was 0.27 and for men was 0.79, corresponding to predicted probabilities of correct response of .57 and .69. Additionally, the probability of correct response varied over students and items ($\sigma^2_{u_{0j}} = 0.38$ and $\sigma^2_{u_{0i}} = 1.14$). Conditional on the same ability, the overall item difficulty varied between women and men ($\sigma^2_{u_{1i}} = 0.168$). The improvement of M2 over M1 in model fit ($\chi^2_2 = 68.92$, $p < .001$), with reduced AIC and BIC, indicated the presence of statistically significant item by gender interaction effects, where items tended to show gender DIF. The negative correlation of the random effects u_{0i} and u_{1i} indicated that, controlling for overall performance, more difficult items were harder for women.

Gender and Item Format

The next step was to examine whether item format was a source of DIF. By incorporating item format as a covariate, the model fit was significantly improved for M3 over M2 ($\chi^2 = 20.13, p < .001$) with decreased AIC and BIC. As shown in Table 19.3 for M3, the mean MC performance for women was 0.50 logits. Men's mean MC performance was 0.40 logits higher. The corresponding predicted probabilities were .62 and .71 for women and men. Women performed worse in CR items than MC items by 0.98 logits. Moreover, the significance of the interaction term ($\gamma_3 = 0.50$, $z = 3.97$, $p < .001$) indicated that compared to women's performance on CR items, men outperformed by 0.50 logits, meaning that CR items favored men more than MC items regardless of the fact that men outperformed women in MC items. More importantly, regarding the reduction of DIF effects, not only were the interaction effects between gender and item format significant, but the variance of gender DIF ($\sigma^2_{u_{1i}}$) was also reduced after the item covariate was introduced. The proportion of gender DIF effects across all the items explained by item format was .29 (= 0.17–0.12/0.17), suggesting that item format contributed to DIF. Because the magnitude of DIF effects was still relatively large, analysis continued with item format remaining in subsequent models.

Gender, Item Format, and OTL

M4 and M5 additionally examined OTL impact and item-by-OTL interaction effects. Both models fit significantly better than the previous ones (see Table 19.4). However, M6 was not found to improve model fit over M5 (χ^2 = 0.02, p = .88); AIC and BIC values both increased. The interaction effects between gender and OTL were not significant ($\gamma_5 = -0.003$, $z = -0.15$, $p = .88$). Thus, M5 was retained as the final model. As indicated in the last column (M5) in Table 19.3, the main effect of OTL was significant ($\gamma_4 = 0.09$, $z = 8.32$, $p < .001$), indicating that a one-point increase in OTL (one additional math topic being studied) corresponded to an increase of 0.09 logits in math performance, holding other variables constant. The improvement of M5 over M4 (χ^2 = 188.49, $p < .001$), with reduced AIC and BIC, revealed that item-by-OTL interaction effects were significant; the random interaction term ($\sigma^2_{u_{2i}}$ = 0.005) showed that item difficulty varied over different levels of OTL. DIF was explained in part by OTL; in addition to a statistically significant main effect for OTL, the effect of OTL varied in a statistically significant way over items. Most importantly, the variance over items between the gender groups ($\sigma^2_{u_{1i}}$) was reduced from 0.12 to 0.09. The proportion of gender DIF that was explained by OTL was .25 (=0.12 − 0.09/0.12). Even though the interaction term between gender and OTL was found not to be significant, OTL mediated the relationship between item difficulty and gender. Therefore, for the U.S. cohort, it was concluded that both item format and OTL contributed to some extent to gender DIF.

Discussion

This study was designed primarily to describe the relationships between mathematics item difficulty, gender, item characteristics (specifically item format) and person characteristics (specifically OTL) within the U.S. cohort of the TEDS-M data set. The study demonstrated how cross-classified models can be used to examine both item and person covariates as potential sources of uniform DIF. Results from this study also provide information about how future secondary teachers' math performance is influenced by gender, item format, and OTL.

Results from the final Model M5 indicated that, overall, men tended to have higher mean math performance than women. Men outperformed women by 0.30 logits. The predicted mean proportion correct for women and men were .63 and .70, respectively. Thus, gender discrepancies in math performance existed in this study. The gender effect in this study is consistent with the findings from other research, where male advantages in standardized math tests have been reported (e.g., Langenfeld, 1997; Liu & Wilson, 2009). Math performance also tended to be higher on MC items than CR items; future teachers performed better on MC items than CR items by 1.00 logit. In addition, the interaction term indicated that, although women performed worse than men on both formats, the discrepancy between genders was even larger on CR items, with an advantage for men. These results are consistent with the previous finding that men have an advantage in MC items, but the results contradict the finding of a female advantage on CR items (e.g., Beller & Gafni, 2000; Bolger & Kellaghan, 1990).

When OTL increased, math performance tended to improve. A one-unit increase in OTL resulted in an increase of 0.09 logits in math performance. In other words, with one more topic studied among the four topics that OTL measures, there was an estimated increase in performance of 0.09 logits. The relationship between OTL and mean performance is in line with the positive correlations reported in Table 19.1 and in previous studies (e.g., Wang & Goldschmidt, 1999).

A DIF effect for gender was defined as the differential item effects of belonging to a specific group. Results indicated the presence of gender DIF effects. Effects were examined in an omnibus test where no specific DIF items were identified. The results indicated that throughout the 76 items, there were some items that functioned differently between the gender groups. Conditioned on overall ability, more difficult items favored men more than women. Thus, it was evident that gender DIF existed.

Item format was tested to determine whether it was associated with gender DIF. Future teachers' overall math performance on the two item formats depended on what gender group they belonged to. The magnitude of gender DIF across all the items was significantly reduced after the covariate item format was incorporated. The proportion of gender DIF effects explained by item format was .29. Results from the literature (e.g., Taylor & Lee, 2012) have confirmed the finding that item format was associated with gender DIF.

This study also revealed that with different levels of OTL, items functioned differentially. Furthermore, the inclusion of OTL resulted in a significant reduction of random gender effects over items. The conclusion was that OTL mediated the relationship between item difficulty and gender for some DIF items. Nevertheless, the interaction between gender and OTL did not improve model fit; thus, the relationship between OTL and overall performance did not differ significantly by gender. These results supported the findings of Albano and Rodriguez (2013) and Cheong (2006), where OTL was found to be related to DIF.

This study is an extension of the original study of Albano and Rodriguez (2013), who examined differential mathematics performance due to gender and OTL using hierarchical generalized linear modeling where person effects were viewed as random and item effects as fixed. To investigate DIF sources, the present study used cross-classified multilevel models, in which both item and person effects were treated as random. Item-level and person-level covariates were then estimated simultaneously. In addition to using OTL as the person-level covariate, this study examined how item format could potentially explain variability in item difficulty and moderate the relationship between gender and item difficulty.

This study demonstrates the application of cross-classified multilevel models in educational research. The cross-classified multilevel model is a flexible tool for explaining potential DIF sources related to item and person characteristics. This approach results in more economical models, where DIF can be detected within an omnibus test. This approach can be helpful in creating and adapting appropriate measurement tools when constructing or translating items. Moreover, with respect to person characteristics, researchers can take variables such as OTL into account and thus improve DIF detection and estimation. By doing so, the validity of group comparisons can be improved.

There are some limitations to the approach demonstrated in this study. Inadequate sample size may have resulted in a lack of precision when estimating complex model terms (e.g., gender-by-OTL interaction effects). A lack of power limits the possibility of incorporating more covariates that could potentially explain variability in item difficulty by gender. Also, the measure of OTL only accounted for self-reported exposure to certain math content. Other important factors in measuring OTL, such as hours in class, quality of teachers' feedback, and level of cognitive demand, are not included in the instrument, which can be problematic. Future studies should seek larger sample sizes and consider more comprehensive measures of OTL.

Additionally, important item features such as item content domain and cognitive subdomain could be explored as potential sources of gender DIF. Past research has shown that men and women tend to adopt different strategies when responding to certain problem characteristics (e.g., Bolger & Kellaghan, 1990; DeMars, 2000). Studies also have indicated that content and cognitive skills required in items are related to gender DIF in math (e.g., Gierl, Bisanz, Bisanz, & Boughton, 2003; Harris & Carlton, 1993). Furthermore, a third level of data nesting (e.g., institution level) could be incorporated to examine social and/or psychological context effects (e.g., Entwisle, Alexander, & Olson, 1994; Van den Noortgate & De Boeck, 2005). Future work could examine other important covariates at the item and person levels while also incorporating additional levels of nesting.

References

Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement, 29*, 67–91.

Adams, R. J., Wilson, M., & Wu, M. (1997). Multilevel item response models: An approach to errors in variables regression. *Journal of Educational and Behavioral Statistics, 22*, 47–76.

Albano, A. D., & Rodriguez, M. C. (2013). Examining differential math performance by gender and opportunity to learn. *Educational and Psychological Measurement, 73*, 836–856.

Baker, D. P., & Jones, D. P. (1993). Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education, 66*, 91–103.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4* [Computer software manual]. (R package version 1.1-8). Retrieved from http://CRAN.R-project.org/package=lme4

Becker, B. J. (1990). Item characteristics and gender differences on the SAT-M for mathematically able youths. *American Educational Research Journal, 27*, 65–87.

Beller, M., & Gafni, N. (1996). 1991 international assessment of educational progress in mathematics and sciences: The gender differences perspective. *Journal of Educational Psychology, 88*, 365–377.

Beller, M., & Gafni, N. (2000). Can item format (multiple choice vs. open-ended) account for gender differences in mathematics achievement? *Sex Roles, 42*, 1–21.
Bielinski, J., & Davison, M. L. (1998). Gender differences by item difficulty interactions in multiple-choice mathematics items. *American Educational Research Journal, 35*, 455–476.
Bielinski, J., & Davison, M. L. (2001). A sex difference by item difficulty interaction in multiple-choice mathematics items administered to national probability samples. *Journal of Educational Measurement, 38*, 51–77.
Blömeke, S., Suhl, U., Kaiser, G., & Döhrmann, M. (2012). Family background, entry selectivity and opportunities to learn: What matters in primary teacher education? An international comparison of fifteen countries. *Teaching and Teacher Education, 28*, 44–55.
Bolger, N., & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement, 27*, 165–174.
Boscardin, C. K., Aguirre-Munoz, Z., Stoker, G., Kim, J., Kim, M., & Lee, J. (2005). Relationship between opportunity to learn and student performance on English and algebra assessments. *Educational Assessment, 10*, 307–332.
Burkes, L. L. (2009). *Identifying differential item functioning related to student socioeconomic status and investigating sources related to classroom opportunities to learn*. Unpublished dissertation, University of Delaware.
Cheong, Y. F. (2006). Analysis of school context effects on differential item functioning using hierarchical generalized linear models. *International Journal of Testing, 6*, 57–79.
Cohen, A. S., & Bolt, D. M. (2005). A mixture model analysis of differential item functioning. *Journal of Educational Measurement, 42*, 133–148.
DeMars, C. E. (1998). Gender differences in mathematics and science on a high school proficiency exam: The role of response format. *Applied Measurement in Education, 11*, 279–299.
DeMars, C. E. (2000). Test stakes and item format interactions. *Applied Measurement in Education, 13*, 55–77.
Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychology of Women Quarterly, 18*, 585–609.
Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103–127.
Entwisle, D. R., Alexander, K. L., & Olson, L. S. (1994). The gender gap in math: Its possible origins in neighborhood effects. *American Sociological Review, 59*, 822–838.
Floden, R. E. (2002). The measurement of opportunity to learn. In A. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 231–266). Washington, DC: National Academies Press.
Gallagher, A., & Kaufman, J. (2005). *Gender differences in mathematics*. New York, NY: Cambridge University Press.
Gamer, M., & Engelhard, G., Jr. (1999). Gender differences in performance on multiple-choice and constructed response mathematics items. *Applied Measurement in Education, 12*, 29–51.
Gierl, M. J., Bisanz, J., Bisanz, G. L., & Boughton, K. A. (2003). Identifying content and cognitive skills that produce gender differences in mathematics: A demonstration of the multidimensionality-based DIF analysis paradigm. *Journal of Educational Measurement, 40*, 281–306.
Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what? *British Educational Research Journal, 28*, 567–590.
Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory (Measurement methods for the social science)*. Newbury Park, CA: Sage.
Harris, A. M., & Carlton, S. T. (1993). Patterns of gender differences on mathematics items on the scholastic aptitude test. *Applied Measurement in Education, 6*, 137–151.
Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In *Test validity* (pp. 129–145). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Husén, T. (1967). *International study of achievement in mathematics: A comparison of twelve countries*. New York, NY: Wiley.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139–155.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321*, 494–495.

Jacobs, J. E., Davis-Kean, P., Bleeker, M., Eccles, J. S., & Malanchuk, O. (2005). "I can, but I don't want to": The impact of parents, interests, and activities on gender differences in math. In A. Gallagher & J. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 246–263). New York, NY: Cambridge University Press.

Kamata, A. (2001). Item analysis by the hierarchical generalized linear model. *Journal of Educational Measurement, 38*, 79–93.

Kim, S.-H., Cohen, A. S., Alagoz, C., & Kim, S. (2007). DIF detection and effect size measures for polytomously scored items. *Journal of Educational Measurement, 44*, 93–116.

Langenfeld, T. E. (1997). Test fairness: Internal and external investigations of gender bias in mathematics testing. *Educational Measurement: Issues and Practice, 16*, 20–26.

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*, 1123–1135.

Liu, O. L., & Wilson, M. (2009). Gender differences in large-scale math assessments: PISA trend 2000 and 2003. *Applied Measurement in Education, 22*, 164–184.

Longford, N. T., Holland, P. W., & Thayer, D. T. (1993). Stability of the MH D-DIF statistics across populations. In H. Wainer (Ed.), *Differential item functioning* (pp. 171–196). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis, 17*, 305–322.

Meece, J. L., Parsons, J. E., Kaczala, C. M., & Goff, S. B. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin, 91*, 324–348.

Mendes-Barnett, S., & Ercikan, K. (2006). Examining sources of gender DIF in mathematics assessments using a confirmatory multidimensional model approach. *Applied Measurement in Education, 19*, 289–304.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543.

Mullis, I. V., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 international mathematics report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Muthén, B. O., Kao, C.-F., & Burstein, L. (1991). Instructionally sensitive psychometrics: Application of a new IRT-based detection technique to mathematics achievement test items. *Journal of Educational Measurement, 28*, 1–22.

Penner, A. M. (2003). International gender × item difficulty interactions in mathematics and science achievement tests. *Journal of Educational Psychology, 95*, 650–655.

R Development Core Team. (2015). *R: A language and environment for statistical computing* [Computer software manual]. (Version 3.2.1). Vienna, Austria. Retrieved from http://www.R-project.org

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danmarks Paedogogiske Institute.

Roussos, L., & Stout, W. (1996). A multidimensionality-based DIF analysis paradigm. *Applied Psychological Measurement, 20*, 355–371.

Rupp, A. A., & Zumbo, B. D. (2006). Understanding parameter invariance in unidimensional IRT models. *Educational and Psychological Measurement, 66*, 63–84.

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin, 135*, 859–884.

Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*, 361–370.

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): Conceptual Framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.

Taylor, C. S., & Lee, Y. (2012). Gender DIF in reading and mathematics tests with mixed item formats. *Applied Measurement in Education, 25*, 246–280.

Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In P. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 67–113). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Travers, K. J., & Westbury, I. (1989). *The IEA study of mathematics I: Analysis of mathematics curricula*. Oxford, UK: Pergamon Press.

Van den Noortgate, W., & De Boeck, P. (2005). Assessing and explaining differential item functioning using logistic mixed models. *Journal of Educational and Behavioral Statistics, 30*, 443–464.

Van den Noortgate, W., De Boeck, P., & Meulders, M. (2003). Cross-classification multilevel logistic models in psychometrics. *Journal of Educational and Behavioral Statistics, 28*, 369–386.

Wang, J., & Goldschmidt, P. (1999). Opportunity to learn, language proficiency, and immigrant status effects on mathematics achievement. *The Journal of Educational Research, 93*, 101–111.

Willingham, W., & Cole, S. (1997). *Gender and fair assessment*. Mahwah, NJ: Lawrence Erlbaum.

Wiseman, A. W. (2008). A culture of (in) equality?: A cross-national study of gender parity and gender segregation in national school systems. *Research in Comparative and International Education, 3*, 179–201.

Chapter 20
Standing the Test of Time: Validating the TEDS-M Knowledge Assessment Against MET II Expectations

Edward A. Silver and Jillian P. Mortimer

Abstract We report the results of a post hoc validation study in which publicly released TEDS-M knowledge assessment items for elementary teachers were validated against the specifications for teacher knowledge found in the Mathematical Education of Teachers II (MET II) report. The validation was based on the expert judgments of two authors of the MET II report. Raters reported a suitable match between specific MET II mathematical content and mathematical practice expectations for almost all TEDS-M items. Inter-rater agreement was very high regarding mathematical content expectations, but much lower for mathematical practices. Expert ratings indicated that TEDS-M released items mapped to some but not all MET II content domains. The findings suggest strong content validity for TEDS-M items in some respects though not in others.

Introduction

Because test scores are often used to make claims about the attributes of test-takers that go beyond what is observable in assessments, close examination of the plausibility of such claims is necessary. In relation to claims about educational assessment, an argument-based approach to validity is used to evaluate plausibility (Kane, 2013). Along with reliability and fairness, validity is generally regarded as one of the most critical features of any educational assessment (AERA, APA, & NCME, 1999; Pellegrino, Chudowsky, & Glaser, 2001).

TEDS-M and the study contained in this chapter were supported by funding provided by a grant from the National Science Foundation Award No. REC – 0514431 (M.T. Tatto, PI). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

E. A. Silver (✉) · J. P. Mortimer
University of Michigan, Ann Arbor, MI, USA
e-mail: easilver@umich.edu; jbpet@umich.edu

M. T. Tatto et al. (eds.), *Exploring the Mathematical Education of Teachers Using TEDS-M Data*, https://doi.org/10.1007/978-3-319-92144-0_20

One component of an argument for the validity of an assessment is the alignment between the assessment and a set of content standards (AERA, APA, & NCME, 1999; Ananda, 2003; Case, Jorgensen, & Zucker, 2004; Resnick, Rothman, Slattery, & Vranek, 2003; Webb, 1997). This is similar to what Messick (1989) referred to as content validity. Expert judgments about what a test item measures or the content domain covered by a test are usually made during test development or soon after, often by subject matter experts involved in the development of the test. Typically, these judgments involve examining the test items in relation to a test specification framework or a set of learning outcomes that the test is intended to assess.

Although expert review of test content to determine alignment is a well-established and almost universally followed procedure for assessments in academic content subject areas such as mathematics, the adequacy of this kind of judgment-based validity evidence has received some criticism (e.g., Guion, 1977). In particular, when the experts making the judgments are also involved in the test development process, the process tends to be highly subjective, and it has a strong confirmatory bias (Kane, 2001). Moreover, even when done well, the expert judgments during the test development process or shortly thereafter can be made quite a long time before the results of an assessment are reported. To the extent that public or professional views of the content domain of interest may be altered during the intervening period, the validity evidence provided by the original expert review is no longer current and is weakened as an argument for the validity of interpretations from the assessment.

While recognizing the important role of expert review during, or shortly after the test development process, we argue that it might be useful to examine the validity argument for an assessment at multiple time points. In particular, post hoc reviews might be useful after the release of assessment results and prior to the revision of an assessment for subsequent administration. We argue that these subsequent reviews could and should be based on views of the content domain of interest that differ from the views that applied during the original content validity review process. In this way, the examination of content validity of a test can be refreshed and updated to enhance the validity of the interpretations of performance on the test.

In this chapter, we report an example of a post hoc review of the mathematics knowledge assessment developed for and used in the Teacher Education and Development Study in Mathematics (TEDS-M). In doing so, we illustrate what a post hoc validation process might look like and how it might be valuable both in interpreting the findings of TEDS-M and in revising the TEDS-M knowledge assessment for future use. We used *The Mathematical Education of Teachers II* (MET II) report (Conference Board of the Mathematical Sciences, 2012) as the referential framework for this post hoc validity analysis.

Across the world there is large-scale consensus about the mathematics needed by secondary school teachers of mathematics; namely, the equivalent of an undergraduate mathematics minor or major. But there is far less consensus regarding the knowledge needed by elementary school teachers of mathematics. In some countries, the expectations are similar to those for secondary school mathematics teachers but in other countries far less mathematics would be expected. For this reason, we focused our attention on the TEDS-M assessment of elementary teachers of mathematics.

We chose the MET II report because it is one of the most recent in a long line of efforts to specify what mathematics teachers need to know to be well prepared to teach at the elementary and middle-grades level, and because it was created and published under the auspices of the Conference Board of Mathematical Sciences (CBMS), which is a consortium of mathematics professional organizations. Thus, we view MET II to be a strong, contemporary representation of the mathematical sciences community's view of the mathematical knowledge needed by elementary and middle-grades teachers. As such, and because the TEDS-M assessment was created and used prior to the publication of the MET II document, we judged the MET II report to be a reasonable choice as a referential framework for a post hoc validation of the TEDS-M mathematics knowledge assessment.

Methods

Representations of Teacher Knowledge

Teacher Education and Development Study in Mathematics (TEDS-M) The TEDS-M assessment was developed, with the support of the International Association for the Evaluation of Educational Achievement (IEA), by the research centers at Michigan State University (MSU) and the Australian Council for Educational Research (ACER), along with national research coordinators (NRCs) in 17 participating countries between 2006 and 2011 (Tatto, 2013). The goal of TEDS-M was to learn more about how primary and lower-secondary teachers are prepared to teach mathematics and the impact of teacher education programs in various countries. To achieve this goal, the researchers drew representative samples of teacher preparation institutions, teacher educators, and future teachers from 17 countries, including over 15,000 primary pre-service teachers, over 9,000 lower-secondary pre-service teachers, and around 5,000 teacher educators in 500 pre-service teacher education institutions (Tatto, 2013).

TEDS-M includes a variety of measures of factors that are thought to influence future teachers' mathematical knowledge. In this study, we focus on the mathematics content knowledge and mathematics pedagogical content knowledge assessment for elementary teachers that was administered near the end of the teacher preparation programs in participating countries. TEDS-M identified three categories for each assessment item: knowledge type (content knowledge or pedagogical content knowledge), content domain (number, geometry, data or algebra) and process (reasoning, knowing, enacting, curriculum/planning, or applying). Figure 20.1 shows examples of the types of questions asked within each knowledge type, content domain, and process.

The assessment consisted of 124 items, from which a subset of 34 items was publicly released. We used only the publicly released items for the analysis reported in this paper. The released items were chosen to be "representative of the different domains in the test and to represent the different types of items used" (M. T. Tatto, personal communication, January 23, 2015). Table 20.1 shows the percent of items in the complete TEDS-M assessment and the percent of released items that were categorized by the TEDS-M staff with respect to content domain, sub-domain or process, knowledge type, item format, and mathematical level.

Validity Evidence for TEDS-M Items While constructing the knowledge assessment items, creators of the TEDS-M assessment went through a process to ensure content validity: "At each stage of the item development process, expert panels examined the content validity and appropriateness of the items. These reviews took into consideration clarity, correctness, cultural relevance, classification within the framework of domains and subdomains, relevance to teacher preparation, and curricular level" (Tatto et al., 2008, p. 35). Items were piloted in 11 countries, and decisions to include or exclude items were made based on item statistics as well as instances of translation-related problems (Tatto, 2013).

The Mathematical Education of Teachers II (MET II) The MET II document was developed by a team of U.S. mathematicians and mathematics educators to inform faculty involved in teacher education, including those in university mathe-

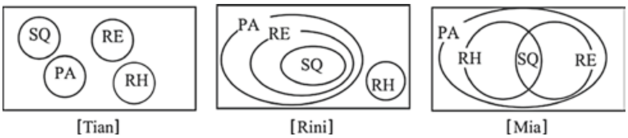
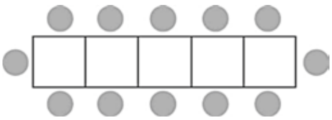
Item, Knowledge Type, Content Domain, Process	
MFC 204 MCK Geometry Knowing	Three students have drawn the following Venn diagrams showing the relationships between four quadrilaterals: Rectangles (RE), Parallelograms (PA), Rhombuses (RH), and Squares (SQ).  Which student's diagram is correct? A. [Tian] 1 B. [Rini] 2 C. [Mia] 3
MFC208A MPCK Number Enacting	[Jeremy] notices that when he enters 0.2 × 6 into a calculator his answer is smaller than 6, and when he enters 6 ÷ 0.2 he gets a number greater than 6. He is puzzled by this, and asks his teacher for a new calculator! (a) What is [Jeremy's] most likely misconception?
MFC308 MCK Algebra Applying	A square table can seat four people, one on each side. When 5 square tables are placed side by side, as shown below, 12 people can sit around them, 5 on each side and 2 on the ends.  How many people can sit around *n* square tables when they are placed side by side? Write your answer to the problem in terms of *n*.
MFC502A MCK Data Reasoning	The following problem was given to children in <primary> school *The graph shows the number of pens, pencils, rules, and erasers sold by a store in one week.*

Fig. 20.1 Examples of TEDS-M assessment items categorized by knowledge type, content domain and process

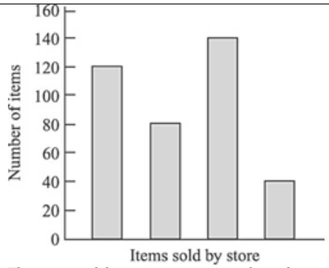
	 *The names of the items are missing from the graph. Pens were the item most often sold. Fewer erasers than any other item were sold. More pencils than rulers were sold.* How many pencils were sold? A. 40 B. 80 C. 120 D. 140
MFC505 MPCK Number Curriculum /Planning	A <Grade 1> teacher asks her students to solve the following four story problems, in any way they like, including using materials if they wish. Problem 1: [Jose] has 3 packets of stickers. There are 6 stickers in each pack. How many stickers does [Jose] have altogether? □ Problem 2: [Jorgen] had 5 fish in his tank. He was given 7 more for his birthday. How many fish did he have then? Problem 3: [John] had some toy cars. He lost 7 toy cars. Now he has 4 cars left. How many toy cars did [John] have before he lost any? □ Problem4: [Marcy] had 13 balloons. 5 balloons popped. How many balloons did she have left? The teacher notices that two of the problems are more difficult for her children than the other two. Identify the TWO problems which are likely to be more DIFFICULT to solve for <Grade 1> children. Problem _______ and Problem _______

Fig. 20.1 (continued)

matics departments, and policy makers about the mathematical knowledge needed by teachers in the elementary, middle, and secondary grades (CBMS, 2012). This updated version of the MET II document was prompted in part by the adoption of the Common Core State Standards (CCSS) by most states. The MET II document identifies items of content knowledge, each called an essential idea (EI), that teachers of mathematics at different grade-level ranges should know. In the document, it is argued that, in addition to knowing EIs associated with specific mathematics content, teachers at all grade levels should know the Common Core State Standards for Mathematical Practice (SMPs), which relate to processes such as problem solving, reasoning, and communication.

In this study, we used only the MET II EIs for elementary and middle-grades teachers. Although this study only considers the TEDS-M assessment for elementary pre-service teachers, we included the expectations for both levels because the

Table 20.1 Percent of items in the primary assessment as compared to the released item set by item characteristic

	All primary assessment items (%)	Released primary assessment items (%)
Content domain		
Number	35.5	29.4
Geometry	28.2	23.5
Algebra	26.6	35.3
Data	9.7	11.8
Sub-domain		
Applying	22.6	23.5
Enacting	18.5	11.8
Knowing	33.9	44.1
Reasoning	11.3	2.9
Curriculum/Plan	13.7	17.6
Knowledge dimension		
MCK	67.7	70.6
PCK/MPCK	32.3	29.4
Item format		
Multiple choice	23.4	44.1
Complex multiple choice	61.3	26.5
Constructed response	15.3	29.4
Mathematical level		
Novice	24.2	38.2
Intermediate	46.0	32.4
Advanced	29.8	29.4

TEDS-M assessment was created to meet international norms and was designed to stretch beyond the specifics required within the target grade span. MET II identified six mathematical knowledge domains for elementary teachers (Counting & Cardinality, Operations & Algebraic Thinking, Number & Operations in Base Ten, Number & Operations-Fractions, Measurement & Data, and Geometry) and six domains for middle-grades teachers (Ratio & Proportional Relationships, The Number System, Expressions & Equations, Functions, Geometry, and Statistics & Probability). Within each domain, MET II specified one or more EIs. Table 20.2 shows the number of EIs in each domain as specified by MET II. Note that Geometry has more EIs than other domains because there are geometry domains for both elementary teachers and middle-grades teachers, and we have collapsed these into one geometry domain with all the corresponding EIs, thereby reducing the total number of domains from 12 to 11.

Table 20.2 Frequency of EIs in each MET II domain

MET2 domain	Number of EIs within domain
Counting and cardinality	1
Operations and algebraic thinking	3
Number and operations in base ten	3
Number and operations-fractions	4
Measurement and data	4
Geometry	8
Ratio and proportional relations	5
The number system	5
Expressions and equations	3
Functions	2
Statistics and probability	4
Total	42

Expert Raters

To explore the content validity of the TEDS-M elementary knowledge assessment with respect to the MET II EIs, we solicited the judgments of two experts in mathematics education who were among the authorship team for the MET II document; one was a mathematics teacher educator, and one was a mathematician.

The Rating Task

The raters worked independently. Each rater was given the released TEDS-M assessment items, descriptions of the MET II EIs for elementary and middle-grades teachers with a naming convention that we created, a list of the SMPs, and a recording tool. For each TEDS-M assessment item the raters were asked to identify, if possible, both a MET II EI and an SMP that best fit the item.

For the EI assignment, raters were given the list of all 42 of the MET II EIs and asked to choose one, if any, that best fit each TEDS-M item. The MET II EIs were organized by domain so that there was a domain at the top of each column in the recording tool and a drop-down list of the EIs that fell under the corresponding domain as well as an "Other" option, which was intended for use in cases when a rater judged that an assessment item fell under a particular domain but did not fit with any of the specific EIs listed for that domain in the MET II document. Also, if a rater determined that additional EIs corresponded to an item, the rater could identify any secondary EIs in the "Notes" column in the recording tool.

For the SMP assignment, raters were given a list of the eight SMPs and asked to choose one, if any, that best fit each TEDS-M item. As with the EIs, if a rater determined that additional SMPs corresponded to an item, the rater could identify any secondary SMPs in the "Notes" column in the recording tool. Because the SMPs were not organized by domain, there was no "Other" option for the SMP assignment.

This process yielded at most one EI and one SMP as the primary EI and SMP for each rater for each TEDS-M assessment item. For some items, secondary EIs or SMPs might also be identified by a rater.

Each rater completed this task individually, and the identity of the other rater was not revealed until after the ratings were completed. Following some preliminary analyses, we conducted follow-up interviews with individuals that were intended to help us understand expert raters' thinking regarding select items, specifically items they identified as Other within a particular domain and items where they identified more than one EI or SMP. Then we conducted a follow-up discussion with both raters to explore items on which the expert raters disagreed, assigning different EIs or SMPs to assessment items. The goal of these discussions was not to have the raters come to a consensus but to understand how different thinking about the mathematical content of assessment items or about the meaning of MET II EIs or SMPs could result in different ratings.

Inter-Rater Agreement

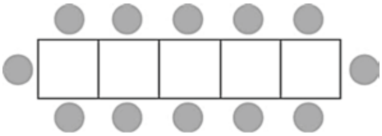
To assess inter-rater agreement we used Cohen's kappa (Banerjee, Capozzoli, McSweeney, & Sinha, 1999). This measure determines the amount of agreement between two raters with a correction for the expectation that some of the agreement is expected purely by chance. Because of differences in the ways in which EIs and SMPs could be matched with items, we separated these judgments in our examination of the strength of agreement between the raters.

MET II Essential Idea Agreement Agreement between raters regarding the MET II EIs could reasonably be based on one or more of three different pairwise comparisons: *identical agreement*, *domain agreement*, and *secondary agreement*.

Identical agreement refers to those instances when the raters assigned the same MET II EI as the primary classification for an item. Domain agreement refers to those instances when the raters differed regarding the specific EI but assigned as the primary classification a different EI from the same MET II content domain.

An example of domain agreement is shown below in Fig. 20.2. Though raters 1 and 2 did not assign the same EI to item MFC308, both agreed that EIs within the MET II domain Expressions and Equations (EAE) mapped to the item. Rater 2 chose EAE1, which is described as "viewing numerical and algebraic expressions as 'calculation recipes,' describing them in words, parsing them into their component parts, and interpreting the components in terms of a context" (CBMS, 2012, p. 42), while rater 1 selected EAE2, which is described as "examining lines of reasoning used to solve equations and systems of equations" (CBMS, 2012, p. 42).

MFC308. A square table can seat four people, one on each side. When 5 square tables are placed side by side, as shown below, 12 people can sit around them, 5 on each side and 2 on the ends.



How many people can sit around *n* square tables when they are placed side by side? Write your answer to the problem in terms of *n*.

Item	Expert rater 1	Expert rater 2
MFC308	EAE2	EAE1

Fig. 20.2 Example of domain agreement

Secondary agreement refers to those instances when raters disagreed regarding the primary EI but were in agreement when a secondary assignment was considered. An example of Secondary agreement for item MFC108 is shown below in Fig. 20.3. Expert rater 1 assigned FUN1—"examining and reasoning about functional relationships represented using tables, graphs, equations, and descriptions of functions in words. In particular, examining how the way two quantities change together is reflected in a table, graph, and equation" (CBMS, 2012, p. 43)—to the item, and expert rater 2 assigned EAE1 to the item. These EIs are neither identical nor drawn from the same domain, but expert rater 2 also indicated FUN1 as a secondary assignment for the item. This was coded as an instance of secondary agreement.

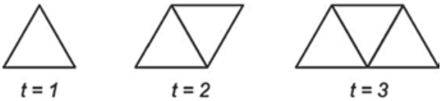
SMP Agreement We similarly calculated instances of rater agreement for SMPs, except that the category of domain agreement was not applicable to SMPs because they are not organized based on domains. Thus, we calculated only identical agreement and secondary agreement.

Representation of MET II EIs in the TEDS-M Knowledge Assessment

Following the agreement calculations, we examined the representation of the different MET II EIs within the released TEDS-M items. We first grouped all of the EIs by domain to give a more succinct and manageable picture of the representation of the kinds of content included in the EIs. Then we counted the number of times expert raters assigned an EI from each domain to a TEDS-M assessment item.

MFC108. [Amy] is building a sequence of geometric figures with toothpicks by following the pattern shown below. Each new figure has one extra triangle.□Variable *t* denotes the position of a figure in the sequence.

In finding a mathematical description of the pattern, [Amy] explains her thinking by saying:



Variable *n* represents the total number of toothpicks used in a figure.□ Which of the equations below best represent [Amy's] statement in algebraic notation?

A. $n = 2t + 1$
B. $n = 2(t + 1) - 1$
C. $n = 3t - (t - 1)$
D. $n = 3t + 1 - t$

			Notes
Expert rater 1	FUN1		
Expert rater 2		EAE1	Also related to FUN1

Fig. 20.3 Example of secondary agreement

Results

The Ratings Task

In 100% of cases the raters were able to assign either a MET II EI or a rating of Other within a domain (e.g., Functions Other) to TEDS-M items. Additionally, expert raters were able to assign SMPs to TEDS-M items in 82% of the cases. When assigning EIs to TEDS-M assessment items, the raters assigned 16 of the possible 42 EIs, and these 16 EIs were drawn from nine of the 11 MET II content domains. For six items, raters assigned a domain but designated Other rather than any of the specific EIs associated with the domain.

We examined the strength of agreement between the raters in assigning EIs to the TEDS-M items. On 56% of the items, the two raters assigned the identical EI as primary, and on an additional 32% of the items, they assigned EIs as primary from within the same MET II content domain. There was only one instance of secondary agreement for the EIs, so we eliminated that from further consideration in our analysis of results. Cohen's kappa for domain-level (i.e., identical EI assignment or assignment of EIs within the same domain) inter-rater agreement was .86; adding secondary inter-rater agreement boosted Cohen's kappa to .89.

We also examined the strength of agreement between the raters in assigning SMPs to the TEDS- M items. On 21% of the items, the two raters assigned the identical SMP as primary, and on an additional 21% of the items, they had secondary agreement on an SMP assignment. Cohen's kappa for SMP inter-rater agreement (either primary or secondary) was .33. Landis and Koch (1977) offer guidelines for interpreting Cohen's kappa: values <0 indicate no agreement, values between .00–.20 indicate slight agreement, .21–.40 indicate fair agreement, .41–.60 indicate moderate agreement, .61–.80 indicate substantial agreement, and .81–1.00 indicate almost perfect agreement. Using these guidelines, the inter-rater agreement for the MET II EIs can be categorized as almost perfect agreement and for the SMPs as fair agreement. This aligns with the raters' own reports in the follow-up interviews that the SMPs were more challenging to assign both because they were not as well specified as the EIs, and, in many instances, more than one SMP could be seen as related to an item depending on the approach taken to solve a task. The uncertainty of interpretation and multiplicity of possibly applicable practices might also be reflected in the relatively frequent incidence of secondary agreement for the SMPs, as compared to secondary agreement for the EIs.

Rater Agreement with Respect to Item Characteristics

Table 20.3 shows the percent agreement between the raters in their EI assignments, either identical agreement or domain agreement, with respect to the dimensions of item variation identified by the TEDS-M assessment creators (content domain, process, knowledge type, item format, and mathematical level). As reported above, the overall percentage of items on which the raters had identical agreement was roughly twice that of the items on which they had domain agreement. As the data presented in Table 20.3 show, this ratio was not observed uniformly for all item characteristics. For instance, raters frequently had identical agreement in assigning EIs for TEDS-M items classified as Geometry or Data, but raters were not as likely to have identical agreement for items classified within the domains of Number or Algebra. Also, with respect to the TEDS-M process classifications of items, raters were more likely to reach domain agreement for items classified as "knowing" than was the case for items within other process categories. Some variations can also be seen in regard to knowledge type or item format categories.

Representation of MET II EI Domains in the TEDS-M Knowledge Assessment

Ratings indicated considerable variation in the representation of the MET II domains within the TEDS-M items. Table 20.4 shows the number of times an EI from each domain was used in assigning EIs to TEDS-M assessment items.

EIs in the domains of Measurement and Data, The Number System, and Expressions and Equations were assigned more than 15 times each. Overall, EIs from these domains comprise 71% of the ratings. In contrast, EIs from the other domains were used infrequently, with EIs from Counting and Cardinality and

Table 20.3 Percent agreement in assigning EIs by item characteristics

	Identical agreement (%)	Domain agreement (%)
Content domain		
Number (10 items)	30.0	60.0
Geometry (8 items)	75.0	12.5
Algebra (12 items)	50.0	41.7
Data (4 items)	75.0	0.0
Process		
Applying (8 items)	62.5	25.0
Enacting (4 items)	75.0	0.0
Knowing (15 items)	40.0	53.5
Reasoning (1 items)	100.0	0.0
Curriculum/Plan (6 items)	66.7	16.7
Knowledge type		
MCK (24 items)	50.0	41.7
MPCK (10 items)	70.0	10.0
Item format		
Multiple choice (15 items)	60.0	26.7
Complex multiple choice (9 items)	44.4	55.6
Constructed response (10 items)	60.0	20.0
Mathematical level		
Novice (13 items)	61.5	30.8
Intermediate (11 items)	54.5	27.3
Advanced (10 items)	50.0	40.0

Table 20.4 Frequency of EI assignment by MET II content domain

MET II domain	EI assignment frequency
Counting and cardinality	0
Operations and algebraic thinking	2
Number and operations in base ten	6
Number and operations-fractions	0
Measurement and data	16
Geometry	5
Ratio and proportional relations	4
The number system	17
Expressions and equations	15
Functions	1
Statistics and probability	2
Instances where no EI was assigned	0
Total	68

Table 20.5 Domains for which "domain agreement" occurred

Domain	Number of Items with domain agreement	Frequency of assignment of EIs from domain	Number of EIs within domain
Geometry	1	5	8
Ratio and proportional relations	1	4	5
The number system	4	17	5
Expressions and equations	5	15	3

Number and Operations-Fractions not being assigned at all. The omission of EIs related to Counting and Cardinality may be explained by the relative simplicity of items that would likely fall into this domain. However, fractions are foundational in mathematics at the elementary grades and are exceptionally difficult for students.

Domain agreement occurred in four of the nine domains in which raters assigned EIs to items: Geometry, Ratio and Proportional Relations, The Number System, and Expressions and Equations. Table 20.5 gives more details about these four domains and the way they were treated within the raters' judgments.

As the leftmost column indicates, the domains of The Number System and Expressions and Equations generated judgments of domain agreement much more frequently than did the domains of Geometry and Ratio and Proportional Relations. This variation in frequency of assignment corresponds roughly to the variation in frequency data provided in the center column showing the total number of times that the raters assigned EIs from each domain to TEDS-M released items. These data seem to suggest that the set of released TEDS-M test items may over-represent certain EIs within the MET II domains of The Number System and Expressions and Equations in comparison with the representation for the domains of Geometry and Ratio and Proportional Relations.

Nevertheless, the ratio of items where domain agreement occurred to the number of items where EIs from the domain were assigned was similar for these four domains. In particular, the Geometry domain contains eight different EIs, and so one might expect there to be more opportunities for domain agreement than in the case of The Number System domain with only five EIs. However, this was not the case. It is unclear whether this is due to the small number of items that were assigned EIs from the Geometry domain or because the Geometry EIs were more distinct and thus easier to assign unambiguously.

An examination of the specific EIs that were assigned in domains where domain agreement occurred more than once could provide information about the distinctiveness, or lack of distinctiveness, among the different EIs. The four instances of domain agreement that occurred in The Number System (TNS) all occurred on a four-part task where one rater assigned TNS5 and the other assigned TNS Other. This likely shows that the item may not have been an exact fit to TNS5 but was at least related to The Number System domain, rather than suggesting a lack of distinctiveness among EIs in The Number System domain.

In the five instances of domain agreement in Expressions and Equations there was more variation in the items and in the EIs assigned. For instance, in three cases one rater assigned EAE1: "Viewing numerical and algebraic expressions as 'calculation recipes,' describing them in words, parsing them into their component parts, and interpreting the components in terms of a context" (CBMS, 2012, p. 42), and the other rater assigned EAE2: "Examining lines of reasoning used to solve equations and systems of equations" (CBMS, 2012, p. 42). The three items where EAE1 and EAE2 were assigned all required the test taker to write an equation or identify an equation that could be used to represent a particular mathematical situation. It makes sense that either of these two EIs could fit reasonably with such tasks, as both an understanding of the parts of the equation and the reasoning used to create and solve the equation are necessary.

Discussion

The results of this study demonstrate that it is both feasible and useful to conduct a post hoc examination of the validity argument for a mathematics assessment. Our expert raters had no prior affiliation or association with the TEDS-M mathematics knowledge assessment, and they were deeply familiar with the more recently published MET II recommendations regarding mathematical knowledge needed for teaching. They were able to complete the validation exercise without any reported difficulty.

Regarding the substance of their judgments concerning the mathematics content assessed by TEDS-M, they assigned MET II essential ideas to almost all of the TEDS-M items, and they did so with a generally high degree of agreement using the criterion of total agreement. This was noted for all the dimensions of variation identified by TEDS-M, including content domain, mental process, knowledge type, item format, and mathematical level. Thus, the overall findings offer a resoundingly affirmative response to the central validation question of interest: Do the TEDS-M items assess mathematical knowledge that is important for teachers to know as judged by contemporary standards?

Is total agreement a reasonable criterion to use? The extent of agreement would certainly be much lower if we used the criterion of identical agreement, but we think that would be an unreasonably restrictive criterion for the purposes of this study. Given that we are conducting a post hoc validation study, rather than a matching of test items to its own design framework or to a set of pre-specified objectives, we think that a less restrictive criterion is in order. The measure of total agreement reflects well the extent to which the raters judged TEDS-M items to be tied to essential ideas within the MET II content domains, and we take this to be a strong indicator of content validation for the purposes of this study.

Though the findings are generally quite supportive with regard to the validity of inferences from TEDS-M mathematics knowledge assessment items about the MET

II essential ideas, there were two findings that were at least somewhat discordant. One was the finding regarding the spotty coverage of MET II content domains, and the other was the finding regarding the standards for mathematical practice.

The raters' judgments revealed unevenness of the correspondence between TEDS-M items and MET II content domains. To the extent that one expects the TEDS-M assessment to be used to make valid inferences about teacher knowledge across the full spectrum of mathematics content knowledge for elementary school teachers, this finding raises some doubts. Raters judged very few items to be related to EIs in the domain of statistics and probability, and the raters judged no items as corresponding to EIs in the MET II domain of Number & Operations-Fractions. Because these two domains are foundational in elementary mathematics education, this finding suggests a need for further examination to determine whether it is attributable simply to omissions in the selection of publicly released items, or, instead, it reveals a limitation in the content coverage of the TEDS-M assessment.

The full set of TEDS-M items was not available to us to extend our analysis and probe this issue more deeply, but the TEDS-M project leaders provided us with descriptions of all the assessment items. The descriptions of several items (none of which were in the set of released items) suggested that they were likely to assess aspects of fraction knowledge. Because these items were not part of our study, we cannot be certain that our raters would have associated these items with MET II EIs related to fractions, so it is unclear how this information should affect interpretation of reported results from the TEDS-M mathematics assessment. Nevertheless, the evidence of potentially spotty coverage of MET II content domains in the TEDS-M assessment items should be useful to those who might wish to revise the assessment for future use.

Regarding the standards for mathematical practice, to which the MET II report refers but does not specify in detail, we found that the expert raters did map these onto TEDS-M items with great frequency. They associated SMPs with more than 80% of the TEDS-M items, but they did so with far less inter-rater agreement. They agreed in less than half of the cases where they assigned one or more SMPs to a TEDS-M item. On the one hand, it is notable that the expert raters assigned an SMP to more than 80% of the assessment items, especially because test items typically provide little opportunity for mathematical processes to emerge. On the other hand, it is equally noteworthy that they seldom agreed with each other about which SMP was associated with a particular item.

It is not possible for us to know from the data collected in this study why there was so much variation between the raters in assigning SMPs to items, but we speculate that this finding may have less to say about the validity of the TEDS-M assessment and more to say about a lack of clarity, even among experts in the field, about the precise meaning of the SMPs. Alternately, it is possible that raters made different judgments about the ways in which test takers would approach TEDS-M items, and this led them to expect distinct solution pathways for items that were in turn associated with different SMPs. More investigation of the possible ramifications of this finding is needed.

Coda

We began this chapter by arguing the potential value of *post hoc* examination of the validity argument for an educational assessment. We reported herein such an analysis of the mathematics knowledge assessment developed for and used in the TEDS-M study. In this chapter, we illustrated how a post hoc validation study might be conducted and how the findings might be interpreted in reasonable ways to provide valuable information both for those who wish to interpret the findings of TEDS-M in light of contemporary views of the mathematical knowledge needed by elementary teachers and for those who might wish to revise the TEDS-M knowledge assessment for future use.

Beyond the possible value of a study like this for purposes of post hoc validation, we would argue that this kind of cross-framework comparison could be broadly useful to the field. Given the proliferation of different conceptualizations and assessments of the knowledge needed to teach mathematics from professional organizations, state and national policy entities, and scholars of teacher education, a careful examination of the similarities and differences among different representations of this knowledge seems useful to inform both researchers and teacher educators.

A clear conceptualization of what teachers need to know to be successful teachers of mathematics would allow researchers to better study this knowledge and the ways in which it affects students. Understanding the similarities and differences among these different conceptualizations would give teacher educators more guidance in deciding what to teach pre-service and in-service teachers in order to positively impact students. The study reported here is a specific example of what we hope will be a larger-scale effort to accumulate and integrate across disparate views of the mathematical knowledge needed by elementary school teachers.

References

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Ananda, S. (2003). Rethinking issues of alignment: A re-examination. *Theory Into Practice, 41*(4), 255–260.

Banerjee, M., Capozzoli, M., McSweeney, L., & Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique, 27*(1), 3–23.

Conference Board of the Mathematical Sciences. (2012). *The mathematical education of teachers II*. Providence, RI /Washington, DC: American Mathematical Society/Mathematical Association of America.

Case, B. J., Jorgensen, M. A., & Zucker, S. (2004). *Alignment in educational assessment*. Retrieved from https://images.pearsonassessments.com/images/tmrs/tmrs_rg/AlignEdAss.pdf?WT.mc_id=TMRS_Alignment_in_Educational_Assessment

Guion, R. (1977). Content validity: The source of my discontent. *Applied Psychological Measurement, 1*, 1–10.

Kane, M. (2001). Current concerns in validity theory. *Journal of Educational Measurement, 38*(4), 319–342.

Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York, NY: American Council on Education and Macmillan.

Pellegrino, J., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Resnick, L. B., Rothman, R., Slattery, J. B., & Vranek, J. L. (2003). Benchmarking and alignment of standards and testing. *Educational Assessment, 9*(1&2), 1–27.

Tatto, M. T. (Ed.). (2013). *The teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries – Technical Report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in mathematics (TEDS-M): Policy, practice, and readiness to teach primary and secondary mathematics. Conceptual framework*. East Lansing, MI: Teacher Education and Development International Study Center, College of Education, Michigan State University.

Webb, N. L. (1997). Determining alignment of expectations and assessments in mathematics and science education. *NISE Brief*, *1*(2). Retrieved from http://files.eric.ed.gov.proxy.lib.umich.edu/fulltext/ED405190.pdf